The Use of Remote Sensing and GIS in the Collection of Survey Data on Households and Land-Use: Example from the Agricultural Frontier of the Brazilian Amazon

Stephen D. McCracken[1]
Eduardo Brondízio[1]
Emilio F. Moran[1]
Donald Nelson[2]
Andrea Siqueira[1]
Carlos Rodriguez-Pedraza[3]

1 Antropological Center for Training and Research on Global Environmental Change (ACT), Indiana University, Student Building Rm. 331, Bloomington, Indiana, 47405, USA. {stmccrac, ebrondiz, moran, asigueir@indiana.edu}
2 Department of Anthropology, University of Arizona, E. Haury Bldg, Tucson, Arizona, 85721, USA. {drn@u.arizona.edu}
3 International Institute of Tropical Forestry, P.O. Box 25000, Rio Piedras, PR, 00928, USA. {iitfgisgps@upr1.upr.clu.edu}

Abstract. This paper explores the use of satellite images and geographical information systems (GIS) for developing a sampling frame, stratifying the sample based on the analysis of remotely sensed information, and as an aid in promoting interview recall on land-use and agricultural strategies during the survey of households. Based on a combination of maps and satellite images (Landsat Thematic Mapper - TM) we have been able to identify 1399 farm properties for a study area along a stretch of the Transamazon highway west of Altamira, Pará, Brazil. A property grid overlay was developed in a GIS and used to extract information on land-cover classes from aerial photographs and maps for 1970 and 1978 (forest/non-forest areas), and from three TM images (1985, 1988 and 1991 for classes of forest, bare soil, pasture, and 3 stages of secondary succession) for each of the 1399 properties identified. In addition to facilitating a more complete sampling frame, the pre-survey analysis of remotely sensed data for the sampling frame permitted stratification of our sample to meet objectives of the project such as identifying cohorts of settlement, levels of deforestation, and pattern of agricultural uses for non-proportional inclusion of households and farms of interest in the sample. During fieldwork, land use/cover maps for each property sampled were presented in visual form and with descriptive statistics (area in each class at each date as well as a transition matrix between classes) for discussion with farm families. These aided interview recall in the development of a year-by-year retrospective land-use history of patterns of deforestation, fallow, crops planted and cattle grazing for the farm. This detailed information on land-use is linked to other survey information on household composition, economic/work activities of household members (age and gender), and retrospective birth and death histories to address questions concerning the demography of families, agricultural strategies and patterns of deforestation.

Keywords: Amazon frontier, demography, sample surveys, multi-temporal analysis.

Introduction

The need for studies of the human dimension of changes in land-use/land-cover (LULC) has emerged as a recurrent theme in recent discussions concerning global environmental change. Much of this research has focused on the use of remotely sensed satellite data and has focused primarily on land cover change (LUCC Science Plan, 1994). Developments in the area of geographical information systems (GIS) have furthered the possibilities of monitoring and analysis of these remotely sensed data. Remotely sensed data analysis with GIS has led to significant advances in the investigation of the distribution and spatial interdependencies of various phenomenon (e.g., soil distribution, topography, hydrological systems and vegetation and ecosystem patterns). Anthropologists and others using case study and ethnographic approaches have made significant headway in integrating satellite images in the study of LULC given their knowledge of the complex ways in which local populations use forest, land and other natural resources (Behrens, 1994; Brondizio et al., 1994, 1996). Increasingly environmental scientists working with remotely sensed data have needed ground measures but also local knowledge to verify, interpret, and analyze patterns that emerge using satellite data. These efforts have made important in-roads in the study of the human dimensions of environmental change but are often limited to certain types of questions that can be addressed and often do not permit quantifiable inferences to the broader landscape in which human activities are affecting environmental change. Outside human geography and anthropology, progress in this attempt at conceptual and methodological integration in environmental and social science research has been slow. In part, this is due to the difficulty of identifying mutually interesting research questions and selecting meaningful units of observation that can be spatially defined and that correspond to methodological and conceptual approaches of the various communities involved [1]. In this paper we discuss how satellite data and GIS can aid in developing a sampling frame for survey interviews on household level land use. The combination of remote sensing, GIS, and sample surveys provides an important set of tools for integrating social and environmental research on the human dimensions of land use and land cover change.

Study Area

Our study area is the region just west of the city of Altamira on the Transamazon Highway in the eastern Brazilian Amazon. For the larger research project fieldwork and analysis will cover an area 20 to 140 kilometers west of the city of Altamira to the town of Medicilandia. Work during this past year has focused on near 1400 properties and households in the eastern section of the study area. The entire study area is composed of about 4,000 properties. During the 1st fieldwork season we carried out survey interviews with 183 households along the main highway and side-roads between km 20 and km 70, in the eastern section of the study area. During the next field season an additional 320 households will be surveyed.

Highway construction in this area began in the late 1960s and by 1970 frontier colonists families began settling the region through the INCRA (Instituto Nacional de Colonização) colonization program. The majority of farm properties were initially settled in the first 5-8 years while later migrants moved further on down the highway or further out along the secondary side-roads. Since INCRA's land distribution program ended municipal governments have continued land distributions further out on the secondary side-roads.

Land distribution in this area followed a gridiron pattern. Individual families received parcels of land ranging from 90 to 120 hectares along the Transamazon Highway and its feeder roads. The primary unit of analysis is clearly the household. Property rights provide discrete and

[1] One exception is the volume People and Pixels (Liverman et al. 1998).

definable boundaries on land access and use. A property grid was developed prior to fieldwork. In a GIS, the property grid was overlaid with the remotely sensed data to extract information on land cover for each of the properties identified. This combination of GIS and satellite data has aided in: 1) pre-field work analysis and sampling, 2) fieldwork and interviews, and is expected to aid in 3) post-field work analysis, modeling and extrapolating to the broader landscape of land-use/land-cover change.

Brief Review of Research Project

During the course of previous fieldwork, Emilio Moran, observed that neighboring farms often had quite different patterns of land-use. Some of these differences could be explained by differences in initial capital of in-coming migrant families, and origin and previous experience of farmer colonists. Access to water, distribution of soil, distance to markets typically are shared among neighbors and provide less insight into these different patterns of land-use (among neighbors) while they appear to be important to the pattern at a broader landscape level. A reoccurring factor that seemed to contribute to these differences is the amount of household labor on these properties. In general large farm families were more likely to be involved in perennial crop production (fruit trees, cocoa or black pepper, and in a very few cases, hard wood trees). Smaller families typically placed greater long-term cash and labor investments in the creation of pasture and the grazing of cattle while only initially involved in annual cash crop production. Furthermore, agricultural strategies involving different combinations of annual crops, perennial crops and pasture appear to be related to the pattern and level of deforestation on the farm property.

During the initial stages of settlement, farmer families are typically involved in clearing forest, 2-5 hectares annually, and growing cash crops (rice, beans and corn). They commonly use these plots for two-to-three years for these annual crops until soil fertility declines. These plots are then either 1) left as fallow to be used again in later years or allowed to return to forest, 2) turned into pasture for cattle grazing, or 3) converted into perennial crop production.

Agricultural strategies that emphasize perennial crops typically involve lower rates of deforestation and, relative to other farms of similar age, lower levels of overall deforestation. Pasture development and cattle raising involve higher rates of deforestation but are smaller than farm activities which rely heavily on annual cash crop production. Families that depend on annual cash crops must continually open new sections of their land for planting; families involved in either cattle raising or perennial crops must devote a large share of their efforts in maintaining orchards and pastures. Maintaining pasture is a seasonal activity often involving the use of temporary labor while perennial crop activities require more continued efforts on the part of family members.

Working with these ideas we developed a conceptual model of household-farm land-use (i.e., annual crops, perennial crops, and cattle grazing) over time based on household labor over the course of the domestic life cycle. While labor supply of the households plays the underlying dimension of the model, other factors such as initial capital, credit availability, distance/transportation to markets, typography, soils and water availability are considered as important contributing and conditioning factors in the land-use strategies pursued.

Initially we anticipate that most settler-families will be involved in the conversion of forest into annual cash crop production. Slowly families will begin to diversify into cattle grazing (often simply as a capital saving strategy) and perennial crops (e.g., cocoa, black pepper, and fruit crops). Perennial crops require a relatively long start up time and substantial labor investment on the part of households – this activity is also limited by the types of soils and the availability of water. Farm families must generate an income from annual cash crops while investing labor and capital in the development of their perennial crops. In contrast, development of pasture requires less regular labor investment – and temporary, seasonal labor may be more important – and appears to be a more viable strategy for families with fewer members and a smaller labor supply. Households who are unable to accumulate enough capital for investment in raising cattle or in perennial crop production, it is anticipated, are faced with continued annual cash crop production or moving on. These households are expected to have higher levels of deforestation as they rotate around various plots on their farms. These households, it is anticipated, effectively deforest much larger areas of their farm early on in the settlement process.

Our objective is to examine the demography of households as these affect agricultural strategies pursued, and how these, in turn, shape farm levels and patterns of deforestation. This focus on households over the course of the domestic life-cycle also suggests that period, age and cohort effects should be considered as we try to distinguish between macro- level processes (e.g., credit policies and market prices) affecting households and age/cohort effects of household. These issues raise concerns for how best to draw a sample of cohorts of settler-households for in-depth interviews. Our sampling design is based on the use of remotely sensed data and a property grid to identify timing of settlement to obtain relatively equal numbers of households/farms from different cohorts arriving on the frontier. The following methods could be developed for arriving at different sampling schemes based on different research agendas. We now will discuss the work on remote sensing analysis, development of a property grid, and pre-field work analysis for survey sampling.

Remote Sensing Analysis, Property Grid Development and Pre-Field Work Analysis

Remote Sensing Analysis

Our current work at the household-property level with the use of sample surveys, a property boundary map/grid, and remotely sensed data draws heavily on previous work of the research team on secondary succession vegetation growth with satellite images and ground truth vegetation and soil data collection. Image classification was based on a hybrid approach and involved the use of previously collected training samples during fieldwork in 1992 and 1993. Training samples representing the proposed land cover classes were evaluated with statistical separability analysis to ensure high accuracy during classification (for more details on the remote sensing work related to this paper see Mausel et al., 1993; Moran et al. 1994). The final classifications were compared to other thematic maps of the area, with an overall accuracy above 85% for all classes. The following seven classes of land-use/land cover for the Thematic Mapper satellite images for 1985, 1988, and 1991 are: 1)Water; 2) Forest; 3) Secondary Succession of less than 6 years (SS1); 4) Secondary Succession of 6 to 10 years of age (SS2); 5) Secondary Succession of 10+ years; 6) Pasture, and 7)Bare Soils.

Two additional sources of forest cover were obtained from aerial photographs acquired in 1970, and from a census bureau map based on aerial photographs of 1978. These various sources provide a time-series of coverages for the analysis of land-use/land-cover change over the course of settlement and farm development for this region.

Development of a Property Grid

We developed the property grid overlay directly from perceived boundaries in individual satellite images and through their comparison. Pre-fieldwork development of the property grid was carried out in three stages: table-digitizing of roads, on screen property definition and digitizing, and property identifier assignment. The Altamira colonization scheme divided land into roughly rectangular lots, of similar spatial extent, distributed around a network composed of feeder or side-roads evenly spaced along the highway. The farm lots average 100 hectares in size (500 m x 2,000 m), and are, therefore, represented by approximately 1,100 pixels in a TM image. These similarities make the definition of properties more apparent than in other areas where it may not be possible to approximate the size and shape of properties ahead of time.

The property grid GIS layer with IDs constitutes a powerful tool for data extraction from the 5 image-coverages (1970, 1978, 1985, 1988 and 1991) for exploratory data analysis at the property level.

Exploratory data analysis of these data permitted the development of a stratified sampling frame for selecting properties and households based on 1) timing of settlement considering the period of initial forest clearing and 2) extent of deforestation. Other sampling methods could be developed with these data based on patterns of land-cover/land-use or for identifying farm properties that have unique or unusual characteristics.

For the purposes of generating a stratified sampling frame congruent with our conceptual model of life cycle of household, agricultural strategies, and deforestation property level analysis of remotely sensed data was very useful. Explicit in our model is the need to disentangle period effects (e.g., credit policies related to cattle and cocoa) from cohort and age effects that may be related to farm development and stages of the domestic life cycle of households. Because in this region, closest to Altamira, the majority of farm properties were settled between 1970 and 1978 we were particularly interested in over-sampling early and late colonists households for comparison and analysis. By stratifying our sampling frame first by timing of initial clearing and subsequently by level of deforestation in 1991 we were able to obtain a sample to address our research questions. With this strategy we will be able to compare households at similar stages of farm development and stages of the household life cycle who arrived during different periods. A series of hypotheses have been developed to understand how households differentially respond to credit policies in their agricultural strategies. Preliminary analysis of land-use data from aerial photography and TM images at the property level was then used to develop a temporally and spatially stratified sampling of farm properties and their households.

First the farm properties were divided into 5 “cohort” groups based on the period in which 5 hectares of forest were initially cleared (pre-1970, 1970-78, 1978-85, 1985-88, 1988-91). As we are interested in the pattern and level of deforestation, each cohort group was further stratified by percent area in forest in 1991 to insure that we selected households across the range of deforestation. We over sampled early and older cohorts as most of the properties in this eastern section of the study area were settled between 1970 and 1978. Deciding on an appropriate number of properties for each group we developed a sampling fraction for each and selected every *n*th case.

<u>Comparison of Pre-Field TM Based Property Grid and Field GPS/Survey Property Grid</u>

During the 1st fieldwork, two teams were involved in data collection. The first team carried out extensive interviews with the male and female heads of households with two survey instruments: one on land-use histories; and another on socio-demographic characteristics of the household. A second team focused on collecting GSP points along side-roads and at property boundaries for correcting the property grid. This GPS data was collected in differential mode increasing the accuracy of data collected to within a few meters. Additional information was collected at each of the points on the farms (Gleba/Lot) bordering the property markers. This often included looking at land titles with the respective farmers and permitted the re-development of a geo-corrected property grid based on very accurate differentially calculated GPS points. The post-field season property grid was developed independently of the earlier pre-field TM property grid based on GPS points collected in the field and the use of a series of more recent land surveys maps obtained from INCRA. Figure 2 provides a comparison of these two property grids for a section of our study area (km 20 to 40). FIG. 2

A comparison of the pre-field and "new" property grids indicates that digitizing property boundaries from satellite images worked relatively well but contains some critical errors along some feeder roads (Figure 2). Also note the errors on the most westerly feeder road. Here our on-screen work led us to digitize side-boundaries of properties diagonally across neighboring properties. Also note the islands created by detours in the roads. In some instances farmers have left these pieces of land idle. Others allow neighboring farmers to use them or have re-negotiated property boundaries. Errors in the pre-fieldwork property grid may be more serious in areas to the west of the "pilot area" which were settled later, and in areas were detours in roads have developed since there initial construction.

For the 398 properties in the pilot study, the average area in forest in 1991 was 35.3 hectares. Between 1988-1991, approximately 3 hectares were cleared annually, and the range of area deforested runs from near "no-deforestation" to 13.6 hectares per year. In regards to forest regrowth, on average the farm properties had approximately 7.5 hectares of regenerated forest, with ranges from none to 25 hectares.

More sophisticated use of remotely sensed data can be incorporated at the property-level using GIS cross-tabulation methods to develop a transition matrix between different classes of land-cover (Brondizio et al. 1994). This involves the pixel-by-pixel comparison of changes between land-cover classes between each image/cover and can be useful in the development of land-cover trajectories for individual properties.

The analysis of remotely sensed data with a property grid can also be used to identify and investigate extreme or 'outlier' patterns of land cover change. During the course of analysis of remotely sensed data for the landscape we noticed that between 30 and 47 percent of the forest between 1978-85, 1985-88, and 1988-91 were being converted into a vegetation structure similar to what had been identified as advanced secondary succession (SS3). This was puzzling. This forest-SS3 conversion suggests that forests have somehow been degraded. At first this pattern of forest-SS3 conversion raised doubts about our remotely sensed data classification methods. Other possible explanations might include 1) selective logging, 2) reduction in ground water or rainfall, or 3) spreading of fires into localized forest areas. Further investigation of this forest-SS3 conversion for previous periods (1978-85; 1985-88) also indicates similar patterns. Furthermore differences between periods could not be explained by rainfall patterns. The dispersion of the forest-SS3 conversion should give us some idea as to the nature of the phenomenon. If the phenomenon is related to a problem with our classification methods, reductions in ground water or rainfall, we might expect this forest-SS3 conversion to be spread out and highly dispersed over the landscape. If it were concentrated, especially on particular properties, then we might suspect that it is related to selective logging and/or localized fires. The latter appears to be the case. Figure 3 illustrates that the phenomena of forest-SS 3 conversion is clustered. During the course of the up-coming field research we will pursue investigation for this important land-cover transformation. FIG. 3

Conclusion

The development of a property grid that can be associated with individual households and their farms permits pre-field work analysis of remotely sensed data to develop a sampling frame that reflects the needs of the research investigation. Apart from aiding in the development of a temporal and spatial sampling frame, the property grid also aided in 1) locating farms during fieldwork, and 2) interviews with farmers in the development of a land-use history for their farms. The property grid can also be used to analyze remotely sensed data at the individual property level for all properties for development of land-cover trajectories associated with different types of farms.

Acknowledgments:

The project "Amazonian Deforestation and the Structure of Households: an Georeferenced Approach" is a collaboration between the Anthropological Center for Training and Research on Global Environmental Change (ACT) at Indiana University and EMBRAPA/CPATU and has been made possible by the support of the National Institute of Child Health and Human Development of the National Institute of Health. The authors are thankful to our collaborators at EMBRAPA, Dr. Adilson Serrão and Dr. Italo C. Falesi, and to the support of EMATER and CEPLAC in Altamira, Brasil Novo and Medicilândia. We acknowledge the extensive help given us by Vonnie Peischl and Cathy Snider. We also thank George Ball and Cliff Hathaway of the Advanced Research Technologies Lab (ART) at the University of Arizona, Tucson for assistance in the development of the property grid. We are equally thankful to the International Institute of Tropical Forestry for providing support for the GPS work done by Carlos Rodriguez-Pedraza. We also thank our collaborators in Altamira, and the farmers of the TransAmazon Highway. The views expressed in this paper are sole responsibility of the authors.

Cited References:

-Behrens C. (**1994**). Recent advances in regional analysis of indigenous land use and tropical deforestation. **Human Ecology** 22(3) (introduction).

-Brondízio E.S., Moran E.F., Mausel P., and Wu Y. (**1996**). Changes in land cover in the Amazon estuary: Integration of thematic mapper with botanical and historical data. **Photogrammetric Engineering and Remote Sensing**, 62(8):921-929;(special issue: *Vegetation & cover analysis*).

-Brondízio E.S., Moran E.F., Mausel P., and Wu Y. (**1994**). Land use change in the Amazon estuary: Patterns of Caboclo settlement and landscape management. **Human Ecology**, 22(3):249-278 (special issue: *Recent Advances on the Regional Analysis of Indigenous Land Use and Tropical Deforestation*).

-Liverman D., Stern P., Rindfus R., and Moran, E. (**1998**) **People and Pixels: Application of Remote Sensing Technology in Social Sciences**. National Academy Press, Washington, DC.

-Mausel P., Wu Y., Moran E., and Brondízio E.S.(**1993**). Spectral identification of successional stages following deforestation in the Amazon. **Geocarto International**, 8(4): 61-71 (special issue: *Global Environmental Change*).

-Moran E.F., Brondízio E.S., Mausel P., and You W. (**1994**) Integrating Amazonian vegetation, land-use, and satellite data. **Bioscience** 44(5):329-338(special issue: *Global Impact of Land Cover Change*).

Figure 1.

Pace of Deforestation – Altamira Pilot Study Area

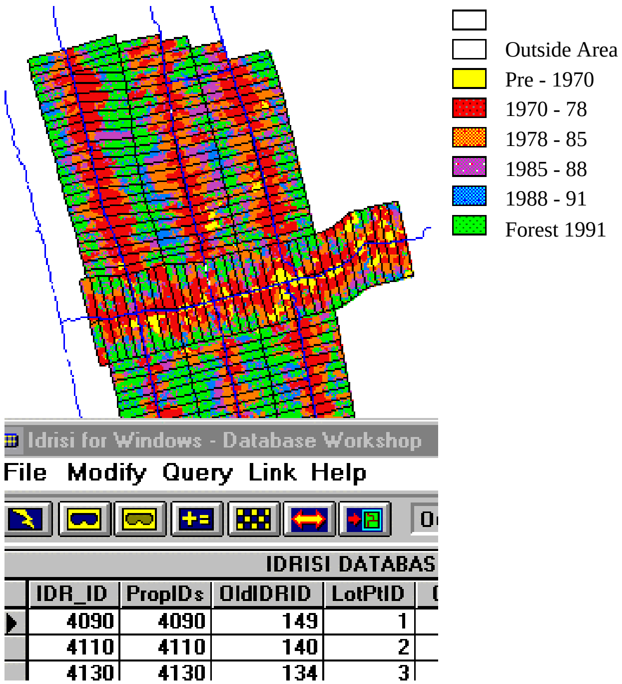

Figure 2.

Pre-Field TM Grid vs. GSP/Survey Map

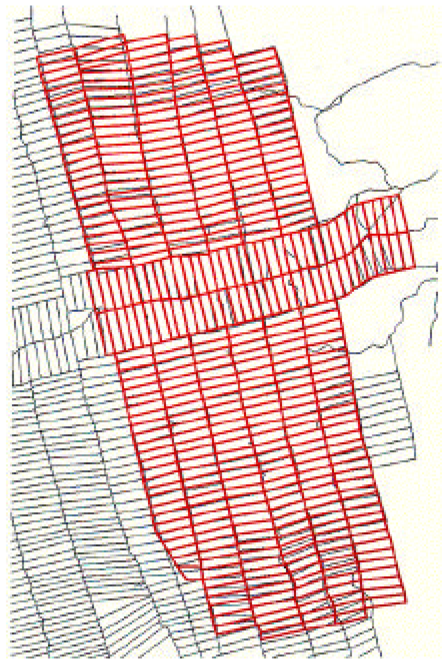

Figure 3.

Forest Change 1988-1991 the Conversion of Forest to SS3