esa

ECOSPHERE

Spatial autocorrelation reduces model precision and predictive power in deforestation analyses

Kristjan D. Mets,[1,†] Dolors Armenteras,[2] and Liliana M. Dávalos[1,3]

[1]*Department of Ecology and Evolution, Stony Brook University, Stony Brook, New York 11794 USA*
[2]*Landscape Ecology and Ecosystem Modelling Laboratory, Department of Biology, Sciences Faculty, Universidad Nacional de Colombia, Bogotá 111321 Colombia*
[3]*Consortium for Inter-Disciplinary Environmental Research, Stony Brook University, Stony Brook, New York 11794 USA*

Citation: Mets, K. D., D. Armenteras, and L. M. Dávalos. 2017. Spatial autocorrelation reduces model precision and predictive power in deforestation analyses. Ecosphere 8(5):e01824. 10.1002/ecs2.1824

Abstract. Generalized linear models are often used to identify covariates of landscape processes and to model land-use change. Generalized linear models however, overlook the spatial component of land-use data, and its effects on statistical inference. Spatial autocorrelation may artificially reduce variance in observations, and inflate the effect size of covariates. To uncover the consequences of overlooking this spatial component, we tested both spatially explicit and non-spatial models of deforestation for Colombia. Parameter estimates, analyses of residual spatial autocorrelation, and Bayesian posterior predictive checks were used to compare model performance. Significant residual correlation showed that non-spatial models failed to adequately explain the spatial structure of the data. Posterior predictive checks revealed that spatially explicit models had strong predictive power for the entire range of the response variable and only failed to predict outliers, in contrast with non-spatial models, which lacked predictive power for all response values. The predictive power of non-spatial models was especially low in regions away from Colombia's center, where about half the observations were clustered. While all analyses consistently identified a core of important covariates of deforestation rates, predictive modeling requires parameter estimates informed by the spatial structure of the data. To inform increasingly important forest and carbon sequestration policy, land-use models must account for spatial autocorrelation.

Key words: Amazon; Andes; Colombia; land protection; posterior predictive check.

Received 29 July 2016; revised 31 January 2017; accepted 17 February 2017. Corresponding Editor: Tobias van Kooten.

† **E-mail:** kristjan.mets@stonybrook.edu

Introduction

Statistical testing and modeling lies at the core of all analyses and projections of deforestation, and land use, but statistical analyses generally assume independence between observations (Kruskal 1988). When data represent measures taken across geographic space, the assumption of independence is not fully met because of spatial autocorrelation (Legendre 1993). Pairs of spatial observations typically exhibit a correlation inverse to the distance between them, and ecological spatial data are no exception (Cliff and Ord 1970).

Modeling spatial patterns under the assumption of independence affects statistical inference in three ways. First, spatial autocorrelation emerges as non-random geographic association of residual errors in regression analyses (Cliff and Ord 1972). Compared against spatially explicit models, then, non-spatial models of spatial patterns have deflated estimates of variance and residual autocorrelation (Legendre and Fortin 1989). This leads to loss of model precision and higher type I error rates (Beale et al. 2010). Second, non-spatial models used for spatial data spuriously internalize spatial autocorrelation into the goodness of fit of the model, undermining comparisons of model performance (Telford and Birks 2005). Lastly, when explanatory variables exhibit different spatial patterns and degrees of autocorrelation, non-spatially explicit models inflate the effect sizes of the more autocorrelated variables (Lennon 2000).

Spatially explicit analyses are increasingly used to address the problems arising from spatial autocorrelation. Accounting for spatial autocorrelation, however, remains relatively uncommon in many disciplines, including the study of land-use change (Brown et al. 2013). The increased complexity these methods present to researchers, as well as a common failure to recognize that spatial autocorrelation in model residuals indicates a violation of independence, prevents the broad adaption of spatially explicit methods (Kühn and Dormann 2012). Within the study of land-use change, the spatial data necessary to reproduce land-use analyses are seldom published along with the models. Most studies will therefore never undergo critical re-analyses by other researchers (Koenig 1999, Hunter et al. 2009). Hence, the consequences of assuming independence of observations in models of land-use change remain unknown.

Tropical deforestation is a prime example of an intrinsically spatial process of change in land use. Although analyses sometimes include a spatial component, linear and generalized linear regressions are two common approaches for identifying relationships between deforestation and a suite of explanatory socioeconomic, environmental, and infrastructure variables (Rudel and Roper 1997, Kaimowitz et al. 2004). With few opportunities to control factors over time, most analyses apply regressions against explanatory variables without regard to spatial structure. Ignoring spatial autocorrelation makes these analyses susceptible to bias from homogenous and spatially concentrated data clusters (Overmars et al. 2003), even as the conclusions of these analyses help shape national and international efforts to curb deforestation by identifying and discouraging enabling factors (Gullison et al. 2007). Improving the predictive power of deforestation analyses has gained urgency as these are components of initiatives aimed at preserving biodiversity or combating climate change such as the Reducing Emissions from Deforestation and Forest Degradation (REDD+) through the United Nation's Framework Convention on Climate Change (UNFCCC), last renewed through the 2015 Paris Agreement (UNFCCC 2015).

Here, we present a series of non-spatially and spatially explicit analyses to quantify the effects of spatial structure when modeling deforestation data. These models build on published data on determinants of deforestation in Colombia (Armenteras et al. 2011*b*, 2013*a*). We compare parameter estimates and residual autocorrelation from deforestation models that do and do not explicitly capture spatial autocorrelation, and evaluate the model performance through Bayesian posterior predictive checks. Bayesian posterior predictive checks allowed us to produce distributions of deforestation values for all municipalities based on estimated coefficients and compare them to observed patterns of deforestation.

Materials and Methods

Material and data

We reanalyzed change in forest cover for the geographically heterogeneous country of Colombia for 1985–2005 for the Andean region and 1990–2005 for the rest of the country (Armenteras et al. 2011*b*, 2013*a*). Data for non-Andean forest cover in 1990 and 2005 were obtained from the report on "Scientific and institutional capacity building to support REDD projects in Colombia" (Montenegro et al. 2011). The original forest cover dataset was measured via remote sensing using over 240 Landsat multispectral satellite images with <10% cloud obstruction from 1990 to 2005. For the Andean region, analyses by the National University of Colombia provided the land-cover data (Armenteras et al. 2011*b*). The annual rate of deforestation of each municipality (R_m, in %) was calculated as (Fearnside 1993):

$$R_m = \frac{A_{m1} - A_{m2}}{FA_1 \times t} \times 100$$

in which A_{m1} and A_{m2} are total forest areas within the municipality at the beginning (1985 or 1990) and end (2005) of remote sensing records, respectively. FA_1 is the total forest area throughout Colombia at the initial year, and t is the time interval in years. Covariates of deforestation included environmental and social variables (Table 1) for 1119 municipalities divided into five socio-environmental regions (Fig. 1).

Table 1. A summary of the potential covariates of forest cover change.

Variables	Units	Spatial resolution	Description	Source(s)
Urban population density	population/ha	Municipality	Change in urban population density between 1985 and 2005	DANE (1985, 1993, 2005)
Rural population density	population/ha	Municipality	Change in rural population density between 1985 and 2005	DANE (1985, 1993, 2005)
Unsatisfied basic needs (NBI)	percentage	Municipality	Percentage of population with unsatisfied basic needs in 2005; includes minimum household connections, access to sanitary services, access to primary education, and minimum household economic capacity as basic needs	DANE (2005)
Crops	ha	30 m	Change in crop area between 1985 and 2005 measured via remote sensing	IDEAM (2007)
Pastures	ha	30 m	Change in pasture area calculated between 1985 and 2005	IDEAM (2007)
Illicit crops	ha	10 m	Area of coca (*erythroxylum coca*) crops	UNODC (2006)
Cattle	number	Municipality	Head of cattle per municipality in 2006	IGAC (2011)
Fire hotspots	number	Municipality	Number of fire hotspots detected per municipality between 2000 and 2005	NASA (2015)
Mining	kg	Municipality	Total gold and silver production in 2005	IGAC (2011)
Protected area	ha	1:100,000	Area under special management as national protected area or indigenous reserve	IGAC (2005)
Road density	km/ha	1:100,000	Density of roads in each municipality	IGAC (2005)
Slope	degrees	90 m	Average maximum slope for each municipality	IGAC (2005)
Water scarcity	index	Municipality	Index of water scarcity in a dry year	IDEAM (2000)
Elevation	~m	90 m	Altitude above sea level	IGAC (2005)
Precipitation	mm	1 km^2	Total annual precipitation	Worldclim (Hijmans et al. 2005)

Note: DANE, National Administrative Department of Statistics; UNDOC, United Nations Office on Drugs and Crime; MADR, Agriculture and Rural Development Ministry; FIRMS, Fire Information for Resource Management System; SIGOT, Geographic Information System for Planning and Territorial Development; IGAC, Agustín Codazzi Geographical Institute; IDEAM, Institute of Meteorology and Environmental Studies, Worldclim global climate data.

Modeling approaches

Previously published analyses using the same data fitted a series of five generalized linear models (GLMs): mainland Colombia as a whole; Andes, Amazon, and Orinoco basins; and the Caribbean region (Armenteras et al. 2013*a*, Rodríguez et al. 2013*a*). Here, we built an initial GLM including data for the entire country and the 15 explanatory variables previously analyzed (Table 2). The percentage of initial forest cover in each municipality was added as a covariate to the previously published data, as research suggests that it is a strong and negative covariate of deforestation (Hargrave and Kis-Katos 2013).

We used four approaches to model deforestation as a function of covariates (Table 3). (1) Generalized linear models were previously used to analyze individual regions (Armenteras et al. 2013*a*, Rodríguez et al. 2013*a*), and served as the baseline lacking any method of accounting for spatial autocorrelation; (2) Bayesian random intercepts and slope models (RIS) were used to model deforestation on a regional level, but without an explicit spatial component. To explicitly account for spatial autocorrelation, we applied (3) linear mixed-effects or hierarchical models with correlation structures (geospatial LMEs); and (4) Bayesian models with conditional autoregressive priors (CARs). The LME model accounts for spatial autocorrelation using distance-based correlation matrices within groups of observations, and the CAR priors use a binary neighbor-based matrix. All models were implemented using the open source R statistical language version 3.0.1 (R Development Core Team 2008).

To reduce the number of explanatory variables in the most computationally demanding of the analyses accounting for spatial autocorrelation, an initial Bayesian CAR analysis was conducted using the CARBayes package (v4.1; Lee 2013). All

Fig. 1. Socio-environmental regions of Colombia.

potential explanatory variables were used as predictors. The 95% high probability density (HPD) of coefficients was examined. Variables with HPD entirely below or above 0 were considered predictive variables, while the variables with credible intervals including 0 were designated as non-predictive. This reduced the potential predictors to seven variables: elevation, legally protected area, fire hotspots, urban population density, unsatisfied basic needs (*necesidades básicas insatisfechas* [NBI]), road density, and initial percentage of forest cover. These variables were then included in subsequent RIS, LME, and CAR analyses as predictors.

Table 2. Coefficients and their P-values are given for an initial GLM using 15 variables as covariates of deforestation for the entire country.

Variables	Coefficient	P
Intercept	3.748	0.003
ln crops	−0.057	0.070
ln pasture	−0.089	0.231
NBI	3.859	**<0.001**
Road density	0.342	**<0.001**
Slope	−0.012	0.784
Mining	−0.004	0.234
Elevation	−0.728	**<0.001**
Precipitation	<0.001	**0.003**
Water scarcity	−0.007	0.959
Illicit crops	0.204	0.114
Fires	−0.062	0.954
ln protected area	−0.147	**<0.001**
ln urban population	0.336	**<0.001**
ln rural density	0.252	0.051
ln cattle	−0.027	0.827

Note: Bold values indicate $P < 0.05$.

Spatial autocorrelation models

The RIS model is an extension of the GLM that captures regional variation without explicitly accounting for spatial autocorrelation. In the RIS model, the intercept and slopes of different covariates vary in the socio-environmental regions of Colombia. This model was applied in a Bayesian framework using JAGS version 3.4 (Plummer 2013) and implemented in R using the R2jags package (v 0.5-6; Su and Yajima 2015).

The LME model included both a grouping factor by socio-environmental region and a geospatial correlation matrix that applies within regions. The geospatial correlation matrix of each LME model assigns correlation weights to pairs of observations determined by the distance between geographic centroids of the corresponding municipalities. The symmetric $n \times n$ correlation matrix is set by a corSpatial object class from the nlme package (v 3.1-109; Pinheiro et al. 2011). The corSpatial object can use one of five different spatial functions to weigh the degree of spatial autocorrelation among residuals (Appendix S1: Fig. S1). We constructed LME models using each spatial decay function with parameters estimated by maximum likelihood. The associated log likelihood of each LME model allowed for Akaike's information criterion comparison with the exponential decay corExp model demonstrating best goodness of fit (Appendix S1: Table S1).

The LME models allowed for estimates of coefficients across all observations (equivalent to those fitted in classical regressions and hereafter called nationwide), alongside coefficients for individual groups, hereafter called group-specific (Gelman 2005, Gelman and Hill 2006). Two variables, forest cover and fire hotspots, had coefficients that changed signs among the different natural regions. For the final LME model encompassing all seven predictive variables, coefficients for forest cover and

Table 3. Summary characteristics of the four models tested.

Characteristics	GLM	RIS	LME with Geospatial matrix	CAR
Implementation	glm	JAGS	nlme, corStruct	CARBayes
Parameters	15	7	7	7
Grouping	No	Yes (7 parameters)	Yes (2 parameters)	No
Autocorrelation	No	No	Yes (distance)	Yes (neighbors)

Note: GLM, generalized linear model; RIS, random intercepts and slopes; LME, linear mixed-effects model; CAR, conditional autoregressive priors.

fires were estimated as specific to regions. The remaining five predictive variables were estimated as nationwide so as to reduce the computational complexity of the LME model when applying the correlation matrix. In every case, the geospatial correlation matrix was applied to residuals of observations within groups.

The second spatially explicit approach used a Bayesian hierarchical model with CAR priors constructed with the CARBayes package (v 4.1; Lee 2013). A symmetric $n \times n$ neighborhood matrix W sets the spatial autocorrelation structure of the CAR priors for N observations. For observations from neighboring spatial units k and i, $w_{ki} = 1$. If the spatial units, municipalities in this study, do not share a border, then $w_{ki} = 0$. The CAR priors used for this study, originally modeled by Leroux et al. (2000), are given by Eq. 2:

$$\phi_k | \phi_{-k}, W, \tau^2, \rho \sim N\left(\frac{\rho \sum_{i=1}^{n} w_{ki}\phi_i}{\rho \sum_{i=1}^{n} w_{ki} + 1 - \rho}, \frac{\tau^2}{\rho \sum_{i=1}^{n} w_{ki} + 1 - \rho}\right)$$

in which ρ is a correlation strength parameter ranging from 0 to 1, and τ is a parameter denoting variance in the correlation weight. This method does not take geographic distance between municipalities into account. Coefficients do not differ between the natural regions because the data are not separated by regional groupings.

Comparing models

We compared the performance of the models by (1) examining coefficients and their significance or HPD, as spatial autocorrelation is expected to inflate effect sizes for autocorrelated variables and increase type I error rates; (2) estimating spatial autocorrelation of residuals from each model to diagnose incorrectly modeled spatial variation; and (3) conducting posterior predictive checks of the models to evaluate model precision. Spatial autocorrelation is expected to increase homogeneity in observed values between contiguous municipalities, especially when distances are small (Ord and Getis 1995), as in the developed central Andes. To estimate residual spatial autocorrelation, we used the ncf package (v. 1.1-3; Bjørnstad 2009) to generate correlograms for each model. The correlograms compared the correlation of pairs of response estimates from municipalities within designated distance intervals from each other. For each 10-km distance interval, 500 valid municipality pairs were randomly sampled. The similarity in municipality pairs among each distance interval was compared to the nationwide similarity in deforestation values. Mean correlations significantly above the nationwide baseline at low-distance intervals indicated the presence of spatial autocorrelation. Posterior predictive checks were completed for the GLM, RIS, and CAR models implemented in Bayesian analyses. The predictive checks used coefficients sampled from the posterior distributions of each model to predict the deforestation response variable. Using 1000 samples for each municipality, we compared the predicted range to the observation.

Results

Covariates of change in forest cover

Statistically significant explanatory variables in the all-variable GLM included demographic, land-use, and geographic attributes (Table 1). The variables unsatisfied basic needs, urban population density, and road density were all significantly associated with deforestation. Fires, protected area, and elevation were associated with forest growth, while precipitation had statistically significant but minor effects on forest growth.

The best-fit geospatial LME model fitted an exponential decay function to the spatial autocorrelation of residuals. All variables with nationwide coefficients had significant effects on change in forest cover ($P < 0.05$). The covariates of deforestation were the same as for the GLM (Table 4). Protected area and elevation were similarly associated with forest growth. Initial forest cover was modeled as a region-specific predictor and changed signs among the five regions (Table 4). In the Amazon, Caribbean, and Pacific, initial forest cover was a covariate of forest growth. In the Andes and Orinoco, initial forest cover correlated with deforestation. Fires was also modeled as region-specific and were consistently opposite in sign to the effect of initial forest cover, albeit to different extents in the five natural regions (Table 4). Fires and initial forest cover had an overall correlation value of −0.77 for the geospatial LME model. Fires had a negative effect on deforestation in the Andes and Orinoco basins. This result may be associated with management and harvesting of agriculture in these areas in addition to forest fires propagated along the colonization front (Armenteras et al. 2011*a*).

The median coefficients of the CAR model matched corresponding nationwide coefficients of the LME model in sign and relative magnitude (Table 4). Fires and forest cover were both linked to forest growth. When compared to these spatial models, the GLM estimated much greater effect sizes for unsatisfied basic needs and urban population density on deforestation.

The coefficients of the RIS model largely agreed with those of the geospatial LME model. Although there were slight differences in the regional coefficients for land protection, elevation, unsatisfied basic needs, urban density, and road density, there was no disagreement in sign. The only notable difference was the coefficient for fires in the Pacific region. Whereas fire was associated with deforestation in the geospatial LME model, it was associated with forest growth in the RIS model.

Residual spatial autocorrelation

We calculated the correlations between the residuals of pairs of municipalities at bins of geographic distance between municipal centroids to obtain correlograms (Fig. 2). Pairs of municipalities in close proximity to each other had an overall significantly positive correlation of residuals in the GLM. This effect persisted for municipality pairs separated by up to 50 km. Municipality pairs in the RIS model shared the general trend of the GLM, with significant correlation of residuals in close proximity, albeit to a reduced extent. The geospatial LME and CAR models displayed different trends, and there was no significant correlation of residuals in pairs of municipalities in close proximity.

Bayesian posterior predictive checks

The CAR and RIS approaches were Bayesian models and thus had posterior distributions for each predictive coefficient. A Bayesian implementation of the GLM had median coefficients (Appendix S1: Table S2) identical to those estimated using the glm function in R. Coefficients were sampled from the posterior distributions of each model and used to predict the rate of deforestation of each municipality. These predictions were compared to the observed rates of deforestation (Fig. 3). The geospatial LME model required computation of an $n \times n$ matrix of distance-based correlation weights for the 1119 municipalities, which presented tremendous computational complexity with a Bayesian implementation. Hence the geospatial LME model was excluded from this analysis.

The models displayed similar trends in predictions for the Andes, Amazon, and Orinoco regions (Fig. 3). The full GLM consistently underestimated deforestation rates and excluded approximately half of all real observations from the predictive ranges. The RIS model similarly underestimated the rate of deforestation. The general trend of underestimating deforestation for both these models held for predictions across all regions. The CAR model predictions, while sharing the tendency to underestimate deforestation, encapsulated most real observations within predictive ranges. The extreme ends of the Andes, for municipalities that experienced the most deforestation or forest growth, fell well outside any predictive posteriors. The Caribbean region displayed the most deviation between predicted and observed deforestation rates. All approaches failed to predict deforestation observed in three-quarters of the municipalities in the Caribbean region.

Table 4. Coefficients for intercept and seven covariates.

Covariates	GLM	CAR	LME with Geospatial matrix	RIS
Intercept	3.041	1.622 (−0.851, 4.155)		
Ama.			1.709	1.790 (−1.035, 4.635)
And.			0.413	1.479 (−1.357, 4.139)
Car.			3.210	2.103 (−0.676, 5.071)
Ori.			0.048	1.698 (−1.318, 4.570)
Pac.			2.836	1.615 (−1.291, 4.387)
Forest cover	−0.457*	−1.166 (−2.309, −0.040)		
Ama.			−1.151	−0.384 (−3.050, 1.701)
And.			0.977	1.012 (−0.305, 2.711)
Car.			−3.655	−1.447 (−5.079, 1.094)
Ori.			1.609	0.337 (−2.255, 3.181)
Pac.			−3.015	−0.355 (−2.392, 1.386)
Fires	−0.623	−1.652 (−3.795, 0.520)		
Ama.			0.800	1.794 (0.312, 5.188)
And.			−0.915	−4.083 (−7.792, −0.507)
Car.			2.408	1.307 (−2.286, 4.661)
Ori.			−1.334	−4.498 (−9.236, −0.212)
Pac.			2.139	−3.403 (−10.694, 2.731)
Protection	−0.131**	−0.161 (−0.231, −0.091)	−0.155**	
Ama.				−0.193 (−0.045, 0.060)
And.				−0.072 (−0.166, 0.021)
Car.				−0.336 (−0.502, −0.172)
Ori.				−0.146 (−0.353, 0.073)
Pac.				−0.194 (−0.324, −0.067)
Elevation	−0.771**	−0.191 (−0.457, 0.071)	−0.290**	
Ama.				−0.387 (−0.695, −0.012)
And.				−0.395 (−0.636, −0.131)
Car.				−0.468 (−0.769, −0.202)
Ori.				−0.364 (−0.671, 0.059)
Pac.				−0.418 (−0.687, −0.133)
NBI	3.674**	1.451 (0.093, 2.816)	2.040**	
Ama.				1.899 (−0.376, 3.732)
And.				1.923 (0.335, 3.399)
Car.				2.914 (0.980, 5.848)
Ori.				1.859 (−0.578, 3.869)
Pac.				2.246 (0.649, 3.951)
Urban density	0.268**	0.081 (−0.044, 0.208)	0.112	
Ama.				0.133 (−0.033, 0.303)
And.				0.108 (−0.035, 0.250)
Car.				0.186 (−0.007, 0.411)
Ori.				0.130 (−0.090, 0.358)
Pac.				0.061 (−0.155, 0.247)
Road density	0.348**	0.257 (0.086, 0.427)	0.362**	
Ama.				0.532 (−0.523, 1.513)
And.				0.261 (−0.081, 0.439)
Car.				0.547 (0.222, 0.867)
Ori.				0.288 (−0.474, 0.942)
Pac.				1.179 (0.751, 1.604)

Notes: GLM, generalized linear model; RIS, random intercepts and slopes; LME, linear mixed-effects model; CAR, conditional autoregressive priors. Ama, Amazon; And, Andes; Car, Caribbean; Ori, Orinoco; Pac, Pacific. Median coefficient values are displayed for the RIS and CAR models. Positive values indicate variables associated with deforestation, and negative values indicate variables associated with forest growth. Median coefficients shown for CAR and RIS with 95% of the highest probability density shown in parentheses.

$^{*}P < 0.05$, $^{**}P < 0.01$ where applicable to nationwide coefficients of GLM and LME.

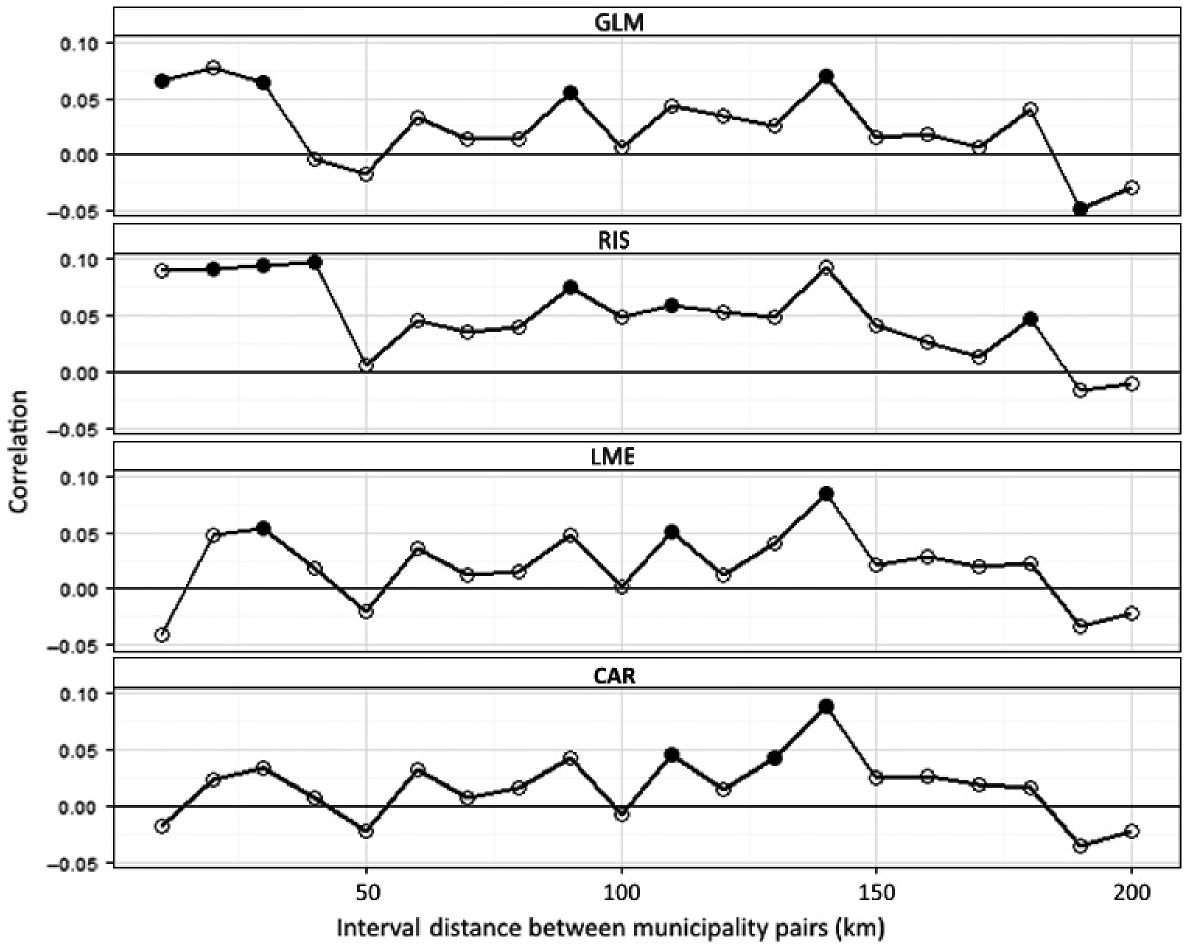

Fig. 2. Correlograms of residual correlation among municipality pairs separated by distance classes defined by km distance between paired municipality centroids. Closed dots indicate average correlation values significantly above the national baseline.

Heterogeneity of deforestation covariates in the Caribbean region

Deviance between observations and the predictive ranges was largest in the Caribbean region for all three models tested using posterior predictive checks. The geographic distribution of forests in the Caribbean region centers in two clusters: the protected Sierra Nevada de Santa Marta in the north and the Serranía de San Lucas at the northernmost end of the central Andes. To test whether deforestation rates in these two clusters have different covariates, we divided the Caribbean region along the Magdalena River into western (containing the Santa Marta area) and eastern (containing the San Lucas area) blocs. A Bayesian model with group-specific intercepts and slopes was then constructed to predict the rate of forest change in the Caribbean municipalities based on the same set of predictive parameters as in the other nationwide models.

The posterior distributions for the group-specific coefficients were compared between the eastern and western portions of the Caribbean natural region. While posterior distributions for the model intercept and most covariates overlay each other (Fig. 4), road density had a stronger association with deforestation in the eastern Caribbean, while fires was considerably more associated with deforestation in the western Caribbean.

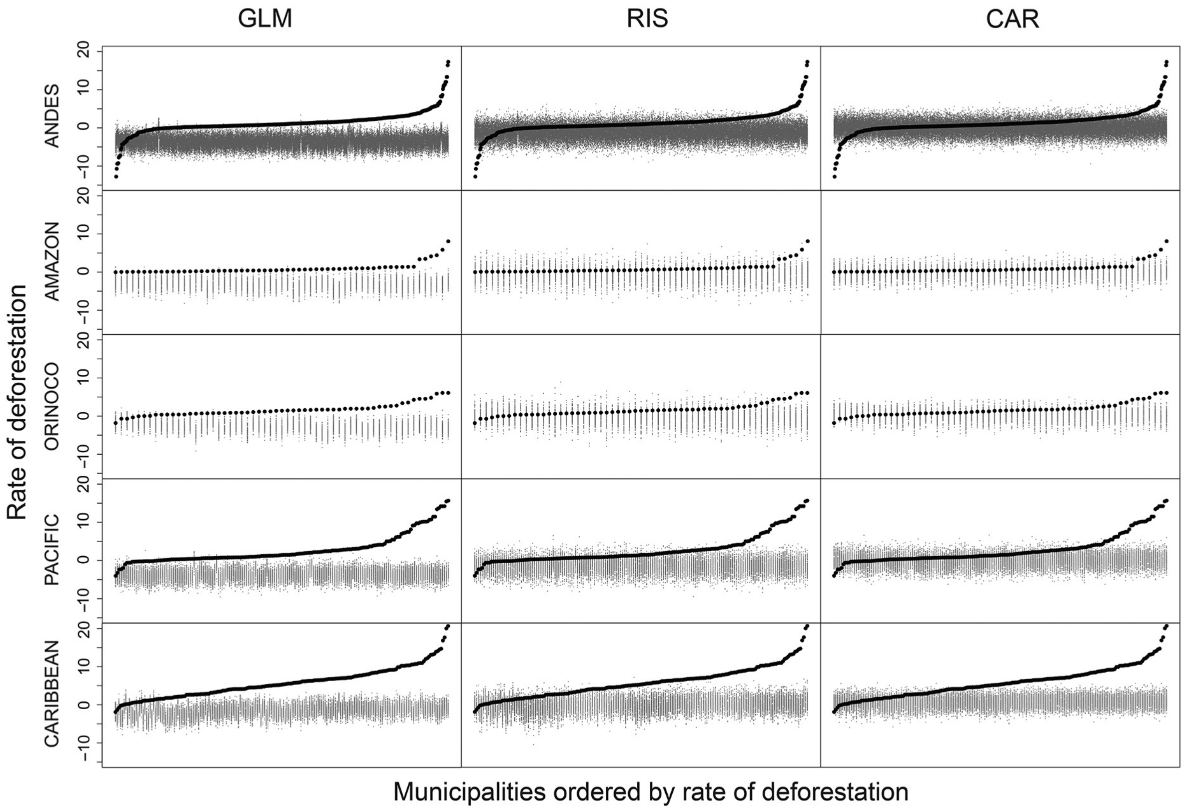

Fig. 3. Posterior predictive checks of three models. For each municipality and model, coefficients are sampled from posterior distributions and used to predict the rate of change in forest cover. Dark points represent observed rates of deforestation for each municipality ordered from least to greatest deforestation. Lighter points represent 1000 predicted rates of deforestation for each municipality as estimated through sampled coefficient values.

Discussion

Our analyses highlight the impact of modeling choices on model coefficients and their statistical significance, as well as the explanatory power of models. As geographic analyses inform concrete actions (Nepstad et al. 2006, Soares-Filho et al. 2010, Aide et al. 2013, Nolte et al. 2013), methodological choices can have profound consequences in shaping international initiatives such as REDD+, or national policies such as designation of land for legal protection (Portocarrero-Aya et al. 2014). Capturing the geographic structure of observations is thus an important requirement to model deforestation for policy decisions. We focus on two key findings: (1) change in model performance, in particular the systemic underestimation of deforestation rates from non-spatial models, and (2) implications for recent analyses of deforestation in Colombia.

Model performance

Posterior predictive checks reveal that autocorrelation not captured by models biases predictions (Fig. 3), and prediction improves with spatially explicit methods such as CAR. But there are many methods to quantify and address residual autocorrelation (Cliff and Ord 1970, Lam 1983, Portocarrero-Aya et al. 2014), and how to interpret results from different approaches is often uncertain (Getis 2007). A comparison between the LME geospatial and the CAR models illustrates the difference between models. The CAR model did not separate the municipalities by natural region, so comparisons are possible for the nationwide coefficients, but not for the region-specific parameters of the

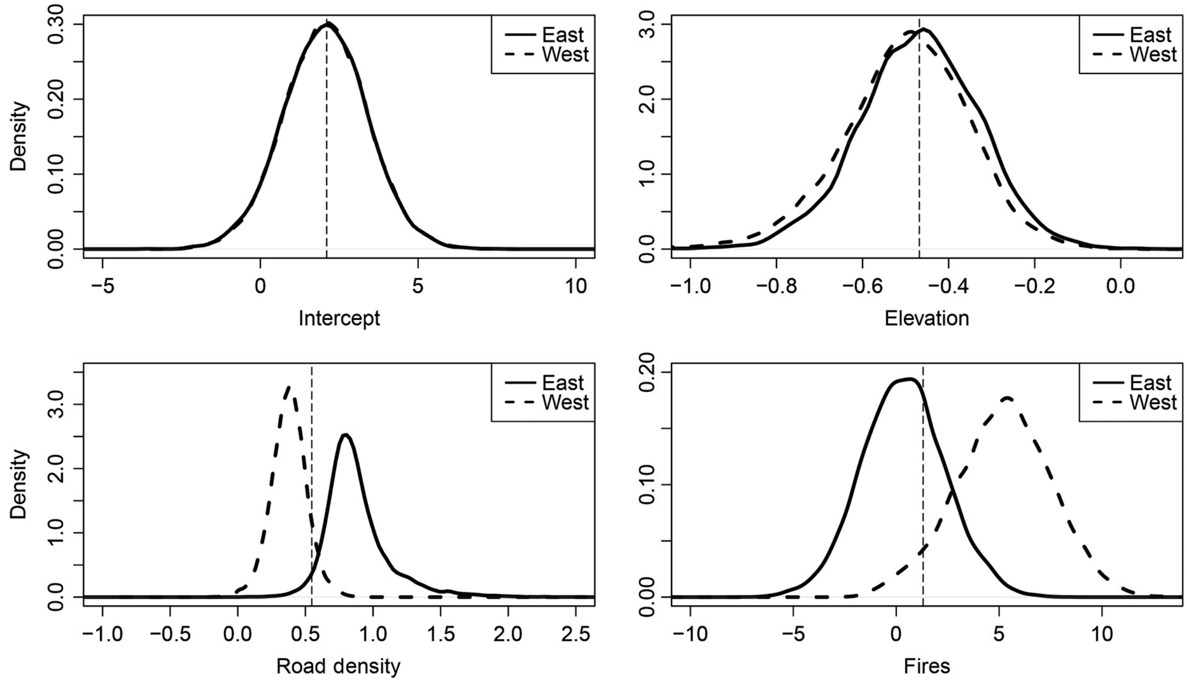

Fig. 4. Density plots of four coefficients differing between the eastern (solid line) and western (dotted line) Caribbean as separated by the Magdalena River. The vertical dotted line indicates the median coefficient value when the Caribbean is modeled as a single region.

LME. Despite using fundamentally different measures of spatial autocorrelation, the two models estimated similar coefficients for the nationwide predictive variables (Table 4). That is, there was no great impact on modeling the effects of covariates whether the correlation matrix was based on distance or on sharing of borders.

Spatial autocorrelation between observations is inversely correlated with distance and thus would be most prevalent in the densely populated cluster of the Andes where distances between municipalities are shortest. This central part of the country is where the Colombian state is most effective and data were anticipated to be most reliable (Mainwaring 2006). In addition, the Andes natural region contains the majority of municipalities, 631 out of 1119. This introduces a twofold problem for nationwide analyses: Model results may be biased toward deforestation determinants unique to the largely settled Andean region, and the spatial clustering of the Andes municipalities may inflate coefficients failing to describe deforestation in the more sparsely populated and heavily forested natural regions.

The effects of spatial autocorrelation with clustered units such as the Andes are best illustrated by comparing the CAR and RIS models. The CAR model outperforms the RIS model in all natural regions, despite the RIS model estimating region-specific coefficients (Fig. 3). Therefore, modeling the effect of spatial autocorrelation in the Andes improves precision for the outlying natural regions as well. The most predictive covariates of deforestation, NBI and road density, proved unidirectional in the RIS model and maintained similar magnitude across the different natural regions (Table 4). Because of the consistency of these covariates, the CAR model did not lose predictive power in outlying regions compared to the RIS, despite estimating only nationwide coefficients.

The optimal corStruct models applied for the LME spatial correlation matrix (Appendix S1: Fig. S1) show that spatial autocorrelation of residuals from deforestation data dissipates rapidly with distance. The best-performing model, an exponential decay function, had the correlation weight of municipality pairs approaching the national baseline within just 30 km distance. In developing regions of Colombia, municipality centroids are typically separated from neighboring centroids by well over 30 km. Observations from those regions where municipalities are larger will not exhibit considerable within-region spatial autocorrelation. This is problematic for the CAR model, which ignores distance in lieu of municipality neighborhood. However, while CAR-determined covariates of deforestation are reduced in magnitude compared to LME coefficients, there is agreement in sign and relative magnitude between variables (Table 4). This suggests that the CAR model, although ignoring physical distance between observations, does not overcompensate in applying correlation weights in regions with large municipalities such as in the Amazon and Orinoco basins.

In short, while LME models have the advantage of estimating group-specific coefficients, the absence of this structure does not undermine inference of whole-sample coefficients in CAR models. Both are preferable to non-spatial models for estimating the statistical relevance and effect size of predictors, and improving model precision. Failing to account for spatial autocorrelation in analyses of deforestation across units with skewed size distributions (as is common in analyses of political units) risks estimating parameters based on clusters of similar neighboring units. Finally, we demonstrated the use of posterior predictive checks to determine model precision as a powerful approach for uncovering region-specific misspecification and improve prediction of deforestation rates.

National determinants of deforestation

All four models estimated similar trends for the five nationwide variables of the LME model: elevation, land protection, urban population density, unsatisfied basic needs, and road density. Unsatisfied basic needs, used in large part as a proxy for areas at the forefront of colonization, was the variable most strongly correlated with deforestation. This index measures the percentage of the municipal population that lacks access to sanitary services, primary education, and minimum household economic capacity. Municipalities with high NBI largely lack state presence and often correspond to the rapidly changing agricultural frontier between settled and relatively developed cores, and newly colonized areas of primarily indigenous occupation (Rudel and Roper 1997, Rodríguez et al. 2012). The coefficient of NBI indicates that increasing unsatisfied basic needs promote deforestation. There is a wealth of literature connecting the agricultural frontier to land-use change (Scherr 2000, Barbier 2012*a*, Pokorny et al. 2013). Briefly, roads and waterways serve as means for colonist migration sometimes promoted through specific projects (e.g., oil development, or allocation of land titles) to abundant forested land (Rudel and Roper 1997). Newly settled areas are then quickly cleared to establish ownership, extract as much of its natural resources as possible, or both (Southgate 1990, Fearnside 2005). Both population growth and poverty contribute to deforesting this frontier, not as driving factors, but as sources of smallholders to colonize newly opened lands (Fearnside 1993, Lambin et al. 2001). The end result of frontier settlement is a largely deforested landscape (Rodríguez et al. 2012), sometimes economically developed but more often not (Barbier 2012*b*). Poor financial returns from smallholder agriculture, policies that effectively promote consolidated landholdings and ranching, and more forested land available discourage long-term sustainable practices and promote further deforestation (Hecht 1993, Coomes et al. 2011). The strong association between NBI and deforestation supports policy initiatives aiming to shift incentives against forest exploitation, and promote sustainable land ownership and practices at the frontier (Rodrigues et al. 2009, Dulal et al. 2012).

Additionally, in Colombia, the forest frontier and high NBI are also associated with the presence of armed groups. Armed conflict resulting in population displacement concentrates poverty into areas where state institutions are least effective (Díaz and Sánchez 2004, Ibáñez and Moya 2010). Municipalities with reduced state presence are also more likely to have forest cleared for coca cultivation (Dion and Russler 2008). Our results did not find illicit crops to be a significant predictor of deforestation (Table 2). Multiple analyses with independent deforestation datasets have found that coca cultivation does not explain deforestation once socioeconomic characteristics are included as covariates (Dávalos et al. 2011, Sánchez-Cuervo and Aide 2013). Violent conflict and displacement partly reflected in NBI and otherwise absent from our models are important predictors of deforestation in Colombia (Sanchez-Cuervo and Aide 2013, Fergusson et al. 2014).

These findings suggest that communities in the process of development exhibit the highest rates of deforestation (Rodríguez et al. 2012). Continued colonization of forested land, whether from exploitation of land resources and/or displacement caused by political instability (López-Carr 2008), will maintain rates of deforestation despite increasing urbanization. The rapidly changing and receding forest frontier of rural colonization is shared across many developing tropical countries with abundant forests (Rudel and Roper 1997, Barbier 2004, López-Carr and Burgdorfer 2013). With surrogates for colonization as the most potent driver of deforestation in our analyses, initiatives to curtail deforestation must stabilize the advancing frontier and provide economic incentives for conservation among rural communities (Blom et al. 2010).

The results demonstrate that urban population density and road density are significantly associated with deforestation. The density of road networks, which allows for easier resource extraction and encourages the conversion of forests to pastures (De Luca 2007, Barber et al. 2014, Newman et al. 2014), is consistently associated with deforestation, but to a lesser extent than NBI. Road development, both paved and unpaved, has been repeatedly identified as a key step to the clearing of forests (Pfaff 1999, Soares-Filho et al. 2004), providing access into otherwise impenetrable forest regions in less developed municipalities (Arima et al. 2005, Pfaff et al. 2007). Forest patches closer to developed urban areas, particularly developing regions, such as the Amazon, are more convenient targets to illegal deforestation (Laurance et al. 2002). Settlement of areas beyond the more developed Andean region required initial spontaneous or directed colonization followed by local resource extraction (Schuurman 1978, Rudel 2007). Urbanization represents a later stage in settlement when demographic cores are more fully established and growing. At this point, local land value increases as economic incentives pull toward converting forest to pasture and industrial agriculture (Rudel et al. 2009, Seto et al. 2010, Dávalos et al. 2014).

The proportion of area protected had a negative association with deforestation. Although protected areas are not necessarily free from deforestation (Armenteras et al. 2011*b*, Rodríguez et al. 2013*b*), our results confirm nationally protected land, including indigenous reserves, fare better against deforestation than territories without legal protection. This has been observed in the Amazon region of Brazil as well (Nepstad et al. 2006). Globally, the establishment of protected areas reduces deforestation although deterrent effects tend to be weaker in areas further from roads and urban areas (Joppa and Pfaff 2010, Geldmann et al. 2013). The association between forest growth and protected areas is strongest in the Caribbean and Pacific regions (Table 4), which are slightly more developed than in the Orinoco and Amazon basins. Legal protection impedes the development of roads, both paved and unpaved, which are critical for supporting deforestation operations (Nepstad et al. 2009). Additionally, insecure property rights promotes deforestation by providing incentive for landholders to clear forests, grow crops, and build structures to claim land (Araujo et al. 2009). Legal protection prevents this ambiguity in property rights (Fearnside 2001). Analyses based in Costa Rica and Thailand have indicated that protected areas promote reduction in local poverty (Andam et al. 2010). The mechanism for this alleviation in poverty is increased tourism (Ferraro and Hanauer 2014), which enforces economic incentives for keeping forests intact.

Increased deforestation rates of lowlands likely drive the negative association between elevation and deforestation. Colombia's colonial-era settlement started along the Andes mountain range and along the Caribbean shore, so that natural forests were already reduced in the densely populated mid-elevation regions of the country by the beginning of the 20th century (Etter et al. 2008). Exploitation of remnant lowland forests has increased on a global scale in the past few decades with rising urban populations and increased international trade (DeFries et al. 2010). Present major targets of deforestation in the Pacific and the Amazon regions are in low-elevation areas (Etter et al. 2006). Additionally, areas of reforestation have been identified in high-elevation parts of the Andes (Sánchez-Cuervo et al. 2012, Sanchez-Cuervo and Aide 2013). Our results corroborate these findings across all natural regions.

Regional determinants of deforestation

Forest cover and fires were modeled as region-specific variables for the geospatial LME model because their sign changed between natural regions in the RIS model. When region-specific effects were not modeled, initial forest cover and fires held a positive association with forest growth (Table 4). Initial forest cover is a measure of how much forest cover was available for exploitation, and fires indirectly measures clear-cutting activity, though the proportion of fires associated with forest clearing may differ among Colombia's natural regions (Armenteras et al. 2011*a*). Deforestation builds upon prior activity in the more well-developed sections of the natural regions (Soares-Filho et al. 2004), and our analyses were consistent with this pattern. In the spatially explicit LME model, municipalities with greater proportions of forest cover experienced reduced deforestation in the Amazon, Caribbean, and Pacific natural regions. The Amazon and Pacific regions had larger proportions of municipalities with forest cover at the start of the study period (Appendix S1: Table S3). In these lowland regions, municipalities with little forest cover are part of the agricultural frontier and experience rapid land-use change. The effect of the forested proportion was the opposite in the Andes and Orinoco basins, both of which generally have low forest cover (Appendix S1: Fig. S2). The Andes are heavily developed and forests have been reduced to smaller remnants decades ago (Etter and Villa 2000, Armenteras et al. 2011*b*). The Orinoco Basin has wide swaths of natural savanna interspersed with gallery forests that make up a smaller proportion of the municipalities there. The deforestation that does occur in both of these regions is located in municipalities with large remaining patches of forest. These results highlight the importance of exploring region-specific effects, as whole-sample analyses of municipalities (e.g., Dávalos et al. 2011) would fail to model these important differences.

Fires exhibited sign change between natural regions with the RIS model. The signs of the coefficient for fires between the RIS and LME models remained the same with the exception of the Pacific natural region (Table 4). The Pacific coast, one of the wettest regions in the world with mean annual precipitation ranging from 3000 to 12,000 mm, records few fires, and thus, the model coefficient for fires had little effect on the prediction of deforestation rates. The strongest signal for fires as a covariate of deforestation comes from the western portion of the Caribbean natural region (Fig. 4). Fires in the western Caribbean region are associated with clear-cutting preceding development of agricultural land, and this association with deforestation is lost for municipalities east of the Magdalena River. Clear-cutting in the western Caribbean coast is a destructive yet locally constrained phenomenon (Chadid et al. 2015). The effect of fires is not as strong elsewhere in the country. This makes the western Caribbean an ideal region for using MODIS remote sensing data to detect forest fires and identify sites of potentially ongoing deforestation (Yu et al. 2005). Prior use of moderate-resolution imaging spectroradiometer, or MODIS data to understand associations between fires and deforestation in the Colombian Amazon has demonstrated the use of fires for conversion of forest to pastures (Armenteras et al. 2013*b*).

While other variables maintained sign consistency in relation to deforestation, the magnitude of coefficient values exhibited differences between natural regions. This is particularly true for the coefficients of protected areas and road density, for which the median values of certain natural regions fell outside the highest probability density interval of other regions (Table 4). This indicates important regional differences in the effects of these variables. Important related attributes that have not been included in the models, such as enforcement of protected areas (Pfaff et al. 2014), and differentiation between settlement types for road density (Perz et al. 2013), may differ between the natural regions and cause the deviation seen in median coefficient values.

Conclusions

We compared four models of deforestation differing in their treatment of spatial autocorrelation. Analyses of autocorrelation show that non-spatial models will incur deflated residual variance, and are thus subject to high type I error for spatially clustered covariates. The failure of RIS models indicates that modeling covariates at a region-specific scale is not enough to correct for the underlying spatial structure of the data, resulting in model bias. Using geospatial LME or CAR models effectively addresses these biases. Posterior predictive checks demonstrate that the spatially explicit CAR model outperforms non-spatial GLM and RIS models.

Compared to previous analyses of deforestation in Colombia, our analyses revealed that differences in coefficients for deforestation covariates do not always correspond to large natural regions. Critical differences between the eastern and western Caribbean regions were identified using posterior predictive checks. Differences in coefficient values were also found between natural regions for protected areas and road density, indicating that these variables may be further affected by unmodeled landscape features. While all models failed to predict outliers of extreme deforestation or relatively high forest growth, these spatially explicit analyses further link frontier development and road construction to deforestation. Additionally, legal protection of land has a demonstrable positive association with forest growth. Continued protection, and further extension to high-risk areas, is a critical component of forest conservation in Colombia.

Acknowledgments

We thank L. Yohe, M. Lim, and E. Lauterbur for helpful comments on the manuscript. We also thank the Instituto de Hidrología, Meterorología y Estudios Ambientales (IDEAM), for providing data maps.

Literature Cited

Aide, T. M., M. L. Clark, H. R. Grau, D. López-Carr, M. A. Levy, D. Redo, M. Bonilla-Moheno, G. Riner, M. J. Andrade-Núñez, and M. Muñiz. 2013. Deforestation and reforestation of Latin America and the Caribbean (2001–2010). Biotropica 45:262–271.

Andam, K. S., P. J. Ferraro, K. R. Sims, A. Healy, and M. B. Holland. 2010. Protected areas reduced poverty in Costa Rica and Thailand. Proceedings of the National Academy of Sciences of the United States of America 107:9996–10001.

Araujo, C., C. A. Bonjean, J.-L. Combes, P. C. Motel, and E. J. Reis. 2009. Property rights and deforestation in the Brazilian Amazon. Ecological Economics 68:2461–2468.

Arima, E. Y., R. T. Walker, S. G. Perz, and M. Caldas. 2005. Loggers and forest fragmentation: behavioral models of road building in the Amazon basin. Annals of the Association of American Geographers 95:525–541.

Armenteras, D., E. Cabrera, N. Rodríguez, and J. Retana. 2013*a*. National and regional determinants of tropical deforestation in Colombia. Regional Environmental Change 13:1181–1193.

Armenteras, D., J. Retana-Alumbreros, R. Molowny-Horas, R. M. Roman-Cuesta, F. Gonzalez-Alonso, and M. Morales-Rivas. 2011*a*. Characterising fire spatial pattern interactions with climate and vegetation in Colombia. Agricultural and Forest Meteorology 151:279–289.

Armenteras, D., N. Rodríguez, and J. Retana. 2013*b*. Landscape dynamics in northwestern Amazonia: an assessment of pastures, fire and illicit crops as drivers of tropical deforestation. PLoS ONE 8: e54310.

Armenteras, D., N. Rodríguez, J. Retana, and M. Morales. 2011*b*. Understanding deforestation in montane and lowland forests of the Colombian Andes. Regional Environmental Change 11:693–705.

Barber, C. P., M. A. Cochrane, C. M. Souza, and W. F. Laurance. 2014. Roads, deforestation, and the mitigating effect of protected areas in the Amazon. Biological Conservation 177:203–209.

Barbier, E. B. 2004. Explaining agricultural land expansion and deforestation in developing countries. American Journal of Agricultural Economics 86: 1347–1353.

Barbier, E. B. 2012*a*. Natural capital, ecological scarcity and rural poverty. World Bank Policy Research Working Paper. https://doi.org/10.1596/1813-9450-6232

Barbier, E. B. 2012*b*. Scarcity, frontiers and development. Geographical Journal 178:110–122.

Beale, C. M., J. J. Lennon, J. M. Yearsley, M. J. Brewer, and D. A. Elston. 2010. Regression analysis of spatial data. Ecology Letters 13:246–264.

Bjørnstad, O. 2009. ncf: spatial nonparametric covariance functions. R package v 1.1-3.

Blom, B., T. Sunderland, and D. Murdiyarso. 2010. Getting REDD to work locally: lessons learned from integrated conservation and development projects. Environmental Science & Policy 13:164–172.

Brown, D. G., P. H. Verburg, R. G. Pontius, and M. D. Lange. 2013. Opportunities to improve impact, integration, and evaluation of land change models. Current Opinion in Environmental Sustainability 5:452–457.

Chadid, M. A., L. M. Dávalos, J. Molina, and D. Armenteras. 2015. A Bayesian spatial model highlights distinct dynamics in deforestation from coca and pastures in an Andean biodiversity hotspot. Forests 6:3828–3846.

Cliff, A. D., and K. Ord. 1970. Spatial autocorrelation: a review of existing and new measures with applications. Economic Geography 46:269–292.

Cliff, A., and K. Ord. 1972. Testing for spatial autocorrelation among regression residuals. Geographical Analysis 4:267–284.

Coomes, O. T., Y. Takasaki, and J. M. Rhemtulla. 2011. Land-use poverty traps identified in shifting cultivation systems shape long-term tropical forest cover. Proceedings of the National Academy of Sciences of the United States of America 108:13925–13930.

DANE [Departamento Administrativo Nacional de Estadística]. 1985. Censo nacional. Departamento Administrativo Nacional de Estadística, Bogotá, D.C., Colombia.

DANE [Departamento Administrativo Nacional de Estadística]. 1993. Censo nacional de población y vivienda. Departamento Administrativo Nacional de Estadística, Bogotá, D.C., Colombia.

DANE [Departamento Administrativo Nacional de Estadística]. 2008. Sistema de Consulta Información Censal: Censo 2005. Departamento Administrativo Nacional de Estadística, Bogotá, D.C., Colombia.

Dávalos, L. M., A. C. Bejarano, M. A. Hall, H. L. Correa, A. Corthals, and O. J. Espejo. 2011. Forests and drugs: coca-driven deforestation in tropical biodiversity hotspots. Environmental Science & Technology 45:1219–1227.

Dávalos, L. M., J. S. Holmes, N. Rodríguez, and D. Armenteras. 2014. Demand for beef is unrelated to pasture expansion in northwestern Amazonia. Biological Conservation 170:64–73.

De Luca, G. D. 2007. Roads, development and deforestation: a review. World Bank Development Research Group Paper, World Bank and CRED, University of Namur, Washington, D.C., USA and Namur, Wallonia, Belgium.

DeFries, R. S., T. Rudel, M. Uriarte, and M. Hansen. 2010. Deforestation driven by urban population growth and agricultural trade in the twenty-first century. Nature Geoscience 3:178–181.

Díaz, A. M., and F. Sánchez. 2004. A geography of illicit crops (coca leaf) and armed conflict in Colombia. Development Research Centre, London, UK.

Dion, M. L., and C. Russler. 2008. Eradication efforts, the state, displacement and poverty: explaining coca cultivation in Colombia during Plan Colombia. Journal of Latin American Studies 40:399–421.

Dulal, H. B., K. U. Shah, and C. Sapkota. 2012. Reducing emissions from deforestation and forest degradation (REDD) projects: lessons for future policy design and implementation. International Journal of Sustainable Development & World Ecology 19:116–129.

Etter, A., C. McAlpine, and H. Possingham. 2008. Historical patterns and drivers of landscape change in Colombia since 1500: a regionalized spatial approach. Annals of the Association of American Geographers 98:2–23.

Etter, A., C. McAlpine, D. Pullar, and H. Possingham. 2006. Modelling the conversion of Colombian lowland ecosystems since 1940: drivers, patterns and rates. Journal of Environmental Management 79:74–87.

Etter, A., and L. A. Villa. 2000. Andean forests and farming systems in part of the Eastern Cordillera (Colombia). Mountain Research and Development 20:236–245.

Fearnside, P. M. 1993. Deforestation in Brazilian Amazonia: the effect of population and land tenure. Ambio-Journal of Human Environment Research and Management 22:537–545.

Fearnside, P. M. 2001. Land-tenure issues as factors in environmental destruction in Brazilian Amazonia: the case of southern Pará. World Development 29:1361–1372.

Fearnside, P. M. 2005. Deforestation in Brazilian Amazonia: history, rates, and consequences. Conservation Biology 19:680–688.

Fergusson, L., D. Romero, and J. F. Vargas. 2014. The environmental impact of civil conflict: the deforestation effect of paramilitary expansion in Colombia. Serie Documento CEDE, 2014-36. Centro de Estudios sobre Desarollo Económico, Bogotá, Colombia.

Ferraro, P. J., and M. M. Hanauer. 2014. Quantifying causal mechanisms to determine how protected areas affect poverty through changes in ecosystem services and infrastructure. Proceedings of the National Academy of Sciences of the United States of America 111:4332–4337.

Geldmann, J., M. Barnes, L. Coad, I. D. Craigie, M. Hockings, and N. D. Burgess. 2013. Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. Biological Conservation 161:230–238.

Gelman, A. 2005. Analysis of variance-why it is more important than ever. Annals of Statistics 33:1–53.

Gelman, A., and J. Hill. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York, New York, USA.

Getis, A. 2007. Reflections on spatial autocorrelation. Regional Science and Urban Economics 37:491–496.

Gullison, R. E., P. C. Frumhoff, J. G. Canadell, C. B. Field, D. C. Nepstad, K. Hayhoe, R. Avissar, L. M. Curran, P. Friedlingstein, and C. D. Jones. 2007. Tropical forests and climate policy. Science 316:985.

Hargrave, J., and K. Kis-Katos. 2013. Economic causes of deforestation in the Brazilian Amazon: a panel data analysis for the 2000s. Environmental and Resource Economics 54:471–494.

Hecht, S. B. 1993. The logic of livestock and deforestation in Amazonia. BioScience 43:687–695.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978.

Hunter, G. J., A. K. Bregt, G. B. Heuvelink, S. De Bruin, and K. Virrantaus. 2009. Spatial data quality: problems and prospects. Pages 101–121 *in* G. Navratil, editor. Research trends in geographic information science. Springer, Berlin, Germany.

Ibáñez, A. M., and A. Moya. 2010. Do conflicts create poverty traps? Asset losses and recovery for displaced households in Colombia. Pages 137–172 *in* R. Di Tella, S. Edwards, and E. Schargrodsky, editors. The economics of crime: lessons for and from Latin America. University of Chicago Press, Chicago, Illinois, USA.

IDEAM. 2000. Informe nacional del agua. Instituto de Estudios Ambientales y Meteorologicos, Bogotá, D.C., Colombia.

IDEAM, IGAC, IAvH, Invemar, Sinchi, IIAP. 2007. Ecosistemas continentales, costeros y marinos de Colombia. Instituto de Hidrología, Meteorología y Estudios Ambientales, Instituto Geográfico Agustín Codazzi, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Instituto de Investigaciones Marinas y Costeras José Benito Vives de Andréis, Instituto Amazónico de Investigaciones Científicas Sinchi e Instituto de Investigaciones Ambientales del Pacífico Jhon von Neumann, Bogotá, D.C., Colombia.

IGAC [Instituto Geográfico Agustín Codazzi]. 2005. Cartografía básica escala 1:500.000 (cubrimiento nacional). Instituto Geográfico Agustín Codazzi, Bogotá, D.C., Colombia.

IGAC [Instituto Geográfico Agustín Codazzi]. 2011. Sistema de información geográfica para la planeación y el ordenamiento territorial nacional SIGOT. Instituto Geográfico Agustín Codazzi, Bogotá, D.C., Colombia.

Joppa, L. N., and A. Pfaff. 2010. Global protected area impacts. Proceedings of the Royal Society of London. Series B: Biological Sciences 278:1633–1638.

Kaimowitz, D., B. Mertens, S. Wunder, and P. Pacheco. 2004. Hamburger connection fuels Amazon destruction. Center for International Forest Research, Bangor, Indonesia.

Koenig, W. D. 1999. Spatial autocorrelation of ecological phenomena. Trends in Ecology & Evolution 14:22–26.

Kruskal, W. 1988. Miracles and statistics: the casual assumption of independence. Journal of the American Statistical Association 83:929–940.

Kühn, I., and C. F. Dormann. 2012. Less than eight (and a half) misconceptions of spatial analysis. Journal of Biogeography 39:995–998.

Lam, N. S.-N. 1983. Spatial interpolation methods: a review. American Cartographer 10:129–150.

Lambin, E. F., B. L. Turner, H. J. Geist, S. B. Agbola, A. Angelsen, J. W. Bruce, O. T. Coomes, R. Dirzo, G. Fischer, and C. Folke. 2001. The causes of land-use and land-cover change: moving beyond the myths. Global Environmental Change 11:261–269.

Laurance, W. F., A. K. Albernaz, G. Schroth, P. M. Fearnside, S. Bergen, E. M. Venticinque, and C. Da Costa. 2002. Predictors of deforestation in the Brazilian Amazon. Journal of Biogeography 29:737–748.

Lee, D. 2013. CARBayes: an R Package for Bayesian spatial modeling with conditional autoregressive priors. Journal of Statistical Software 55:1–24.

Legendre, P. 1993. Spatial autocorrelation – Trouble or new paradigm. Ecology 74:1659–1673.

Legendre, P., and M. J. Fortin. 1989. Spatial pattern and ecological analysis. Vegetatio 80:107–138.

Lennon, J. J. 2000. Red-shifts and red herrings in geographical ecology. Ecography 23:101–113.

Leroux, B. G., X. Lei, and N. Breslow. 2000. Estimation of disease rates in small areas: a new mixed model for spatial dependence. Pages 179–191 *in* M. E. Halloran and D. Berry, editors. Statistical models in epidemiology, the environment, and clinical trials. Springer, New York, New York, USA.

López-Carr, D. 2008. Population and deforestation: why rural migration matters. Progress in Human Geography 33:355–378.

López-Carr, D., and J. Burgdorfer. 2013. Deforestation drivers: population, migration, and tropical land use. Environment: Science and Policy for Sustainable Development 55:3–11.

Mainwaring, S. 2006. The crisis of representation in the Andes. Journal of Democracy 17:13–27.

Montenegro, E. C., D. M. V. Galvis, G. G. García, M. C. García Dávila, M. F. Ordóñez Castro, L. K. Vergara, A. M. Pascagaza, J. C. Rubiano, and P. G. Rodriguez. 2011. Memoria Técnica de la Cuantificación de la Deforestación Histórica Nacional Escalas Gruesa y Fina. Instituto de Hidrología, Meteorología, y Estudios Ambientales, Bogotá, Colombia, USA.

NASA. 2015. Fire Information for Resource Management System (FIRMS). http://earthdata.nasa.gov/data/near-real-time-data/firms/active-fire-data

Nepstad, D., S. Schwartzman, B. Bamberger, M. Santilli, D. Ray, P. Schlesinger, P. Lefebvre, A. Alencar, E. Prinz, and G. Fiske. 2006. Inhibition of Amazon deforestation and fire by parks and indigenous lands. Conservation Biology 20:65–73.

Nepstad, D., B. S. Soares-Filho, F. Merry, A. Lima, P. Moutinho, J. Carter, M. Bowman, A. Cattaneo, H. Rodrigues, and S. Schwartzman. 2009. The end of deforestation in the Brazilian Amazon. Science 326:1350–1351.

Newman, M. E., K. P. McLaren, and B. S. Wilson. 2014. Assessing deforestation and fragmentation in a tropical moist forest over 68 years; the impact of roads and legal protection in the Cockpit Country, Jamaica. Forest Ecology and Management 315:138–152.

Nolte, C., A. Agrawal, K. M. Silvius, and B. S. Soares-Filho. 2013. Governance regime and location influence avoided deforestation success of protected areas in the Brazilian Amazon. Proceedings of the National Academy of Sciences of the United States of America 110:4956–4961.

Ord, J. K., and A. Getis. 1995. Local spatial autocorrelation statistics: distributional issues and an application. Geographical Analysis 27:286–306.

Overmars, K., G. De Koning, and A. Veldkamp. 2003. Spatial autocorrelation in multi-scale land use models. Ecological Modelling 164:257–270.

Perz, S. G., Y. Qiu, Y. Xia, J. Southworth, J. Sun, M. Marsik, K. Rocha, V. Passos, D. Rojas, and G. Alarcón. 2013. Trans-boundary infrastructure and land cover change: highway paving and community-level deforestation in a tri-national frontier in the Amazon. Land Use Policy 34:27–41.

Pfaff, A. S. 1999. What drives deforestation in the Brazilian Amazon?: evidence from satellite and socioeconomic data. Journal of Environmental Economics and Management 37:26–43.

Pfaff, A., J. Robalino, E. Lima, C. Sandoval, and L. D. Herrera. 2014. Governance, location and avoided deforestation from protected areas: Greater restrictions can have lower impact, due to differences in location. World Development 55:7–20.

Pfaff, A., J. Robalino, R. Walker, S. Aldrich, M. Caldas, E. Reis, S. Perz, C. Bohrer, E. Arima, and W. Laurance. 2007. Road investments, spatial spillovers, and deforestation in the Brazilian Amazon. Journal of Regional Science 47:109–123.

Pinheiro, J., D. Bates, S. DebRoy, and D. Sarkar. 2011. R Development Core Team. 2010. nlme: linear and nonlinear mixed effects models. R package version 3.1-97. R Foundation for Statistical Computing, Vienna, Austria.

Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing. Technische Universität Wien, Vienna, Austria.

Pokorny, B., W. de Jong, J. Godar, P. Pacheco, and J. Johnson. 2013. From large to small: reorienting rural development policies in response to climate change, food security and poverty. Forest Policy and Economics 36:52–59.

Portocarrero-Aya, M., G. Corzo, A. Diaz-Pulido, M. F. González, M. Longo, L. Mesa, A. Paz, W. Ramírez, and O. L. Hernández-Manrique. 2014. Systematic conservation assessment for most of the Colombian territory as a strategy for effective biodiversity conservation. Natural Resources 5:981.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rodrigues, A. S., R. M. Ewers, L. Parry, C. Souza, A. Veríssimo, and A. Balmford. 2009. Boom-and-bust development patterns across the Amazon deforestation frontier. Science 324:1435–1437.

Rodríguez, N., D. Armenteras, and J. R. Alumbreros. 2013*a*. Land use and land cover change in the Colombian Andes: dynamics and future scenarios. Journal of Land Use Science 8:154–174.

Rodríguez, N., D. Armenteras, R. Molowny-Horas, and J. Retana. 2012. Patterns and trends of forest loss in the Colombian Guyana. Biotropica 44:123–132.

Rodríguez, N., D. Armenteras, and J. Retana. 2013*b*. Effectiveness of protected areas in the Colombian Andes: deforestation, fire and land-use changes. Regional Environmental Change 13:423–435.

Rudel, T. K. 2007. Changing agents of deforestation: from state-initiated to enterprise driven processes, 1970–2000. Land Use Policy 24:35–41.

Rudel, T. K., R. Defries, G. P. Asner, and W. F. Laurance. 2009. Changing drivers of deforestation and new opportunities for conservation. Conservation Biology 23:1396–1405.

Rudel, T., and J. Roper. 1997. The paths to rain forest destruction: crossnational patterns of tropical deforestation, 1975–1990. World Development 25:53–65.

Sanchez-Cuervo, A. M., and T. M. Aide. 2013. Identifying hotspots of deforestation and reforestation in Colombia (2001–2010): implications for protected areas. Ecosphere 4:art143.

Sánchez-Cuervo, A. M., and T. M. Aide. 2013. Consequences of the armed conflict, forced human displacement, and land abandonment on forest cover change in Colombia: a multi-scaled analysis. Ecosystems 16:1052–1070.

Sánchez-Cuervo, A. M., T. M. Aide, M. L. Clark, and A. Etter. 2012. Land cover change in Colombia: surprising forest recovery trends between 2001 and 2010. PLoS ONE 7:e43943.

Scherr, S. J. 2000. A downward spiral? Research evidence on the relationship between poverty and natural resource degradation. Food Policy 25:479–498.

Schuurman, F. J. 1978. From resource frontier to periphery agricultural colonization east of the Andes. Tijdschrift voor economische en sociale geografie 69:95–104.

Seto, K. C., R. Sánchez-Rodríguez, and M. Fragkias. 2010. The new geography of contemporary urbanization and the environment. Annual Review of Environment and Resources 35:167–194.

Soares-Filho, B., A. Alencar, D. Nepstad, G. Cerqueira, V. Diaz, M. del Carmen, S. Rivero, L. Solórzano, and E. Voll. 2004. Simulating the response of land-cover changes to road paving and governance along a major Amazon highway: the Santarém-Cuiabá corridor. Global Change Biology 10:745–764.

Soares-Filho, B., P. Moutinho, D. Nepstad, A. Anderson, H. Rodrigues, R. Garcia, L. Dietzsch, F. Merry, M. Bowman, and L. Hissa. 2010. Role of Brazilian Amazon protected areas in climate change mitigation. Proceedings of the National Academy of Sciences of the United States of America 107: 10821–10826.

Southgate, D. 1990. The causes of land degradation along" spontaneously" expanding agricultural frontiers in the Third World. Land Economics 66:93–101.

Su, Y. S., and M. Yajima. 2015. R2jags: a package for running jags from R. http://cran.r-project.org/package=R2jags

Telford, R., and H. Birks. 2005. The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance. Quaternary Science Reviews 24:2173–2179.

UNFCCC [United Nations Framework Convention on Climate Change]. 2015. Adoption of the Paris Agreement -/CP21. UNFCCC, Paris, France.

UNODC, Government of Colombia. 2006. Colombia: Coca cultivation survey. United Nations Office on Drugs and Crime, Bogotá, Colombia. http://www.unodc.org/pdf/andean/Colombia_coca_survey_2005_eng.pdf

Yu, L., N. Wang, and X. Meng. 2005. Real-time forest fire detection with wireless sensor networks. Pages 1214–1217 *in* H. Zhou, editor. Proceedings of International Conference on Wireless Communications, Networking and Mobile Computing, Wuhan, China, September 23–26, 2005. Institute of Electrical & Electronics Engineers, New York, New York, USA.

Supporting Information

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1824/full