

**Time Series Analysis of Satellite Data:  
Deforestation in Southern Mexico**

Jacqueline Geoghegan  
Department of Economics  
Clark University  
Worcester MA

Julie Hewitt  
National Center for Environmental Economics  
U.S. Environmental Protection Agency  
Washington DC

Colin Vance  
German Aerospace Center  
Institute of Transportation Research  
Berlin, Germany

*Paper prepared for presentation at the American Agricultural Economics Association  
Annual Meeting, Montreal, Canada, July 27-30, 2003*

## Introduction

Tropical deforestation is significant to a range of themes that have relevance for the study of environmental change and economic development, including global warming, land degradation, species extinction, and sustainability issues. Scientific and public concerns about these and other potentially massive ecological disruptions have incited a growing number of studies that aim to quantify the social and biophysical determinants of deforestation processes, as well as their interactions over time and space. Recognition that both the *location* and *pattern* of forest clearance are often as important as its *magnitude* has motivated an increasing number of studies that combines high-resolution satellite imagery, geographic information systems (GIS), and socioeconomic and geophysical data to model the human-environment interactions that drive land-use change (*e.g.* Liverman et al., 1998; Nelson and Geoghegan, 2002).

In the economics literature, the initial focus of this research has been on identifying the socioeconomic forces that explain the location and spatial pattern of tropical deforestation in a number of countries: Belize (Chomitz and Gray, 1996); northern Mexico (Nelson and Hellerstein, 1997); Brazil (Pfaff, 1999); Thailand (Cropper, Griffiths, and Mani, 1999; Cropper, Puri, and Griffiths, 2001); and Panama (Nelson, Harris, and Stone, 2001), in general, using one set of observations on land use change. However, more recent attention has been given to capturing the temporal dynamics from which these patterns emerge, using time series data in a number studies in different countries: Mexico (Vance and Geoghegan, 2002; Geoghegan, Schneider, and Vance, 2003); in China (Seto and Kaufmann, 2003); and Honduras (Munroe, Southworth and Tucker, 2002).

In all these deforestation studies, there are modeling questions concerning the definition of deforestation as well as defining the unit of observation. The first issue arises because in many of the studies, land use change is often the result of swidden agriculture, so “deforestation” could be defined as only the initial clearing of primary forest, or “deforestation” could include the clearing of secondary forest that has already become part of the fallow cycle. For the second issue, the unit of observation would optimally be the economic decision maker, but information on that is not easily accessible in data sparse conditions as is often the case in developing countries (Nelson and Geoghegan, 2002). Therefore what most researchers have resorted to is to either aggregate approach, where the data on land use change is aggregated to some political unit, or using the unit of observation from the satellite itself, the pixel, which for the data used in this paper is Thematic Mapper (TM) satellite data, and is approximately 28 meters squared. In the former aggregated case, the dependant variable of deforestation becomes less spatial and the model can be estimated by panel regression techniques. In the latter case, with the unit of observation as the pixel, the data remain spatially disaggregated, and some form of discrete choice approach is used.

The definition of the unit of observation also entails determining whether the dependent variable describes the direct satellite image results or changes from one satellite image to the next. For three satellite images, as is the case here, the first approach results in three observations per pixel, while the second results in two. In principle, no information is lost by taking the differences approach. In practice, however, there are three possible outcomes between satellite images in the differencing approach: no change, reforestation and deforestation. The data is still panel data, and to employ panel data techniques in estimation, one would have to use a dichotomous choice dependent variable, where the obvious definition is deforested in this period or not. In this case, the information distinguishing pixels with no change from those reforested over the period is lost. This may not seem important from the standpoint of analyzing deforestation, but when three or more satellite images are available, it is possible for a pixel to be observed as not-deforested than deforested (the differences) based on the pixel transitioning from not-forest, to forest, to not-forest in succession. This succession is likely due to swidden practices on the same plot, rather than new deforestation. This is a particular concern for the relatively short time frames involved in this and similar studies, where the not-forest, forest, and not-forest succession could not involve primary forest. For example, four percent of the observations in Munroe et al. (2002) fall into this category. One might be tempted to dismiss this as an issue, since four percent is not a large portion of the data, but these observations have an effect on the estimated model coefficients. Even with a non-differencing approach, great care must be taken in the definition of forest to be able to draw inferences about factors influencing deforestation.

However, there are econometric issues using spatially disaggregated and discrete time series data on land use, specifically in how the observations are linked through time. One approach is to use hazard (also known as survival) models, as has been done in previous related work in Southern Mexico (Vance and Geoghegan, 2002) and is also used in suburbanization models in the United States (Irwin and Bockstael, 2002; Irwin, Bell and Geoghegan, 2003). A second approach is to take advantage of panel data techniques. As noted above, panel approaches are only readily available for dichotomous choice dependent variables (there are, however, nested logit panel models). This suggests that researchers are confronted with the choice between addressing the panel nature of the data, or the more complex nature of the dependent variable that defies dichotomous categorization. The panel approach has the following attractive feature that has led to its popularity. The random error can be separated into two terms: one which is randomly distributed over all observations, and one which is randomly distributed over pixels. The former varies over time while the latter does not. This model imposes a certain type of structure on the time variability, in that the correlation across time periods is equal (Greene, 2002). Given that a transition from forest to non-forest depends critically on how long the pixel had been in forest (e.g., whether it is a swidden or deforestation transition), the equal correlation across time assumption likely imposes more structure than is truly warranted.

Our study focuses on deforestation in an agricultural frontier spanning the southern Mexican states of Campeche and Quintana Roo. This region contains one of the largest and oldest expanses of tropical forests in the Americas outside of Amazonia and

has been identified as a “hot spot” of forest and biotic diversity loss. Over the past 30 years, these forests have been under sustained pressure following the construction of a highway in 1972 that opened the frontier to settlement. The road was part of a larger development effort to promote agricultural colonization and has contributed to a prolonged period of land transformation that has been captured by the TM satellite imagery. We capture these landscape dynamics by assembling a spatial database that links the pixels from four TM images spanning the years 1986-1997 and other spatial environmental and GIS-location derived data with government censuses for the region, which include a time series of basic information on population, education, language and housing characteristics at the village level.

Following a brief overview of the study region, our analysis takes as its point of departure a simple utility-maximizing model that suggests many possible determinants of forest clearance. We subsequently test the significance of these determinants using a probit model, where data from all three satellite images are pooled. The probit model is run for two distinct interpretations of the dependent variable of interest. The first dependent variable measures whether a pixel is forested or not, while the second measures whether a pixel has been developed or not. We compare these simple models to show that our concerns regarding the definition of the dependent variable are valid. We conclude this section with a brief discussion of our next steps.

## **The Study Region**

The southern Yucatán peninsular region occupies roughly 22,000 km<sup>2</sup> of southwestern Quintana Roo and southeastern Campeche, north of the Mexican-Guatemala border, see Figure 1.. A rolling karstic terrain of semi-deciduous tropical forests covers the landscape, with elevations in the center reaching a peak of about 250 to 300 meters. The zone corresponds to what was once a portion of the Maya lowlands, and was nearly completely deforested one thousand years ago during the Classic Period of Lowland Maya domination (A. D. 100-900) (Turner, 1983). Following the collapse of Maya civilization in A. D. 800-1000, the region experienced a period largely free of settlement that, continuing past the birth of the Mexican nation-state in 1821, allowed the return of the forests. By the first half of the 20<sup>th</sup> century, human intervention here re-emerged but was primarily limited to the selective logging of tropical woods, particularly mahogany (*Swietenia macrophylla*) and cedar (*Cedrela odorata*), as well as the extraction of chicle, a tree resin (from *Manilkara zapota*) used in the production of chewing gum.

More extensive deforestation followed with the construction of a two-lane highway across the center of the region in 1972, which opened the frontier to agricultural colonization primarily via the extension of *ejido* land grants from the federal government. The *ejido* sector was created following the Mexican Revolution (1910-1917), a political and social upheaval with roots in inequitable land distribution. Within *ejido* communities, an elected committee communally regulates land, but in this area of southern Mexico, *ejido* members (*ejidatarios*) typically enjoy usufruct access to a single parcel that is permanently allocated to their use. The size of these parcels varies considerably, ranging from 10 to 350 hectares, with an average size in the sample of

roughly 110 hectares. Our study focuses specifically on the *ejido* sector as it historically has been the predominate form of land tenure in the study region.

In the region as a whole, deforestation continued unabated through the decades of the 80s and 90s, with available satellite imagery revealing a total of 617 square kilometers of mature forest cut between 1987 and 1997. In 1989, the Calakmul Biosphere Reserve (723,185 ha) was established in the middle of the region, partly in response to extensive deforestation along the highway and associated international pressures to impede further clearance. Various land-uses surround the reserve, predominately: *ejidos*, on which slash-and-burn subsistence production has prevailed but increasingly is giving way to commodity production; *ejidos* coupled with NGO-sponsored agricultural and forest projects; and a small number of private lands largely devoted to livestock production. Cumulatively, these activities have earned the region a designation as a “hot-spot” of forest and bio-diversity loss by various sources (Achard et al., 1998). Whether this prognosis is supported by future trends will depend in large part on current plans surrounding the *Mundo Maya*, an international development scheme to create an ecotourism-archaeological tourist economy stretching across portions of southern Mexico, Belize, Guatemala, El Salvador and Honduras. This plan has led to recent hotel and road expansion in the region.

Tourist sector development notwithstanding, the region retains many of the features of an agricultural frontier. Population density is low, infrastructure is poorly developed, ties to outside markets remain few, and the social relations of production are, by and large, organized around the family farm. These features motivate the theoretical and empirical approach taken in this paper, which emphasize the importance of household-specific characteristics, particularly demographic composition and farm capital, in explaining land-use decision-making.

### The Theoretical Model

The theoretical model focuses on both the *location* and *timing* of land-use change and is based on Irwin and Bockstael (2001). Let  $A(i, t)$  be the net benefits to agricultural use for each time period after pixel  $i$  is cleared in time period  $T$ . Let  $F(i, t)$  be the net benefits to the farmer for leaving pixel  $i$  in forestry use in each time period, and let  $C(i, T)$  be the one-time clearing costs associated with clearing the pixel in time period  $T$  and let  $\delta$  be the discount rate. Then the net benefits to the farmer of clearing pixel  $i$  in period  $T$  are:

$$\sum_{t=0}^{\infty} A(i, T+t)\delta^{T+t} - \sum_{t=0}^{\infty} F(i, T+t)\delta^{T+t} - C(i, T) \quad (1)$$

For  $T$  to be the optimal time period for clearing, the following two conditions must hold:

$$\sum_{t=0}^{\infty} A(i, T+t)\delta^{T+t} - \sum_{t=0}^{\infty} F(i, T+t)\delta^{T+t} - C(i, T) > 0 \quad (2a)$$

$$\begin{aligned} \sum_{t=0}^{\infty} A(i, T+t)\delta^{T+t} - \sum_{t=0}^{\infty} F(i, T+t)\delta^{T+t} - C(i, T) > \\ \sum_{t=1}^{\infty} A(i, T+1)\delta^{T+t} - \sum_{t=1}^{\infty} F(i, T+1)\delta^{T+t} - \delta C(i, T+1) \end{aligned} \quad (2b)$$

The first condition, (2a), requires that net benefits to clearing are positive. The second condition (2b) considers that although clearing may yield net positive benefits at time  $T$ , there may still be benefits to waiting because of the potential for even higher benefits at some future date. Such a circumstance could arise, for example, in anticipation of improved technologies that reduce clearing costs.

Let the characteristics of pixel  $i$  be  $X(i)$ . The optimal time for clearing this pixel then is the first time period in which the following holds:

$$A(X(i), T) - F(X(i), T) - \delta C(X(i), T+1) \geq 0 \quad (3)$$

Add an error term to Equation 3 to account for unobservable characteristics:

$$A(X(i), T) - F(X(i), T) - \delta C(X(i), T+1) - \varepsilon(i) \geq 0 \quad (4)$$

This model clearly abstracts away from fallow cycle dynamics, which is planned to be the focus of further research.

## Data

The econometric model presented in this paper is estimated using Landsat TM satellite data on land cover as the dependent variable and government census and other biophysical spatial data for the independent variables. The data sources for each variable are discussed briefly. The unit of observation for the model is the TM pixel, an admittedly arbitrary unit of analysis but one that is about half the size of an average agricultural plot in the region. The satellite images were obtained across four contiguous zones spanning the study region, the dates for each of which are given in Table 1, with the boundaries of each image approximately given in Figure 1. The process of imagery classification included the normal preparatory steps of geo-referencing, haze removal, adding NDVI information, and principal component analysis. These steps were followed by texture analysis, which lead to the creation of a 7-band image for signature development and classification. Signature development was facilitated by a combination of ground truth data derived from GPS assisted field visits and topographic, vegetation and land-use maps. Maximum likelihood supervised classification methods produced six land cover classes. Excluding clouds, shadows and water, these include: mature lowland and upland forest older than 15 years of age; one stage of upland successional growth – predominantly secondary forest – between 7 and 15 years of age; agriculture (including

pasture); an invasive fern that occurs after agricultural use; and inundated savannas. For further detail on these methods see Turner, Geoghegan, and Foster (2003).

Other spatial data collected or created for use in the models include: elevation and slope from a digital elevation model; soil types digitized from a 1:250,000 Mexican government (INEGI) map; digitized road network from INEGI 1:50,000 topographic maps; rainfall data (interpolated to cover the region) from the Mexican government (Secretary of Agricultural and Hydrological Resources); and socio-demographic data from the Mexican government 1990 population census. The soil maps for the region are complex, offering many geographical “clusters” of soils based on taxonomic order and type, plus other characteristics. For modeling purposes, the many types and subtypes have been aggregated into two simple categories that reflect the basic conditions facing farmers: rendzina and other mollisols (“upland soils”, preferred by farmers); litosols and vertisols (all other soils). With the use of the road layers, distance from each individual pixel to roads and markets were calculated. Only four precipitation-temperature stations exist in the region. To these, 16 others lying just outside the region were added to create 20 stations or precipitation points and used to create an interpolated map (ordinary kriging analysis, see Isaacks and Srivastava, 1989) of region rainfall (annual dry season, averaged over thirty years). Using a GIS map of *ejido* boundaries, the census data were linked with the pixels associated with each of the *ejidos* via a uniform distribution, so that each pixel in the *ejido* was assigned the value of the variable from the census data for that *ejido*. The data include information on number of households, language and literacy, and structural characteristic of houses, such as electricity and water service.

## **Econometric Results and Discussion**

Our plan is to eventually estimate a model with a carefully measured dependent variable that is likely a polychotomous choice model. To take advantage of the polychotomous choices exhibited over time in the various satellite images, we hope to develop a panel approach to a polychotomous choice model. This would allow us to separate swidden effects from the type of deforestation that matters most for biodiversity. We begin with a simple test of the notion that the definition of the dependent variable matters. Our test compares models estimated using two competing dichotomous definitions of land use. The first definition, Forested, denotes whether a pixel is observed to be forested (1) or not (0). In this definition, the non-forested pixels include pixels that have been deforested, as well as pixels that are unlikely to be forested for a variety of reasons. The second definition, Disturbed, denotes whether a pixel is observed in a land use that has been clearly disturbed by humans from the forested state (1), or not disturbed (natural state) (0). Land uses which are clearly human disturbed are secondary forest, agriculture and fern. In this definition, the non-disturbed pixels includes pixels that are in primary forest (the forested pixels of the first dependent variables) as well as pixels that have not been disturbed because they were neither suited for forests nor for agriculture or other developed land uses.

We report the results of a pooled probit model using each dependent variable in Table 2. The explanatory variables include socio-demographic variables as well as

geospatial variables. The socio-demographic measures are from the Census, and vary by the ejido in which a pixel is situated, while the geo-spatial variables vary with the pixel. We first note that most of the explanatory variables are significant in both models (the exception is the percent of households served by a sewer system), and have the expected sign, such as the larger the *ejido*, the greater the probability for land to remain in forest use, the better the soil is for agriculture, the less likely the pixel is in forested use, etc. In general, we would expect that the Census variables will be less important in explaining the dependent variable; as they do not measure the characteristics of the household associated with the pixel, but of a representative household in the ejido, and are thus measured with error. Indeed, the most significant explanatory variables are the geo-spatial variables, in particular, distance to the nearest agricultural land and whether the pixel is in primary forest. Primary forests are associated with higher fertility soils, but have much higher clearing costs than secondary forests, so this dummy variable was included to control for this difference.

In comparing the two simple models, the first thing to note is that the dependent variables measure two concepts that are opposites, though not quite mirror-image opposites. One measures whether a pixel is observed in forest, and the other measures whether it is in a developed state rather than the historical forested state. Thus, we would expect that the signs of the coefficients (except for the constant term) would reverse sign from one model to the other. This is indeed the case. The two models have similar log likelihoods. Care should be taken in reading too much into the ad hoc comparison that shows a higher log likelihood for the model explaining the Forest dependent variable; the model explaining Developed has slightly greater predictive value. The two models also have similarly extremely significant chi-squared statistics (owing to the large number of observations).

Some of the socio-demographic Census variables are more instrumental in explaining the Developed model. We would expect this to be the case, as it is for the population variables. Others do a better job explaining the Forested model: higher education, and percent served by services. Of the spatial variables, the distance variables are more important to the Developed model than the Forest model, as we might expect. Although the two models are fairly similar—which is reassuring for the near mirror-image of the two approaches—it is clear that there are subtle differences in the coefficient estimates. These differences are likely to be more important in a polychotomous dependent variable model, as well as in a model that addresses the panel nature of the data. Our plan for proceeding is to estimate these variations on the model and include them in our conference presentation.



## **Acknowledgments**

This work was undertaken through the auspices of the Southern Yucatán Peninsular Region (SYPR) project with core sponsorship from NASA's LCLUC (Land-Cover and Land-Use Change) program (NAG 56406) and the Center for Integrated Studies on Global Change, Carnegie Mellon University (CIS-CMU; NSF-SBR 95-21914). Additional funding from NASA's New Investigator Program (NAG5-8559) also supported the specific research in this article. SYPR is a collaborative project of El Colegio de la Frontera Sur (ECOSUR), Harvard Forest—Harvard University, the George Perkins Marsh Institute-Clark University and CIS-CMU. The views expressed in this article are those of the authors and do not necessarily represent those the institutions with which they are affiliated. We thank Laura Schneider for her assistance with data preparation for this paper.

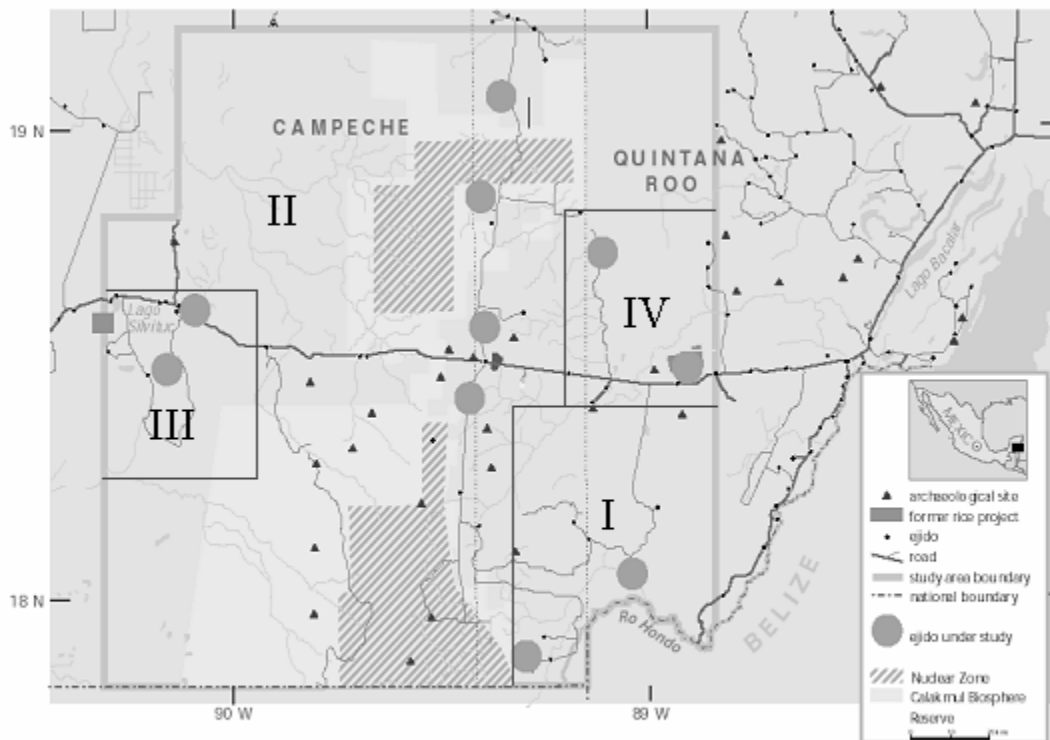
**Table 1: Dates of Thematic Mapper Satellite Imagery**

ZONE I	ZONE II & III	ZONE IV
Nov. 11, 1984	Apr. 01, 1987	Jan. 14, 1985
Feb. 21, 1993	Oct. 29, 1994	Nov. 07, 1994
Jan. 31, 1997	Feb. 05, 1996	Jan. 31, 1997

**Table 2: Econometric Results**

Dependant Variable:	FORESTED	DISTURBED
Variable	Coefficient (t-stat in parenthesis)	
<b>1. Socio-economic (varies by ejido)</b>		
constant	0.8231 (23.14)	-0.5419 (-15.37)
male pop	0.0018 (16.54)	-0.0025 (-23.19)
female pop	-0.0030 (-24.87)	0.0034 (28.60)
spanish	-0.0001 (-1.91)	0.0001 (5.01)
higher ed	0.0049 (31.90)	-0.0040 (-26.30)
ejido size	0.0000 (69.60)	0.0000 (-63.33)
% hh water	0.0030 (35.20)	-0.0023 (-27.43)
% hh elect	-0.0009 (-11.69)	0.0007 (9.71)
% hh sewer	0.0004 (2.07)	-0.0002 (-0.95)
good soil	-0.1471 (-34.85)	0.2224 (52.58)
<b>2. Distance (varies by pixel)</b>		
roads	0.0003 (65.30)	-0.0003 (-67.47)
ag land	0.0005 (125.30)	-0.0005 (-123.76)
market	0.0000 (-13.18)	0.0000 (17.64)
<b>3. Biophysical (varies by pixel)</b>		
primary	1.9630 (424.35)	-1.8060 (-400.61)
elevation	0.0012 (30.33)	-0.0010 (-27.11)
rain	-0.0017 (-58.05)	0.0012 (43.00)
slope	-0.0103 (-14.29)	0.0115 (16.13)
freq_ag	-0.0054 (-14.40)	0.0125 (33.48)
Number of Obs:	1,207,965	1,207,965
Log likelihood:	-312,752	-319,389
Chi-Squared:	475,606	441,537

**Figure 1: The Southern Yucatán Peninsular Region Study Area**



Adapted from: Evaluation of the Conservation of Selva Maya. Wildlife Conservation Society, 1995

Note: The thick gray line represents the boundaries of the general region of study, comprising about 22,000 km<sup>2</sup>. Roman numerals indicate the approximate locations of the TM satellite zones listed in Table 1.

## References

- Achard, F., Eva, H., Glinni, A., Mayaux, P., Richards, T., Stibig, H. J. (Eds.), 1998. *Identification of Deforestation Hot Spot Areas in the Humid Tropics*. Trees Publication Series B. Research Report No. 4. Space Application Institute, Global Vegetation Monitoring Unit. Joint Research Centre, European Commission. Brussels.
- Butler, J.S., Moffitt, R., 1982. A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model. *Econometrica* 59(3):761-764.
- Chamberlain, G., 1980. Analysis of Covariance with Qualitative Data. *Review of Economic Studies* 47(1):225-238.
- Chomitz K., Gray, D., 1996. Roads, Land Use and Deforestation: A Spatial Model Applied to Belize. *World Bank Economic Review*. 10, 487-512.
- Cropper, M., Griffiths, C., Mani, M., 1999. Roads, Population Pressures, and Deforestation in Thailand. *Land Economics*. 75, 58-73.
- Cropper, M., Puri, J., Griffiths, C., 2001. Predicting the Location of Deforestation. *Land Economics*. 77, 172-186.
- Geoghegan, J., Schneider L., Vance C., 2003. Temporal Dynamics and Spatial Scales: Modeling Deforestation in the Southern Yucatan Peninsular Region. *GeoJournal*, (forthcoming).
- Greene, W.H. 2002. *Econometric Analysis (5<sup>th</sup> Edition)*. Prentice Hall.
- Hsiao, C. 1986. *Analysis of Panel Data*. New York: Cambridge University Press.
- INEGI, 1985-87. *Cartas topograficas*. 1:50,000. Instituto Nacional de Estadística Geográfica e Informática, Aguascalientes.
- Irwin, E.G., Bell, K.P., Geoghegan, J., 2003. Modeling and Managing Urban Growth at the Rural-Urban Fringe: A Parcel-Level Model of Residential Land Use Change. *Agricultural and Resource Economics Review*, 32(1):83-102.
- Irwin, E., Bockstael, N., 2002. Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Residential Land Use Pattern. *Journal of Economic Geography*, 2, 31-54.
- Isaacks E.H., Srivastava R.M., 1989: *Applied Geostatistics*. Oxford: Oxford University Press.

Liverman, D., Moran, E F., Rinfuss, R. R., Stern, P. C. (Eds.), 1998. *People and Pixels: Linking Remote Sensing and Social Science*. National Academy Press, Washington, D.C.

Maddala, G.S., 1983. *Limited-Dependent and Qualitative Variables in Econometrics* Cambridge University Press.

Munroe, D.K., Southworth, J., Tucker, C.M., 2002. The Dynamics of Land-Cover Change in Western Honduras: Exploring Spatial and Temporal Complexity. *Agricultural Economics* 27:355-369.

Nelson, G., Geoghegan, J., 2002. Modeling Deforestation and Land Use Change: Sparse Data Environments. *Agricultural Economics* 27(3):201-216.

Nelson, G., Harris, V., Stone, S., 2001. Deforestation, Land Use and Property Rights. *Land Economics*. 77, 187-205.

Nelson, G., Hellerstein, D., 1997. Do Roads Cause Deforestation? Using Satellite Images in Econometric Analysis of Land Use. *American Journal of Agricultural Economics*. 79, 80-88.

Pfaff, A., 1999. What Drives Deforestation in the Brazilian Amazon? *Journal of Environmental Economics and Management*. 37, 26-43.

Seto, K.C., Kaufmass R.K., 2003. Modeling the Drivers of Urban Land Use Change in the Pearl River Delta, China: Integrating Remote Sensing with Socioeconomic Data. *Land Economics* 79(1):106-121.

Turner II, B.L., Geoghegan, J. Foster, D.R. (Eds.), 2003. *Integrated Land-Change Science and Tropical Deforestation in the Southern Yucatan: Final Frontiers*. Oxford Geographical and Environmental Studies. Clarendon Press of Oxford University Press (forthcoming).

Vance C., Geoghegan, J., 2002. Temporal and Spatial Modeling of Tropical Deforestation: A Survival Analysis Linking Satellite and Household Survey Data. *Agricultural Economics*. 27(3): 317-332.