Published in final edited form as:

Ecol Modell. 2012 March 24; 229: 64–75. doi:10.1016/j.ecolmodel.2011.08.020.

Space versus Place in Complex Human-Natural Systems: Spatial and Multi-level Models of Tropical Land Use and Cover Change (LUCC) in Guatemala

David López-Carr^a, Jason Davis^a, Marta Jankowska^b, Laura Grant^c, Anna Carla López-Carr^b, and Matthew Clark^d

^aDepartment of Geography, University of California, Santa Barbara, Human-Environment Dynamics Lab, 4836 Ellison Hall UC Santa Barbara (UCSB) Santa Barbara, CA 93106-4060

^bDepartment of Geography, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182

^cDepartment of Economics and Bren School of the Environment, UC Santa Barbara (UCSB) Santa Barbara, CA

^dDepartment of Geography and Global Studies, Sonoma State University, Stevenson Hall 3066, 1801 E. Cotati Avenue, Rohnert Park, CA 94928

Abstract

The relative role of space and place has long been debated in geography. Yet modeling efforts applied to coupled human-natural systems seemingly favor models assuming continuous spatial relationships. We examine the relative importance of placebased hierarchical versus spatial clustering influences in tropical land use/cover change (LUCC). Guatemala was chosen as our study site given its high rural population growth and deforestation in recent decades. We test predictors of 2009 forest cover and forest cover change from 2001-2009 across Guatemala's 331 municipalities and 22 departments using spatial and multi-level statistical models. Our results indicate the emergence of several socio-economic predictors of LUCC regardless of model choice. Hierarchical model results suggest that significant differences exist at the municipal and departmental levels but largely maintain the magnitude and direction of single-level model coefficient estimates. They are also intervention-relevant since policies tend to be applicable to distinct political units rather than to continuous space. Spatial models complement hierarchical approaches by indicating *where* and to what magnitude significant negative and positive clustering associations emerge. Appreciating the comparative advantages and limitations of spatial and nested models enhances a holistic approach to geographical analysis of tropical LUCC and human-environment interactions.

Keywords

land use/cover change (LUCC); human-environment dynamics; deforestation; coupled human-natural systems

© 2011 Elsevier B.V. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

1.1 Modeling space and place in coupled human-natural systems and land change research

A “coupled human-natural systems” (CHANS) approach builds on complexity theory to incorporate feedbacks, multiple equilibria, and unstable dynamics in coupled human-environment dynamics (Liu et al., 2007; Manson, 2003; Turner and Robbins, 2008). Land use/cover change (LUCC) is a central research initiative of CHANS research, with important implications to both human and natural systems (Turner et al., 2007). A large LUCC literature has investigated LUCC drivers (Lambin et al., 2003; Carr, 2004). Recent empirical studies have uncovered feedbacks within migration and LUCC (Carr, 2009), population change and LUCC (Liu et al., 2001), and socio-economic development, consumption, and LUCC (Rudel et al 2009). All of these prior studies, however, imperfectly capture the influence of space and place in modeling land use outcomes.

Spatial and hierarchical models may usefully complement contemporary human-environment research tools such as agent-based modeling (ABM). To investigate coupled natural-human linkages, researchers have increasingly turned to ABM (An et al., 2002) while hierarchical approaches remain scarce (Rindfuss et al., 2008; Verburg et al., 2004). While ABM can model cumulative effects of actor-environment interactions, it fails to explicitly integrate multiple spatial scales. Modeling spatial behavior at the individual level may fall victim to the atomistic fallacy, where the contextual surroundings of individual behaviors are ignored (Alker, 1969; Diez-Roux, 2002). A hierarchical approach, in contrast, offers the potential to examine structure as manifested in space versus agency by explicitly locating the agents of environmental change within larger meaningful geographical units and contexts (Chowdhury and Turner, 2006). A further strength of multi-level analysis is its ability to overcome inherent bias due to autocorrelation by separating out spatial or temporal similarity at various levels of analysis. However, a potential drawback of the hierarchical modeling approach is its dependence on a priori definitions of discrete spatial units (Fotheringham and Brunson, 1999). Lastly, a nested model allows us to distinguish one place from another rather than to model space continuously (Nelson, 2001). This is appropriate when units of interest, such as a village or a county, are naturally place based. In turn, since policy is made at village, county, state, and national levels – not continuously over space – policy implications of hierarchical models can be particularly applicable.

ABM and other CHANS tools may also be usefully accompanied by spatial regression approaches. The modifiable areal unit problem (MAUP) and associated ecological fallacy have been both a bane and a *raison d'être* for geographers for many years. MAUP is a potential source of variation associated with aggregating data into arbitrarily bound groups resulting in summary values and statistical properties of variables that are influenced by the boundaries themselves as much as by on-the-ground phenomena (Openshaw, 1984; Fotheringham and Wong, 1991; Greenland, 2002). Since entirely objective boundaries are impossible to derive, researchers are challenged to justify the most feasible and appropriate areal unit to analyze. At times, this is done for them, especially when data are available only for pre-defined units (e.g. county data). As a result, modeling of LUCC has been plagued by the natural occurrence of spatial or temporal autocorrelation—the fact that areas that occur closer in space or time to one another respond more similarly to a stimulus than areas that are further apart (Getis, 2010). Such dynamics violate the independence assumptions of many statistical analyses, particularly regression analysis (Geary, 1954). While multi-level models can reveal relationships of nested parameter measures, they cannot reveal a parameter's spatial distribution.

Relationships that are liquid over space, in other words exhibit spatial non-stationarity, will not be modeled well by global parameter estimates (Fotheringham and Brunsdon 1999), as found in ordinary least-squares (OLS) and hierarchical regression techniques. Measures of LUCC are often inherently spatially dependent, and necessitate consideration of how geography will influence model parameters (Overmars et al., 2003; Kupfer and Farris, 2007). Furthermore, incorporating parameter variability over space will lead to better understanding of the connections between geography and LUCC drivers. Geographically Weighted Regression (GWR), is a regression method that accounts for spatial non-stationarity by running a local regression at each spatial unit while incorporating weighted values of surrounding spatial neighbors (Brunsdon et al. 1998). While GWR has been applied in LUCC studies, its use in deforestation research, particularly in Latin America, is rare (Witmer, 2005; Pineda Jaimes et al., 2010).

This paper examines space versus place in a comparison of GWR (i.e., space) versus multi-level models (i.e., place) applied to tropical land change in Guatemala. Of Central America's seven nations, Guatemala (Figure 1) is the most populous (just under 14 million people) and supports the second highest population density (131 individuals/sq. km.) within a land area of 108,890 sq. km (World Bank 2011). Its 2009 GDP (purchasing power parity) was estimated at \$64.1 billion, ranking it 139th in the world. Guatemala is also a highly rural and impoverished nation. Over 60% of its inhabitants reside in rural communities and 74% of them are poor. Additionally, its annual population growth rate was 2.4%, its total fertility rate was 3.27 and its infant mortality rate was 40 per 1000 live births in 2009 (World Bank 2011).

The case study examines spatial and nested predictors of change in woody cover from 2001-2009 and absolute woody cover in 2009 for Guatemala's 331 municipalities nested within 22 departments. We first describe the importance of the study site for the examination of LUCC and the related data sets, followed by a discussion of multi-level and spatial analytical methods in coupled human-environment research. We then describe our data sources and analytical methods before reporting results for predicting woody cover in Guatemala using OLS, multi-level, and GWR approaches and compare the way these procedures model data and the estimates they generate. Following the discussion of similarities and contrasts between nested and spatial approaches, we elaborate potential implications for enhanced modeling of tropical LUCC and the significance for the theoretical, methodological, and political dimensions of human-environment dynamics.

1.2 Tropical LUCC in Latin America and Guatemala

According to Houghton et al. (1991), between 1850 and 1985, 28% (370×10^6 ha) of Latin America's forests were replaced by some form of agricultural operation, or fragmented to provide building and fuel supplies for nearby settlements. Forest conversion to pasture represents 44% of this change while conversion to cropland, fragmented land, and fallow each represents 25, 20 and 10% respectively. The rate at which Latin American forests conversion took place remained low and stable prior to 1940, but rapidly accelerated thereafter (Houghton et al., 1991). In the 1990s, forest conversion for agriculture and settlement construction continued to alter large tracts of forested land. Two sets of authors, Achard et al. (2002) and Mayaux et al. (2005), report Latin American deforestation rates of 2.2 and 2.5×10^6 ha per year, respectively, during the 1990s. These rates equate to a 0.33 and 0.38% loss of forest cover per year.

Deforestation trends in Central America conform to those seen throughout Latin America. Following human settlement, Central America has suffered an estimated 82% loss of total forest cover and an 89% loss in primary forest cover (Myers, 1991). In the 1990s, Achard et al. (2002) and Barbier (1997) reported continued deforestation throughout Central America

with rates ranging from 0.8 to 1.5%. This is down from the 1.8% deforestation rate (3,000km²/year) for the region reported by Myers (1991) for the late 1980s.

While much of Central American forests continues to be exploited for timber, fuelwood, and agricultural resources, some are experiencing small but significant re-growth. For example, the mountainous region of La Champa, Honduras experienced a shift from slash and burn agriculture to more sedentary and intensive farming practices from 1987 to 1996, resulting in a net increase in reforestation as marginal lands were abandoned (Southworth and Tucker, 2001; Monroe et al., 2004). Additional examples can be found in Panama. Wright and Samaniego (2008) reported a 0.36%/year increase in total forest cover in Panama, but a 1.3/ year loss of primary forests for the entire nation between 1992 and 2000. Kleinn et al. (2002) described a declining rate of deforestation in Costa Rica between the 1940s and the 1980s that eventually flattens out and starts to curve upward in the early 1990s).

Within Guatemala, rapid population growth and shifting agricultural practices have inscribed visible impacts on its biologically diverse landscape. According to Loening and Markussen (2003), Guatemala's total forest cover shrank from 65% to 26% during the second half of the twentieth century (1950-2000). Deforestation rates were especially high in the 1990s (1.7% per annum) (Carr et al, 2009, Brooks *et al.*, 2002), where most deforestation can be directly attributed to rural-frontier migration by cattle ranchers (Colchester, 1991) and subsistence farmers (Carr, 2004; 2005) who often convert their land to cattle ranches following a few years of row cropping (Sader *et al.*, 1997). The extreme concentration of landholdings and underemployment, combined with the country's very high fertility rate has led to a fragmentation of Guatemalan farm plots and to rural poverty, thus stimulating rural out-migration to cities, international destinations, and the forested frontier (Bilsborrow and Stupp, 1997; Cincotta *et al.*, 2000; UN, 2003).

A vast majority of Guatemala's current deforestation is occurring within the frontier regions of the north, in the Department of Petén. This deforestation is mainly undertaken by economically and demographically displaced subsistence farmers migrating from southeastern and southwestern rural areas of Guatemala who desire to own farmland (Carr, 2008). Some have argued that providing land title to the “landless” who reside in frontier regions like the Petén would attenuate deforestation by incentivizing stability over the more nomadic practice of swidden agriculture (Alston *et al.*, 1995; Pichon, 1997; Katz, 2000). Others, relying on empirical evidence, make the counter argument that providing land title improves access to credit and ultimately leads to higher rates of forest conversion to agriculture (Feris, 1999; Carr, 2004b). Gould (2005) addresses these arguments for Northwestern Petén by concluding that given the many imperfections of regularizing land tenure in frontier regions—not opening up credit markets or increasing market efficiencies—that providing a process for obtaining tenure to one's *de facto* land is not a panacea for solving the many socio-economic issues that plague landless farmers there. By extension, one can also conclude that Gould would argue the regularization of the land tenure process would not have a substantial effect on deforestation in the Petén either.

2. Methods

2.1 Data

Data used for this analysis come from three sources: (1) 2000 Guatemalan Living Standards Measurement Survey, (2) 2003 Guatemala National Agriculture Census, and (3) 2001-2009 Forest Cover Change database for all municipalities in Latin America. The three data sources provide detailed information regarding land use and management at the municipality and department levels. A description of each dataset follows:

2000 Guatemala Living Standards and Measurement Survey—This rich source of household and community data covers 20,969 (11,302 rural) households and over 100,000 individuals. The surveys were implemented by the Guatemalan National Statistics Institute (INE), with technical guidance from the World Bank Living Standard Measurement Surveys (LSMS) research team. The nationally representative survey includes demographic, economic, and community characteristics modules. For our project, we aggregate mean household fertilizer use and tractor ownership to the municipality and department levels (Table 1).

2003 Agriculture Census—Due to political conflict that ravaged Guatemala during its multi-decade civil war, national-level agricultural and land use data had not been gathered since 1979. Therefore, the 2003 Agricultural Census represents the first large-scale collection of agricultural land use data in 24 years. These data were collected by the INE and contain information from 822,188 farmers—roughly 90% of all producers within the country. Of the many subject areas covered by the agricultural census, our project used information gathered on 2003 crop yields for four of Guatemala's main commodities (coffee, sugar, white corn, yellow corn) and for fallowed land (Table 1). These were also aggregated to the municipality and department levels.

Forest Cover Change 2001-2009—The dependent variables are woody forest cover and change in woody forest cover. A NSF-funded project produced LUCC estimates for all of Latin America for the years 2001 and 2009, based on methods developed in Clark et al., (2010) for the Dry Chaco ecoregion of South America and expanded to include Latin America using methods in Clark and Aide (2011a and b). Eight land cover types were mapped at a 250-m pixel resolution using the MODIS MOD13 Q1 (16-day L3 Global 250 m) Vegetation Indices product (Huete et al., 2002), using statistics of mean, standard deviation, minimum, maximum and range for 4-, 6-, and 12-month temporal windows within a calendar year from the red, near-infrared, and mid-infrared reflectance and Enhanced Vegetation Index (EVI) bands. Classes included agriculture, herbaceous vegetation, built-up areas, bare areas, water, plantations, woody vegetation, all defined as having ≥80% cover in the respective class, as well as mixed woody vegetation with <80% woody cover with bare, herbaceous or agriculture (Clark et al., 2010; Clark and Aide, 2011b). Maps were produced for eastern moist forests and western conifer and dry forests separately, with boundaries defined using the majority cover of biomes (Olson et al., 2001) within municipalities. Reference data for classifier training and accuracy were collected across Central America with visual interpretation of Google Earth high resolution imagery within a custom web-based tool (Clark and Aide, 2011b). There were 2,082 and 1,422 reference samples for the moist and conifer/dry forests, spanning years 2001 to 2009. These reference samples were connected to MODIS predictor variables (statistics) by year and used in a Random Forests classification routine (Clark and Aide, 2011a). Based on 1/3 cross-validation of reference samples within Random Forests (“out of bag” samples), overall map accuracies for moist and conifer/dry forest maps with eight classes were 77.6% and 81.9%, respectively. However, with a post-classification reclassification of agriculture and herbaceous vegetation, plantations and mixed woody, and bare and urban areas, our resulting five-class maps had overall accuracies of 85.4% and 85.8% for moist and conifer/dry forest maps, respectively (Table 2).

We focused our analyses on the more conservative measure of forest in our maps, the woody vegetation class (≥80% cover). Using Random Forests cross-validation, woody vegetation class had 99.8% and 70.6% producer (due to omission errors) and 93.6% and 80.3% user (due to commission errors) accuracies for moist and conifer/dry forests, respectively (Table 2). It should be noted that with 250-m pixels as the mapping unit, fine-scale land use, such as small farms and subsistence land clearing (Sader et al 1997), and subtle vegetation forest

regeneration, may not be detected. Although not tested here, we would expect that fine-scale clearing within a woody vegetation pixel through time would transition the pixel to mixed-woody vegetation (20 to 80% woody cover) and then to agriculture or herbaceous vegetation (<20% woody cover). Likewise, we would expect woody encroachment of an abandoned pasture or agricultural field to transition to mixed woody and then to woody vegetation classes through time. However, as with all discrete classifications, it is unrealistic to expect all pixels to respect the hard thresholds between our defined classes, especially for classes that are spectrally mixed, such as mixed woody and plantations. For the conifer/dry forest map, there was a tendency for mixed woody and plantation pixels to be confused with woody vegetation (Table 2), and so there should be more caution in interpreting results in this region of Guatemala. However, to minimize the effects of pixel-scale error, we analyzed the area of woody vegetation from 2001 and 2009 at the municipality scale. The area of woody vegetation was summarized for pixels within each of Guatemala's 331 municipalities (i.e., a zonal sum) and used to calculate our two dependent variables: (1) percent woody cover in 2009 and (2) change in woody cover from 2001 to 2009 calculated as the ratio of woody cover in 2009/woody cover in 2001. Ratio values were converted to percent change for all model interpretations. Furthermore, we used the project's 1990 and 2000 population census and density information for each of Guatemala's 331 municipalities (Table 1).

Our independent variables presented in Table 1 were selected based on our prior work in Guatemala (Carr, 2008), in Latin America more generally (Carr et al., 2009; Aide and Grau, 2004), and following the literature on proximate and underlying causes of tropical deforestation (Geist and Lambin, 2002). Two sets of variables are tested; demographic and technological-economic. We recognize additional household, municipal, and departmental level socio-economic, political and environmental variables as potentially influential as well but we have insufficient data richness to pursue all of these potential LUCC drivers. In prior models we tested a host of demographic and socio-economic variables that were insignificant in bivariate regressions and trimmed from our final models. Under demographic factors, population density from 1990-2000 and its squared form are examined to probe potential linear and non-linear relations between demographic change in the decade prior to the land cover analysis. Population density in 2000 is also examined in relation to LUCC outcomes. The technological-economic variables tested are directly related to land use intensity in the cases of household fertilizer and tractor use, and to subsequent output in the cases of corn production. The converse of intensification is also tested as measured by the percentage of land in fallow.

2.2 Multi-level Statistical Analysis

For this investigation, we employ 2-level random intercept models to separate out inherent spatial error in land cover at the municipal level from that found at the department level. Instead of assuming that the regression line for land cover change at the department level passes through the same intercept, a random intercept formulation allows this higher level variable to conform to different regression intercepts in order to more accurately model the outcome of interest and to overcome some issues of spatial dependence (Vance and Iovanna, 2006). To estimate the models, the event categories are treated as a multivariate response vector using dummy variables with no variation at level 1, rather, with a level 1 covariance specified at level 2. In addition to analyzing variance across nested scales, i.e., the municipality and the department scales, we also explore an independent dummy variable in the models to test potential significant differences between experimental versus control regions (i.e., 0=experimental municipalities/departments, 1=control municipalities/departments). Adequate estimation of cross-level interactions and tests of random effects has generally been found by simulation studies to require 30 groups with 30 observations, 60 groups with 25 observations, or 150 groups with 5 observations (Kreft and de Leeuw, 1998).

Our sample generally meets these criteria (Van der Leeden et al., 1997). The two-level random-intercept linear regression model is written as follows:

$$y_{ij} = \beta_1 + \beta_2 x_{2ij} + \dots + \beta_p x_{pij} + \zeta_j + \varepsilon_{ij}$$

For model 1, y_{ij} is the percentage of land identified by satellite imagery as woody vegetation in 2009, where i represents municipalities within j^{th} departments. For model 2, y_{ij} is the ratio of municipal woody vegetation from year 2001 to 2009 (natural log transformed to normalize the long-tailed distribution of these data), where i represents municipalities within j^{th} departments. β_1 is the intercept along with its independent error term ζ_j while β_2 through β_p are regression coefficients with corresponding explanatory variables x_{2ij} through x_{pij} with their independent error term ε_{ij} .

We employ municipalities nested within departments as the two multi-level categories and use four specifications of heterogeneity among the groups: ordinary least squares (not multilevel), maximum-likelihood estimator, least squares random effects, and least squares fixed effects. The distinction between maximum-likelihood and generalized least squares is important: the former is more flexible when the underlying intercept distribution remains unknown. There are also important differences between random and fixed effects models. In the random effects models, each intercept is modeled as random deviations from a common mean intercept. Therefore, estimating random effects requires many groups, so that the draws from the hypothetical distribution of intercepts can identify the parameters. In these circumstances, the estimates are more efficient than fixed effects. Another property of the random effects model is the ability to estimate explanatory covariates that do not vary over time by group, such as a region's size. The most important factor when using random effects is the relationship between the categories/groups and the other explanatory variables. First, the model requires numerous covariates; otherwise we have concern that omitted variables are biasing the remaining coefficients. Second, group effects need to be uncorrelated with the other independent variables for appropriate estimation of random effects. In other words, this model assumes the observations pertain to the entire population regardless of the group characteristics.

In fixed effects, we estimate a parameter for each group's intercept as an average deviation from the remaining groups. Because explanatory variables that do not vary across time in each unit will be perfectly collinear with the fixed effects we cannot estimate their effects. While estimation removes all cross-section variation in the dependent and independent variables, the advantage is controlling for unavailable, omitted variables. Furthermore, if group effects by department are correlated with the covariates, then the fixed effects model is more consistent than the random effects model. By employing a Hausman test, we ensure that random effects models are not mis-specified compared with the fixed effects model.

2.3 Spatial Analysis and Modeling

GWR is employed to examine spatial distributions of parameter estimates for the two models detailed in section 2.1. As commonly available software, with easily replicable results, we utilize ArcMap 10.0 for the GWR model (documentation, parameters options, and code sample for the GWR tool can be found at <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//005p00000021000000.htm>). The ArcMap GWR tool results in global and local R^2 values, local parameter estimates, standard errors, and standardized residual values. However, ArcMap GWR has drawbacks including a black-box approach. We attempt to spell out the process and parameters used below to increase the transparency of the tool. However the tool does not give certain calculations such as bandwidth range.

Additionally the tool does not provide global parameter estimates or statistical significance measures. While available from the GWR software created by Fotheringham, Charlton, and Brunsdon, the problem of global p-values for parameter estimates is discussed in the corresponding software white paper (Charlton et al., 2005). For exploring spatial distributions, this study is interested in local variation in parameter estimates rather than in global estimates.

The two dependent variables, woody cover in 2009 and change in woody cover from 2001 to 2009, were assessed for global clustering using Moran's I with inverse distance weighting and Euclidean distance calculation method. Local cluster was examined using the Getis-Ord G_i^* statistic for local spatial autocorrelation hot and cold spot analysis (Getis and Ord, 1992). To assess spatial stationarity of variables utilized in the multi-level model, two OLS regression models predicting the dependent variables were run at the municipality spatial scale. Resulting standardized residuals were measured for clustering using Moran's I (Cliff and Ord, 1972).

To account for residual spatial non-stationarity in the multi-level model, GWR was performed. GWR, a variation on the basic linear regression model, accounts for variations in space throughout the parameter surfaces. The GWR for this study is stated as:

$$y_i = \beta_{i0} + \sum_{k=1}^n \beta_{ik} x_{ik} + \varepsilon_i$$

where β_{i0} is the intercept at location i with latitude and longitude coordinates that correspond to the centroid of each municipality and department, n is the number of model independent variables, and β_{ik} is value of β_k at point i (Fotheringham et al., 2000). For this section y_i is the percentage of land identified by satellite imagery as woody cover in 2009 in model 1, and the change in woody cover from 2001 to 2009 in model 2. β_k is n regression parameter estimates displayed in Table 1 with corresponding explanatory variables x_k , with independent error term ε at location i . It is important to note that parameters are estimated with a weighting function based on distance, resulting in locations closer to the estimated point having more influence on the projected value than locations farther away. Using different weighting functions can have some impact on the model, however this impact is often small (Brunsdon et al., 1998). A matrix form parameter estimate equation is shown (Cahill and Mulligan, 2007):

$$\hat{b}_{ik} = (X^T W_i X)^{-1} X^T W_i y$$

where $\hat{b}_{ik}$ is the estimate of the location-specific parameter b_k , and W_i is a spatial weights matrix.

ArcGIS 10.0 allows for the user to provide a weights matrix, or creates one for the regression model. Due to the varying size and density of municipalities throughout Guatemala, the adaptive kernel option in the GWR Tool was used (the tool uses only Gaussian kernels). However, Fotheringham et al., (2000) note the minor impact of kernel calculation method on the overall model, which allows each point's spatial context to vary by feature density. Therefore, the weighting matrix varies by k . An AIC bandwidth was applied for both models, which allows the GWR Tool to find the optimal distance for the varying kernel throughout the calculation (a discussion of appropriate bandwidth and kernel

choice can be found in Fotheringham et al., (2000)). The ArcGIS GWR tool performs a local multicollinearity diagnostic, and outputs a 'condition number' based on VIF. Condition numbers higher than 30 indicate local multicollinearity and the necessity to adjust the model. Based on condition numbers the percentage of population density change from 1990 to 2000 variable was dropped due to high local multicollinearity within the GWR models only.

While GWR has significant advantages for addressing common violated assumptions of global regression models, there are limitations such as spatial error dependence (Fotheringham et al., 2002), local multicollinearity (Wheeler and Tiefelsdorf, 2005), and extreme coefficients (Cho et al., 2009). These limitations are minimized in our model as there are no large areas with sparse data points. We tested for local multicollinearity in the software tool, and we chose an adaptive rather than fixed kernel and bandwidth.

3. Results and Discussion

3.1 Multi-level/Grouped Models

The model results for percent woody cover in 2009 (model 1) are given in Table 3. The number of groups is 22, which is a relatively small amount for determining random effects parameters. The number of observations per group varies from 5 to 27, allowing sufficient averaging for the fixed effects estimator. As for time-invariant covariates, we do not have data available over time with sufficient variation in the independent variables to consider estimating invariant factors. Finally, the covariates that are available are likely correlated within groups; for example, population density is related to region. These aspects point to using fixed effects.

Multilevel formulations are significantly more descriptive than single-level formulations as shown by the highly significant Likelihood Ratio Test for the MLE model and F-Test of Model Significance for the fixed effects model. However, the highly significant Hausman Test suggests the random effects models are mis-specified compared with the fixed effects model. Therefore, further discussion of model 1 results will pertain to the fixed-effects multilevel model only.

Multiple coefficients are significant but exhibit values below 0.00001—below the level of practical significance (2000 population density, café production, sugar production, and white corn production). Percentage of land in fallow was highly significant for model 1 with a coefficient of -0.479, meaning for every percentage increase in fallow land at the municipality level, there was a 0.48% decline in woody cover in 2009. Percentage of households owning a tractor was marginally significant with a coefficient of -0.242, meaning for every percentage increase in households that own a tractor at the municipality level, there was 0.24% less woody cover in 2009.

The results of change in woody cover between 2001 and 2009 (model 2) are shown in Table 4. As in model 1, a multilevel formulation was significantly more descriptive than a single-level model. Furthermore, the Hausman Test again indicates mis-specified random effects models; therefore, subsequent interpretations of model 2 pertain to the fixed effects multilevel model only.

As in model 1, 2000 population density and café production remain significant but below the level of practicality with coefficients below 0.00001. Also consistent with model 1 findings, percentage of households that own a tractor remains significant and negatively correlated with the change in woody cover from 2001 to 2009. However, unlike 2009 woody cover, for woody cover change we observe a significant and positive relationship between both the

percentage of households that use fertilizer and percentage of the municipality in fallow. Specifically, every one percent rise in the number of households that use fertilizer was correlated with a 0.07% increase in woody cover between 2001 and 2009. Furthermore, each percent rise in fallow land at the municipality level was correlated with a 2.76% increase in woody cover between 2001 and 2009. A likely explanation for such a large corresponding increase in woody cover with a rise in fallowed land is that the “fallow” category captures abandoned land that is reverting back to forest as argued under Forest Transition Theory (Mather and Needle 1998).

3.2 Spatial Analysis and Modeling

Both the global and localized clustering measures indicate significant spatial dependence for the two variables of interest. Figure 2 shows percent woody cover in 2009 and change in woody cover between 2001 and 2009. Table 5 displays the Moran's I and Z scores for the cover variables, with larger significant Z scores indicating greater clustering. Both dependent variables show statistically significant clustering at the 99% confidence interval. Percent woody cover in 2009 has more global clustering with a Z score of 12.59 as opposed to the change in woody cover with a Z score of 3.78. Standardized residuals from OLS regression for model 1 and model 2 were assessed using Moran's I (Table 5). Both model residuals have significant clustering, indicating spatial stationarity, and potential utility for running GWR.

Results from the GWR showed improvement in global adjusted R^2 values as compared to the OLS results, from 0.251 to 0.471 for model 1, and from 0.016 to 0.024 for model 2. By accounting for spatial non-stationarity, the GWR model outperformed both OLS models, although better results were seen for model 1 than model 2 likely due to the higher spatial clustering of OLS residuals from model 1 that were accounted for by the GWR. Standardized residuals from GWR for model 1 and model 2 were measured for spatial autocorrelation using Moran's I, and results confirm the lack of clustering (Table 5).

A global pattern of woody cover clustering for 2009 is confirmed when examining local clustering using G_i^* (Figure 3) hot and cold spot analysis. Hot spots are spatial units of high values surrounded by other units of high values (with cold spots the reverse), which are statistically significant (measured by Z scores) when compared to a random distribution. Percent woody cover in 2009 exhibits three hot spot regions with high forest cover for diverse reasons. The northeast portion of the country forms part of the Maya Biosphere Reserve and is devoid of roads and settlements; the area in the northwest was severely affected by civil war until the late 1990s and experienced some forest regrowth as a result of out-migration and farm abandonment. The central hotspot is covered by steep volcanoes and has been a priority conservation area by The Nature Conservancy during recent years. There is also a significant cold spot in the southeastern area of the country. Change in woody cover from 2001 to 2009 shows one hot spot in the central area of the western region of the country, indicating an area that experienced significantly clustered increase in woody cover from 2001 to 2009 in association with increasing migration remittance receipts and expanding shade-grown coffee.

Local GWR R^2 values for models 1 and 2 are displayed in Figure 4. In model 1, values range from 0 in the western-central area of the country where model variables did not predict any woody cover in 2009, to 0.70 in the south-eastern area of the country, where it can be concluded that the variables selected for the model were best able to predict percent of woody cover in 2009. Local R^2 values were equal to 0 in a much larger portion of the country in model 2, indicating that the selected variables are not good explanatory variables for areas where re-forestation is occurring. Comparatively good model fit was found in the

northern part of the country, where deforestation is observed, although the highest it achieved was an R^2 value of 0.15.

Figure 5 presents coefficient estimates from the GWR model 1 of percent population density change from 1990 to 2000 on the left and percent of land in fallow on the right. Population density change has a negative impact in the northern parts of Guatemala. This result logically follows results on intensification. In southern areas increase in population density is likely related to urbanization and thus some land abandonment, whereas in the north the population is overwhelmingly rural and thus increasing population density here adds to the number of farm workers and subsistence consumers, leading to less woody cover. Percent of fallow land has a negative impact on percent of woody cover in 2009 country-wide; however, this effect grows moving east through the country. Eastward and northward the countryside is more rural; the practice of semi-subsistence agriculture and the presence of more forest cover which can be converted to fallow cause a negative relationship between percent fallow land and forest cover. Conversely, the more densely populated and urban areas of the southwest are largely deforested, and so fallow land is more likely to be associated with regrowth.

Figure 6 illustrates the parameter estimates for percentage of households using fertilizer for both models. Models 1 and 2 show similar spatial patterns of coefficient estimates increasing moving northward through the country. However in model 1 the percent of households using fertilizer has an increasingly negative impact moving north, and in model 2 the percent of households using fertilizer has an increasingly positive impact moving north. The southern portion of the country is intensively managed in export crops such as sugar cane with high fertilizer inputs but low forest cover at the initial period and low change in forest cover during the period. Conversely, moving northward, agriculture is generally more small scale, characterized by semi-subsistence maize cultivation with relatively fewer inputs of fertilizers. Within this context, the more wooded remote northern regions are more likely to have recently cleared forest or secondary growth in order to intensify and extensify production simultaneously.

Figure 7 displays mapped coefficient estimates for percent of households owning a tractor. Coefficient estimates for model 2 are higher, and follow a similar spatial pattern to the percent of households using fertilizer coefficient estimates, with more negative values northward through the country. The pattern follows the simultaneous extensification and intensification relationship observed with fertilizer inputs described above. Model 1 provides an interesting spatial view of the households owning tractors variable. Here, coefficient estimates have lower values both in the north and southwestern parts of the country. However, there is a patch of *positive* coefficient estimates in the central and northwestern areas demarcated in a cyan-green color. In these areas, the little forest that does remain may be preciously coveted for fuelwood. Thus, agriculture intensification with tractors may help increase production and relieve pressure on forests.

4. Conclusion

This paper examines GWR versus multi-level models as applied to tropical LUCC and considers theoretical and practical implications to the two approaches. Together, these models allow us to draw a more complete picture of land cover change patterns in Guatemala and to elaborate more comprehensive implications for theory, empiricism, and policy. While the data sources remain the same, each model provides comparative advantages, highlighting both possibilities and limitations when determining an appropriate statistical approach.

OLS regression suggested that agricultural intensification in the form of fertilizers and tractors and higher levels of fallowed land are negatively associated with forest cover in 2009 while increased population density is positively associated with higher forest cover. Examining forest change during the first decade of the 2000s, we observe that areas that increasingly rely upon mechanized equipment—which is relatively rare—and/or fertilizers have more thoroughly captured and put into agricultural production available agricultural land—especially white and yellow corn. By taking into account spatial heterogeneity at the municipal and department levels, the multilevel model results modify our OLS findings to allow the analysis to better comply with the spatial independence assumptions of OLS regression. Thus, the fixed effects multilevel model results differ from the OLS only by finding negative relationships between both tractor ownership and percent fallowed land with 2009 woody cover. However, in addition to previously mentioned OLS relationships with tractor ownership and fertilizer use, multilevel model results for change in woody cover from 2001 to 2009 capture a positive trend between percentage of land in fallow with woody cover change.

Multi-level effects and spatial clustering were significant. Results from the two approaches corroborated each other synergistically yet also revealed important differences among nested versus spatial predictor and outcome variables. Our use of the hierarchical model separated the effects of municipal versus departmental variables, allowing us to distinguish the scale at which variables were most influential. The GWR indicated *where* these associations of changes are most salient. The GWR model controlled for spatial autocorrelation in the geographic distribution of woody cover and woody cover change across the Guatemalan landscape. It indicated that estimated coefficients vary over space, and it visually displayed where these coefficients positively influence the outcome in some areas, while negatively influencing the outcome in others. A clear trend emerges: The southwest to northeast gradient of decreasing population density, higher but decreasing forest cover, and lower but increasing technological inputs are particularly illuminated by the GWR.

The use of both types of models illustrates the importance of considering both space and place when crafting policies for LUCC. For example, fertilizer use leads to higher LUCC at both municipality and department levels. However, spatial analysis shows differences between rural and mixed urban-rural departments. Therefore, in rural departments a uniform fertilizer policy may be enacted, while in mixed urban-rural departments a municipality level policy approach might work better.

The debate in geography and cognate sciences over the importance of space vs. place is often investigated with quantitative or qualitative research, respectively. It need not be so. Place can be quantified and measured. Doing so reveals information that is lost with the assumption of continuous space in spatial regression analysis (Nelson, 2001). Nested or hierarchical effects are often more descriptive of the phenomena of interest. Importantly, they are also more policy-applicable. Policy is not formulated or applied in continuous space but within a nested structure of political units from neighborhoods, to towns, counties, states, and nations. No one model is correct. Rather, appreciating the comparative advantages and limitations of each enhances a holistic approach to geographical analysis of tropical LUCC and human-environment interactions.

Great care should be taken in constructing any model, whether OLS, MLE, random, fixed, multi-level, or spatial in order to ensure the most parsimonious and sensible final model and interpretation of model coefficients. The results of this study indicate that future research would benefit by considering spatial versus nested effects and combining both for a more holistic measurement and interpretation of complex coupled human-natural systems.

Acknowledgments

The authors would like to acknowledge the following sources of funding for the development of this research: University of California Pacific Rim Fellowship, National Science Foundation Dynamics of Coupled Natural and Human Systems (CNH) (EF 0709627, 0709645 and 0709598), National Institutes of Health Career Development Award, K01 (HD049008), National Science Foundation Geography and Regional Science grant (BCS-0525592), and a UCSB Faculty Career Development Award.

References

- Achard F, Eva HD, Stibig H, Mayaux P, Gallego J, Richards T, Malingreau J. Determination of deforestation rates of the world's humid tropical forests. *Science*. 2002; 297:999–1002. [PubMed: 12169731]
- Aide TM, Grau HR. Globalization, migration, and Latin American ecosystems. *Science*. 2004; 305:1915–16. [PubMed: 15448256]
- Alston LJ, Libecap GD, Schneider R. Property rights and the preconditions for markets: the case of the Amazonian frontier. *Journal of Institutional and Theoretical Economics*. 1995; 151(1):89–107.
- Alker, HS. A typology of ecological fallacies. In: Dogan, M.; Rokkan, S., editors. *Quantitative Ecological analysis*. Boston: MIT Press; 1969. p. 64–86.
- An L, Lupi F, Liu J, Linderman MA, Huang J. Modeling the choice to switch from fuelwood to electricity Implications for giant panda habitat conservation. *Ecological Economics*. 2002; 42(3): 445–57.
- Barbier E. The economic determinants of land degradation in developing countries. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. 1997; 352(1356):891–9.
- Bilsborrow, RE.; Stupp, P. Demographic processes, land, and environment in Guatemala. In: Bixby, LR.; Pebley, A.; Mendez, AB., editors. *Demographic diversity and change in the Central American Isthmus*. Santa Monica, CA: RAND; 1997. p. 581–624.
- Brooks TM, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Rylands AB, Konstant W, Flick P, Pilgrim J, Oldfield S, Magin G, Hilton-Taylor C. Habitat loss and extinction in the hotspots of biodiversity. *Conservation Biology*. 2002; 16(4):909–23.
- Brunsdon C, Fotheringham S, Charlton M. Geographically weighted regression-modelling spatial non-stationarity. *The Statistician*. 1998; 47:431–43.
- Cahill M, Gordon M. Using geographically weighted regression to explore local crime patterns. *Social Science Computer Review*. 2007; 25(2):174–93.
- Carr DL. Proximate population factors and deforestation in tropical agricultural frontiers. *Population and Environment*. 2004a; 25(6):585–612. [PubMed: 19672475]
- Carr DL. Ladino and Q'eqchi' Maya land use and land clearing in the Sierra de Lacandon National Park, Petén, Guatemala. *Agriculture and Human Values*. 2004b; 21:171–9.
- Carr DL. Population, land use, and deforestation in the Sierra de Lacandón National Park, Petén, Guatemala. *The Professional Geographer*. 2005; 57(2):157–68.
- Carr DL. Migration to the Maya Biosphere Reserve, Guatemala Why place matters. *Human Organization*. 2008; 67(1):37–48. [PubMed: 19657470]
- Carr DL. Population, rural development, and land use among settler households. *Journal of Global and International Studies*. 2009; 1(1):51–69.
- Carr, DL.; Pan, W.; Bilsborrow, RE.; Suchindran, C.; Whitmore, T.; Barbieri, AB.; Murtinho, F. Un análisis de multi-nivel de población y deforestación en El Parque Nacional Sierra de Lacandón. Petén, Guatemala: Documents d'Anàlisi Geogràfica; 2008. p. 49–67.
- Carr DL, Lopez CA, Bilsborrow RE. The population, agriculture, and environment nexus in Latin America: country-level evidence from the latter half of the 20th century. *Population and Environment*. 2009; 30:222–46.
- Clark, ML.; Aide, TM. An analysis of decadal land change in Latin America and the Caribbean mapped from 250-m MODIS data; 34th International Symposium on Remote Sensing of Environment; Sydney, Australia. April 10–15; 2011a.

- Clark ML, Aide TM. Virtual interpretation of Earth Web-Interface Tool (VIEW-IT) for collecting land-use/land-cover reference data. *Remote Sensing*. 2011b; 3:601–620.
- Clark ML, Aide TM, Grau HR, Riner G. A scalable approach to mapping annual land-cover at 250 m using MODIS time-series data: a case study in the Dry Chaco ecoregion of South America. *Remote Sens of Env*. 2010; 114(11):2816–32.
- Charlton, M.; Fotheringham, S.; Brundson, C. Geographically Weighted Regression. *NCRM Methods Review Papers*. 2005. <http://ncg.nuim.ie/ncg/GWR/index.htm>
- Cho S, Lambert DM, Kim SG, Jung S. Extreme coefficients in geographically weighted regression and their effects on mapping. *GIScience and Remote Sensing*. 2009; 46(3):273–288.
- Chowdhury RR, Turner BL II. Reconciling agency and structure in empirical analysis: smallholder land use in the Southern Yucatán, Mexico. *Annals of the Association of American Geographers*. 2006; 96(2):302–22.
- Cincotta RP, Wisniewski J, Engelman R. Human population in the biodiversity hotspots. *Nature*. 2000; 404(6781):990–1. [PubMed: 10801126]
- Cliff A, Ord K. Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*. 1972; 4:267–84.
- Colchester M. Guatemala: the clamour for land and the fate of the forests. *The Ecologist*. 1991; 21(4): 177–85.
- Diez-Roux AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *Am J Public Health*. 1998; 88:216–22. [PubMed: 9491010]
- Faris, R. [Accessed on June 7, 2011] Deforestation and land use on the evolving frontier: an empirical assessment. Harvard Institute for International Development. 1999. <http://www.cid.harvard.edu/hiid/678.pdf>
- Fotheringham AS, Brunsdon C. Local forms of spatial analysis. *Geographical Analysis*. 1999; 31:340–58.
- Fotheringham, AS.; Brundson, C.; Charlton, ME. *Quantitative Geography*. London: Sage; 2000.
- Fotheringham, AS.; Brundson, C.; Charlton, ME. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. West Sussex: John Wiley & Sons Ltd; 2002.
- Fotheringham AS, Wong DW. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*. 1991; 23:1025–44.
- Getis, A. Spatial autocorrelation. In: Fischer, MM.; Getis, A., editors. *Handbook of applied spatial analysis: Software tools, methods and applications*. New York: Springer; 2010. p. 255-75.
- Geary RC. The contiguity ratio and statistical mapping. *The Incorporated Statistician*. 1954; 5(3):115–45.
- Geist HJ, Lambin EE. Proximate causes and underlying driving forces of tropical deforestation. *BioScience*. 2002; 52(2):143–50.
- Getis A, Ord JK. The analysis of spatial association by use of distance statistics. *Geographical Analysis*. 1992; 24:189–206.
- Goldstein, H. *Multi-level statistical models*. 4th. West Sussex UK: Wiley; 2011.
- Gould KA. Land regularization on agricultural frontiers: the case of Northwestern Petén, Guatemala. *Land Use Policy*. 2005; 23(4):395–407.
- Greenland S. A review of multi-level theory for ecologic analyses. *Statistics in Medicine*. 2002; 21:389–95. [PubMed: 11813225]
- Hartzel J, Agresti A, Caffo B. Multinomial logit random effects models. *Statistical Modelling*. 2001; 1:81.
- Houghton RA, Lefkowitz DS, Skole DL. Changes in the landscape of Latin America between 1850 and 1985: I. Progressive loss of forests. *Forest Ecology and Management*. 1991; 38:143–72.
- Katz EG. Social capital and natural capital: a comparative analysis of land tenure and natural resource management in Guatemala. *Land Economics*. 2000; 76(1):114–32.
- Kleinn C, Corrales L, Morales D. Forest area in Costa Rica: A comparative study of tropical forest cover estimates over time. *Environ Monit Assess*. 2002; 73:17–40. [PubMed: 11878627]
- Kreft, I.; de Leeuw, J. *Introducing Multi-level Modeling*. London: Sage; 1998.

- Kupfer JA, Farris CA. Incorporating spatial non-stationarity of regression coefficients into predictive vegetation models. *Landscape Ecology*. 2007; 22:837–852.
- Lambin EF, Geist HJ, Lepers E. Dynamics of Land-Use and Land-Cover Change in Tropical Regions. *Annual Review of Environment and Resources*. 2003; 28:205–41.
- Liu J, Dietz T, Carpenter SR, Alberti M, Folke C, Moran E, Pell AN, Deadman P, Kratz T, Lubchenco J, Ostrom E, Ouyang Z, Provencher W, Redman CL, Schneider SH, Taylor WW. Complexity of coupled human and natural systems. *Science*. 2007; 317(5844):1513–16. [PubMed: 17872436]
- Loening, LJ.; Markussen, M. Poverty, deforestation and biodiversity loss in Guatemala. Ibero-America Institute for Economic Research Discussion Paper No. 91; 2003;
- Mather AS, Needle CL. The forest transition: a theoretical basis. *Area*. 1998; 30(2):117–24.
- Mayaux P, Holmgren P, Achard F, Eva H, Stibig H, Branthomme A. Tropical forest cover change in the 1990s and options for future monitoring. *Philosophical Transactions of the Royal Society*. 2005; 360:373–84.
- Manson SM. Simplifying Complexity: A Review of Complexity Theory. *Geoforum*. 2001; 32(3):405–14.
- Monroe DK, Southworth J, Tucker CM. Modeling spatially and temporary complex land-cover change: the case of western Honduras. *The Professional Geographer*. 2004; 56(4):544–59.
- Myers N. Tropical forests: present status and future outlook. *Climatic Change*. 1991; 19:3–32. [PubMed: 12317221]
- Nelson A. Analysing data across geographic scales in Honduras: detecting levels of organisation within systems. *Agriculture, Ecosystems and Environment*. 2001; 85:107–31.
- Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR. Terrestrial ecoregions of the world: A new map of life on earth. *Bioscience*. 2001; 51:933–38.
- Openshaw, S. Concepts and techniques in modern geography number 38: The modifiable areal unit problem. Norwick: GeoBooks; 1984.
- Overmars KP, de Koning GHJ, Veldkamp A. Spatial autocorrelation in multi-scale land use models. *Ecological Modelling*. 2003; 164(2-3):257–70.
- Pichon FJ. Colonist land allocation decisions, land use, and deforestation in the Ecuadorian Amazon frontier. *Economic Development and Cultural Change*. 1997; 45(4):707–44.
- Pineda Jaimes NB, Bosque Sendra J, Gómez Delgado M, Franco Plata R. Exploring the driving forces behind deforestation in the state of Mexico (Mexico) using geographically weighted regression. *Applied Geography*. 2010
- Rindfuss RR, Entwisle B, Walsh SJ, An L, Badenoch N, Brown DG, Deadman P. Land use change: complexity and comparisons. *Journal of Land Use Science*. 2008; 3(1):1–10. [PubMed: 20354590]
- Rudel TK, DeFries R, Asner G, Laurance W. Changing drivers of deforestation and new opportunities for conservation. *Conservation Biology*. 2009; 23:1396–405. [PubMed: 20078640]
- Sader SA, Reining C, Sever T, Soza C. Human migration and agricultural expansion, a threat to the Maya tropical forest. *Journal of Forestry*. 1997; 95(12):27–32.
- Southworth J, Tucker C. The influence of accessibility, local institutions, and socioeconomic factors on forest cover change in the mountains of western Honduras. *Mountain Res Dev*. 2001; 21(3): 276–83.
- Turner, BL., II; Lambin, E.; Reenberg, A. The Emergence of Land Change Science for Global Environmental Change and Sustainability. *Proceedings, National Academy of Sciences of the United States of America*; 2007; p. 20666-71.
- Turner BL II, Robbins P. Land-Change Science and Political Ecology: Similarities, Differences, and Implications for Sustainability Science. *Annual Review of Environment and Resources*. 2008; 33:295–316.
- Wheeler D, Tiefelsdorf M. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*. 2005; 7:161–187.
- Witmer F. Simulating future global deforestation using geographically explicit models. IIASA Interim Report IR-05-010. 2005

- Wright SJ, Samaniego MJ. Historical, demographic, and economic correlates of land-use change in the Republic of Panama. *Ecology and Society*. 2008; 13(2):17.
- UN (United Nations). World Fertility Report 2003. New York, NY, USA: 2003.
- Vance C, Iovanna R. Analyzing spatial hierarchies in remotely sensed data: Insights from a multi-level model of tropical deforestation. *Land Use Policy*. 2006; 23:226–36.
- Van der Leeden R, Busing F, Meijer E. Applications of bootstrap methods for two-level models. Paper, Multilevel Conference. 1997
- Verburg PH, Schot PP, Dijst MJ, Veldkamp A. Land use change modelling: current practice and research priorities. *GeoJournal*. 2004; 61(4):309–24.
- World Bank World Development Indicators. [Accessed on June 8, 2011] 2011. <http://data.worldbank.org/data-catalog/world-development-indicators>

Figure 1.
Guatemala regional boundaries.

Figure 2. Percent woody cover in 2009 (left) and percent change in woody cover from 2001 to 2009 (right), with increase in cover in greens and decrease in cover in browns. The 22 departments are outlined in black

Figure 3. Getis-Ord G_i^* maps for woody cover in 2009 and change in woody cover from 2001 to 2009 with significant clusters of high values in red and significant clusters of low values in blue.

Figure 4.

GWR global adjusted R^2 values as compared to the OLS results show improvement, from 0.251 to 0.471 for model 1, and from 0.016 to 0.024 for model 2. However local GWR R^2 values shown above for model 1 (left) and model 2 (right) demonstrate a more complicated story of where the models predict woody cover best.

Figure 5. GWR model 1 coefficient estimates for percent population density change from 1990 to 2000 (left), and percent of land in fallow (right).

Figure 6. GWR coefficient estimates for percent households using fertilizer for model 1 (left) and model 2 (right).

Figure 7. Coefficient estimates for percent households owning tractors for model 1 (left) and model 2 (right).

Table 1
Explanatory Variables Used in the Models

Explanatory Variable	Definition	Mean	Standard Deviation	N
Population Density 1990	Year 1990 Municipal Population Density in Persons per Square Kilometer	185.91	311.54	329
Population Density 2000	Year 2000 Municipal Population Density in Persons per Square Kilometer	275.14	483.42	329
% Population Density Δ (1990 to 2000)	Percentage of Population Change by Municipality from 1990 to 2000	29.77	12.47	329
% Population Density Δ (1990 to 2000)²	Percentage of Population Change by Municipality from 1990 to 2000 Squared	886.25	155.50	329
% HHs Using Fertilizer	Percentage of Municipal Farming Households that Use Fertilizer in 2000	83.02	35.60	330
% HHs Owning a Tractor	Percentage of Municipal Farming Households that Own a Tractor in 2000	3.80	8.87	300
Café Production	Municipality Café Production in Kilograms per Hectare in 2003	536.95	1243.40	331
Sugar Production	Municipality Sugar Production in Kilograms per Hectare in 2003	2885.79	14192.98	331
White Corn Production	Municipality White Corn Production in Kilograms per Hectare in 2003	85.14	125.09	331
Yellow Corn Production	Municipality Yellow Corn Production in Kilograms per Hectare in 2003	28.39	67.69	331
% Fallow Land	Percentage of Municipal Farmland in Fallow in 2003	20.53	12.51	330

Table 2

Error matrices for A) the moist forest biome and B) the combined conifer-dry forest biome maps.

A Classification	Ag/Herb	Bare/Built	Reference MW/Plant	Water	Woody	Total	User
Ag/Herb	646	25	167	0	0	838	77.1%
Bare/Built	4	62	2	0	0	68	91.2%
MixedWoody/Plantations	60	0	281	0	1	342	82.2%
Water	2	0	0	161	0	163	98.8%
Woody Veg	14	0	29	0	628	671	93.6%
Total	726	87	479	161	629	2082	
Producer	89.0%	71.3%	58.7%	100.0%	99.8%		85.4%
B Classification	Ag/Herb	Bare/Built	Reference MW/Plant	Water	Woody	Total	User
Ag/Herb	344	3	35	0	1	383	89.8%
Bare/Built	7	129	0	0	0	136	94.9%
MixedWoody/Plantations	55	2	477	0	62	596	80.0%
Water	0	0	0	119	0	119	100.0%
Woody Veg	3	0	34	0	151	188	80.3%
Total	409	134	546	119	214	1422	
Producer	84.1%	96.3%	87.4%	100.0%	70.6%		85.8%

Table 3

Multi-level results for modeling percent woody cover in 2009 with socio-economic variables. The estimates are from 298 data points grouped by 22 departments.

	OLS (Not Multilevel)	Maximum Likelihood Estimator	Random Effects	Fixed Effects
PopDen 2000	-0.000 (0.005) **	-0.000 (0.017) *	-0.000 (0.014) *	-0.000 (0.030) *
PrctChgPopDensity1990-2000	0.565 (0.001) **	0.324 (0.053) ⁺	0.401 (0.015) *	0.185(0.282)
PrctChgPopDensity19902000²	-0.131 (0.003) **	-0.077 (0.075) ⁺	-0.094 (0.030) *	-0.046 (0.293)
% HHs Using Fertilizer	-0.059 (0.073) ⁺	-0.058 (0.072) ⁺	-0.059 (0.070) ⁺	-0.054 (0.121)
% HHS Owning a Tractor	-0.404 (0.004) **	-0.293 (0.032) *	-0.326 (0.019) *	-0.242 (0.097) ⁺
Café Production	0.000 (0.000) **	0.000 (0.000) **	0.000 (0.000) **	0.000 (0.000) **
Sugar Production	-0.000 (0.008) **	-0.000 (0.025) *	-0.000 (0.019) *	-0.000 (0.058) ⁺
White Corn Production	-0.000 (0.005) **	-0.000 (0.001) **	-0.000 (0.002) **	-0.000 (0.001) **
Yellow Corn Production	0.000 (0.184)	0.000 (0.166)	0.000 (0.177)	0.000 (0.181)
% Fallow Land	-0.565 (0.000) **	-0.524 (0.000) **	-0.543 (0.000) **	-0.479 (0.000) **
R-squared	0.286		0.282	0.236
Log likelihood		77.51		
Random Intercept Standard Deviation (ζ_j)		0.08	0.05	0.10
Residual Standard Deviation (e_{ij})		0.18	0.18	0.18
Interclass Correlation (ρ)		0.16	0.07	0.23
Likelihood Ratio Test of ζ_j		17.71		
Chi-squared		(0.000)		
F-test of Model Significance				3.27
Chi-squared				(0.000)
Hausman Test			24.32	
Chi-squared			(0.002)	

p values in parentheses:

⁺ significant at 10%;

* significant at 5%;

** significant at 1%

Table 4

Multi-level results for (natural log) change in woody cover from 2001 to 2009 with socioeconomic variables. Estimates are from 278 data points grouped by 22 departments.

	Not Multilevel (OLS)	Maximum Likelihood Estimator	Random Effects	Fixed Effects
PopDen_2000	0.001 (0.000) **	0.000 (0.003) **	0.000 (0.003) **	0.000 (0.011) *
PrctChgPopDensity1990-2000	-0.243 (0.312)	0.093 (0.658)	0.086 (0.688)	0.122 (0.575)
% HHs Using Fertilizer	1.226 (0.000) **	0.623 (0.001) **	0.641 (0.001) **	0.536 (0.005) **
% HHS Owning a Tractor	-2.399 (0.009) **	-2.777 (0.000) **	-2.755 (0.000) **	-2.890 (0.000) **
Café Production	0.000 (0.000) **	0.000 (0.003) **	0.000 (0.003) **	0.000 (0.009) **
Sugar Production	-0.000 (0.464)	-0.000 (0.814)	-0.000 (0.820)	-0.000 (0.793)
White Corn Production	0.003 (0.001) **	0.001 (0.049) *	0.001 (0.045) *	0.001 (0.114)
Yellow Corn Production	0.001 (0.308)	0.001 (0.231)	0.001 (0.240)	0.001 (0.257)
% Fallow Land	-0.221 (0.757)	1.134 (0.079) ⁺	1.094 (0.096) ⁺	1.325 (0.048) *
R-squared	0.307		0.252	0.158
Log likelihood		-406.32		
Random Intercept Standard Deviation (ζ_j)		0.96	0.87	1.07
Residual Standard Deviation (ϵ_{ij})		0.94	0.96	0.96
Interclass Correlation (ρ)		0.51	0.45	0.56
Likelihood Ratio Test of ζ_j		86.32		
<i>Chi-squared</i>		(0.000)		
F-test of Model Significance				9.65
<i>Chi-squared</i>				(0.000)
Hausman Test			22.25	
<i>Chi-squared</i>			(0.002)	

p values in parentheses:

⁺ significant at 10%;

* significant at 5%;

** significant at 1%

Table 5

Moran's I values for cover measures, OLS standardized residuals, and GWR standardized residuals.

Variable	Moran's I	Z Score	P Value	Autocorrelation
Woody Cover 2009	.09	12.59	< .0001	Strongly Clustered
Change in Woody Cover 2001 – 2009	.02	3.77	< .0001	Strongly Clustered
OLS Model 1 Standardized Residuals	.07	9.42	< .0001	Strongly Clustered
OLS Model 2 Standardized Residuals	.01	1.78	.07	Clustered
GWR Model 1 Standardized Residuals	.00	.72	.47	Random
GWR Model 2 Standardized Residuals	.01	1.16	.25	Random