Preprint of: Vanclay, J.K., D. Kaimowitz, A. Puntodewo and P.Mendez, 1999. Spatially explicit model of deforestation in Bolivia. *In*: Y. Laumonier, B.King, C.Legg and K. Rennolls (eds) *Data Management and Modelling using Remote Sensing and GIS for Tropical Forest Land Inventory*. Rodeo, Jakarta, p. 371-382. ISBN 979-95696-0-5.

Spatially Explicit Model of Deforestation in Bolivia

Jerome Vanclay[1], David Kaimowitz[1], Atie Puntodewo[1], Patricia Mendez[2]

1. Center for International Forestry Research, Bogor, Indonesia, Fax +62 251 622100,
cifor@cgiar.org, http://www.cgiar.org/cifor

2. Land Use Planning Technical Office (OT-PLUS), Departmental Government (Prefectura), Santa Cruz, Bolivia

Abstract

A GIS compiled by the Departmental Government of Santa Cruz, Bolivia offers data that may help to resolve some competing theories of tropical deforestation. The GIS contains many attributes relating to land use at two points in time, 1989 and 1994, and allow us to address questions like:

- What has been the impact of past road construction on deforestation and land use?
- What impacts might be expected from future road construction?
- What impact do zoning policies such as forest concessions and protected areas have?
- What influence do cultural factors have on forest clearing and fragmentation?

We discuss our methodology and report interim results. We seek to provoke discussion on appropriate statistical procedures for such analyses.

Introduction

Tropical deforestation is topical and controversial, and many researchers and agencies would like to better understand when, where, and more importantly, why it occurs. Kaimowitz and Angelsen (1998) reported the existence of 150 deforestation models, most of which were developed since 1990. They found that in all, more than115 different variables had been used in these attempts to explain deforestation and that major uncertainties continue to exist about how most of these influence deforestation. We take this to be an indication of the inherent complexity of the task, the scarcity of decisive indicators, and the limitations of proxy variables used in these studies.

Among different possible modeling approaches, Kaimowitz and Angelsen (1998) concluded that household and regional-level studies are likely to be more productive than national and global studies. They expressed particular enthusiasm for the potential of the growing availability of spatial data bases providing insights into the role in deforestation processes of such spatial variables as access to markets, land use zoning policies, and ecological conditions. They note that such models use relatively reliable data, involve large sample sizes that give model makers more degrees of freedom to work with, and are particularly suited for predicting where deforestation is likely to occur. In addition, the model's robustness can often be tested by measuring what percentage of the time they accurately predict where deforestation will occur.

This paper presents a spatial econometric model of one particular Latin American region, the department of Santa Cruz in eastern Bolivia. Several reasons inspired us to study deforestation patterns in that region:

- Deforestation in the Bolivian tropics has historically been limited but has increased rapidly in recent years, and it is important to understand why;

- Deforestation patterns in Bolivia differ significantly from other areas in Latin America in that the expansion of large-scale mechanized agriculture has been more important in the former;
- CIFOR has an on-going international project comparing the effects of different policies and social trends on tropical forests in Bolivia, Cameroon, and Indonesia; and
- The Santa Cruz government (prefectura) had compiled a GIS with much of the data needed to examine the influence of different geographic variables on deforestation trends.

Our objectives in pursuing the present study were three-fold. We wanted to
1) Test some established theories of deforestation,
2) Improve the capacity to formulate land use policies within the Department of Santa Cruz, and
3) Contribute to a better understanding of factors that determine land use in locations similar to those of study area.

Previous Spatial Econometric Deforestation Models

Spatial regression models measure the correlation between land use and other geo-referenced variables such as:
- Transportion costs (distance from markets and road, railways, and rivers),
- Zoning categories (national parks, forest concessions, colonization areas, indigenous territories), and
- Ecological conditions (topography, soil quality, precipitation, and forest fragmentation).

The models focus on land use in a single time period or the change in land use over two or more periods. The majority relate the state of the independent variables in the first period to the probability that the forest in that location is removed between the first and second periods.

Unlike the Santa Cruz model presented below, most previous models have drawn their data from a random sample of locations (points) within a selected region or country. Sample sizes are typically several thousand points or more. Chomitz and Gray (1995) used a multinomial maximum likelihood model with a random sample of 10,000 data points. Tom Tomich (pers comm) examined deforestation rates within a study area of 4.9 million hectares in Jambi, Sumatra, by sub-sampling with a 1 km square grid and using a binomial-probit transformation. Gerald Nelson (pers comm) made a similar study with raster data by taking a systematic 1% sample that yielded 25,000 sample points. He claimed results were "meaningful" because some parameters were found to be statistically significant while others were not.

Some models include all types of locations, others just locations covered with forest during the first time period. Typically, the land use information comes from national forest inventories, remote sensing and aerial photographs.

The models show land holders are more likely to convert forest to agricultural use where good access to markets and favorable conditions for farming make agriculture more profitable and the government has not restricted forest conversion (Table 1). Forests close to roads in physical distance and traveling time are more likely to be cleared (Chomitz and Gray, 1995; Liu *et al.,* 1993; Ludeke *et al.,* 1990; Mertens and Lambin, 1997; Nelson and Hellerstein, 1995; Sader and Joyce, 1988; Rosero-Bixby and Palloni, 1996). Most studies show that forest clearing declines rapidly beyond distances of two or three kilometers from a road, although Liu *et al.* (1993) report significant forest clearing up to around 15 km from the nearest road for the Philippines. Similarly, Chomitz and Gray (1995) found that locations closer to urban markets have less remaining forest in Belize and Mertens and Lambin (1997) reported that deforestation drops off dramatically beyond 10 kilometers from the nearest town in Eastern Cameroon.

Forest fragments have a higher risk of being lost than forests in large compact areas, with forests close to the forest edge especially likely to be cleared (Brown *et al.,* 1993; Liu *et al.,* 1993; Ludeke *et al.,* 1990; Mertens and Lambin, 1997; Rosero-Bixby and Palloni, 1996). In addition, areas with higher quality soils (flat, adequate drainage, and high soil fertility) and drier climates are also more likely to be cleared (Chomitz and Gray, 1995; Gastellu-Etchegorry and Sinulingga, 1988; Sader and Joyce, 1988; Rosero-Bixby and Palloni, 1996).

The effect of roads and environmental conditions may interact. Thus roads may induce greater deforestation in areas with good soils and favorable climatic conditions. In Belize, Chomitz and Gray (1995) showed that the probability of an area being used for agriculture (rather than being retained as natural vegetation) on high quality land next to a road was 50%, whereas lands next to roads with marginal soils had only a 15% probability of being deforested.

Mertens and Lambin (1997) noted that variables affect forest clearing differently depending on the type of deforestation process. In peri-urban deforestation, forest clearing exhibits a circular pattern around the towns, and distance to towns and roads strongly affects forest clearing but proximity to forest edge does not. Roads may exhibit a “corridor pattern” of deforestation where proximity to roads and forest edges are significant determinants of forest clearing, but distance to towns is not. Finally, in areas where diffuse shifting cultivation dominates, proximity to forest edge increases the probability of forest clearing, whereas distance to roads and towns is less important.

Deforestation in Santa Cruz, Bolivia

Department of Santa Cruz extends some 900 by 800 kilometers, and occupies some 35 million hectares (Figure 1). Forest cover estimates based on Landsat data are available for 1989 and 1994 (Morales 1993 and 1996). The total accumulated area of forest cleared by humans prior to 1994 was about 2.1 million hectares or 6% of the land area, most of it is concentrated within about 200 km of the capital city, Santa Cruz. In 1994, some 15 million hectares of forest remained, along with some 1.9 million hectares of agriculture and 3.2 million hectares of pasture or savanna (an increase of 281 thousand hectares since 1989). Much of the savanna and pasture is natural, especially in areas of the Chiquitano Shield, the Pantanal, the Quimome area, and in the sub-Andean zone.

Annual deforestation rates have been increasing rapidly since the mid-1980s. Between 1986 and 1990, CUMAT (1992) found that 38,000 hectares of forest were cleared annually in the Amazonian portion of Santa Cruz. That region covers only 61% of Santa Cruz, but accounts for a much higher percentage of forest clearing. Approximately 78,000 hectares were cleared annually in all of Santa Cruz between 1989 and 1992, rising to 117,000 hectares annually between 1992 and 1994 (Morales 1993 and 1996).

Most deforestation in Santa Cruz is by large mechanized soybean and wheat farmers, small agricultural colonists who practice mainly slash and burn rice and maize cultivation, and large cattle ranchers (Pacheco 1998). The mechanized farm sector has grown rapidly over the last fifteen years, and now accounts for a majority of forest clearing. Most of this growth has been in the area east of the Rio Grande River, known as the “expansion zone”. Small agricultural colonists have expanded into moister forest areas suitable for rice growing to the north and west of the city of Santa Cruz. Forest clearing for pastures is concentrated in northeastern Santa Cruz.

Data

Our data were drawn from a GIS produced by the ‘Santa Cruz National Resource Protection Project’ implemented by the Government of Santa Cruz with funding from KFW and technical assistance from a consortium composed of the IP, SCG, and KWC consulting companies. The initial objective of that GIS was to develop a land use plan (PLUS) for the entire department of Santa Cruz. Hence forth, we will refer to it as the PLUS GIS.

The PLUS GIS was compiled from several sources. Most data were digitized from 1:250,000 maps, but some layers were captured at other scales and obtained from other sources. Many layers obtained were based on the UTM elipsoid IU661967.

GIS layers of particular significance for our study included:

- Land use in 1989, 1992, and 1994 (i.e., several classes of urban, agriculture, forest, etc)
- Vegetation, soil types and rainfall data using a standard classification,
- Details of the road and rail network (including logging/mining roads),
- Administrative data including urban areas, forest concessions, colonization areas, indigenous territories, protected areas, etc.

The land use, vegetation type and soils data were provided in raster form, and were converted to vector format. The forest concession boundaries were obtained from the Sustainable Forestry Management (BOLFOR) Project. The 1989 land use data were compiled from Earthsat Data analyzed by the CUMAT consulting company, and were considering "quite reliable" by Ivan Morales (pers comm), the expert who analyzed the 1992 and 1994 Landsat data.

The 1989 land use data delineate forests, deforested areas, savanna and pastures, areas with little or no vegetation, water, and urban areas. The 1994 land use data further sub-divides the deforested areas into traditional agriculture, commercial agriculture, mixed agriculture, agriculture with forests and forests with agriculture. The 1989 data had a resolution of 1 x 1 km (100 ha), whereas the resolution in 1994 was 250 x 250 m (about 6 ha). Some areas were omitted from the north-west in the 1989 data and from the east in the 1994 data and these areas have been excluded from our study. The cloud cover was minimal in the 1994 images used to assess land use, so this assessment is considered more comprehensive than the previous assessments.

Despite considerable care and attention to detail in compiling the GIS, there were some anomalies that we could not reconcile. Deforestation estimates obtained by calculating the area in agricultural land in 1994 that had been forest in 1989 inexplicably provided different estimates than when we combined all agricultural lands in 1989, 1992, 1993 and 1994, and then subtracted the land already agriculture in 1989. Although the latter approach provided estimates consistent with independent estimates by BOLFOR (namely 552,985 ha), the discrepancy is unsettling.

Methods

An analysis of deforestation of this kind poses many interrelated questions:

- What should we try to predict: deforestation rate 1989-94 or total deforestation to 1994?
- Should we use a binomial (forest, non-forest) or a multinomial model that considers the various end-uses of former forest land?
- How should we transform the dependent variable to make analyses tractable and results meaningful: is it better to use a logarithm, logistic or probit transformation?
- What explanatory variables should we consider in our analysis, and how can we minimize the correlation between these variables?
- How can we efficiently transfer the data between the GIS and the statistics package, while minimizing spatial autocorrelation?[1]

---

[1] Spatial autocorrelation is a common problem with geographic data, since nearby locations are more likely to be similar than distant ones. This can lead to inaccurate measures of statistical significance. Several methods exist for partially correcting for spatial autocorrelation, although none is fully satisfactory (Rosero-Bixby and Palloni, 1996; Chomitz and Gray, 1995).

- How can we tell if we have a problem with multiple or spatial autocorrelation?
- How can we discriminate endogeneous and exogeneous variables?

Somewhat surprisingly, prior studies offer little guidance on these issues.

Theory, initial hypotheses and response variable

Based on economic theory and previous modeling exercises, we hypothesized that the more productive land (i.e., Soil type I, with rainfall exceeding 1000 mm) and land with better access to markets (lower transportation costs) would be cleared first. We anticipated that zoning an area as a forest concession or protected area would impede deforestation, while zoning it as a colonization areas would encourage deforestation. In addition, we hypothesized that indigenous people have cultural attributes that lead to less deforestation.

Although both the total deforestation to date and the recent deforestation rate (1989-94) are of interest, it is the latter that is of most interest, as it is the best indicator of current trends and responses to existing policies. Similarly, although the end-use of deforested land is of interest, the binomial model is more tractable and simplifies analyses. We examined a simple binomial model that considered only land forested in 1989: if deforested during 1989-94 the response variable was coded 0, otherwise it was coded 1. This provides a model that could be used to predict deforestation during 1994-99, and could be checked by making empirical tests of predictions for 1999. To ensure unambiguous regarding deforestation during 1989-94, we deleted from our data set all areas that were not forest in 1989, including areas that were ambiguously defined in the 1989 classification (e.g., cloud, “no data”, etc.).

Economists often favour the use of logarithmic transformations, as parameter estimates can then be interpreted directly as elasticities (i.e., predictor variables are multiplicative, so that a unit change in a predictor variable always causes the same percentage change in the response variable). This may be helpful when all predictor variables are expressed in the same units, but becomes less relevant when the nature of the predictors varies greatly. Statisticians tend to prefer logistic and probit transformations for binomial data because standard assumptions are better satisfied, and predictions are constrained correctly. The probit and logistic transformations are similar in many respects, but my previous experience (Vanclay 1994) inclines me to favour the logistic transformation (weighted for polygon area). Fortunately for economists, the logistic is very similar to the logarithmic transformation if rates of change do not exceed 0.25, so provided deforestation rates remain modest, parameter estimates may still be interpreted as elasticities.

Sampling

Although some statistics packages claim to be able to interface directly with GIS, it is convenient to extract data from the GIS as a simple text file, so that it can be used with any statistics package. However, this raises the question of how best to extract the data: should a sample of selected points be taken, should polygons form the basis for analysis, or should some other alternative be adopted (Figure 2)?

Systematic selection of sample points ensures a compact data set and simplifies analyses, but fails to make full use of the available information. A small sample may be statistically inefficient, but computationally convenient. A larger sample size makes more efficient use of information, but also increases the potential for spatial autocorrelation. Some researchers prefer to sample tiles rather than points, arguing that these are more representative in fragmented landscapes where correct alignment of the various GIS layers may be problematic. If these tiles tessellate the study area (e.g., square tiles rather than circular plots), a complete census may be analysed, but this again introduces the possibility of spatial autocorrelation. One way to make better use of information while minimizing autocorrelation is to stratify (e.g., forested versus deforested, and close to versus distant from town/road), sample strata with different intensities (i.e., sample more intensively in strata of particular interest), and weight regression analyses accordingly (sometimes called choice-based sampling). Another alternative is to use polygons occurring spontaneously within the GIS. Although this may reduce the unnecessary proliferation of sample units, layers contributing to the tessellation need to be chosen carefully to ensure meaningful polygons and avoid problems with omitted variables. Additional problems may occur if some polygons become excessively large, since they may no longer be homogeneous (especially with regard to distance to roads and towns) and this may mask relationships. We chose to adopt the polygon approach, using polygons based on variables identified in our a priori hypotheses (Land use in 1994, Concession, Indigenous, Protected, Colony, Soil type). Other attributes (e.g., rainfall, distances to roads, towns, etc) of polygons are assessed at the polygon centroid. Large polygons were further fragmented using a regular grid to improve homogeneity with respect to distances.

Although spatial autocorrelation is a significant concern, it is only one factor to be considered. One aspect of spatial autocorrelation can be minimized by taking care to avoid omitting key variables. Even where all relevant variables are included in the model, error terms may remain spatially autocorrelated: although this will cause unreliable test statistics, the parameter estimates will remain unbiased (see Nelson nd).

Another issue canvassed in the literature is selection bias, and the effect it may have on parameter estimates. The classic example of this is the observation that high school grades show little impact on college grades, if a sample is restricted to current college students - because students into college despite poor school grades probably had something else going for them. And it may be that in our study there are some observations on good soils close to a road that for some reason remained forested in 1989, perhaps because they are subject to some informal protection about which we have no data. It is not unexpected that they remained forested in 1994, and this may bias our estimates of the impact of roads. There are some techniques that attempt to adjust for the selection bias to gauge the real effect of roads (e.g., the 'heckit' model, Ken Chomitz, pers comm), but the real issue is the definition of the population to which the model applies.

Explanatory variables

Two issues are of particular relevance in the selection of possible explanatory variables:

Multicollinearity: there are many variables of potential interest, and since most are highly correlated with other explanatory variables, a parsimonious selection is necessary be made to ensure a tractable analysis. We chose potential explanatory variables according to our a priori hypotheses, examined the correlation matrix, and monitored parameter estimates to avoid problems with multicollinearity.

Endogeneity: how do we know that roads and towns are not located in places which are good to deforest, rather than deforestation occurring because of the proximity of roads and towns? Although it is reasonable to assume that major highways and logging roads are exogenous with respect to local suitability for agriculture, it seems likely that tertiary rural roads are endogeneous. The problem may be reduced by controlling for agricultural suitability and using independent variables from a time period prior to the dependent variables.

We computed 20 potential explanatory variables (Table 2) each polygon in the GIS, and transferred these to the statistics package S-plus for further analysis. Some transformations of these variables were considered, but appeared to contribute little improvement in predictions. Table 3 reveals the correlations within the data used in our analyses. It is noteworthy that the single best predictor, distance to Santa Cruz, is highly correlated with most other variables.

Results

A series of models were fitted with generalized least squares using a logistic model weighted by polygon area. At each step, we selected the model that offered the greatest reduction in deviance, until at step 6 (see Table 4), no further significant improvements in the model could be obtained ($P<0.01$). All parameter estimates in Table 3 are significantly different from zero ($P<0.01$). The relative stability of parameter estimates as additional terms were included in the model confirms that multicollinearity is not problematic in the models presented.

The parameter estimates obtained are not surprising, and are consistent with a priori hypotheses. In one sense, this is disappointing - a large amount of work has not revealed anything that we did not know already. Perhaps the most interesting result is that indigenous territories are not statistically significant, in contrast to our initial expectation that this tenure class might help to preserve forest. There is empirical evidence to support our contention that in Bolivia, forest concessions do in fact, help to minimize deforestation, in contrast to colonization areas where deforestation is higher (Figure 3).

Although we believe our results to be reliable, we stress that these results are preliminary, and subject to further examination of our GIS data and to further refinements to our method. Care is also required in interpreting the coefficients reported in Table 3. An economist who examines on the size of the regression coefficient will observe the apparently high elasticity of forest fragmentation, as reflected in the parameter estimate for the Matheron Index. However, a graphical analysis suggests a different conclusion - namely that the distance to deforestation is the most important factor. These interpretations differ because the Matheron Index varies between 0 and 0.05, whereas the distance to deforestation varies between 0 and 100 km. Comparisons of our results with parameter estimates for the Matheron Index in other studies should be made cautiously, since we computed the index for a circular neighbourhood within 1 km of the polygon centroid, whereas many other studies compute the index within a 9 x 9 pixel neighbourhood.

Discussion

The results we have obtained warrant further examination, in part because of concerns about the integrity of the GIS, the adequacy of the sampling scheme, and the high correlation between the distance to Santa Cruz and other variables. Two other important factors that may have been usurped by the distance to Santa Cruz include rainfall and soils type, and we wish to re-examine alternative models which exclude the distance to Santa Cruz. We also wish to examine other aspects of soil type, in addition to the 8-point USDA classification. We have details of a number of specific soil characteristics, including erosion susceptibility, drainage, salinity, alkalinity, depth, nutrient status, presence of hard pans, and other factors. We are also interested to explore alternative sampling schemes (see Figure 2).

References

Anselin, L. SPACESTAT for large datasets

Kaimowitz, David and Angelsen, Arild. 1998. Economic Models of Tropical Deforestation, A Review. Bogor: Center for International Forestry Research (CIFOR).

Brown, S., L.R. Iverson, and A. Lugo, 1993. Land Use and Biomass Changes in Peninsular Malaysia During 1972-1982: Use of GIS Analysis. In *Effects of Land-use Change on Atmospheric CO2 Concentrations: Southeast Asia as a Case Study*, ed. Virginia H. Dale (New York: Springer Verlag).

Chomitz, Kenneth M. and David A. Gray. 1995. Roads, Lands, Markets, and Deforestation, A Spatial Model of Land Use in Belize. Policy Research Working Paper 1444, Washington D.C. The World Bank, Policy Research Department, April, 50 p.

CUMAT (Capacidad de Uso Mayor de la Tierra). 1992. “Desbosque de la Amazonia Boliviana”. La Paz: CUMAT.

Gastellu-Etchegorry, J.P. and A.B. Sinulingga. 1988. Designing a GIS for the Study of Forest Evolution in Central Java. *Tijdscrift voor Economische en Sociaale Geografie* 79(2), 93-103.
Limdep manual http://wuecon.wustl.edu/limdep/limmanual.html, chapter 22, bivariate probit models.
Liu, Dawning S., Louis R. Iverson, and Sandra Brown. 1993. Rates and Patterns of Deforestation in the Philippines: Applications of Geographic Information System Analysis. *Forest Ecology and Management* 57, 1-16.
Ludeke, Aaron Kim, Robert C. Maggio, and Leslie M. Reid. 1990. An Analysis of Anthropogenic Deforestation Using Logistic Regression and GIS. *Journal of Environmental Management* 31, 247-59.
Mertens, Benoit and Eric F. Lambin. 1997. "Spatial Modeling of Deforestation in Southern Cameroon" *Applied Geography* 17(2), 1-19.
Morales, I. 1996. Memoria explicativa del monitoreo preliminar del desbosque en el Departamento de Santa Cruz, 1994. CORDECRUZ, Santa Cruz.
Morales, I. 1993. Monitoreo del bosque en el Departamento de Santa Cruz, Periodo 1988/89 - 1992/93. CORDECRUZ, Santa Cruz.
Nelson, Gerald and Daniel Hellerstein. 1997. Do Roads Cause Deforestation? Using Satellite Images in Econometric Analysis of Land Use. Staff paper 95-E488, *American Journal of Agricultural Economics* 79, 80-8.
Pacheco, Pablo. 1998. *Estilos de desarrollo, deforestación y degradación de los bosques en las tierras bajas de Bolivia*. La Paz: CIFOR / CEDLA / TIERRA.
Rosero-Bixby, Luis and Alberto Palloni. 1996. Population and Deforestation in Costa Rica. Paper presented at the Annual Meeting of the Population Association of America, New Orleans.
Sader, S.A. and A.T. Joyce. 1988. Deforestation Rates and Trends in Costa Rica. *Biotropica* 20, 11-9.
Vanclay, J.K., 1994. *Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests*. CAB International, Wallingford, U.K. xvii+312 p. ISBN 0-85198-913-6.

Acknowledgements

This paper represents a collaborative effort between the Center for International Forestry Research (CIFOR) in Bogor, Indonesia, the Natural Resources Department of the Government (Prefectura) of Santa Cruz, Bolivia, and the Bolivian Sustainable Forest Management Project (BOLFOR). The authors wish to express their gratitude to individuals in each of these institutions that have contributed to our efforts, including Sergio Antelo, Andreas Carstens, Francisco Kempff, John Nittler, Christian Vallejos, and Roderich von Offen. Other useful comments and suggestions have come from Ken Chomitz, Ivan Morales, Gery Nelson, and Tom Tomich.

***Table 1. Predictors** of Deforestation*

Study	**Country**	**More roads**	**Closer to markets**	**Better soils &/or drier**	**Nearer forest edge**
Brown *et al.* (1993)	Malaysia	n.a.	n.a.	n.a.	increase
Chomitz & Gray (1995)	Belize	increase	Increase	increase	n.a.
Gastellu-Etchegorry & Sinulingga (1988)	Indonesia	n.a.	n.a.	increase	n.a.
Liu *et al.* (1993)	Philippines	increase	n.a.	n.a.	increase
Ludeke *et al.* (1990)	Honduras	increase	n.a.	increase	increase
Mertens & Lambin (1997)	Cameroon	increase	Increase	n.a.	increase
Nelson & Hellerstein (1995)	Mexico	increase	Increase	n.a.	n.a.
Rosero-Bixby & Palloni (1996)	Costa Rica	increase	n.a.	increase	increase
Sader and Joyce (1988)	Costa Rica	increase	n.a.	increase	n.a.

n.a. = not applicable

Table 2. *Variables extracted from GIS for statistical analysis*

Polygon-ID
Area of the polygon (ha)
Forest-94 (1=still forest, 0=not forest)
Concession (1=inside, 0=outside)
Indigenous territory (1=inside, 0=outside)
Protected area (1=inside, 0=outside)
Colonization area (1=inside, 0=outside)
USDA Soil group (1-8)
Precipitation (mm, for centroid of polygon)
Distance to nearest paved road (km)
Distance to nearest unpaved road (km)
Distance to nearest forestry/mining road (km)
Distance to nearest railroad (km)
Distance to nearest category 1 town (Santa Cruz, km)
Distance to nearest category 2 town (km)
Distance to nearest category 3 town (km)
Distance to nearest category 4 town (km)
Distance to nearest category 5 town (km)
Distance to nearest category 6 town (km)
Distance to nearest non-forest land (km)
Distance to nearest land that was deforested prior to 1989 (km)
Matheron's index for forest/non-forest within 1 km radius of centroid

Table 3. *Correlation matrix*

	Forest	**Conc**	**Indig**	**Prot**	***Colon***	**Soil**	**Rain**	**DR1**	**DR2**	**DLRb**	***DLRo***	**DRR**	**DAT**	***DT1***	**DT4**	**DNF**	***DDF***	**FA1**	***MI***
Forest	1.00	0.14	-0.01	0.08	*-0.19*	0.13	-0.03	0.23	0.23	0.24	0.24	0.16	0.20	*0.27*	0.19	0.27	**0.34**	0.01	0.08
Conc	0.14	1.00	-0.01	0.10	-0.19	-0.01	**0.37**	0.21	0.12	0.10	0.30	0.24	0.25	0.19	0.08	0.18	0.16	0.10	0.11
Indig	-0.01	-0.01	1.00	-0.06	-0.08	0.07	-0.06	-0.10	-0.09	-0.08	-0.06	-0.01	-0.05	0.06	0.07	-0.11	-0.08	-0.03	-0.05
Prot	0.08	0.10	-0.06	1.00	-0.07	0.08	0.05	0.31	0.19	0.20	**0.37**	0.20	0.29	0.12	0.02	0.30	0.24	-0.01	0.00
Colon	-0.19	-0.19	-0.08	-0.07	1.00	-0.14	0.09	-0.20	-0.18	-0.18	-0.18	-0.15	-0.15	-0.34	-0.28	-0.21	-0.22	-0.10	-0.13
Soil	0.13	-0.01	0.07	0.08	-0.14	1.00	0.02	0.04	0.14	0.16	0.15	0.16	0.17	0.23	0.18	0.07	0.13	-0.09	-0.08
Rain	-0.03	**0.37**	-0.06	0.05	0.09	0.02	1.00	0.05	-0.04	-0.09	0.11	0.14	0.08	-0.06	-0.18	-0.05	-0.06	-0.15	-0.15
DR1	0.23	0.21	-0.10	0.31	-0.20	0.04	0.05	1.00	**0.46**	**0.46**	**0.67**	0.04	0.22	**0.48**	0.32	**0.59**	**0.48**	0.13	0.17
DR2	0.23	0.12	-0.09	0.19	-0.18	0.14	-0.04	**0.46**	1.00	**0.97**	0.22	**0.42**	**0.69**	0.23	0.26	**0.56**	**0.60**	0.15	0.19
DLRb	0.24	0.10	-0.08	0.20	-0.18	0.16	-0.09	**0.46**	**0.97**	1.00	0.26	**0.43**	**0.71**	0.26	0.28	**0.57**	**0.62**	0.16	0.19
DLRo	0.24	0.30	-0.06	**0.37**	-0.18	0.15	0.11	**0.67**	0.22	0.26	1.00	0.21	0.27	**0.42**	0.04	**0.37**	0.32	0.08	0.12
DDR	0.16	0.24	-0.01	0.20	-0.15	0.16	0.14	0.04	**0.42**	**0.43**	0.21	1.00	**0.79**	0.16	0.10	0.21	**0.34**	0.01	0.05
DAT	0.20	0.25	-0.05	0.29	-0.15	0.17	0.08	0.22	**0.69**	**0.71**	0.27	**0.79**	1.00	0.21	0.17	**0.35**	**0.45**	0.07	0.10
DT1	0.27	0.19	0.06	0.12	-0.34	0.23	-0.06	**0.48**	0.23	0.26	**0.42**	0.16	0.21	1.00	**0.88**	**0.45**	**0.45**	0.14	0.17
DT4	0.19	0.08	0.07	0.02	-0.28	0.18	-0.18	0.32	0.26	0.28	0.04	0.10	0.17	**0.88**	1.00	**0.45**	**0.43**	0.15	0.17
DNF	0.27	0.18	-0.11	0.30	-0.21	0.07	-0.05	**0.59**	**0.56**	**0.57**	**0.37**	0.21	**0.35**	**0.45**	**0.45**	1.00	**0.67**	0.20	0.25
DDF	**0.34**	0.16	-0.08	0.24	-0.22	0.13	-0.06	**0.48**	**0.60**	**0.62**	0.32	**0.34**	**0.45**	**0.45**	**0.43**	**0.67**	1.00	0.14	0.21
FA1	0.01	0.10	-0.03	-0.01	-0.10	-0.09	-0.15	0.13	0.15	0.16	0.08	0.01	0.07	0.14	0.15	0.20	0.14	1.00	**0.85**
MI	0.08	0.11	-0.05	0.00	-0.13	-0.08	-0.15	0.17	0.19	0.19	0.12	0.05	0.10	0.17	0.17	0.25	0.21	**0.85**	1.00

Table 4. *Parameter estimates*

Model	Intercept	Santa Cruz (km)	Deforest.n (km)	Log Road (km)	Colony	Matheron	Concession
1	1.10	0.0144					
2	0.63	0.0059	0.31				
3	0.41	0.0036	0.32	0.0082			
4	0.58	0.0031	0.32	0.0082	-0.46		
5	0.41	0.0029	0.32	0.0081	-0.44	14	
6	0.35	0.0029	0.32	0.0076	-0.39	14	0.35

Figure 3. *Effect of land tenure on deforestation rate*

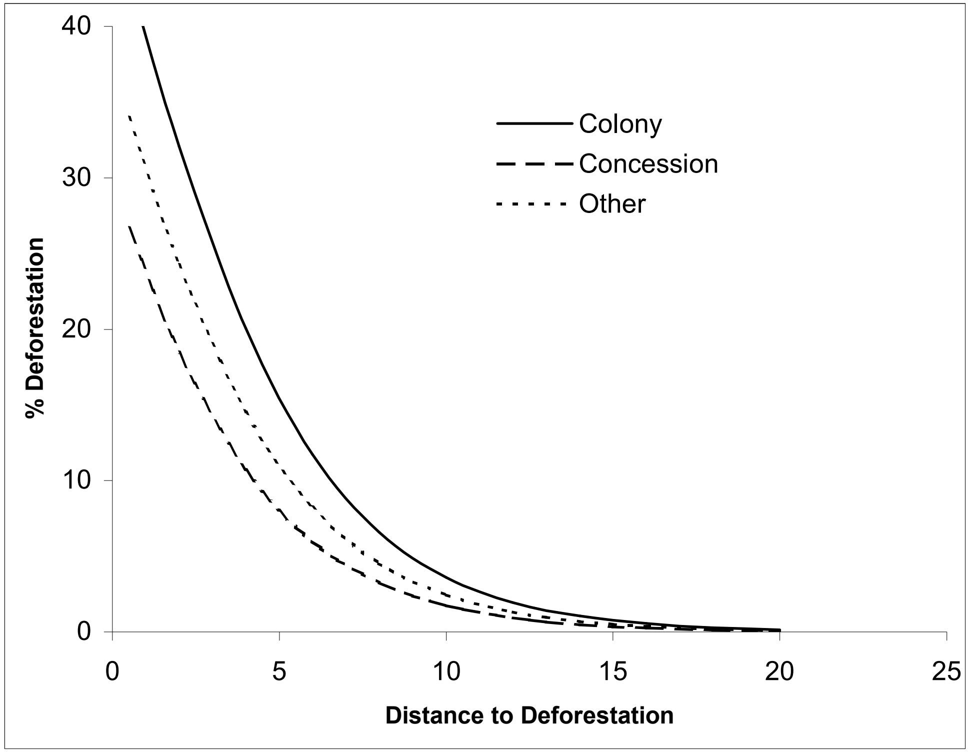

Figure 2. Sampling options