Preprint of: Kaimowitz, D., P. Mendez, A. Puntodewo and J. Vanclay, 2002. Spatial regression analysis of deforestation in Santa Cruz, Bolivia. *In:* C.H. Wood and R. Porro (eds) *Land Use and Deforestation in the Amazon.* University Press of Florida, p. 41-65. ISBN 0-8130-2464-1.

Spatial Regresion Analysis of Deforestation in Santa Cruz, Bolivia

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

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

This paper applies a spatial economic regression model to analyze the relation between deforestation in the period from 1989 to 1994 and access to roads and markets, ecological conditions, land tenure, and zoning policies in Santa Cruz, Bolivia. The data comes from a Geographic Information System (GIS) compiled by the Natural Resources Department of the Santa Cruz Government. Locations closer to roads and the City of Santa Cruz and that have more fertile soils and higher rainfall have a greater probability of being deforested. The same also applies to colonization areas. National parks and areas occupied by indigenous people do not have significantly less deforestation than sites with similar acess and ecological conditions. Forest concessions, on the other hand seem to protect forests.

1. Introduction

International interest in the issue of tropical deforestation has grown rapidly during the last twenty years, but major uncertainties persist regarding when, where, and why it occurs. In a recent survey by Kaimowitz and Angelsen (1998), the authors identified more than 150 quantitative models that researchers have developed to answer these questions, most of them since 1990. These models assess the impact of over 115 variables that potentially influence deforestation, but in many instances the direction and magnitude of their effect on deforestation remains uncertain. This reflects the inherent complexity of the issue and limited data availability, as well as various methodological weaknesses.

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, Osvaldo Escalante Saldaña, 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.

The Kaimowitz and Angelsen study evaluated the potential of different modeling approaches for improving our understanding of tropical deforestation and concluded household and regional-level studies showed more promise than national and global models. It expressed particular enthusiasm for the opportunity the growing availability of spatial databases presents for examining the relation between deforestation and spatial variables such as access to markets, land tenure, land use zoning policies, and ecological factors. Most spatial models use relatively reliable data and large sample sizes and are particularly well suited for predicting where deforestation will occur. In addition, researchers can often test the models' robustness by measuring what percentage of the time they accurately predict which areas will be deforested.

This paper presents a spatial econometric model of deforestation in the Department of Santa Cruz, Bolivia between 1989 and 1994. We chose Santa Cruz for several reasons:

- The Bolivian tropics have historically had low deforestation, but this has changed rapidly in recent years, particularly in Santa Cruz. Spatial models may be able to help us understand why;
- Unlike many Latin America regions, large-scale mechanized agriculture plays a major role in forest loss in Santa Cruz. Thus, the causal pathways that influence deforestation there may differ greatly from other locations;
- CIFOR has an on-going project comparing the effects of different policies and social trends on tropical forests in Bolivia, Cameroon, and Indonesia; and
- The Government of Santa Cruz (prefectura) had an existing Geographic Information System (GIS) with much of the data we needed to model the influence of spaital variables on deforestation trends.

This paper first briefly describes spatial econometric deforestation models and reviews the conclusions of previous models made for other tropical areas. Then, we give background information on deforestation in Santa Cruz. Next, we present the model and data we used in this study and finally we discuss our results.

2. *Spatial Econometric Deforestation Models*

Spatial econometric 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),
- Ecological conditions (topography, soil quality, precipitation, and forest fragmentation), and
- Land tenure and land use zoning categories (national parks, forest concessions, colonization areas, and indigenous territories).

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

Some model makers include only locations covered with forest during the initial period in their samples. Others include both locations that had forest cover in the initial period and those that did not.

Based on economic theory, model makers hypothesize that farmers decide whether to deforest an area based on if the net present value of the returns they receive from putting the land into agriculture outweighs the cost of forest clearing and any benefits they might obtain from forest products and services. Areas with soils, rainfall, and topography more suited to agriculture should provide higher returns to farming and thus farmers would be more likely to deforest them. Likewise, government subsidies for deforestation in colonization areas increase the likelihood farmers will clear forests there. On the other hand, higher transportation costs reduce the net returns from converting an area to agriculture. Similarly, government land use restrictions, such as the designation of a location as a protected area or forest concession, should raise the expected cost to farmers of deforesting the area and make forest clearing there less likely. Some researchers also argue that indigenous people are less likely to deforest their land because of their distinct belief systems and traditions or their more limited access to capital.

The data used in spatial econometric models comes from maps. Most models have obtained that data by selecting a random sample of locations (points on the map) and determining the state of each variable in that location. They then treat the characteristics of that point as one observation, as if it were one individual in a household survey.

Typically, modelers use samples of several thousand points or more. Chomitz and Gray (1995), for example, used a random sample of 10,000 data points for their Belize study. Tom Tomich of the International Center for Agroforestry Research (personal communication 1998) worked with 49,000 points in his model of Jambi, Sumatra. Gerald Nelson from the University of Illinois (personal communication 1998) made a similar study with raster data that yielded 25,000 sample points.

Most land use information comes from national forest inventories, remote sensing, aerial photographs, and ground truthing. GIS programs generate the information on distance to roads and markets using the maps in their databases. The remaining information comes largely from local government departments.

3. Previous Model Results

Previous models have generally confirmed the hypotheses that land holders convert more forest to agricultural use in locations that have better access to markets, ecological characteristics more favorable for farming, and no government restrictions on forest clearing. (See Table 1.)

Table 1. Conclusions from Previous Spatial Regression Models About the Effects of Different Variables on Deforestation

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

* Only in Cameroon. No effect in Zaire

Forests are more likely to be cleared when they are closer to roads in physical distance and traveling time (Chomitz and Gray, 1995; Deininger and Minten 1996; Liu *et al.,*1993; Ludeke *et al.,* 1990; Mamingi *et. al.* 1996; Mertens and Lambin, 1997; Nelson and Hellerstein, 1995; Sader and Joyce, 1988; Rosero-Bixby and Palloni, 1996). Most studies show forest clearing declines rapidly beyond distances of two or three kilometers from a road. Liu *et al.* (1993) and Mamingi *et. al.* (1996), however, report that forest clearing and distance to roads remain strongly correlated at much greater distances in Cameroon, the Philippines and Zaire. Proximity to railroads is also positively associated with deforestation in Cameroon and Zaire (Mamingi *et. al*. 1996).

Chomitz and Gray (1995) found locations near urban markets have less remaining forest in Belize and Mertens and Lambin (1997) say most deforestation occurs less than ten kilometers from the nearest town in Eastern Cameroon. Nelson and Hellerstein (19997) found that distance to villages had a much more significant effect on land use than distance to urban areas.

Farmers are also more likely to clear areas with higher quality soils (flat, adequate drainage, and high soil fertility) and drier climates (Chomitz and Gray, 1995; Gastellu-Etchegorry and Sinulingga, 1988; Sader and Joyce, 1988; Rosero-Bixby and Palloni; 1996).

Forest fragments have a higher risk of being lost than forests in large compact areas, with those 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).

The effect of roads and environmental conditions interact, so roads induce greater forest clearing in areas with good soils and favorable climatic conditions. In Belize, for example, Chomitz and Gray (1995) show the probability of an area being used for agriculture (rather than natural vegetation) on high quality land next to a road was 50%, while lands next to roads with marginal soils had only a 15% probability of being deforested. Mamingi et. al. (1996) obtained similar results in Cameroon and Zaire.

Mertens and Lambin (1997) note 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. Along roads you get 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.

Deininger and Minten (1996) show protected areas were deforested less than other locations with similar attributes in Mexico. We were unable to locate any previous spatial regression studies that examine the statistical relations between other zoning or tenure classifications and land use.

4. Deforestation in Santa Cruz, Bolivia

The department of Santa Cruz extends some 900 by 800 kilometers, and occupies 36.4 million hectares. Of these, 30.7 million hectares (84.3%) were still in forest in 1994 and 3.2 million hectares (8.8%) was in pasture or savanna, most of which is natural and has never been forested. Farmers used 2.1 million hectares (5.8%) for agriculture and the remaining 1.4 million hectares was dedicated to other uses (Morales 1993 and 1996).

As late as 1950, Santa Cruz had less than 60,000 hectares of land under cultivation. Construction of a road connecting Cochabamba and Santa Cruz in the early 1950s and government policies encouraging agricultural colonization and sugar and rice production in the 1960s gradually changed that. These trends accelerated in the 1970s, as did subsidized credit for local farmers (Pacheco 1998).

Most deforestation prior to 1980 occurred in the 'integrated zone' composed of the provinces of Andrés Ibañez (where the city of Santa Cruz is located) and Warnes, and the eastern portions of Ichilo, Obispo Santiesteban, and Sara provinces (Pacheco 1998).

The eastern portion of the 'integrated zone', located roughly between the Pirai and Grande rivers, was the traditional center of Santa Cruz' large-scale commercial agriculture. In the 1970s, large Bolivian landholders and agricultural colonies populated by Japanese, Okinawan, and Mennonite settlers grew sugar cane, maize, cotton, and sorghum and grazed cattle there. More recently, cotton has become less important and wheat and soybeans have emerged as major crops (CAO 1996). This area has moderate rainfall (1,000 – 1,500 mm), some of the department's best soils, and the most complete transportation infrastructure.

Small farm colonists dominate in the west and north of the integrated zone (Thiele 1995). Government – sponsored settlements in those areas date back to the 1960s. This area has a more humid climate (with an average annual rainfall of 1,500-2,000 mm), better suited for rice production, and the soils tend to be poorer.

Since the mid-1980s, the central focus of forest clearing has shifted to the east, into the so-called 'expansion zone', which covers the western portion of Chiquitos province and the south of Nuflo de Chavez province. Construction of a large bridge over the Grande river and the implementation of structural adjustment policies that encouraged non-traditional exports such as soybeans greatly encouraged that process (Kaimowitz, Thiele, and Pacheco 1999).

Like the 'integrated zone', the 'expansion zone' has two sub-regions. Large mechanized soybean and wheat farmers dominate in the south, which has lower rainfall (900-1,000), while small agricultural colonists have rice – based systems farther to the north. The soils of the expansion zone are moderately fertile, but highly susceptible both to compaction to soil erosion (Barber 1995).

The 'integrated' and 'expansion' zones combined account for only 19% of the area of Santa Cruz, but 70% of the area deforested prior to 1994 (Pacheco 1998).

As one moves farther east and north-east from the 'expansion zone' one hits the 'Brazilian' shield, characterized by infertile and acidic soils. Cattle ranching and logging are the main productive activities there, although a certain amount of small-farm agriculture exists as well. A large portion of the land was in timber concessions prior to 1996. Ranchers have probably cleared a moderate amount of forest for pasture, but it is difficult to distinguish these areas from natural savanna in the satellite images.

To the South, in the Chaco region, where annual precipitation often falls below 800 mm, the dry climate discourages crop production. Dry forest and scrub makes up a large portion of the natural vegetation.

In the provinces of Caballero, Florida, and Valle Grande, the area southwest of Santa Cruz City, the steep slopes have limited agricultural development. This area has some of the department's oldest settlements but lacks economic dynamism and has weak transportation infrastructure (Davies 1994).

Overall, annual deforestation rates have increased 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 (the area north of the 18° parallel). That region covers 61% of Santa Cruz, but accounts for a much higher percentage of forest clearing, as most of the rest of the department is in the dry south and mountainous south-west. Between 1989 and 1992, this rose to around 78,000 hectares annually in the entire department of Santa Cruz, and from 1992 to 1994 the yearly total reached 117,000 hectares (Morales 1993 and 1996).

5. The Data

The model we created to estimate the influence of different spatial variables on deforestation in Santa Cruz uses information on:

- Land use in 1989 and 1994
- "Land use potential" (soil quality and topography)
- Rainfall
- Roads and railways
- Urban areas
- Forest concessions
- Colonization zones
- Indigenous territories
- Protected areas

We drew our data from a Geographic Information System (GIS) produced by the 'Santa Cruz National Resource Protection Project'. The Government of Santa Cruz implemented that project 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. Land use and potential land use data were provided in raster form, and were converted to vector format. Many GIS layers obtained were based on the UTM elipsoid IU661967.

The 1989 land use data delineate forests, deforested areas, savanna and pastures, areas with little or no vegetation, water, and urban areas. The data are based on Earthsat satellite images and have a resolution of 0.5 x 0.5 km (25 hectares). They were produced by Ivan Morales, a remote sensing expert contracted by the Santa Cruz Natural Resource Protection Project, who revised a previous analysis of the same images by the CUMAT consulting company (Morales 1993). The data omits about 10,337 square kilometers from northwest Santa Cruz.

We omitted all areas from our study that were not classified as either forests or deforested areas in 1989, including those areas for which there was no data. As mentioned previously, although some areas classified as savanna and pasture in 1989 were once forest, the great majority is natural savanna that was apparently never forest and we have no way of distinguishing that natural savanna area from previously forested areas (Ivan Morales, personal communication 1997).

The 1994 land use data is based on Landsat images analyzed by Ivan Morales. It sub-divides deforested areas into traditional agriculture, commercial agriculture, anthropogenic pastures, mixed agriculture, and agriculture with forests (Morales 1996). As its name implies, agriculture with forests includes some forest area, but Morales classified it as ‘deforested’. The resolution was also 0.5 x 0.5 square kilometers. Cloud cover was minimal in the 1994 images used to assess land use. Nevertheless, no data is available for some areas in eastern Santa Cruz.

The land use data set we used in our analysis combines all land that classified as forest in 1989 for which information was also available in 1994. This covers a total area of 302,000 kms or 82% of Santa Cruz.

The relatively low resolution of the data means that deforested areas smaller than 12.5 hectares cannot be detected. Thus, our conclusions better reflect the factors influencing forest clearing for mechanized agriculture and ranching for shifting cultivation.

The land use potential data follows the United States Department of Agriculture (USDA) classification, which classifies land on a scale from I to VIII, with one being areas with the highest agricultural potential and eight being the lowest. This classification takes into account soil fertility, depth, and texture, slope, salinity, and chemical toxicity. The Santa Cruz Natural Resource Protection Project team assembled the data using secondary sources, satellite interpretation, aerial photography and observation, ground truthing, and soil sampling (Prefectura del Departamento – Consorcio IP/CES/KWC 1996).

Rainfall refers to average annual precipitation and has been divided into discrete classes by rounding off to the nearest 100 millimeters. We have no information regarding the specific sources of information the government of Santa Cruz used to prepare its rainfall map.

The information on roads, trails, and railroads lines was assembled by the Santa Cruz Natural Resource Protection Project based on secondary source and ground truthing. The road data include all classified roads. Although it would have been preferable to use road data from 1989, the only data available were for 1993. The ‘trails’ data includes temporary petroleum exploration and logging roads.

The PLUS divided Santa Cruz’s 41 urban areas into four categories based on their population and infrastructure (education, health and banking facilities, agricultural markets, and telephones). We have considered these categories as proxies for the size of the towns’ markets for foods, with lower numbers used to refer to the towns with larger markets (Prefectura del Departamento – Consorcio IP/CES/KWC 1996). The first category includes only the City of Santa Cruz. The second and third categories include medium sized towns and the fourth category smaller towns.

The ‘forest concession’ boundaries used in the analysis were obtained from the GIS of the Sustainable Forestry Management Project (BOLFOR) and come from Bolivia’s Forestry Development Center (CDF). We have no information regarding the specific years the Bolivian government allocated these concessions, but we understand that it allocated practically all, if not all, of them prior to 1989.

Santa Cruz has government – sponsored small farm colonization zones and Mennonite and Japanese agricultural colonies. Our colonization zone data includes both and comes from a map produced by the Center for Research on Management of Renewable Natural Resources (CIMAR). The map has a scale of 1/1,000,000 and should be considered a first approximation (de Vries 1994).

The lands classified as 'indigenous territories' in our analysis refer only to areas that indigenous communities had legal titles for or similar instruments in the early 1990s. This represents only a small portion of the area traditionally utilized by indigenous communities or to which they have made legal claims. The Santa Cruz Natural Resource Protection Project compiled this data from primary sources.

In early 1990s, Santa Cruz had three protected areas with forest: the Amboro National Park, the Noel Kempff Mercado National Park and Biological Reserve, and the Rios Blanco y Negro Wildlife Reserve. The Bolivian government established both Amboro and Noel Kempff Mercado prior to 1989, although it expanded Amboro in 1991. The Rio Blanco and Negro Wildlife Reserve was created in 1990. In our analysis, we classified the buffer zone of the Noel Kempff Mercado National park as a protected area.

6. Our Sample - The Polygon Approach

As noted earlier, most previous studies using spatial econometric models have worked with a set of systematically selected points taken from a GIS grid as their sample. 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. Spatial autocorrelation is a common problem with geographic data, since nearby locations are more likely to be similar than distant ones. This can leads to biased coefficients, inefficient parameters, and inaccurate measures of statistical significance (Chomitz and Gray 1995; Rosero-Bixby and Palloni 1996).

Some researchers prefer to sample tiles rather than points. They argue that these are more appropriate in fragmented landscapes where correct alignment of the various GIS layers may be problematic. In this case, researchers typically use the proportion of the land area still in forest as their dependent variable, rather than a discrete (forest or non-forest) variable. If the tiles cover the entire study area, a complete census may be analyzed, but this again increases the probability of spatial autocorrelation.

One way to make better use of information while minimizing autocorrelation is to first stratify the sample into different relevant segments (e.g., forested versus deforested, and close to versus distant from town/road) and then sample more intensively in strata of particular interest. One can then correct for the difference in sampling intensities used in each strata when doing the regression analysis by weighting the data drawn from each strata by the sampling intensity used.

Another alternative to working with a sample based on points or tiles is to use polygons occurring spontaneously within the GIS. To the extent we are able to take into account the main spatial variables that influence deforestation when we construct these polygons, this will also reduce spatial autocorrelation. This approach also reduces the unnecessary proliferation of sample units since it combines areas that share the same major characteristics. It is important, however, to carefully select the layers chosen to determine the polygons 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.

In this study we chose to adopt the polygon approach. We first divided Santa Cruz into areas that were deforested between 1989 and 1994 and those that remained in forest. Next, we separated out all of the areas inside colonization areas, protected areas, forest concessions, and indigenous territories from those outside such areas. Then, we broke up the polygons based on USDA soil quality classifications. To eliminate slivers, with combined any polygons smaller than 0.1 ha or less than meters wide with the adjacent polygon. Large polygons were further fragmented using a regular grid to improve homogeneity with respect to distances.

Using this method, we obtained a total of 24,208 polygons. That is our sample. For each polygon in the GIS we computed 20 potential explanatory variables and transferred these to the statistics package S-plus for further analysis. To calculate the distance from the polygons to roads, railroads, and markets we took as a reference the geometrical center ("centroid").

7. The Model

Our model analyzes the determinants of deforestation in Santa Cruz between 1989 and 1994. The dependent variable we used is the probability that a given location covered with forest in 1989 still had forest in 1994. We do not explicitly analyze what determined deforestation prior to 1989, although some of the evidence discussed below provides insights into that question.

The independent variables we examined include: distance to roads, railroads, trails, and markets, “land use potential” (soils and topography), rainfall, whether or not a location falls within a protected area, colonization zone, forest concession, or indigenous territory, and the distance to the nearest area deforested prior to 1989.

The initial specification of the model, based on theoretical considerations and data availability, was:

$$\text{Forest} = \text{Intercept} + ß_1\,\text{Conc} + ß_2\,\text{Indig} + ß_3\,\text{Prot} + ß_4\,\text{Colon} + ß_5\,\text{Soil} + ß_6\,\text{Rain} + ß_7\,\text{Rain}^2$$

$$+ß_8\,\text{DR} + ß_9\text{DT1} + ß_{10}\,\text{DT4} + ß_{11}\,\text{DLRo} + ß_{12}\,\text{Conc:DLRo} + ß_{13}\,\text{DRR} + ß_{14}\,\text{DDF} + \text{error term}$$

Where:

Forest	(1 if forested in 1989 and still forested in 1994, 0 if forested in 1989 and not forest in 1994)
Conc	(1 is inside forest concession, 0 if outside)
Indig	(1 if inside indigenous territory, 0 if outside)
Prot	(1 if inside protected area, 0 if outside)
Colon	(1 if inside colonization zone, 0 if outside)
Soil	(USDA soil group, from I-VIII)
Rain	(precipitation, millimeters)
Rain2	(precipitation squared)
DR	(Distance to classified roads, kilometers)
DT1	(Distance to Santa Cruz, kilometers)

DT4	(Distance to nearest small town, kilometers)
DLRo	(Distance to temporary logging or mining trails outside a forest concession)
Con:DLRo	(Distance to nearest trail inside a forest concession)
DRR	(Distance to railroad)
DDF	(Distance to nearest area deforested prior to 1989)

We did not include medium sized towns in our analysis since practically all of them are very close to the City of Santa Cruz. Hence, it would be practically impossible to distinguish between their separate effects.

We included the quadratic rainfall term because we expected that deforestation would be highest at medium rainfall levels. Thus, we were looking for a functional form that would allow deforestation to first rise with additional rainfall and then fall.

We fitted the model as a logistic model weighted by polygon area using generalized least squares. Economists generally favor the use of logarithmic, rather than logistic, transformations since parameter estimates can then be interpreted directly as elasticities (i.e. a unit change in an independent variable always causes the same percentage change in the dependent variable). While this is quite helpful when all independent variables are expressed in the same units, it becomes less relevant when the nature of the independent variables varies greatly, as in our case.

Statisticians 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 previous work by Jerry Vanclay led us to favor the logistic transformation.

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.

Besides spatial autocorrelation, two additional common problems in these models are multicollinearity and endogeneity. Variables such as distance to different types of roads and markets and soil quality tend to be highly correlated with each other and this can make it difficult to distinguish between their separate effects. We know of no undisputed test for multicollinearity for situations such as ours. To avoid potential multicollinearity, we have tried to be parsimonious in our use of explanatory variables and to monitor how the inclusion or exclusion of particular variables affects our results.

The endogeniety problem arises because it is hard to distinguish situations where human settlements, productive activities, and infrastructure are located in certain places because those places have environmental conditions that make them good to deforest from those where deforestation occurs because people settle, build roads, or make specific zoning decisions. We have attempted to reduce this problem by controlling for agricultural suitability and using independent variables from a time period prior to the dependent variables.

8. Results

Before discussing the fitted model, it is worth briefly analyzing the simple correlations that exist between deforestation and the independent variables and among the independent variables themselves. (See table 2.)

Table 2. Correlation matrix

	Forest	Conc	Indig	Prot	Colon	Soil	Rain	DR	DLRo	DRR	DT1	DT4	DDF
Forest	1.00												
Conc	0.14	1.00											
Indig	-0.01	-0.01	1.00										
Prot	0.08	0.10	-0.06	1.00									
Colon	-0.19	-0.19	-0.08	-0.07	1.00								
Soil	0.13	-0.01	0.07	0.08	-0.14	1.00							
Rain	-0.03	0.37	-0.06	0.05	0.09	0.02	1.00						
DR	0.23	0.12	-0.09	0.19	-0.18	0.14	-0.04	1.00					
DLRo	0.24	0.30	-0.06	0.37	-0.18	0.15	0.11	0.22	1.00				
DDR	0.16	0.24	-0.01	0.20	-0.15	0.16	0.14	0.42	0.21	1.00			
DT1	0.27	0.19	0.06	0.12	-0.34	0.23	-0.06	0.23	0.42	0.16	1.00		
DT4	0.19	0.08	0.07	0.02	-0.28	0.18	-0.18	0.26	0.04	0.10	0.88	1.00	
DDF	0.34	0.16	-0.08	0.24	-0.22	0.13	-0.06	0.60	0.32	0.34	0.45	0.43	1.00

The simple correlations support most of our a priori hypotheses. They show the forests that existed in 1989 were more likely to be deforested over the next five years if they were closer to roads, trails, railways, markets, and areas that had already been deforested in 1989 and had better soils. Colonization areas were more likely to be deforested and forest concessions and protected areas less likely.

As expected, the table also shows that locations near Santa Cruz tend to have better soils and road infrastructure and to be closer to colonization areas and areas already deforested in 1989. Protected areas and forest concessions tend to be located in places farther away from roads, towns, trails, railways, and previously deforested areas.

In general, distance to areas deforested prior to 1989 is more highly correlated with the other variables than deforestation between 1989 and 1994. This may reflect a weakening of the association between the different variables over time or a higher level of endogeneity of areas deforested prior to 1989 than of more recent deforestation.

Contrary to expectations, indigenous territories do not have less deforestation than other areas and deforestation increases with rainfall; issues we discuss more fully below.

The first regression model we ran gave us the following results:

Variable	Coefficient	t-value	significance
Intercept	0.859	1.95	-
Concessions	0.959	3.88	.001
Indigenous territories	0.036	0.13	-
Protected areas	0.235	0.44	-
Colonization areas	-0.342	2.42	.05
Soil quality	0.069	1.83	-
Rainfall	0.000	0.83	-
Rainfall 2	0.000	0.00	-
Distance to roads	- 0.004	1.88	-
Distance to trails (outside concession)	0.007	4.56	.001
Distance to trails (in concession)	0.003	1.64	-
Distance Santa Cruz	0.005	2.55	.05
Distance small town	-0.003	1.34	-
Distance to deforest	0.314	16.20	.001

The regression model as a whole was significant at a probability of .001 and explained about 35% of the total variation in deforestation. (This can be thought of as our "r2".) The only significant explanatory variables at a 95% confidence level are distance to Santa Cruz, areas deforested prior to 1989, and trails outside forest concessions and whether the location is in a forest concession or falls within a colonization area; each of which have the expected sign.

The fact that neither protected areas nor indigenous territories have any significant effect on deforestation in this model is particularly noteworthy. In both cases, the coefficients have the expected sign, but are rather small. In particular, if we take these coefficients as valid, both variables seem to protect areas much less from deforestation than do forest concessions.

The results with regards to distance to roads are even more surprising. In addition to showing that distance to roads has no significant effect on deforestation, the coefficient estimate calculated in regression analysis is negative. In other words, the closer we are to roads, the less likely an area is to be deforested!

Despite the fact that distance to deforestation in 1989 was the most important explanatory variable in this initial regression, it is also the variable that per se has least economic content. We know that locations close to place deforested in the past have a greater chance of being deforested in the future, but that may very well be because they share some of the same unspecified characteristics in both cases. Moreover, including deforestation prior to 1989 may introduce substantial multicollinearity since the variable is a function of many of the same things as deforestation in the later period. Hence, we decided to estimate our model again without the DDF variable. We also eliminated the quadratic rainfall, protected area, indigenous territory, and distance to small town terms because they did not seem to contribute any thing to the previous model. Once we did this, we got the following results:

Variable	Coefficient	t-value	significance
Intercept	0.833	3.72	.001
Concessions	1.771	5.99	.001
Colonization areas	-0.600	4.23	.001
Soil quality	0.150	4.53	.001
Rainfall	-0.001	4.61	.001
Distance to roads	- 0.051	9.61	.001
Distance to trails (outside concession)	0.008	6.04	.001
Distance to trails (in concession)	0.003	2.21	.05
Distance Santa Cruz	0.008	12.33	.001

This model explains only about 23% of our total variation but it still significantly significant at a 999% confidence level. In general, the coefficients in the model are larger, the significance levels are higher, and the results correspond more closely to our initial expectations. Locations farther from roads, trails, and Santa Cruz, which have poorer soils and less precipitation, have a better chance of retaining their forest cover. Colonization areas are more likely to be deforested and forest concessions less so. Even though the results presented above do not include the protected areas variable, we reran the same model including that variable and once again found that it had no significant effect.

To get some feeling for the relative magnitude of the coefficients in this model, we looked at the impact of changing each one in a "typical" case. The case we chose was a location 200 kilometers from Santa Cruz, 18 kilometers from the nearest classified road, and 100 kilometers from a trail, with average soils (USDA category = 4) and an annual precipitation of 1,100 mm that was neither a forest concession or a colonization area. In that context:

- Making the location a colonization area would have increased the likelihood of deforestation by 80%;
- Converting it into a forest concession would have reduced it by half;
- Reducing the distance to Santa Cruz by 50 kilometers would have made deforestation 50% more likely;
- Improving the soil class by one increases the likelihood by 20%; and
- Either increasing rainfall by 200 mm or reducing the distance from the nearest road to 16 kilometers both increase the likelihood by 10%.

9. Discussion

We consider the previous results preliminary, yet suggestive. We are currently in the process of validating the results using smaller polygons and alternative sampling procedures. We expect to test for spatial autocorrelation, which we have so far been unable to do because our initial data set did not include the x and y coordinates. We also wish to examine other aspects of soil type besides the 8-point USDA classification, including erosion susceptibility, drainage, salinity, alkalinity, depth, nutrient status, and presence of hard pans.

Given these caveats, our most notable initial conclusions are that protected areas did not seem to reduce the probability of an area being deforested in Santa Cruz in the period we studied, while forest concessions did. Although protected areas have less deforestation than do other areas in Santa Cruz, this appears to be because they are located farther from roads and towns, not due to protection *per se*.

Two possible explanations emerge for why locations closer to classified roads seem to have lower deforestation in the model where distance to areas deforested prior to 1989 is also included but higher deforestation when we omitted that variable. Putting both variables in the same model may produce multicollinearity and hence make the coefficient estimates unreliable. It is also possible that locations still in forest in 1989 that were close to both roads and to areas deforested prior to 1989 were protected by some unknown factor that we omitted from the model and hence were less likely to be deforested.

Logging and mining trails do appear to increase the probability of deforestation, but much less so when these trails are located within forest concessions.

The few indigenous territories with legal recognition may not be unrepresentative of indigenous territories in general and this may have influenced our result that indigenous territories do not have lower deforestation. We hope to clarify that issue through further research.

References

Barber, R.G. 1995. 'Soil Degradation in the Tropical Lowlands of Santa Cruz, Eastern Bolivia' *Land Degradation and Rehabilitation*. Vol. 6:95-107.

Brown, S., L.R. Iverson, and A. Lugo "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), 1993.

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.

Davies, P. 1994. Bosquejo Socioeconómico de Santa Cruz, Bolivia. Informe Técnico No. 16. May. Santa Cruz: Centro de Investigación Agrícola Tropical (CIAT) / Misión Británica en Agricultura Tropical (MBAT)

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.

Deininger, K. and Minten, B. 1997. Determinants of Forest Cover and the Economics of Protection: an Application to Mexico. World Bank, Washington D.C.

De Vries, A. 1994. Uso actual de la tierra en el departamento de Santa Cruz. #10. Santa Cruz, Bolivia: Centero de Investigacion y Manejo de Recursos Naturales Renovables / Servicio Holandes de Cooperacion al Desarrollo.

Kaimowitz, D., Thiele, G., and Pacheco, P. 1999. The Effects of Structural Adjustment on Deforestation and Forest Degradation in Lowland Bolivia. *World Development*.

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

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, 1990:247-59.

Mamingi, N., Chomitz, K.M., Gray, D.A., and Lambin, E.F. 1996. Spatial Patterns of Deforestation in Cameroon and Zaire. Working Paper 8. Research Project on Social and Environmental Consequences of Growth-Oriented Policies. Policy Research Department, World Bank, Washington D.C.

Mertens, Benoit and Eric F. Lambin. 1997. “Spatial Modeling of Deforestation in Southern Cameroon” Applied Geography Vol. 17, No. 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, P. 1998. *Estilos de desarrollo, deforestación y degradación de los bosques en las tierras bajas de Bolivia*. La Paz: CIFOR / Centro de Estudios para el Desarrollo Laboral y Agraria (CEDLA) / Taller de Iniciativas y Estudios Rurales y de Reforma Agraria (TIERRA).

Prefectura del Departamento – Consorcio IP/CES/KWC. 1996. Memoria del Plus, Plus de Uso del Suelo del Departamento de Santa Cruz, Bolivia. Santa Cruz, Bolivia: Prefectura del Departamento – Consorcio IP/CES/KWC

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.

Thiele, G. 1995. The Displacement of Peasant Settlers in the Amazon: The Case of Santa Cruz, Bolivia. *Human Organization*. Vol. 54, No. 3: 273-82.