CHAPTER 1: INTRODUCTION

1.1 BACKGROUND

Deforestation was one of the first environmental topics discussed at a world level when the term sustainability was introduced in the UN Conference on Environment and Development in Rio de Janeiro in 1992[1]. In that conference the management of the Amazonian forest was a key issue in discussion (Kolk, 1996). Unfortunately deforestation is still an ongoing environmental problem. According to estimates of the Global Forest Resources Assessment, (UN FAO, 2001) on average the world's natural forests decreased by 16.1 million hectares (ha) per year during the 1990s, which represents a loss of 42% of the natural forest that existed in 1990. On the other hand FAO estimated that tropical regions lost 15.2 million ha of forest per year during the 1990s (FAO, 1995).

Deforestation is threatening tropical forests and their capacity to provide economic and ecological services not only for sustainable local livelihoods but also for ecosystems and climate equilibrium on the Earth. During last four decades there has been increased interest in tropical deforestation since forest loss is linked with disruption of hydrological regimes (Shukla et al 1990), degradation of soil (Hecht 1981; Buschbacher et al 1988), loss of species (Myers, 1980, Wilson, 1989) and changes in green house gases in the atmosphere, which induce climate change (Houghton et al 1983; Post et al 1990; Dale et al 1991). Brazil has the potential to provide for global warming benefits if deforestation is stopped or slowed down, because the tropical forest

[1] Sustainability: The concept of meeting the needs of the present without compromising the ability of future generations to meet their needs (*Our common future*, 1987). In the context of forests, this term applies to the use of ecosystems and their resources in a manner that satisfies current needs without compromising the needs or options of future generations.

can act as a sink of CO_2 and because CO_2emissions caused by deforestation are large, about 250-350x10^6 tons annually as compared to approximately 60x10^6 tons from fossil fuels (Fearnside, 1999; Laurance, 2000). Moreover, there is a growing concern that more frequent and catastrophic El Niño events will occur as a result of the combination of massive deforestation, forest fragmentation, logging, and forest burning from thousands of small farmers (Laurance, 2000).

1.1.1 <u>Deforestation in the Brazilian Amazon and the role of small farmers</u>

The Brazilian Amazon, the largest remaining rainforest, covers approximately one-third of the total rainforest area in the world. Brazil's "Legal Amazon" region covers 5 million km^2 of which 4 million km^2 was originally forested[2]. Approximately 3.5 million km^2 (87%) of this originally forested area was still standing as of 1997 (Fearnside, 1999, Molofosky et al, 1985).

During the last four decades the Brazilian Amazon has experienced increasing rates of deforestation. According to satellite estimates of deforestation by Brazil's National Institute for Space Research (INPE), deforestation in the Legal Amazon has increased from a total land area of 155,200 km^2 in 1978 to 551,782 km^2 in 1998 (INPE, 2001). Such figures correspond to 4.4 % and 15.6%, respectively, of all land originally forested in the Brazilian Amazon. This implies an annual average increase of 6.2% in the land area deforested (INPE, 2001). However, the Brazilian Amazon is still the largest tract of forest and now one of the "hot-spots" for bio-conservation and land cover research.

[2] The Legal Amazon is made up of the entire North region (the states Acre, Amapa, Amazonas, Para, Rondõnia, Roraima, and Tocatins) plus parts of the states of Maranhao, Mato Grosso, and Goias. The southern edge is the 16th parallel, and the eastern edge is the 44th meridian.

The tremendous land cover change (LCC) in the Brazilian Amazon has been linked to many forces, in particular to the extension of the Brazilian highway network and large-scale development projects for mineral extraction, hydroelectric development and logging[3]. Research has grown around the study of social and biophysical changes linked to colonization projects in the Amazon (Browder, 1988, Evans, 2001, Binswanger, 1991). Research has focused on depicting the social evolution and landscape change of this "new" land, often called the Amazon Frontier and special interest has been given to agricultural colonization projects[4]. Regarding the landscape change, most of the land allocated to the colonists was mature forest prior to settlement but it turned into a mosaic of pasture, croplands and different stages of forest re-growth associated with fallowed agricultural land (Evans, 2001).

Small farmers are regarded by many researchers and policy makers as "major deforesters" in the Amazon mainly because of their clearing practices for agricultural and cattle ranching purposes[5]. Small farmers are commonly defined as farmers with properties having an area less than or equal to 100 ha (Fearnside, 1999). Research estimates indicate that these farmers are responsible for about 30% of the deforestation in the Brazilian Amazon (Fearnside, 1999). They are important deforestation agents because of their large number, their incentives to clear land to prove land occupation, and their mobility.

---

[3] Land cover: refers to the "attributes of a part of the Earth's land surface and immediate subsurface, including biota, soil, topography, surface and groundwater, and human structures." One type of land cover is forest. (Turner et al, 1993 in Lambin et al, 2000)

[4] Frontier: a land that was not accessible before and an area sparsely populated by indigenous communities.

[5] Small farmers are commonly defined as farmers with properties, which area is less or equal to 100 ha (Fearnside, 1999).

Population growth, colonization of the Amazon frontier and deforestation are interrelated. However, theorists still disagree on what is the effect of population changes on deforestation trends. In the 1970's, population in the Legal Amazon grew at almost 4 percent per year showing also net migration gains of almost 20,000 people per year (Perz, 2001). This population growth was suggested as one factor that favored increased deforestation among farms, since family labor was readily available. During the 1980's population growth in the Legal Amazon slowed down to 3 percent per year, showing also net migration losses of about 40,000 people per year. In the 1990's, Brazilian Amazon's population growth was 2 percent per year (Perz, 2001). Some observers consider this slow down in population the result of the aging process in the colonists' families. Moreover this household aging process is linked to the change in land use strategies from annual crops to cattle ranching, under the premise that annual crops require larger number of young laborers than cattle ranching.

1.1.2 <u>Modeling deforestation and small farmers</u>

Simulation models of anthropogenic land-cover changes can be broadly divided into two groups, according to their capability for spatial representation and the nature of the data entered into the model. The first group is composed of models developed in Physical Geography during last decade making use of the advancements in satellite images and Geographic Information Systems (GIS) in order to provide spatially explicit results through maps of future land cover scenarios. Most of these models use the pixel as the unit of analysis and predict the future classification of the pixel based on a set of landscape variables that can be "seen" through satellite images or can be calculated using GIS (i.e., distance to road, distance to water bodies, density of forest patch, measures of forest aggregation, etc) (Irwin and Geohegan, 2001, Geohegan, 2001, Mertens, 2002)[6]. Other statistical models in the Geography field use these spatial variables to predict land cover changes through the use of linear regression techniques (i.e. Mertens and Lambin, 1997; Andersen, 1996; Ludeke et al, 1990) without a spatial representation.

The second group of models includes a large number of statistical analyses that apply linear regression to survey data with the aim to explain general land cover changes without providing spatially explicit results (i.e. Godoy, 1997; Aspinal, 2004). These models are mostly developed by Human Geographers, sociologists or development researchers.

Debate exists about which approach, the one using landscape variables only or the one using survey data, provides better explanatory and predictive results in terms of accuracy and regional applicability. It is a special challenge to mix both approaches to produce spatially explicit predictions while at the same time linking these landscape changes to explanatory social, economic and demographic variables. This hybrid alternative would require vast and appropriate empirical data, both cross sectional and longitudinal/panel data[7]. These hybrid models among other research possibilities make the future of land-cover change modeling a fertile ground for more accurate and user-friendly predictive models, given the improved availability of survey and landscape data, and the computational capabilities nowadays.

[6] A pixel or pixel element, is the smallest addressable unit on a display screen or bitmapped image. In a Land Sat TN image or GIS map, a pixel in the image represents a 3m x 3m cell. in real area.

[7] Cross sectional data is composed of measures or variables taken at a certain point on time.
Panel data involves the measure of the same variables over time on the same subjects.

This debate expands to the simulation of small farmers' LCC practices, where survey data collection in the rural Amazon is often questioned due do the extensive resources demanded (i.e. time, human and financial capital, transportation). On other hand, there is an increasing need for more specific simulation models to predict small farmers' land cover change and deforestation patterns, given their important role in shaping the agricultural Frontier. Furthermore, changes in landownership and land use strategies among small farmers are occurring at a fast pace. Thus, the initial assumption, that all small farmers would deforest at the same rate and can be treated all as a homogeneous group based only on their farm size, may not reflect deforestation on the Frontier..

New lines of research should focus on modeling land use/land cover changes taking into consideration the smallholders' household characteristics that affect their decisions, while making use of the detailed and accurate landscape information that GIS provides. The emphasis on demographics, socioeconomic variables and land cover variables at the household level relies on past research indicating that small farmers are more affected by endogenous variables when compared to large farmers that are more sensitive to exogenous variables (Fearnside, 1999).

The research presented in this dissertation addresses the debate among human geographers and landscape modelers about the utility of collecting and using household data in addition to landscape data when predicting land cover change, specifically deforestation. The two modeling approaches -- using only land cover change variables compared to using demographic, socioeconomic and land cover data at the farm level – are compared based on their explanatory and predictive capacity. Two common assumptions used when modeling small farmers' land cover change strategies are tested. The first assumption refers to the "poolability" of small farmers, where they are considered as a homogeneous group with regard to their probabilities for specific land cover changes. The second assumption relates to the probabilities of land cover change, often considered to be constant through time, without further proof. Finally, the research results are translated into some planning lessons, considerations for future predictive models, and ideas for development strategies with sustainable conservation considerations and, rather than competing strategies.

1.2 RESEARCH QUESTIONS AND HYPOTHESES.

The present dissertation research has three purposes: the first one is to predict anthropogenic deforestation caused by small farmers first using only pathways of past land cover change and second using demographic, socioeconomic and land cover data at the farm level. The second purpose is to compare the explanatory and predictive capability of both approaches at identifying areas at high risk of deforestation among small farms in Rondônia, Brazil. The third purpose is to test the assumptions of stationary probabilities and homogeneous subjects, both commonly used assumptions in predictive stochastic models applied to small farmers' deforestation decisions[8].

The above research goals can be stated in terms of objectives, research questions, and hypotheses as follows:

[8] Stochastic models are used to simulate stochastic processes, which are phenomena that vary to some degree unpredictably as time goes on. The main characteristic of stochastic processes is the use of probabilities.

OBJECTIVE 1:

Describe patterns of the transitional probabilities of land cover change among small farmers on Rondônia, Brazil, especially different patterns among subdividing, expanding and stable owner types[1]. Test the assumption of homogeneous probabilities among subjects, a commonly used assumption in stochastic predictive models applied to small farmers' deforestation decisions.

RESEARCH QUESTION 1:

- Are there significant differences in the transitional probabilities of LCC among small farmers in Rondônia, especially among subdividing, stable and expanding farmers?

In other words, does the assumption of homogeneous probabilities among farmers hold for small farmers in Rondônia?

HYPOTHESIS FOR OBJECTIVE 1:

H_1: Probabilities of land cover change are not constant (homogeneous) among small farmers in Rondônia, so they cannot be treated all as one homogeneous sample.

H_2: There is a significant difference among the LCC transitional probabilities of stable, subdividing and expanding farmers (owner types 1, 2 and 3), a classification based on how the area of their total landholdings changes through time.

[1] Owner types: stable (same total area owned in 1992 and 2002), expanding (more area owned in 2002), subdividing (less area owned in 1992).

OBJECTIVE 2:

Describe temporal patterns of the transitional probabilities of land cover change among small farmers in Rondônia and test the assumption of stationary probabilities through time, which is a broadly used assumption in stochastic predictive models of LCC.

RESEARCH QUESTION 2:

- Are the probabilities of land cover change (land cover transitional probabilities) constant through time among small farmers in Rondônia?

In other words, does the assumption of stationary probabilities through time hold true for the LCC process among small farmers in Rondônia?

HYPOTHESIS 2:

H3: Probabilities are not constant through time for the sample of farmers as a whole.

H4: Probabilities could be assumed constant through time for a short study time period for homogeneous group of farmers, specifically owner types (stable, subdividing, expanding).

OBJECTIVE 3:

Compare the explanatory and predictive capacity of two modeling approaches to predict land cover change. The first one uses only pathways of past land cover change and the second uses household survey data (demographic, socioeconomic and land cover data at the farm level).

RESEARCH QUESTION 3:

- Can the use of demographic, socioeconomic and land cover data at the farm level significantly change the explanatory and predictive power of models of land-cover change to identify areas at risk of deforestation among small farms in Rondônia?

HYPOTHESIS 3:

H5: The model using household survey data has more explanatory and predictive power than the model using only historical land cover changes.

H6: Land cover changes among small farmers are better explained and predicted when classifying farmers according to how their total landholdings area changes through time (owner typology).

1.3 OVERVIEW OF THE METHODOLOGY

The present dissertation research uses data generated by the NSF-funded project BCS-0136965, Patterns and Processes of Landscape change in the Brazilian Amazon: A Longitudinal, Comparative Analysis of Smallholder Land Use Decision-Making. The project was directed by Profs. John O. Browder and Randolph Wynne (Virginia Tech), and Robert Walker (Michigan State University)

This dissertation uses the following data: household surveys, maps, satellite images and their land cover classification at the pixel level, and pathways of past land cover change for each farm. This data is available for a panel sample of farms in three municipios in Rondônia, Brazil

(Alto Paraiso, Nova União, and Rolim de Moura) and covers a ten-year period of study (1992-2002). Survey data have been integrated with Land satellite Thematic Map (Landsat TM) imagery of each study site for the entire ten-year period. Land-cover pixel classification of the satellite images was performed by M.S. Katherine A. Budreski in her Master's thesis, under the supervision of Prof. Randolph Wynne.

Pathways of past land cover change are graphic representations in the form of flow charts that depict land cover changes (LCC) in each farm during the ten-year period of study. Pathways were constructed using satellite images, survey data and maps, and a set of interviews performed in a sub-sample of 70 farms.

A research approach was designed for each of the research objectives and is described below. The influence of differences among farmers on the LCC probabilities is studied in objectives 1 and 3. Thus, the panel data analysis is mentioned in the two corresponding sections of the methodology and the discussion of results.

RESEARCH APPROACH FOR OBJECTIVE 1:

- Estimate empirical transition probabilities per each LCC, per year per farmer, using the pathway analysis.
- Conduct descriptive statistics analysis (analysis of difference of means) and graphs of probabilities versus time for each owner type sub-sample and for the whole sample
- Perform panel data analysis of transition probabilities (group effects fixed model)
- Conduct a poolability test by group with a partial F-test

- Test the use of other classifications besides owner type (OT), for example property type (PT), farmer type (FT) and municipio (MUN).

RESEARCH APPROACH FOR OBJECTIVE 2:

- Estimate empirical transition probabilities
- Conduct descriptive statistics analysis (analysis of difference of means) and graphs of probabilities Vs time for each LCC.
- Perform panel data analysis of transition probabilities (time effects)
- Conduct a poolability test by time with a partial F-test

RESEARCH APPROACH FOR OBJECTIVE3:

- Panel data analysis is performed to describe trends on the transitional probabilities of land cover change (previously estimated through pathway analysis) and to test two common assumptions used in stochastic predictive land cover change models. The first assumption relates to the homogeneity of probabilities among small farmers. The second assumption is about the stationary (constant) probabilities across time.
- Panel data analysis is used to predict future probabilities of land cover change by using past probabilities of land cover change.
- A multinomial logit regression model is used to explain and predict land cover change through the use of household survey data.
- The model that best explains/ predicts the land cover in the farmer pixels for next time period (year) is used in the Markov chain model to predict LCC in the short term future (several time periods later, five years later).

1.4 CONTRIBUTIONS OF THE RESEARCH

The contribution of this research is both, methodological and theoretical. Methodologically, the importance of this research lies in the estimation and analysis of empirical transition probabilities. To the best of my knowledge, probabilities per farmer, per year, per LCC and their temporal trends in a panel sample have not been formally tested in previous research. Another contribution to simulation methology for deforestation is the testing of underlying assumptions about the transition probabilities through the comparison of predictive accuracy when probabilities calculated under different assumptions are used in a Markov chain model.

Theoretically, this work tests variables at the farm and household level that have been suggested by the literature as causes of LCC and deforestation in the Amazon. The rationale supporting the selection of these variables relies on Frontier and deforestation theories. This dissertaion focused on exploring the variables in the context of their contribution in a LCC predictive model.

1.5 OVERVIEW OF THE CHAPTERS

This dissertation contains six chapters. Chapter two provides a literature review of the agents of deforestation in the Brazilian Amazon, from the global forces to national, regional, and local agents. The literature review lands on a description of the deforestation scenario and colonization process in Rondônia, Brazil and the specific role of small farmers as agents of deforestation. Chapter 3 presents a literature review of mathematical models to predict land cover change. This chapter evaluates which approach or combination of models best addresses the research questions as presented in this dissertation. Chapter 4 describes the study sites and research methodology. Methodological choices are described, including the statistical models selected and the way the Markov chain model is modified to accommodate the LCC probability functions. Chapter 5 presents and discusses relevant results in the context of the research questions. Results follow the same order presented in the methodology section. Chapter 6 summarizes conclusions derived from the key findings of this doctoral dissertation

CHAPTER 2: LITERATURE REVIEW: CAUSES AND AGENTS OF DEFORESTATION

2.1 CAUSES AND AGENTS OF DEFORESTATION IN THE BRAZILIAN AMAZON

At the landscape level, deforestation can be reduced and studied as a single land-cover change or the result of a chain of land-cover changes. However, the forces, agents and interrelations behind the deforestation process form a complex network that is constantly changing over time (Perz, 2002). The forces or actors affecting land conversion in the Amazon include: human actors, social groups, state agencies, market conditions, political conditions, and natural disasters, among others (Kaimowitz and Angelson, 1998; Brown and Pearce, 1994).

Causes of deforestation can be broadly divided into proximate causes and underlying causes (Turner, et al, 1995). Proximate causes of deforestation are more localized and include among others: individual decisions to clear vegetation, micro processes at the household or firm level. Underlying causes include among others: state policies to support Amazon colonization by farm families, macro scale processes in Brazil's national economy and forces at the global and international level (Wood et al, 1996). Proximate causes are suggested as the main forces affecting anthropogenic deforestation at local level among small farmers in the agricultural frontier.

On other hand, the causes of deforestation can be studied at different scales, ranging from household to local, to regional, to national and to global levels (Angelsen and Kainmowitz, 1999). This section will review the literature on the diverse causes and drivers of deforestation in the Brazilian Amazon, from the global level to the household level.

2.1.1. Economic globalization’s effects on land use change

Globalization processes are not direct drivers of deforestation, but instead underlying forces that amplify or attenuate driving forces of land-use/land-cover change (Lambin, 2003). Globalization forces have affected land-cover change and deforestation in the Amazon through three key changes at the global level: trade liberalization and reforms in the agro-industrial sector, the emergence of global environmental politics (Perz, 2002), and development projects funded by multilateral banks.

2.1.1.1 Trade liberalization and reforms in the Brazilian agro-industrial sector

Changes in agricultural practices in the Amazon provide a good example of the direct and indirect impacts of globalization on land-cover change and deforestation. For example, a way in which land use change is affected by economic globalization is through trade liberalization and reforms to open up the agro-industrial sector to international markets and investors (Barbier, 2000). These reforms, along with economic incentives may affect small and large landowners’ decisions on how to invest or use their land. Direct impacts on land degradation occur as increased agricultural activity leads to conversion of forests and increased use of “unsustainable” production methods. However, there may also be indirect effects if agro industrial development displaces landless, near-landless and rural poor, who then migrate to marginal agricultural lands and forest frontier regions. (Barbier, 2000)

> “This ecological marginalization usually follows population growth, agricultural modernization- associated with mechanization and consolidation- inequalities in land tenure in the most fertile and accessible agricultural regions, or other pressures of social or political origin. It leads to migration of poor farmers into areas with high ecological sensitivity for which existing management practices may be inadequate.” (Lambin et al, 2003, p. 230)

A vicious cycle occurs when poor rural households abandon degraded land for frontier forested lands, deforestation and cropping of poor soil lead to further degradation, which in turn leads to land abandonment and additional forestland conversion, and so on. (Barbier, 1997) The main question is if economic liberalization reforms have further increased such processes of rural resource degradation (Barbier, 2000) and at which extent.

International demand for Brazilian Agricultural exports has also influenced to a significant extent changes in Amazon farming strategies, land cover-change and its intensity. During 1990s, soybeans played an important role in export earnings from the European Community (OECD, 1997). Soybean cultivation in the Amazon increased from 33ha in 1975 to 858,000ha in 1985 to 1.66 million ha in 1996 (IBGE, 1979, 1990, 1998a) and the expansion of the soybean cultivation reflected on the expanded exports of soybeans.

The soybean phenomena of the 1990’s in the Brazilian Amazon has been described as the “Soybean story”, by (Perz, 2002):

> "High prices in OECD countries (global level) and new bank loans to Brazil, which emerged as a large soybean exporter (national level), to fund new infrastructure projects in the Amazon to open new land for soybean cultivation (regional level), raised value of the land close the roads and stimulate investment (local level), also encouraging landowners to clear land for the prospective returns (household/firm level)" (Perz, 2002, p 232)

2.1.1.2 International environmental politics.

In the international arena, deforestation in the Amazon has become not only an environmental concern but also a political issue. International interest focused on the Amazon when satellite-based estimates of deforestation were released in the late 1980s (Perz, 2002). The Brazilian government received pressure from governments and NGOs in the OECD to manage its forests more responsibly. The Group of Seven organized meetings to discuss the topic and environmental groups in OECD countries mobilized media coverage. (Kolk, 1996) Debate existed over the question of Brazil's sovereignty to continue deforesting in order to achieve economic growth (Hurrel, 1991). During the UN Conference on Environment and Development in Rio de Janeiro in 1992, the management of the Amazonian forest was a key issue in discussion (Kolk, 1996).

Brazilian and international NGO's formed alliances with local grassroots and mobilized in the fight for tenure rights and forest conservation. These alliances proliferated especially after the murder of rubber taper- turned into environmentalist Chico Mendes in 1988. The Chico Mendes case was one of many in which local grass roots groups looked for national and international NGO support in order to protect the forest and their livelihoods (Perz, 2002). Brazilian capitalists and ranchers enjoyed better organization, funding and good relationships with the state government, and therefore a better position to fight for their rights to exploit the forest in their pursuit of economic development. This unequal fight to protect and to exploit the forest moved to the national and international arena thanks to the alliances formed between local and international NGO's, who found donors and pressed to make legislative changes.

Another example of international NGO's intervention to protect the Amazon is their participation in the re-shaping of the logging industry. Mainstream environmental organizations (NGOs), such as Greenpeace, the Worldwide Fund for Nature (WWF) and Friends of the Earth (FoE) directed their efforts to the timber industry based on the assumption that timber consumption in the North caused deforestation in the South. During the 1980s the strategy was to boycott tropical timber. During the 1990s the strategy focused on using the idea of "sustainable development" to convince loggers to improve their forestry practices. The Forest Stewardship Council (FSC) was created by U.S. tropical hardwood importers to develop a certification scheme (Zhouri, 2004).

International mobilization by environmental NGOs to protect the Brazilian Amazon followed the launching of development projects funded by multilateral development banks (MDB) in 1983 (Zhouri, 2004). Projects, such as the POLONOROESTE in Rondônia --a large colonization and infrastructure project which advanced the frontier towards the northwestern parts of the Amazon-- have been the subject of protests by grassroots movements in Brazil.

NGO's efforts focused also on the Grande Carajas project, aimed at the construction of transport infrastructure and mining facilities to exploit large mineral reserves in the southeastern Amazon.

NGOs used intense media documentation as a strategy to increase public awareness of the negative environmental and social effects related to MDB project lending (Kolk, 1998). Their goal was to make clear the links and responsibility of decisions of Northern governments and financing institutions on the environmental degradation in the Soth. (Keck and Sikkink, 1998; Kolk, 1996; Rich 1994). NGO's campaign emphasized the modes of living of indigenous people, the so-called forest people, as sustainable alternatives to development (Zhouri, 2004).

Transnational activism and international interest in protecting the Amazon has focused on different environmental threats since it first started in the 1970's, when environmental issues gained momentum (Keck and Sikkink, 1998). Efforts have focused on wildfires, logging, mining, cattle ranching, and development and colonization projects, among other causes of deforestation. The environmental concern has moved from environmental pollution to loss of biodiversity and more recently to climate change. Brazil has a potential for global warming benefits if deforestation is stopped or slowed down, because the tropical forest can act as a sink of CO_2 and because emissions caused by deforestation are pretty large, about 250-350x10^6 ton C annually as compared to approximately 60x10^6 ton C from fossil fuels (Fearnside, 1999; Laurance, 2000). Moreover, there is a growing concern that more frequent and catastrophic El Nino events will occur as a result of the combination of massive deforestation, forest fragmentation, logging, and slash and burn cultivation. (Laurance, 2000).

2.1.1.3 Development projects funded by multilateral banks.

Multilateral banks funding for both, development and conservation projects in the Amazon, has generated controversy in terms of the projects' goals, implementation and results. Moreover, the magnitude and direction of their contributions to the deforestation process is also subject of debate.

Since 1980, most of the funding from multilateral banks for development projects in the Amazon has been used for highways construction and infrastructure projects. Highways have been regarded not only as direct drivers of deforestation and landscape fragmentation, but also as conduits for landless farmers, loggers, miners and other deforesters to penetrate the Amazon forest (Nepstad, et al 2001). Research shows that more than two-thirds of Amazon deforestation has taken place within 50 km of major paved highways (Alves, 1999, Nepstad et al, 2001). The Brazilian government's national development economic plan *Avança Brasil* has as goals to pave, recuperate or construct 6245 km of roads in the Amazon, which would nearly double the area of forestland accessible by paved highways, including the 192,000km^2 of fire prone forest (Laurance et al, 2001; Nepstad et al, 2001). Nepstad's analysis (Nepstad et al, 2001) found that these roads would stimulate 120,000-270,000 km^2 of additional deforestation and forest impoverishment through logging and fire.

Other examples of development projects during the 1980's in the Amazon include: Grande Carajas and Polonoroeste. The Grande Carajas Project (PGC) was a US$61.2 billion project for construction of extraction, processing, and export facilities for the world's largest iron deposit (Perz, 2002). The iron mine and 890 km railroad to the coast raised land values and led to land speculation, deforestation and conflicts (Hall, 1989). The Northwest Development Pole (Polonoroeste) was a U$1.5 billion project, with $500 million from the World Bank. This project sought to pave BR364, a key highway in Rondônia and Mato Grosso, and regularize land settlement via accelerated titling and credit programs. Polonoroeste led to a massive in-migration and rapid deforestation in both states (Millikan, 1992). Again, the magnitude and direction of these projects' contributions to the deforestation process is subject of debate.

During the 1990's there were also projects that involved state support and foreign investment in large infrastructure projects. The Northern Corridor Highways (ECN), a U$603 million project with US$220 million from the World Bank (Perz, 2002) had as its objective the creation and improvement of highways in order to open "unsettled" land for soybean cultivation along the Amazon forest margins. The ECN led to new land conflicts, deforestation and soil erosion (Hageman, 1996: chapter 2). Another example is the Agricultural and Forestry Plan for Rondônia (Planaforo), a US$229 million project successor to Polonoroeste, with US$167 million from the World Bank. The project aimed to help the state aid agriculture and agro forestry activities via improved infrastructure, with ecological zoning to focus development on already settled areas (Perz, 2002). Implementation of Planaforo began in 1993, and by 1994, local nongovernmental associations (NGOs) participated in a forum for the management and evaluation of the project. NGOs heavily criticized the lack of participation by local groups, the emphasis in spending for road improvements and credit for cash crops (Hageman, 1996: chapter 4). NGOS complained also because many local governments would have reduced their revenues because they were zoned for conservation rather than deforestation and local people needed a livelihood. Since then, there has been increasing pressure from local and international NGO's to change zoning guidelines to allow for more deforestation (Mahar, 2000)

2.1.2 National-level driving forces of land-cover change.

A challenge when studying land use/land cover changes such as deforestation is that there are complex interrelations among the driving forces and causes of deforestation and it is difficult to distinguish them as endogenous or exogenous. Some forces that can be regarded as national forces are very sensitive to international politics and funding. On other hand, local and regional forces may re-shape the way national policies affect land-cover change. Some examples include: national demand for land, policies to develop the forest frontier, capital investment in logging and agricultural activities, population movements, comodification of the economy, the development of urban markets and infrastructure expansion.

This section will describe the following relevant driving forces of land-cover change and deforestation at the national level: national land economy, Brazilian government's development projects and policies, and national conservation projects.

2.1.2.1 National land economy.

At the national level some federal policies and national economic conditions have worked to encourage deforestation. More specifically, the effects of inflation reduction on land markets have encouraged land speculation in the Amazon and investments in cattle. (Perz, 2002)

With the agricultural frontier expansion of the 1970's, the demand for land in the Amazon rose; land values increased and landowners had an incentive to clear land in order to prove land occupation (Perz, 2002). During the 1980's, with the Brazilian economic crisis, the government froze bank assets to prevent capital flight. Such action made money unavailable for land acquisition and for investments and then some reduction in deforestation was observed around 1990 (INPE, 2000). However, in 1994 the Real Plan, a fiscal stabilization plan that created the "Real" a new currency with equal value to the U.S. dollar, reduced inflation and stimulated investment. Thus, people had money to invest in the Amazon, which may explain the spike in deforestation in the region during the period 1994-1995 (Fearnside, 2000) From 1995 to 1997 deforestation declined as land prices declined, which suggests again that speculation had been a significant driver of deforestation (Fearnside, 1999)

Deforestation can be associated with macroeconomic variables and provide some insight about the agents that are driving deforestation. During the economic recession, from 1987 through 1991, there was a decline in deforestation rates. Ranchers did not have money to invest in expanding their clearings as quickly as they did in the past and the government lacked funds to build more highways and establish settlement projects. (Fearnside, 1999) On the other hand a major peak in deforestation was associated with the Plano Real and the economic recovery that made suddenly large amounts of money available for investment in cattle ranches. Some researchers have suggested that this association of macroeconomic variables, such as money availability and inflation rates, and deforestation indicates that deforestation is more associated with investments in medium and large cattle ranches than with small farmers using family labor (Fearnside, 1999).

2.1.2.2 Brazilian government's development projects and policies

During the 1960's and under a military government, large-scale colonization projects started in the Amazon frontier (Andersen and Reis, 1997; Kolk, 1998). One concern of the government at that time was national security and the ability to defend its borders. The occupation of that "empty" space in the Amazon fitted the development model to ensure national security and territorial integrity. (Kolk, 1998)

Government colonization projects in the Brazilian Amazon were driven by several economic, social and political intentions (Foresta, 1991). The economic intention was to use the vast natural resources in the Amazon to provide livelihood to the growing population. The social intention was to relieve pressure in the populated coastal cities. (Machado, 1991) The political intention was to populate and protect the national borders with Brazilian citizens. The social intention was not accomplished, even when a large number of farmers came from the southwest part of Brazil (Pedlowski & Dale, 1992). The economic goal produced limited results because most of the land was used in unsustainable ways that make questionable the long-term livelihood for rural families. (Dale et al, 1994)

After 1964, roads and electric plants were built to open the region accompanied by colonization and land titling projects. (Andersen and Reis, 1997; Kolk, 1998; Pfaff, 1999) Dams were constructed and a free trade zone was created in Manaus. (Pfaff, 1999) Subsidized credit was offered and income taxes were forgiven if the funds went to approved development projects (Andersen and Reis, 1997; Pfaff, 1999). Furthermore, regional development plans attempted to attract foreign and national investors and the exploitation of the rainforest for its resources and land was encouraged. (Kolk, 1998) Moreover, private Brazilian corporations also encouraged the government to favor large-scale land development, particularly cattle ranching (Browder 1988, Mahar, 1989, Hecht, 1989, Hecht and Cockburn 1989). The actions taken appear to have stimulated occupation of the Amazon. The total population more than doubled between 1970 and 1991 and cleared forest increased significantly. (Pfaff, 1999)

In the 1970's attempts to economically integrate the Amazon continued. Government's efforts focused on the construction of highways, offering of fiscal incentives and the support of colonization projects. (Browder, 1988) Such policies were key factors in facilitating in-migration land settlement and deforestation in the Amazon (Perz, 2002). Many felt that such empty land was an ideal "safety valve" for pressure arising from a growing population. (Pfaff, 1999) Rapid agricultural modernization in the south of Brazil in the 1970s had left many farm workers and peasants without jobs or land (Barraclough and Ghimire, 2000; Skole et al, 1994) The government polices focused on settling as many as possible of these and other landless people in the "empty" forested areas of the Amazon through incentives for small-scale agriculture and cattle-ranching (Barraclough and Ghimire, 2000). With the economic crisis of the 1980's, more and more people bought land in the frontier for speculative purposes and then farmers had incentives to sell their land to the new comers and move on to settle and deforest new frontiers (Almeida and Camapri, 1995; Andersen and Reis, 1997).

2.1.2.3 Brazilian government's conservation projects

The Brazilian government has also developed policies and created regulations to prevent deforestation, but such efforts have been ineffective to stop forest conversion. In 1988, after severe international criticism over major deforestation events during the previous year, a new environmental policy, Our Nature (*Nossa Natureza*) was announced. Such policy also instituted the creation of a National Environmental Institute (IBAMA) in 1989. (Perz, 2002) Furthermore, Brazil modified its 1965 Forestry Code to increase the legal requirement for the amount of private land to be left forested from 50% to 80%. (Hall, 1997: chapter 2) However, a main problem is the lack of law enforcement designated to protect forests (Schomberg, 1998b).

A chronic problem is that some governmental departments create policies to protect the forest while other departments encourage colonization and economic development in the Amazon. For example, agencies like INCRA are in charge of promoting and planning colonization and development activities in the Amazon, without considering other legislation and programs to protect forests (Laurance, 2000). Another example is the federal program *Brasil Avançã* that will improve and add roads to increase access to forest. This along with a "new waterway transport systems (hidrovias) and the extension of paving of roads to the Caribbean (from Amazonas North to Venezuela) and the Pacific (from Acre west to Peru) will make the Amazon forests more accessible, more valuable and potentially more prone to clearing." (Lovejoy, 2000, p 56)

2.1.3 Local and regional-level driving forces of land-cover change.

This section will describe the following relevant driving forces of land-cover change and deforestation at the local and regional level: population growth, property rights, logging, mining and cattle ranching.

2.1.3.1 Population growth and composition.

The Amazon's rapid population growth has been associated with deforestation since the large scale migration of a population with high rates of fertility into the Amazon led to expanding populations, new settlements and more clearing of land (Perz, 2002). However, overall population growth no longer accounts alone for land cover conversions in rural areas of the Amazon.

In the 1970's, the population in the Legal Amazon grew at almost 4 percent per year. (Perz, 2001). This population growth was suggested as one factor that increased deforestation among farms, since family labor was readily available. During the 1980's, population growth in the Legal Amazon lowered to 3 percent per year, In the 1990's, Brazilian Amazon's population growth was relatively low, with just 2 percent per year. In fact between 1991 and 1996, the Amazon's rural population declined by 1 percent per year, whereas urban population grew at almost 4 percent annually. (Perz, 2001) Despite this drastic demographic transition, deforestation in the Amazon continued, suggesting that population growth alone was not anymore a key driver of deforestation during this stage of the frontier development in the Amazon (Perz, 2002).

Population growth's effect on deforestation has changed not only through time, but also across new frontiers in both rural and urban areas. Furthermore, deforestation practices may differ among old and recent immigrants. It seems there is a trend where first migrants have greater impact on deforestation rates than later migrants (Pfaff, 1999).

2.1.3.2 Property rights

The Amazon frontier is notorious for rural violence in conflicts over property rights. However, the role of land ownership on deforestation practices is still unclear, since research has shown ambivalent effects across the Amazon (Rudel and Horowitz, 1993; Schmink and Wood, 1987, 1992). Both, lack and presence of clear property rights have been said to encourage deforestation rates so the real effect is still unclear. It has been argued that without clear property rights, colonists and ranchers are encouraged to clear land in order to prove land occupation (Schmink and Wood, 1992). However, during the 1990's deforestation continued to grow despite the increase in formal property rights, most likely because tenure security encouraged investments in cattle pasture (Perz, 2002)

2.1.3.3 Logging

Expansion of timber extraction in the 1990s also contributed to deforestation in the Amazon despite the slowdown in population (Perz, 2002). During the 1980's the demand for Brazilian high quality hardwoods increased and extraction did as well. The impact was so vast that the area logged between 1996 and 1997 was between 9,730 and 15,090 km^2 (Nepstad et al, 1999), an area comparable in size to the area completely deforested the same year. (INPE, 2000)

Industrialized logging, logging in farmers' plots and illegal logging cause deforestation both, directly and indirectly. Selective logging damages forests directly because of the removal of trees and the formation of small patches of forest which are less viable than large areas of forest. On other hand, logging causes deforestation indirectly because the roads and trails to access the forest also make easier farmers migration and colonization. Moreover, logging increases the risk of fires due to the leaf and tree litter on the ground and the disruption of forest canopy that allow sun radiation to dry forest floor. (Uhl and Kauffman, 1990; Holdsworth and Uhl, 1997)

Industrialized logging is still a large driver of deforestation and it is increasing dramatically in central Amazonia. (Laurance, 2000) During the 1990's, Asian multinational companies bought large forest tracks of timber concessions (long-term forest lease) and purchased interests in Brazilian timber firms. Just in 1996, Asian companies invested more than $500 million in the Brazilian timber industry. Multinational corporations were estimated to own or control about 4.5 million ha of the Brazilian Amazon (Laurance, 2000). On other hand, landowners often allow sawmills to extract timber in order to pay for forest cleaning or pasture remediation (Perz, 2002). This is an option for farmers with reduced family labor to be able to clear large portions of land.

Although plantations are considered one option to fight global warming and deforestation of tropical forest at the same time, homogeneous single-species plantations typically are ecologically unsustainable in the Amazon. Furthermore, plantations are benefiting mostly large companies. For example the Brazilian FLORAM project of agro-forestry envisioned small plots where local population would have enough space for food production in the areas between the agro-forestry blocks. However, this was not the case and there are now companies with over 200,000 ha of continuous plantations (Fearnside, 1999).

Another issue is that most Amazonian timber operations are not being effectively managed. Most legal operations by the nearly 1000 Brazilian timber companies are virtually unregulated due to lack of inspection government resources (Laurance, 2000). No enforcement ensures that these companies are using accepted practices to limit harvest damage.

Finally, illegal logging is a common practice in the Amazonian. In 1997, the Brazilian government estimated that 80% of the Amazonian logging was illegal (Abramovitz, 1998, Zhouri, 2004). Repeated extraction, legaly and illegally is becoming more common, but the illegal one is more likely to use high impact logging techniques or inadequate techniques and to remove nearly all remaining individuals of marketable species including those deemed previously to be too small. (Gerwing, J.J., 2002)

2.1.3.4 Cattle ranching

Among the regional economy activities, cattle ranching is one of the key growing drivers of deforestation (Perz, 2002). From 1985 to 1996 the Amazon's overall herd nearly doubled, from 18.7 to 35.5 million (IBGE, 1990, 1998). Consequently, from 1985 to 1996 a transition of land under annual and perennial crops towards pasture was observed. In fact pasture rose from 42.3 to 51.0 million ha (IBGE, 1990, 1998). Concerns about the expansion of cattle ranching arise from the fact that approximately 50 % of the Amazon's pastures are degraded (Serrao and Homma, 1993), which calls for more clearings.

In the Brazilian Amazon cattle ranching is expanding so rapidly that concerned researchers have adopted the term "*pecuarizaçaõ*" (bovinization) "to depict the extraordinary increase in the cattle herd among landowners of all sizes" (Mertens, 2002). In the past, cattle ranching was only associated with large owners during the initial stages of the frontier development, but now even small farmers have incentives to clear and they do to establish pasture for cattle often to the exclusion of other land use options (Veiga et al, 2001). Credit and fiscal policies for livestock and crops seem to have stimulated investment in cattle ranching and deforestation (Andersen and Reis, 1997). Andersen and Reis have looked at the relationship between deforestation and development policies for the period 1970-1985, and concluded that subsidized credit was more correlated with economic growth and deforestation, followed by large road building projects (Mertens et al, 2002).

At the regional level, crop diseases have caused farmers to look for other income resources and expand cattle (Perz, 2002). The increasing national and international demand for beef and milk encouraged the Brazilian government to create credit programs for investment in cattle during the 1990's, and even small farmers had incentive to clear land and substitute annual crops by cattle. Furthermore, deforestation for cattle pasture was seen as a cheap way to show that owners were making productive use of their land and thus they could claim land ownership (Fearnside, 1999). Many landowners prefer cattle ranching because it constitutes a high value product, which can be marketed at any time and thus; it is a capital reserve to cover unforeseen costs. (Perz, 2002) Moreover, its low demand of labor is favorable for aging households or for those seeking to free labor for off-farm work. (Walker, Moran, and Anselin, 2000)

Several lines of research support the idea that ranchers (both, medium and large) are the main agents of clearing in terms of area deforested (Fearnside, 1999). Moreover, it was observed that behavior of large landowners reacts more to exogenous variables, such as the interest rates, government subsidies, inflation and price of land. On other hand small landowners react more directly in response to endogenous variables at the household and local levels (Fearnside, 1999).

An interesting phenomena is that cattle ranching has expanded in relative terms faster among small to medium sized landholders in terms of growing rates of deforestation (Perz, 2002). The relative proportion of small farmers versus large landowners is constantly changing as a result of changing economic and demographic pressures. The distribution of clearings in the nine states of the Legal Amazon indicates that most of the clearing took place in states that are dominated by ranchers. (Fearnside, 1999) For example Mato Grosso, a state dominated by ranches of 1000 ha or more, alone accounted for 26% of the 11.1x 10^3 km^2 total deforestation. On other hand, Rondônia is a state that has become notorious for its deforestation by small farmers who arrived on the BR364 highway that was paved in part with World Bank financing in the early 1980s. By contrast with Mato Grosso, Rondônia accounted for only 10% of the 1991 deforestation total. (Fearnside, 1999)

Differences among large and small farmers, with regard to their deforestation patterns and rationale, present policy challenges for the Brazilian government, since measures to reduce the net amount of area deforested may not reduce the rate of deforestation in all states of the Legal Amazon. Thus, specific strategies to target both small and large landowners are needed. According to some studies, small farmers (properties <100 ha in area) accounted for 30% of the deforestation in the Legal Amazon from 1990 to 1991, while medium and large ranchers accounted for the remaining 70% (Fearnside, 1999, p. 182). However, although small farmers account for only 30% of the deforestation activity, the intensity or rate of deforestation within the area they occupy is greater than for the medium and large ranchers that hold 89% of the legal Amazon's private land (Fearnside, 1999). Large ranchers and small producers have often been held out in opposition, as two sets of fundamentally different actors. But Walker, et al (2000) emphasized the expansion of the cattle economy across all sectors; large producers are specialized in cattle production, but small producers are moving in this direction as well.

2.2 THE HOUSEHOLD LEVEL: SMALL FARMERS AS AGENTS OF DEFORESTATION

Small farmers are regarded by many researchers and policy makers as "major deforesters" in the Amazon mainly because of their clearing practices for agricultural and cattle ranching purposes. Small farmers are commonly defined as farmers with properties that are less than or equal to 100 ha (Fearnside, 1999). Research estimates indicate that these farmers are responsible for about 30% of the deforestation in the Brazilian Amazon (Fearnside, 1999). They are important deforestation agents because of their large number, their incentives to clear land to prove land occupation, and their mobility.

Since the 1970's, when aggressive colonization projects started in the Amazon and large-scale forest clearing was observed, land use practices have changed substantially. Migrant families that arrived on the frontier followed land use practices that encouraged land degradation. This situation resulted in a vicious cycle that encouraged deforestation when farmers cleared forest to cultivate annuals, subsequent land degradation and poor crop yields forced farmers to move and clear forest in other areas of the frontier (Moran, 1981). Moreover, the processes of subdivision, clearing and agglomeration of parcels are not only reconfiguring the landscape but also influencing the land use strategies in the Amazon

In the past, it was common for farmers to clear land to prove land occupation and claim land titles. This triggered the widespread deforestation in the 1970s (Pedlowski et al, 1997). Presently, the laws that required farmers to keep at least 50% of the property as forest have changed this figure to 80%. However, most farmers continue clearing more forest in order to guarantee continuous sources of income. Typical farmers will clear small patches of primary forest to establish a mix of annuals and perennial crops and pasture (Pedlowski and Dale, 1992)[10].

Another common practice among farmers is the slashing of existing secondary growth to increase the area in production, which allows the use of more land without having to clear more primary forest. Clearing of secondary forest is common because it is easier to slash and there are not environmental laws to prevent it. Furthermore, farmers prefer to establish pasture in older

[10] Primary forest or Mata Virgen: the original, pristine forested area in the frontier.

areas of their plots and clear new areas to cultivate annual and perennials. Thus, secondary growth areas are not allowed to regenerate. (Pedlowski et al, 1997)

Land use in the Brazilian Amazon follows a typical pattern, although farming methods might vary depending on the land quality and farmer's background (Coy 1987, Millikan 1988). Among farmers, the main trend is to cut forest and burn the slash to provide nutrients to the soil in order to plant annual and perennial crops. Typically, annuals such as rice, corn, beans, or mandioc, and perennials such as coffee, cocoa and rubber are planted during about the first four years. Farmers burn the fields every year to reduce weeds and to fertilize the land. However, land quality degrades quickly, crops become more susceptible to pests and diseases and annual agricultural production eventually declines. As result of the production decay and other clearing forces, farmers continue to clear land and start to plant pasture grasses and raise a small number of cattle (Tucker et al, 1984, 1986; Duncan et al, 1990, Frohn et al, 1990). According to Dale, fully cleared land under the current land use practices does not sustain cattle ranching or any other type of farming beyond 6 to 8 years. After that time, "the farmer either cuts more forest and begins the land degradation process again or abandons the land and moves elsewhere" (Dale et al, 1994). Although these figures may vary across the entire Amazon, they portray a generalized pattern of the farming-clearing-degradation process.

2.3 DEVELOPMENT HISTORY AND DEFORESTATION SCENARIO IN RONDONIA, BRAZIL.

Since Rondônia, Brazil is an agricultural frontier, mainly composed of small farmers, this research focuses on studying the factors influencing small farmers LCC decisions, assuming this can be inferred from past LCC decisions and survey data.

Development policies funded by the World Bank in Rondônia have been linked to the massive immigration of peasant peoples and the rapid disappearance of the rain forest (Brown, 2001). Local and international NGO's have expressed their criticism through a campaign to hold the World Bank responsible for the negative consequences of its projects.

In the late 1970's, the Brazilian military tried to integrate Rondônia and the rest of the Amazon into the national economy through the paving of the highway BR-364 and other development projects. Rondônia was considered by the Brazilian government as one of the development poles of the Amazon and the region received millions of dollars of government investment, including support for agricultural colonization. (Brown, 2001)

After rudimentary infrastructure was established in Rondônia, the promise of vast "inhabited" land and federal money for settlement attracted a large number of migrants to the new frontier (Millikan, 1988; Hecht and Cockburn 1990). However, it was difficult to establish viable small farms in the frontier, mostly because of the lack of services and assistance, the agronomic constraints of tropical soils, the high incidence of malaria and low prices of farm produce (Brown, 2001). Moreover, land speculation made many colonists sell their lots and move further into the forest, invading indigenous land that were poorly protected by the Brazil's National Indian Foundation, FUNAI (Cultural Survival Inc. 1981)

In the early 1980s the Brazilian government, with the support of the World Bank, launched the Northwest Brazil integrated Development Program (POLONOROESTE) with the goal to bring order and economic development to Rondônia and western Mato Grosso. The U$1.6 billion plan was focused in paving the BR-364 highway, which accounted for 50% of the entire budget (Brown, 2001). Despite the hope in this project the situation of Rondônia only worsened. Population exploded due to massive immigration through the newly paved highway, world market prices of the crops fell and the amount of land used for cattle ranching and annual cropping continue to expand (Brown, 2001). The **Brazilian** government's National Indian Foundation (**FUNAI**) did not complete the demarcation and protection of vast areas of indigenous lands before intrusions by colonists, loggers and miners occurred (Coy, 1986; FAO-CP 1987; BRA/87/037 1989). POLONOROESTE failed to stabilize the situation in Rondônia and deforestation just accelerated. Until 1980, even after more than 10 years of development, only 3.12% of the state's total area had been deforested. However, in the next five years, deforestation grew to 11.37% (Fearnside, 1989). Unfortunately, by 1997 22.8% of Rondônia's forests had been destroyed (Millikan, 1998; Fig 2)

The World Bank responded to the POLONOROESTE experience with a new loan for Rondônia, incorporating many of the NGO's concerns expressed after the failure of POLONOROESTE. The Rondônia Natural Resources Management Project (PLANAFORO) was a US$228.9 million project, with US$167 million from the World Bank (Brown, 2001).

2.3.1 Agents of deforestation in Rondônia.

In Rondônia the main drivers of deforestation are loggers, miners, small farmers and cattle ranchers, with the latter two being the most important (Pedlowski et al, 1997). Logging has been well studied as one cause of deforestation in the Amazon (Uhl and Buschbacher, 1985; Browder, 1985, 1986; Nepstand et al, 1992; Verissimo et al, 1992). Gold-mining has become the most common non-agricultural activity among farmers that failed in their agricultural enterprises. Mining causes deforestation because it requires the removal of land cover and the use and release of toxic mercury. (Pedlowski, 1997)

Small farmers' land-use strategies tend to clear forest. During the first colonization projects in the 1970's, clearing of land was considered a way to prove land occupation, which encouraged the spread of deforestation. Generally during the first years farmers clear forest to cultivate annual and perennial crops, and establish pasture (Pedlowski and Dale, 1992). Farmers generally choose land-use strategies on empirical bases and trial-and-error approaches, with limited technical assistance (Browder, 1996). Farmers clear both, primary and secondary forest in order to increase the area in production, regardless the laws to preserve 50% of primary forest in the property. As discussed in earlier sections there are not regulations to prevent the clearing of secondary forest. Since the quality of the soils decline quickly, farmers continue to clear land and replace annual and perennial crops by pasture. "Simulation models predict that the current approaches to cropping will lead to almost complete deforestation in 20 years." (Dale et al, 1994)

Cattle ranching is expanding among small and large owners alike. In Rondônia, the cycle of cattle ranching usually starts with small farmers clearing few hectares of land to cultivate annuals and perennials (Pedlowski et al, 1997). However, in most cases, these small farmers are obligated to move since financial returns are not high enough. The land is then bought by another small farmer or by a more capitalized cattle rancher (Coy, 1987; Millikan, 1988). Thus, there is a clear trend towards land aggregation (Pedlowski, 1997). These large farmers are more likely to clear land for cattle ranching because it is the most profitable option. Moreover, cattle ranchers have political influence on top state administrators, and state environmental agencies have low power to enforce the 50% rule of the Brazilian forestry code, where owners can clear only 50% of their land ownership. As a result, most of large properties have much of their area under pasture. The total impact of cattle ranching in deforestation is bigger than that of small farmers because they tend to control larger areas of land. (Pedlowski, 1997)

An interesting phenomenon is that cattle ranching has expanded in relative terms faster among small to medium sized landholders, which can be explained by the changing patterns in agricultural land use. However, although small farmers account for only 30% of the deforestation activity, the intensity or rate of deforestation within the area they occupy is greater than for the medium and large ranchers. (Fearnside, 1999) Walker, et al (2000) emphasize the primacy of the cattle economy across all sectors; large producers are specialized in cattle production, but small producers show an evolution in this direction (Mertens, 2002)

2.3.2 Research challenges and needs

Both, small and large farmers have been linked to deforestation processes, the first mostly because of their agricultural practices and the second ones because of their cattle activity. Large ranchers and small producers have often been held out in opposition, as two sets of fundamentally different actors. But Walker, et al (2000) emphasized the expansion of the cattle economy across all sectors; large and small producers.

Research suggests that the behavior of large landowners reacts more to exogenous variables, such as the national commodity prices, interest rates, government subsidies, inflation and price of land. On other hand small landowners react more directly to endogenous variables at the household and local levels. (Fearnside, 1999) The relative proportion of small farmers versus large landowners is constantly changing as a result of changing economic and demographic pressures and it also reflects the land aggregation and subdividing processes in the Amazon frontier.

CHAPTER 3: LITERATURE REVIEW OF LAND-COVER CHANGE MODELS AND DEFORESTATION

The research field of land-use/land-cover change has grown rapidly during last three decades thanks to the technological improvements in Geographic Information Systems, remote seinsing and aerial photography (Lambin et al, 2000; Baker, 1989; Irwin, 2001). Numerous approaches to model land-use/land-cover change have been attempted. Several authors have classified previous modeling approaches and have also discussed their efficiency, limitations and potentials. Some of the most comprehensive classifications include the ones by Lambin et al, 2000 and Baker, 1989. Lambin classifies models based on the research question, known data and the methods used. Baker's classification focuses on the level of aggregation of the landscape and the use of discrete or continuous mathematics. Elements of both classifications relevant to this dissertation will be discussed in detail in this chapter.

3.1 MODELS TO PREDICT LAND COVER CHANGE CLASSIFIED ACCORDING TO THE RESEARCH QUESTION AND KNOWN DATA.

Most of the research on land-use/land cover change has focused on issues of land cover conversion (Lambin, 1997; Kaimowitz and Angelsen, 1998), i.e. deforestation. Different modeling approaches have been used in land-use/land-cover change research.

Lambin (Lambin, 2000) classified land-cover/ land –use models, based on how models address the following research questions:

1. Which variables best explain land-cover changes? (Why a change occurred or will occur?)
2. Where do land-cover changes occur? (Where?)
3. At what rates do these land-cover changes occur? (When?)

Empirical-statistical models

Empirical-statistical models attempt to identify the variables that cause land-cover changes through multivariate analysis, using in most cases multiple linear regression techniques (Mertens and Lambin, 1997; Andersen, 1996, Ludeke et al, 1990, LaGro and DeGloria, 1992, Godoy, 1997; Aspinal, 2004, Geohegan, 2001, Mertens, 2002). Empirical statistical models are useful to describe past events and to find proximate causes. However, it is important to acknowledge the limitations of the results to the data used and the context of the study site. A limited extrapolation capability is reduced to very similar samples. Most of the empirical-statistical models attempted until now are based on cross-sectional analysis.

Stochastic models

Stochastic models, or transition probability models, describe processes that happen in a sequence of steps, as in the transition from one kind of land cover to another (Thornton and Jones, 1998, Finn, 1985, Jahan, 1986, Aaviksoo, 1995, Weng, 2002, Lopez, 2001). Transition probabilities from one cover to another can be approximated from a sample of transitions occurring during a time interval. These models include only transitions that have been observed in the past.

Optimization models.

Many land-use/ land-cover change models apply optimization techniques, either at the microeconomic or macroeconomic level. The first type of models use linear programming techniques to find the best allocation of resources at the farm level that maximizes profit, assuming land will be used in the way that maximizes its rent (Bockstael, 1996, Chomitz and Gray, 1996, Pfaff, 1999, Nelson and Hellerstein, 1997, Landis, 1995, Landis and Zhang, 1998a,b, Walker, 2004, Walker, shortcoming). The second type of models is composed by general equilibrium models (Kaimowitz and Angelsen, 1998). In this approach, optimization models attempt to find the best possible allocation of resources in order to achieve the maximum of a specified goal under certain constraints or restrictions. The land rent theory of von Thűnen, 1966, and Ricardo's theory are the underlying theories of most of the optimization land-use models (Lambin, 2000). Under these theories it is assumed that landowners will manage their land in the way that earns the highest rent, and will allocate resources accordingly to achieve that goal. Some of the limitations of the optimization models include: "the somewhat arbitrary definition of objective function and non-optimal behavior of people, e.g., due to differences in values, attitudes and cultures" (Lambin, 2000), as well as extrapolation limitations when individual behavior is aggregated at a regional scale.

Dynamic (process-based) simulation models

Dynamic simulation models or process-based models attempt to reproduce the biophysical and socioeconomic processes that cause land-use/ land-cover changes (Dale et al, 1994, Portela and Rademacher, 2001, Evans et al, 2001). They go beyond mimicking the processes; they follow their evolution. They require a prior understanding of the systems and the driving force that cause changes in the system, and then these interactions are reduced into differential equations. These models are very mechanistic and attempt to incorporate all single variables that participate in the system. The understanding of the driving force in the systems comes most of the time from relying on theory that explain the landowner's behavior in the rural or urban settings. One limitation of the dynamic models is the scale issue. Although some models can be parameterized using local observations, the relationships used in process-based models cannot be used straightforwardly to model aggregate behavior. Therefore, most of the dynamic models use randomly generated data and later the models are calibrated with empirical data from secondary sources, i.e. governmental data at aggregated level.

Integrated modeling approaches

Some approaches combine elements of different modeling techniques to predict land-use/ land-cover changes, these models are called integrated models (Berry, et al, 1996, Flamm and Turner, 1994, Hazen and Berry, 1997, Veldkamp and Fresco, 1996, 1997a,b). However, when the level of integration is not high, the models are called "hybrid models". Multiple combinations of model types are possible. For example, Wassenaar et al (1999) applied a dynamic, processed-based crop model at the regional scale using statistical relationships for a rural land-use intensity model. The problem of limited extrapolation power of statistical functions is avoided since new statistical relationships are calculated each time the dynamic program is run. Another example is White et al (1997), who used a land-use model that combines a stochastic, cellular automata approach with dynamic systems models of regional economics[11]. Integrated models can be very complex, such as the integrated Model to Predict European Land Use (Rounsevel et al., 1998)

[11] Cellular automata: A cellular automaton (CA) is a collection of cells arranged in a grid, such that each cell changes state as a function of time according to a defined set of rules that includes the states of neighboring cells.

that combines dynamic models with optimization techniques. Integrated models often require multidisciplinary and interdisciplinary research teams, due to the large volume of data and the complexity of the models.

3.2 MODELS TO PREDICT LAND COVER CHANGE CLASSIFIED ACCORDING TO THE LANDSCAPE UNIT AND MATHEMATICS USED

Baker's literature review of models of landscape change (Baker, 1989) covers models used in a broad range of disciplines, ranging from geography to ecology and economics. This classification is based on the level of aggregation (level of detail with which the landscape change is modeled) and the use of continuous or discrete mathematics. Baker's classification of models of landscape includes three broad categories: whole landscape models, distributional landscape models, and spatial landscape models.

Whole landscape models work for the landscape as a whole (aggregated level) and provide as output an aggregate number for a determined characteristic. Distributional landscape models provide a distribution of the land area among classes of landscape phenomena, for example land cover types. Spatial landscape models provide numbers for different variables in each cell of a grid and provide maps as well. A combination of these models at different scales is also possible. The choice of the model depends on the research question we want to answer and the unit of analysis and observation of the land cover change.

As mentioned before, deforestation is basically a land cover change from forest to non-forest. If we study this transition from one land cover to another as occurring in steps of discrete time intervals, then differential equation models using discrete time are the ones that best resemble my statement of the research problem. These models will be described in detail in the present section. Distributional models used at the farm level do not provide information on the location of the states in the landscape. But when these models are used at the pixel level and in association with the capabilities of satellite images and Geographic Information Systems they share the advantages of the landscape spatial models. Then, distributional models can be used as part of a landscape spatial model or a hybrid model. Deforestation can be regarded at its simplest level as a land cover change from a forest cover to a non-forest cover. This dissertation will follow this approach to test the main assumptions that LCC change models have been using. Moreover, the unit of observation and classification of LCC available for the study site make possible only the analysis at the level of land cover, compared with a more detailed research at the land use level.

Differential equation models in discrete time.

All difference equation, distributional models using discrete state spaces can be expressed in their simplest form, in matrix notation as:

$$n_{t+1} = Pn_t \qquad \text{(Eq 1)}$$

Where n_t is a column vector, $n = (n_1 \ldots n_m)$ whose elements are the fraction of land area in each of m states at time t (where states are land cover types), and P , called the transitional probability matrix, is an m x m matrix, whose elements p_{ij} have the transition probabilities among states during the time interval from t to t+1 (in the case of land cover change will be the transition probabilities from one cover to another cover)[12]. When the unit of analysis is the pixel, the input into the model is the initial classification of the pixel in one of the m states (land covers), and the output is the probability of that pixel being in any of the land covers at time t+1 obtained using the transition matrix P. The matrix used in this dissertation will be explained in detail in the methodology chapter, with the 9 feasible land-cover change transitions that result from the 3 LC classes.

Markov chain models

Markov chain models are one type of differential equation models that describe a system through the use of stages and states, the latter of which can be continuous or discrete[13]. For example, in the case of deforestation, the states are the land cover classification that can be assigned to a pixel and the stages may be years, months, days, etc , defining the time interval in which a measure is taken to determine the current state of the system.

Markov chain models are stochastic because the output—distribution of land among states—is obtained using a probabilistic transition matrix between one land cover (state i) and other land cover (state j). Transition probabilities can be approximated from a sample of transitions occurring during a time interval. Two main assumptions are commonly used in Markov chain theory, which can be relaxed in several ways as it will be explained later in this chapter.

[12] Transition probabilities of land cover change represent the probability of each pixel changing from one land cover to another land cover.
[13] Stage: the length of the time interval at which we record the current state of the system, i.e. satellite images are collected every year.
State: the possible conditions of the system, i.e. the land cover types in which we can classify every pixel

Assumption 1.

The definition of Markov models based on the first assumption, also called the Markov property (first order dependency or first order behavior) is that the Markov chain is a first-order process. The future (state) of a process depends just on the present (state) and not on the previous states of the process. Thus, the conditional probability of the distribution among land covers in time t+1 depends only on the current distribution at time t and on the transition probabilities. In terms of conditional probabilities this is usually expressed as:

$$Pr[X_n = i_n \mid X_1=i_1; \ldots ; X_{n-1} = i_{n-1}] = Pr[X_n = i_n \mid X_{n-1} = i_{n-1}] \qquad \text{(Eq 2)}$$

The first order characteristic of a Markov chain is rarely proved in published studies, since it is assumed an intrinsic feature of a Markov chain. Moreover, to test this assumption, vast empirical panel data is needed. However, even when the Markov property does not correspond to the reality of the physical system, a Markov chain model can be used to simulate "what if" scenarios and provide useful information for decision makers.

Assumption 2.

The second assumption, called the homogeneous or stationary property states that transition probabilities are stationary or constant through time. If the assumption of stationary probability transitions is not satisfied, then Markov models can only provide answers to "what if" questions. Testing of this assumption becomes difficult due to the general unavailability of the detailed panel data necessary to estimate the transition probabilities per year, per subject, per transitional change. This assumption can be relaxed if a switch between transition probability matrices is included, i.e. a matrix of transition probabilities can be used for farmers that just arrived to the frontier and have stable size properties and later when the household dynamics and land use strategies change, another matrix of transition probabilities can be used that better fits the current expanding or subdividing behavior of the farmer's landholdings.

The strict use of the two Markov chain assumptions has created misconceptions about a limited applicability of the Markov chains. In particular, their capability to accommodate higher-order effects, the influence of exogenous or endogenous variables, spatial effects and heterogeneity has been questioned. These misconceptions have limited the application of Markov chains to social research.

Markov chains originated in the mathematics and engineering fields and applications have been amply studied there (Kemeny and Snell 1960; Feller 1968; Bhat 1984). Markov chains have been also applied in other fields for various purposes (Collins et al 1974; Collins 1975; Hulst 1979; Pickles, 1980), in particular in ecological studies, for example modeling biological successions in a plant-plant replacement process (e. g. Anderson, 1966; Horn 1975), changes in diameter distributions of forest trees (e. g. Roberts and Hruska 1986) and migration of people (e.g., Brown 1970).

However, the use of Markov chains in land cover change modeling has been based mostly on transition probabilities determined by landscape variables. Further research is needed to better portray the social and economic aspects of land cover change. This is particularly important for land cover conversions such as deforestation, which is mostly a human-driven process. Linking landscape changes with socioeconomic and demographic variables will help to improve not only the prediction of future land changes but also the prevention of such changes deemed socially undesirable.

Through the extensive use of Markov chains in many fields, methods to overcome many of the original limitations have been developed. Some cases of relaxed assumptions of the Markov chain models include: higher order effects, exogenous and endogenous variables, spatial dependence, and heterogeneity.

Higher order effects.

A first-order process was defined in Assumption 1 or Markov property. In a Markov chain with higher order effects the future state (land cover at time t+1) of a process depends not only on the present state (land cover at time t) and transition probabilities, but also on past states (land covers at time t-1, t-2, t-3, etc). It is possible to model higher order effects, changing the definition of the state to include present and previous land covers. For example in a second-order Markovian chain the definition of the state would include the land cover in the present and previous time period. Additional data from at least two time intervals following the initial observation is needed in that case.

Exogenous and/or endogenous variables.

In order to include the influence of exogenous (e.g. macroeconomic variables) or endogenous variables (e.g. household level variables), such variables can be included in the matrix of transitional probabilities. Thus, the transition probability (i.e. probability of changing from one land cover to other land cover) becomes a linear or non-linear function of the endogenous and exogenous variables, where the function can be theoretical or empirical. In a linear function, each element of the transition probability matrix can be redefined as:

$$p_{ij} = b_1x_1 + b_2x_2 + \ldots + b_sx_s + u_i \qquad \text{(Eq 3)}$$

In the new equation p_{ij} is an element of the matrix P (the matrix of transitional probabilities among land covers), $b_1 \ldots b_s$ are the parameters that relate p_{ij} to the endogenous or exogenous variables $x_1 \ldots x_s$, u_i 's are the portion of the transition probabilities not explained by the x's variables. The variables $x_1 \ldots x_s$ can be exogenous variables (e.g. macroeconomic variables) (Ginsberg 1972) or endogenous variables (e.g. household variables) (Conlisk 1976). Another way to model non-stationary transition probabilities is to switch between different stationary transition matrices (Haray et al. 1970; Horn 1975; Rejmanek et al, 1987), as it was explained under Assumption 2.

Spatial dependence

Spatial dependence occurs when the land cover of a specific pixel is affected by LCC in surrounding pixels[14]. In order to account for spatial dependence in the transitions, models have been developed using Markov chains in each cell of a spatial landscape model. Then, transitions are modeled as endogenous functions, using equation 2, with a specific function for a pre-defined or varying "window" around each grid cell or pixel (Turner, 1987).

[14] Spatial dependence occurs when the probability of LCC in a specific portion of land is strongly affected by the current LC or LCC in the surrounding parcels of land.

Heterogeneity.

Landscapes may be so heterogeneous that we cannot estimate and use a unique transition probability for a group of pixels, farms or municipios. Then, extrapolation of the results of the predictive models may result in inaccurate predictions, reflecting the fact that the underlying assumption of poolability of subjects may not hold for the study site. If we could find a pattern for the transition probability matrix of certain group of subjects (e.g. expanding owner type) or a single probability number that could be used for all those subjects, then extrapolation of results for regional planning would be a real possibility. It would rely only on discerning which characteristics make this group or cluster homogeneous. Then we could use the same probability or criteria to calculate that probability to any other farmer in the Amazon that we could fit in this cluster of farmers.

Several solutions to the heterogeneity problem have been proposed. First, we can disaggregate the land area into homogeneous subunits and calculate individual transition matrices for each subunit (i.e. a lot, plot, pixel, a contiguous set of pixels). A second option is to explicitly model how the land area is distributed over the transition probability values (Ginsberg, 1973). To make the model more explicit we need to find a function that depicts the behavior of the u's (residuals) in equation 2, since they are not homogenous among the units of analysis (pixels). This can be accomplished either by assuming that heterogeneity has a particular distribution, for example a beta distribution (Massy et al, 1970) or finding the true distribution of heterogeneity (Ginsberg, 1973)[15]. Heterogeneity is closely related to spatial dependence and

[15] Beta distribution: The probability distribution of a random variable with density function $f(x) = [x^{\alpha-1}(1-x)^{\beta-1}]/B(\alpha,\beta)$, where B represents the beta function, α and β are positive real numbers, and $0<x<1$. The beta distribution is also known as Pearson Type I distribution.

therefore, careful screening of the data is recommended before specifying advanced statistical or stochastic models.

3.3 RESEARCH CHALLENGES AND NEEDS.

As mentioned in section 2.3.2, much research has been done to address the different deforestation patterns of small and large farmers. This research is dominated by empirical statistical models using linear regression. However, specific research is needed to study the differences among small farmers under the property expansion and fragmentation processes of the evolving Frontier. Among current predictive LCC models, it is a common practice to consider all small farmers as homogeneous LCC decision makers based solely on the size of the plot under current study. This reductionist scenario is not likely to be the case given that farmers may not share the same socioeconomic, farming and decision making background prior to arriving on the Frontier. Differentiation among farmers results when some "succeed", prevail and evolve on the Frontier, while others face economic or unsustainable situations that force them to fragment their holdings or to leave. There is a need for research that acknowledges the diversity of backgrounds, land cover strategies, success or failure in farming endeavors, and the general evolution of farmers on the Frontier.

The relative proportion of small farmers versus large landowners is constantly changing, which can be explained due to some small farmers becoming larger landowners through the property subdivision/aggregation processes. Moreover, changes in the internal composition of small farmers' household (economic, social, demographic variables) affect the predictive accuracy of simulation models based on assumptions of homogeneous subjects.

Two common assumptions are broadly used in empirical statistical and stochastic simulation models. The first assumption refers to the above mentioned homogeneity of subjects which may affect not only the accuracy of the predictions, but also the policy recommendations drawn from the model's results. Another common assumption is to consider that the transition probabilities of land cover change are stationary through time. This assumption may lead to the search for a magical probability number that can be used at local or even regional level to make predictions over a large period of time. Such attempts may result in disappointing predictions that are not due to the model's main specifications, but instead due to omitted temporal patterns in the LCC probabilities. Both assumptions are commonly accepted without proof due to the vast amount of time series and panel data needed to prove the assumptions under the particular context of each research project. However, formal testing will not only provide valuable contributions to the body of methodology, but also to the accuracy of the model's predictions.

CHAPTER 4: STUDY SITE AND RESEARCH METHODOLOGY

4.1 STUDY SITE: RONDONIA, BRAZIL.

Rondônia, Brazil is one of the Amazon states experiencing large-scale social and biophysical change related to rapid in-migration and deforestation. Since Rondônia, Brazil is an agricultural frontier, mainly composed of small farmers, this research focuses on studying the effects of small farmers' LCC decisions, assuming these can be inferred from past LCC decisions and survey data. This dissertation research studies three municipios in the state of Rondônia , Brazil: Alto Paraiso, Nova Uniao and Rolim de Moura, which are shown in Figure 1.

Colonization projects have played an important role in shaping the frontier's landscape in Rondônia. Most of the land allocated to the colonists was mature forest prior to the settlement but it turned into a mosaic of pasture, croplands and different stages of forest re-growth associated with fallowed agricultural land (Evans, 2001).

This dissertation is part of a larger research project, under the leadership of P.I. Dr. John O. Browder who has studied deforestation, agricultural and land use practices in Rondônia, Brazil since 1984. More detailed description of available longitudinal and cross sectional data will be provided in section 4.3.

Figure 1. Map of the study sites

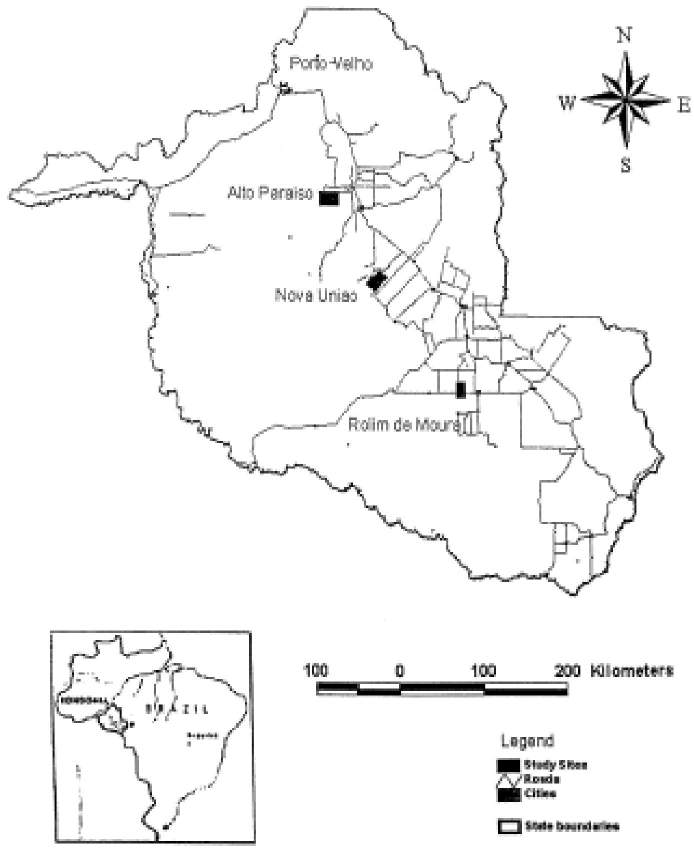


*Figure used with permission of Dr. John O. Browder

4.2 OVERVIEW OF THE RESEARCH STRATEGY

A research approach was designed to address the three research objectives described in chapter 1. A detailed description of the research objectives, research questions, hypotheses, and steps of the research approach was also provided in chapter1. An overview of the dissertation's research strategy is summarized in Table 4.1 presents an overall view of the research strategy.

Table 4.1. Overview of the research strategy

PRELIMINARY STEPS	Testing assumptions	2 modeling approaches	Comparing accuracy
- Estimation of empirical transition probabilities - Descriptive statistical analysis, (analysis of difference of means) of the probabilities of LCC using different classification typologies criteria to group the farmers. - Graphs showing probability trends across time. - descriptive statistical analysis (analysis of difference of means) of relevant LCC variables.	A- Testing homogeneity of subjects: - fixed effects group model - poolability test by group - relevance of the OT and PT typologies will be tested in the multinomial logit model B- Testing stationary probabilities - fixed effects time model - poolability test by time	A. Panel data analysis as the "only-LCC variables" approach - LSDV1-farmer - LSDV1-year - OLS pooled - Predicted probabilities are used in the Markov matrix multiplication B. Multinomial logit model as the "survey-data variables" approach	A. Assign future LC using predicted probabilities from the Markov model - compare against CART pixel classification - find % cells correct B. Compare predicted LC for each pixel against CART pixel classification - find % cells correct

4.3 DATA

This study uses the following data: household surveys, maps, satellite images and their land cover classification at the pixel level, and pathways of past land cover change for each farm. This data is available for a panel sample of farms in three municipios in Rondônia, Brazil (Alto Paraiso, Nova União, and Rolim de Moura) and covers a ten-year period of study (1992-2002).

The paneled dataset is composed of 192 farms in 1992 and 228 farms in 2002 in an attempt to capture a longitudinal view of land cover change. This longitudinal data reflects the land aggregation and subdivision processes, new land ownership and migration that are experienced in the study sites. Survey data contains detailed household level information on demographic, socioeconomic and land-use/ land-cover variables, as well as maps that locate main land cover types within the plot. Survey data were collected during the months of June-July in 1992 and 2002 by a Brazilian research team under the supervision of Dr. Marcos Pedlowski, from the *Universidade Estadual de Norte Fluminense*, and PhD student Percy Summers, from Virginia Tech.

Survey data have been integrated with Land satellite Thematic Map (Landsat TM) imagery of each study site for the entire ten-year period. Satellite images were obtained during the months of June and July during each year between 1992 and 2002. Land-cover pixel classification of the satellite images was performed by M.S. Katherine A. Budreski in her Master's thesis, under the supervision of Dr. Randolph Wynne, both researchers at Virginia Tech. This collaborative effort with the Department of Forestry at Virginia Tech was part of the NSF founded project mentioned above. Pixels were classified into three land-cover classes: forest (Primary Forest, PF), re-growth (Secondary Forest, SF) and cleared land (Non-forest, NF). Land cover classification was performed for Nova Uniao and Rolim de Moura since images for Alto Paraiso had excessive cloud cover and reduced visibility.

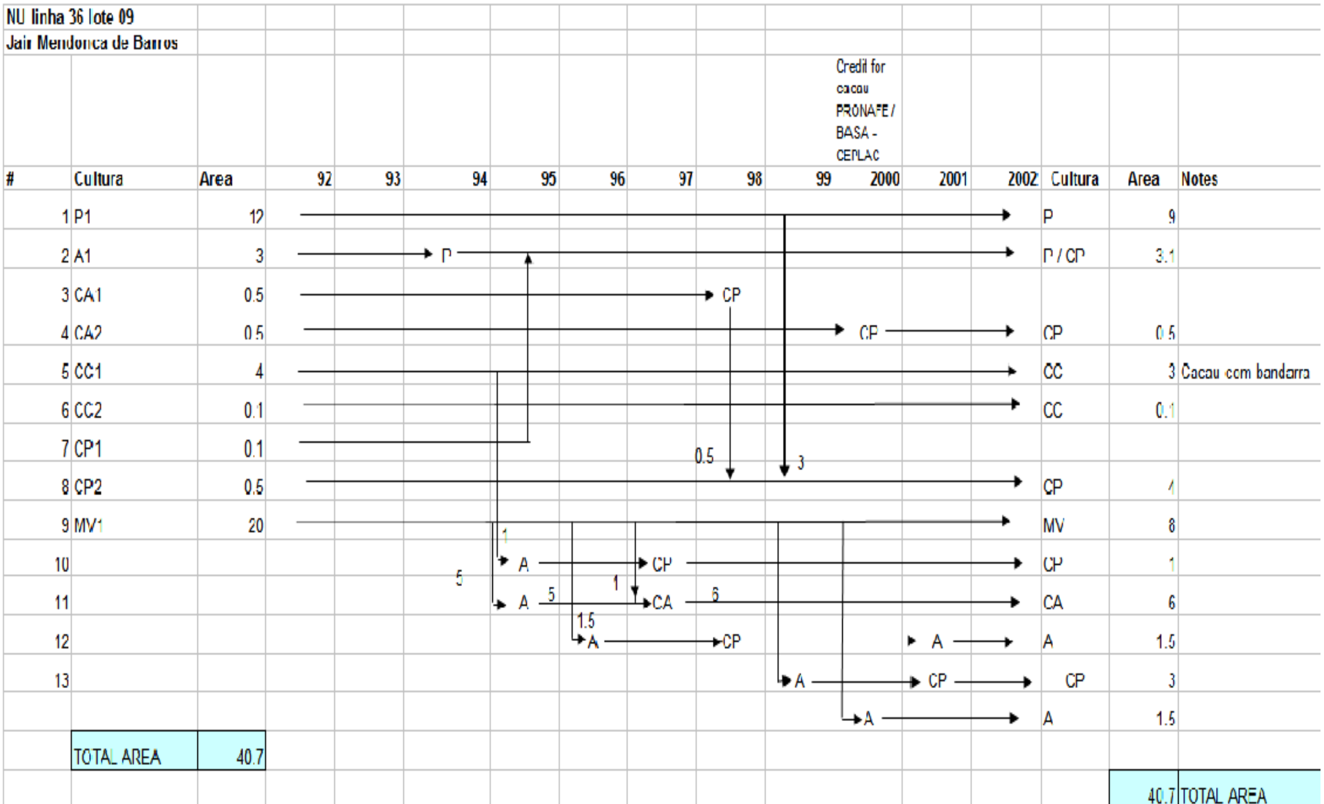
Pathways of past land cover change are graphic representations in the form of flow charts that depict LCC on each farm during the ten-year period of study. Pathways were constructed using satellite images, survey data and maps, and a set of interviews performed in a sub-sample of 70 farms. Interviews and detailed data for each of the 70 farms were collected in 2003 by a research team supervised by Percy Summers. . Pathways of land cover change in the form of flow diagrams were constructed by Percy Summers and Nancy Becerra-Cordoba in 2005.

These graphic reconstructions are an approximation of past land cover transitions that follow up each *alqueire* of land from its land cover in 1992 to its final one in 2002. Figure 2 shows an example of a pathway and the land area calculations. Pathways provide not only the percentage of land in each land cover at a given year, but also the previous-year and the following-year land-cover classification for each alqueire of land[16]. The capacity to identify how much area of a given land cover (e.g. primary forest) in a given year (e.g. 1994) shifts to any of the other two possible land covers or remains in the same land cover (land covers: NF, SF, PF) in the following year (e.g. 1995) proved to be very useful in estimating the empirical probabilities of LCC. Calculations for the land area and LCC probabilities will be explained in section 4.4.1. The pathways were originally elaborated using five main land covers: Primary forest (*Mata Virgen*, MV), Secondary Forest (*Capoeira*, CP), Perennials (PE), Annuals (A), and Pasture (PA). Several more specific land covers were identified, e.g. Agroforestry (AF), Coffee (Café, CA), Pomar (PO, among others. Later these land covers were converted to NF, SF, PF to make possible the validation of the models in this dissertation using the land cover pixel classification described above.

[16] Alqueire: Brazilian unit of area used in Rondõnia. The corresponding equivalence to English units is 1 alq = 2.4 ha.

Figure 2. Example of a pathway of past land-cover changes.

(Alto Paraiso, Line 40, Farm #9)



	P	12	12	15	16.1	16.1	16.1	16.1	13	12	12	12
Areas per year	PE	5.1	5.1	5.1	4.1	4.1	10.1	9.5	9	9	9	9
	MV	20	20	20	15	13.5	12.5	12.5	9.5	6	6	6
	A	3	3	0	6	6.5	1.5	0	3	6	3	3
	CP	0.35	0.35	0.35	0.25	1	1	3	6	6.5	8.5	8.5

											TOTALS
Number of shifts	0	0	1	3	1	2	2	1	2	2	0	14
Percentage of shifts per year	0.0	0.0	7.1	21.4	7.1	14.3	14.3	7.1	14.3	14.3	0.0	100.00
area shifted per year	0	0	3	7.1	1.5	7	2	6.75	2	4.5	0	33.85
area shifted in yr i / total area shifted in 10 yrs	0.0	0.0	8.9	21.0	4.4	20.7	5.9	19.9	5.9	13.3	0.0	100
% of area shifted per year of total lot area	0.0	0.0	7.4	17.4	3.7	17.2	4.9	16.6	4.9	11.1	0.0	83.2

The land covers included in this pathway are: Pasture (P), Annuals (A), Coffee (CA), Cacao (CC), Capoeira (CP), Mata Virgen (MV). In the section of "Area per year", each row has the total area per year in each of the five main land covers.

4.3.1 Unit of analysis and unit of observation

Since land cover change decisions are made at the individual or farm level (Irwin, 2001), the unit of analysis (unit at which analysis is conducted and conclusions formulated) in this research will be the individual property. The methodology uses survey data, interviews and land-cover maps at the farm level. However, the unit of observation (unit at which the land cover classification is obtained or observed) for predicting land-cover change will be the pixel. Each pixel was classified into one of the three categories: forest (primary forest), re-growth (secondary forest) and cleared (non-forest) using the CART methodology[17]. Later, each pixel was classified in a land cover type and associated with the survey variables of the farm in which the pixel is located. Thus, land cover change could be predicted for each pixel but the results can then be aggregated to make conclusions at the farm level. The multinomial logit regression uses the pixel as unit of analysis and all pixels that belong to the same farm will have the same value of demographic, socioeconomic and land cover variables at the farm level. For the scope of this dissertation, we will start with the assumption that all pixels in certain land cover state on a farm will have the same transition probabilities of changing to another land cover in the next time period and conclusions at the farm level can be drawn from aggregated pixels. This is assumed since the focus of the present research is to study LCC decisions at the farm level according to

[17] CART (Classification and Regression Trees) method is a very modern decision tree algorithm developed by Breiman (Breiman *et al*. 1984) used widely to develop classifiers. The CART methodology is technically known as binary recursive partitioning.

the expanding, subdividing and stable total landholdings of each farmer. Table 4.2 shows the units of analysis and observation used in each step of the methodology.

Table 4.2. Units of analysis and observation in each step of the methodology.

Step of the methodology	Unit of analysis	Unit of observation
Estimation of the empirical transition probabilities	Farmer	Farm, pixel.
Panel data analysis of transition probabilities	Farm	pixel
Multinomial logit regression, Markov chain	Pixel	Farm, pixel

4.3.1.1 Typologies used to classify small farmers

Several typologies were used to classify small farmers data into homogeneous groups: property type (PT), owner type (OT), farmer type (FT) and municipio (MUN). These classifications are explained below.

The property type classification is based on changes in the area of the lot under study (primary lot). A lot classified as property type 1 (PT1), a stable lot, is a lot that had the same area in 1992 and 2002. A lot that has a smaller area in 2002, compared with its baseline area in 1992, is classified as a subdividing lot or property type 2 (PT2). This change usually happens when a part of the lot is sold or conveyed to others. Lots that have larger area in 2002, compared with the baseline area in 1992, are classified as expanding or enlarging and are called property type 3 (PT3).

The owner type classification is based on changes in the farmer's total landholdings, regardless of where the other lots are (contiguous or not to the primary lot, in other rural or urban area of Rondônia, or in another state). In this classification it is the farmer who is classified as stable, enlarging or subdividing. This is the main difference with respect to the property type classification, which classifies the property or lot under study. An owner that has the same total owned area in 1992 and in 2002 is classified as stable or owner type 1 (OT1). Farmers with more landholdings in 2002, compared with the baseline in 1992, are classified as enlarging owners or owner type 3 (OT3). Farmers with less total landholdings in 2002, compared with land owned in 1992, are called subdividing or owner type 2 (OT2).

The farmer type classification is a more detailed version of the owner typology. Now farmers are not only classified based on changes on total landholdings, but also on their relationship with the owner in the baseline year 1992. There are three possible scenarios, when comparing the owner in 2002 against the owner in 1992: they are the same owner, the new owner is related (a relative or family member) or unrelated (not a family member or relative). Combining these three categories with the three landholding categories (stable, expanding, subdividing), the new classification has nine possible classes: same owner with stable landholdings (farmer type 1, FT1), same owner with subdividing properties (farmer type 2, FT2), same owner with enlarging properties (farmer type 3, FT3), new owner related and stable properties (farmer types 4, FT4), new owner related and subdividing properties (farmer type 5, FT5), new owner related and enlarging properties (farmer type 6, FT6), new owner unrelated and stable properties (farmer type 7, FT7), new owner unrelated and subdividing properties (farmer type 8, FT8), and new owner unrelated with enlarging properties (farmer type 9, FT9).

Farmers were divided based on the municipio where their primary property is located. The municipio (MUN) typology included 3 municipios in the state of Rondônia: Alto Paraiso, Nova Uniaõ, and Rolim de Moura.

4.3.2 Variables

The starting premise is that demographic, socioeconomic and land cover survey data at the farm level significantly affect our capacity to predict future land cover. The rationale behind the premise is that small farmers are more sensitive and responsive to changes in endogenous variables compared with exogenous variables.

Variables from the long list of variables in the survey questionnaire were selected based on previous research suggesting their relevance to deforestation, the land –cover change event of main interest for this dissertation. The list of the relevant variables used in the multinomial logit regression model is shown in Table 4.3.

Table 4.3. List of variables for the final multinomial logit regression model

Dependent variable:	LC of the pixel
Independent variables:	
Municipio (1,2,3)	Checking account (0/1)
Owner Type (1,2,3)	Urban properties (0/1)
Multifamily household (0/1)	Other rural properties (0,1)
Land title (0/1)	Paid farm workers (0/1)
Planted native tree species (0/1)	Off-farm work (0/1)
Property type (1,2,3)	Syndicate (0/1)
Credit (0/1)	Cooperative (0/1)
Savings account (0/1)	Extracts forest products (0/1)
Area of the lot in alqueires	Mutual help group (0/1)
Percentage area in annuals	Interest in planting native tree species (0/1)
Percentage area in forest	Interaction of the variables: Total cattle owned now and number of people living in the lot.
Number of people living in the lot	Total cattle owned now
	F1 dependency ratio

4.4 DETAILS OF THE RESEARCH METHODOLOGY

4.4.1 Estimation of the empirical transition probabilities using pathway analysis of past land cover changes.

Section 4.3 discussed the definition of a pathway of past land cover change and how it was produced. This section will describe the pathway analysis to estimate empirical proximate probabilities of land cover change.

Pixels can be classified in one of three possible land-cover classes, based on the satellite images and the pixel classification developed by Katie Budreski. Therefore, the matrix of transitional probabilities is a matrix of 3 x 3, with 9 possible transitions, which are not necessarily changes in the landscape as it will be explained later. Table 4.4 shows the matrix of 3 x3 transitional probabilities and the 9 possible transitions, with the nomenclature that will be used in the rest of this dissertation.

Table 4.4 Matrix of LCC transitions among land cover classes

		Final land cover		
		PF	SF	NF
Initial land cover	PF	**PF →PF**	**PF →SF**	**PF →NF**
	SF	**SF →PF**	**SF →SF**	**SF →NF**
	NF	**NF →PF**	**NF →SF**	**NF →NF**

Note:Since there are three feasible land cover classifications studied in this research, the initial land cover and the final land cover have to fall in one of the three categories PF (Primary Forest), SF (Secondary Forest or re-growth) and NF (non forest). The nine possible combinations or land cover change transitions are highlighted.

Even before doing any calculation we know that SF-PF and NF-PF are events with probability equal to zero since altered forest can never go back to the pristine stage of land (*Mata Virgen*) before the colonists arrived. These two transitions will be called "non-events". PF-PF, SF-SF and NF-NF will be named "persistence events" for the scope of this dissertation. PF-PF represents the probability of the land in primary forest remaining in that state (primary forest) in the next stage (year). NF-NF and SF-SF may or may not involve a change in the landscape, for example, NF-NF may reflect a pixel that started as cacao in year1 and remained as cacao in year 2 (even when the cacao is older, it has a different appearance in the landscape and it produces a different reflection of land in the satellite image). A land cover shift from annuals to pasture will produce a NF-NF transition for the scope of the present research, such transition could be studied in more detail as $NF_1 - NF_2$ if pixel classification is feasible in the future. The last two statements apply to SF-SF as well, even though the re-growth class is fuzzier and it basically includes everything that looks in between forest and non-forest from the satellite image. For example, Agroforestry (AF), a mix of trees and perennial crops, was classified as SF along with re-growth from abandoned pasture. SF imposes classification and computational challenges not only at the satellite-image level, but also at the farmer-survey level. Prof. Randy Wynne's research found that farmers tend to overestimate the amount of forest and they tend to classify re-growths as secondary forest. NF-SF and PF-SF are considered "transitional shifts" and may represent a shift from abandoned pasture to re-growth, or a re-growth after timber extraction (work in progress).

The two land cover changes that can be classified as deforestation are PF-NF and SF-NF, the second one being the more likely as it will be shown later in the results. PF-NF may involve the change from *Mata Virgen* (Primary Forest) to annual crops or cleared land for housing purposes, while SF-NF may represent re-growth that is being slash-and-burnt to be used in annuals again.

From the pathway diagram we can obtain not only the percentage of land in each land cover at a given year, but also the previous-year and the following-year land-cover classification for each unit of land (in *alqueires*, the Brazilian unit of area). Figure 3 shows an example of a pathway diagram and the calculations of areas in each land cover class in each year, in this case for farm 9 in Nova Uniao, line 40. The initial amounts of land in 1992 in each land cover are: Primary Forest (20 alq, from Mata Virgen 1), Secondary Forest (0.6 alq, from the two

Capoeiras) and Non-Forest (20.1 alq, from the areas in pasture, annuals, coffee and cacao). The first LCC event, after the baseline in 1992 occurred in 1995, given that the transition from annuals to pasture in 1994 is a NF-NF transition. From 1994 to 1995, 5 of the 20 alq in MV became NF and 0.1 of the 0.6 alq of SF became NF. After these changes happened we can recalculate total areas in each of the three LC classes, now the new totals are 15 alq of PF, 0.5 alq of SF, and 25.2 alq of NF. These new subtotals take into account all land that ends in each of the three possible land covers, regardless of the original land-cover and the land-cover change transition that led to the final land cover observed. In 1996, 1.5 of the remaining 15 alq in PF became NF and the new subtotals after this LCC happened are: 13.5 alq of PF, 0.5 alq of SF, 26.7 of NF. We proceed in the same fashion, recalculating areas in each year.

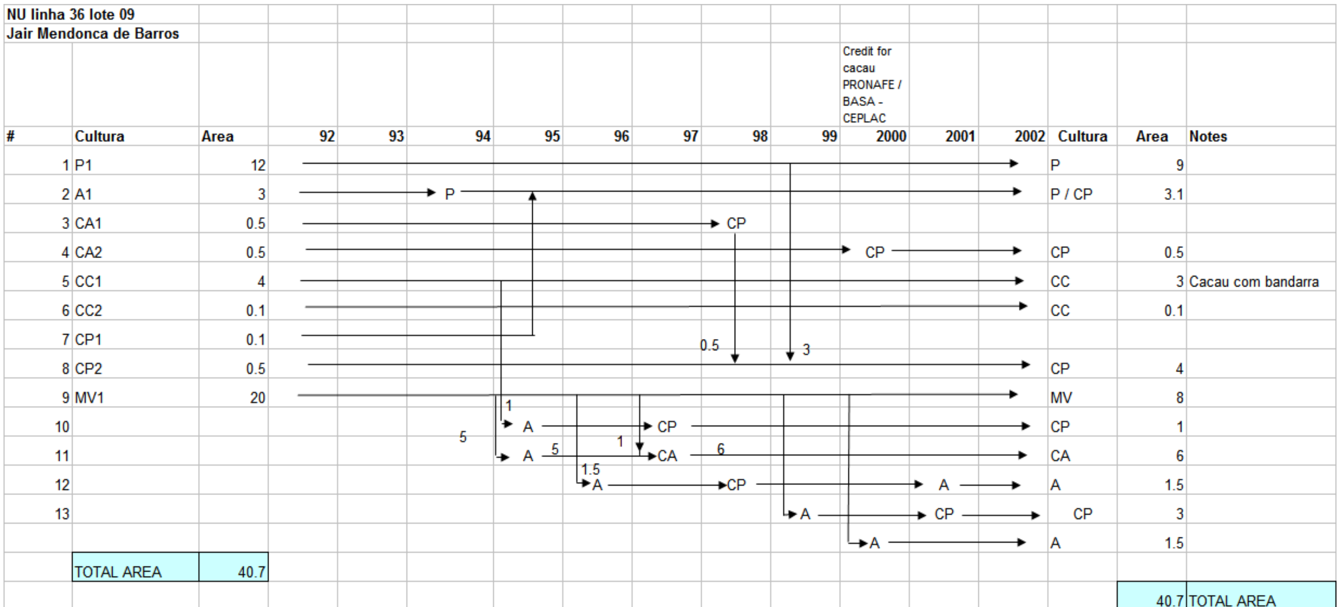
After the distribution among areas has been calculated for each year and for each transition we can proceed to estimate the probability of the LCC that just happened. The probability for the transition PF-NF, for the year 1994-1995 for farmer 9 in Nova Uniao is equal to 0.25 given that 5 of the original 20 alq in PF became NF (5/20 = 0.25). Thus, the probabilities of LCC are not equal to the percentage of total farm area in each of the three possible land covers. Instead the probabilities are equal to the proportion of area -- of the new recalculated subtotal areas after land-cover changes have occurred and new classifications have been assigned -- that shift to any of the three possible LC in the next time period. In other words, land area is allocated among the 9 possible land cover changes, subtotals for each final land cover are recalculated after the transitions and these new subtotals are used as the baselines (100%) for the next year transitions. The rest of the calculations are shown in Figure 4.

Therefore, we can obtain sets of 9 probabilities (one for each LCC) for each of the years (10 years), for each of the farmers (n=70) resulting in a pool of 6300 probabilities. These probabilities are, in fact, panel data since the probability for each of the 9 LCC is measured every year for the same farmer, providing cross- sectional and longitudinal data for each farmer.

The rationale behind this procedure is that although events already happened, we can estimate proximate probabilities if the event is repeated over and over. It is similar to the event of flipping a coin. If we flip a coin over and over, after a large number of trials the proportion of times that we get tails divided by the total number of trials will be a number that in the long run will be approaching the true probability of getting tails. This phenomenon relies on the law of large numbers and it allows estimating the proximate-empirical probabilities given that we cannot try an event the infinite number of times that it will require to obtain the true probability.

Figure 3. Example of pathway analysis for areas in different LCs.

(Alto Paraiso, Line 40, Farm #9)



		92	93	94	95	96	97	98	99	2000	2001	2002
	P	12	12	15	16.1	16.1	16.1	16.1	13	12	12	12
Areas per year	PE	5.1	5.1	5.1	4.1	4.1	10.1	9.5	9	9	9	9
	MV	20	20	20	15	13.5	12.5	12.5	9.5	6	6	6
	A	3	3	0	6	6.5	1.5	0	3	6	3	3
	CP	0.35	0.35	0.35	0.25	1	1	3	6	6.5	8.5	8.5

CALCULATING TRANSITION PROBABILITIES

AREAS

LC	Aggregated Area 1992 inicio	LCC	Disaggregated area 1993	Aggregated Area 1993	Disaggregated area 1994	Aggregated Area 1994	Disaggregated area 1995	Aggregated Area 1995	Disaggregated area 1996	Aggregated Area 1996	Disaggregated area 1997	Aggregated Area 1997
		PF-PF	20.00		20.00		15.00		13.50		12.50	
PF		PF-SF	0.00		0.00		0.00		0.00		0.00	
		PF-NF	0.00		0.00		5.00		1.50		1.00	
subtotal	20			20.00		20.00		15.00		13.50		12.50
		SF-PF	0.00		0.00		0.00		0.00		0.00	
SF		SF-SF	0.60		0.60		0.50		0.50		0.50	
		SF-NF	0.00		0.00		0.10		0.00		0.00	
subtotal	0.6			0.60		0.60		0.50		0.50		1.50
		NF-PF	0.00		0.00		0.00		0.00		0.00	
NF		NF-SF	0.00		0.00		0.00		0.00		1.00	
		NF-NF	20.10		20.10		20.10		25.20		25.70	
subtotal	20.1			20.10		20.10		25.20		26.70		26.70
TOTAL	40.7			40.70		40.70		40.70		40.70		40.70

LC	Aggregated area 1997 inicio	LCC	Disaggregated area 1998	Aggregated Area 1998	Disaggregated area 1999	Aggregated Area 1999	Disaggregated area 2000	Aggregated Area 2000	Disaggregated area 2001	Aggregated Area 2001	Disaggregated area 2002	Aggregated Area 2002
		PF-PF	12.50		9.50		8.00		8.00		8.00	
PF		PF-SF	0.00		0.00		0.00		0.00		0.00	
		PF-NF	0.00		3.00		1.50		0.00		0.00	
subtotal	12.5			12.50		9.50		8.00		8.00		8.00
		SF-PF	0.00		0.00		0.00		0.00		0.00	
SF		SF-SF	1.50		3.50		6.50		5.50		8.50	
		SF-NF	0.00		0.00		0.00		1.50		0.00	
subtotal	1.5			3.50		6.50		7.00		8.50		8.50
		NF-PF	0.00		0.00		0.00		0.00		0.00	
NF		NF-SF	2.00		3.00		0.50		3.00		0.00	
		NF-NF	24.70		21.70		24.20		22.70		24.20	
subtotal	26.7			24.70		24.70		25.70		24.20		24.20
TOTAL	40.7			40.70		40.70		40.70		40.70		40.70

Figure 4. Example of the estimation of empirical LCC probabilities using pathway analysis.

(Alto Paraiso, Line 40, Farm #9)

TOTAL AREA	40.7											
		92	93	94	95	96	97	98	99	2000	2001	2002
	P	12	12	15	16.1	16.1	16.1	16.1	13	12	12	12
Areas per year	PE	5.1	5.1	5.1	4.1	4.1	10.1	9.5	9	9	9	9
	MV	20	20	20	15	13.5	12.5	12.5	9.5	6	6	6
	A	3	3	0	6	6.5	1.5	0	3	6	3	3
	CP	0.35	0.35	0.35	0.25	1	1	3	6	6.5	8.5	8.5

CALCULATING TRANSITION PROBABILITIES

AREAS

LC	Aggregated Area 1992 inicio	LCC	Disaggregated area 1993	Aggregated Area 1993	Disaggregated area 1994	Aggregated Area 1994	Disaggregated area 1995	Aggregated Area 1995	Disaggregated area 1996	Aggregated Area 1996	Disaggregated area 1997	Aggregated Area 1997
		PF-PF	20.00		20.00		15.00		13.50		12.50	
PF		PF-SF	0.00		0.00		0.00		0.00		0.00	
		PF-NF	0.00		0.00		5.00		1.50		1.00	
subtotal	20			20.00		20.00		15.00		13.50		12.50
		SF-PF	0.00		0.00		0.00		0.00		0.00	
SF		SF-SF	0.60		0.60		0.50		0.50		0.50	
		SF-NF	0.00		0.00		0.10		0.00		0.00	
subtotal	0.6			0.60		0.60		0.50		0.50		1.50
		NF-PF	0.00		0.00		0.00		0.00		0.00	
NF		NF-SF	0.00		0.00		0.00		0.00		1.00	
		NF-NF	20.10		20.10		20.10		25.20		25.70	
subtotal	20.1			20.10		20.10		25.20		26.70		26.70
TOTAL	40.7			40.70		40.70		40.70		40.70		40.70

LC	Aggregated area 1997 inicio	LCC	Disaggregated area 1998	Aggregated Area 1998	Disaggregated area 1999	Aggregated Area 1999	Disaggregated area 2000	Aggregated Area 2000	Disaggregated area 2001	Aggregated Area 2001	Disaggregated area 2002	Aggregated Area 2002
		PF-PF	12.50		9.50		8.00		8.00		8.00	
PF		PF-SF	0.00		0.00		0.00		0.00		0.00	
		PF-NF	0.00		3.00		1.50		0.00		0.00	
subtotal	12.5			12.50		9.50		8.00		8.00		8.00
		SF-PF	0.00		0.00		0.00		0.00		0.00	
SF		SF-SF	1.50		3.50		6.50		5.50		8.50	
		SF-NF	0.00		0.00		0.00		1.50		0.00	
subtotal	1.5			3.50		6.50		7.00		8.50		8.50
		NF-PF	0.00		0.00		0.00		0.00		0.00	
NF		NF-SF	2.00		3.00		0.50		3.00		0.00	
		NF-NF	24.70		21.70		24.20		22.70		24.20	
subtotal	26.7			24.70		24.70		25.70		24.20		24.20
TOTAL	40.7			40.70		40.70		40.70		40.70		40.70

PROBABILITIES

LC	Aggregated perc 1992 inicio	LCC	Transition probab 92-93	sum=1	Transition probab 93-94	sum=1	Transition probab 94-95	sum=1	Transition probab 95-96	sum=1	Transition probab 96-97	sum=1
		PF-PF	1.00		1.00		0.75		0.90		0.93	
PF		PF-SF	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00
		PF-NF	0.00		0.00		0.25		0.10		0.07	
		SF-PF	0.00		0.00		0.00		0.00		0.00	
SF		SF-SF	1.00	1.00	1.00	1.00	0.83	1.00	1.00	1.00	1.00	1.00
		SF-NF	0.00		0.00		0.17		0.00		0.00	
		NF-PF	0.00		0.00		0.00		0.00		0.00	
NF		NF-SF	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.04	1.00
		NF-NF	1.00		1.00		1.00		1.00		0.96	

LC	Aggregated perc 1997 inicio	LCC	Transition probab 97-98	sum=1	Transition probab 98-99	sum=1	Transition probab 99-2000	sum=1	Transition probab 2000-2001	sum=1	Transition probab 2001-2002	sum=1
		PF-PF	1.00		0.76		0.84		1.00		1.00	
PF		PF-SF	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00
		PF-NF	0.00		0.24		0.16		0.00		0.00	
		SF-PF	0.00		0.00		0.00		0.00		0.00	
SF		SF-SF	1.00	1.00	1.00	1.00	1.00	1.00	0.79	1.00	1.00	1.00
		SF-NF	0.00		0.00		0.00		0.21		0.00	
		NF-PF	0.00		0.00		0.00		0.00		0.00	
NF		NF-SF	0.07	1.00	0.12	1.00	0.02	1.00	0.12	1.00	0.00	1.00
		NF-NF	0.93		0.88		0.98		0.88		1.00	

In our case, each farmer is a coin and every year we flip the coin to know the probability of the LCC. We have 70 coins (70 farmers) of twenty-five cents each and we flip them at the same time every day during 10 days (June of every year during 10 years). We could assume that all the 25-cent coins will have the same probability of getting tails. However, some may be new, others may be worn out and some may be just defective, etc. At the end, not all the twenty-five cents coins are identical. This is similar to the case of the small farmers studied in this research. All farmers started with a farm of 100 ha of primary forest but they developed different land-cover change practices regardless of being all classified as “small farmers”.

One of the premises of this dissertation research is that small farmers cannot be treated all the same since they are experiencing differently the aggregation and subdividing land processes in the Frontier, depending on their household attributes and changes in total landholdings. Some smallholders are evolving into large landowners, while others are selling land and others keep the same amount of land constant. Moreover, the probabilities for LCC in a given farm may be affected not only by the total landholdings (present anywhere), but also by other variables at the farm and household level (i.e. available family labor, credit, land title). This is one of the three premises in this dissertation, which is stated in research objective 1 and hypothesis 1 and 2.

The probabilities calculated from with the pathways analysis were used in further steps of the methodology. First, analysis of difference of means was conducted using different typologies to cluster the farmers, and graphs of probabilities versus time for each typology were produced as well. Second, panel data analysis was performed to test differences among subjects (using the fixed group effects model) and across years (using the fixed-time-effects models). Third, poolability tests by subject and by time were conducted. Predicted land cover classifications from the panel data analysis were compared against logistic regressions using survey data, with the purpose of comparing accuracy of prediction using the "LCC-only variables" and the "survey-data variables" modeling approaches.

4.4.2 Modified Markov chain model applied to LCC leading to deforestation.

Deforestation is basically a land cover conversion of a forested area to a non-forested one. In this dissertation, we are inferring the LCC decisions based on the predicted land cover change. Relevant explanatory variables for that predicted LCC are then interpreted as proximate causes of deforestation. If we study this transition from one land cover to another as occurring in steps of discrete time intervals, then a Markov chain model can be used to portray the deforestation process. The literature suggests that most tropical deforestation processes seem to be compatible with the Markov property of first-order dependency (explained in section 3.2 under *Assumption 1*) (Lambin, 1997). Even where the first order assumption may not hold always, it provides a good approximation of the real process, since farmers will make decisions about changing the land use-cover of certain piece of land based on the present condition of the land (i.e. fertility, access to water, current land cover, etc) without taking into consideration all sequence of land cover/uses in the past. Markov chains have been used in many other fields and relaxed assumptions have been tested.

The stochastic simulation model will use a Markov chain to predict changes from one land cover to another using a transitional probably matrix. The process is assumed to follow a first-order behavior (Markov property) following the rationale explained above. Probabilities are assumed to be stationary during a short period of time (relative to the time periods used in predictive models). The models used discrete stages and discrete states[18].

Both, probabilities based only on land cover changes and probabilities calculated with different sets of endogenous and/or exogenous variables will be tried separately and results will be compared. When endogenous variables are considered, Markov chain model's basic equations are modified to take into account endogenous (household and pixel-related) variables that can cause deforestation.

Thus, transition probabilities will be initially assumed stationary through time. Later, transition probabilities will be calculated as a function of non-stationary socioeconomic and demographic variables. Such function may be a linear or non-linear function. Panel data analysis and logistic regression are the two functions that will be used in conjunction with the Markov chain.

The Markov chain then will be expressed as:

$$n_{t+1} = P[p_{ij}\ f(t)]\ n_t \qquad \text{(Eq 5)}$$

where p_{ij} is an element of the matrix of transitional probabilities P, and the function for the probabilities of LCC is the fixed effect model used in the panel data analysis.

$$p_{ij} = f\ (\text{panel data of past LCC probabilities for the possible LCCs, except } p_{ij}) \qquad \text{(Eq 6)}$$

[18] Discrete stages: The units of time (stages) are discrete if the system is observed in a countable number of times. Discrete states: When the system can be classified in a countable number of states, in the case of these dissertation a countable number of land covers.

Equation 6 is the general representation of the following 7 equations:

$$P_{PF\text{-}PF} = \beta_0 + P_{PFSF}\,\beta_1 + P_{PFNF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 7)}$$

$$P_{PF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFNF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 8)}$$

$$P_{SF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 9)}$$

$$P_{NF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{SFNF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 10)}$$

$$P_{PF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 11)}$$

$$P_{SF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 12)}$$

$$P_{NF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{SFNF}\,\beta_5 + P_{NFSF}\,\beta_6 \qquad \text{(Eq 13)}$$

These equations use the panel data set of empirical probabilities calculated through the pathway analysis. The rationale behind the equations is that the probability of a given land cover change depends on probabilities of other land-cover changes occurring in the present or previous years, in the specific farm under study and in the other farms as well. Basically these functions depict the interdependence among the LCC probabilities, and the temporal patterns in the probabilities series. Similar analyses have been used to study inflation and interest rates, and financial probabilities in the stock market. No equations for P_{SF-PF} and P_{NF-PF} were calculated since we know ahead that these events have probability zero -- as it was explained in section 4.4.1—and to avoid perfect multicollinearity. For the same reasons, the list of dependent variables is reduced to 6, instead of 8 variables. The fixed effects model to analyze panel data allows for correlation among the independent variables and it is often used merely for that reason.

4.4.2.1 Stages

In the Markov chain model proposed the stages or steps are the discrete time interval size in which we will divide time to measure it in "units of time". In other words, every time we take a measurement or observation of the system to assign a land-cover classification to the pixels. For this research the stages are 1 year apart since LC for each pixel was recorded every year based on the satellite images and also for the probabilities calculated from the pathway analysis.

4.4.2.2 States

States are the possible land cover types in which we can classify every pixel. The classification to be used in this study is: PF (primary forest), SF (re-growth or secondary forest) and NF (cleared land or non forest). Such classification has been proven to be feasible (for secondary forest that is less than 10 years old) from analysis of the satellite images of the study sites using the CART pixel classification[19]. Classification of the pixels into more refined land use categories (e.g. annual, perennial crops, pasture, etc) was not feasible due to the limited capability of Landsat data to reliably differentiate pixels associated with different land use

[19] Pixel classification using the CART methodology was performed by M.S. Katie Budreski (Virginia Tech).

classes (e.g. the difficulty to identify same vegetal cover in different ages and to classify mixed land uses).

4.4.2.3 Predicting land cover change in the short future using the Markov Chain

After the analysis of transition probabilities described above, each of the probability sets were used to find future transition probabilities for the second half of the ten-year period of the study. Under the stationary assumption, transition probabilities for future time periods were calculated multiplying the matrix of one-step transition probabilities by itself in order to validate the model for predictions of LC in year 2002). Satellite images can be manipulated in GIS and a database can be obtained with all pixels associated with a land cover classification and to the farm number. Using this number we can connect the pixels with the household survey data to be used in the multinomial logit regression.

4.4.3 Testing modeling assumptions.

Two assumptions commonly used in models predicting LCC will be tested in this research: homogeneity of subjects and stationary probabilities. Before formally testing these assumptions, screening of data (panel data of LCC probabilities and household survey data) consisting of descriptive statistical analysis was performed. This exploratory analysis provided insight of general trends in the data, described differences among farmers using diverse typologies, and aided the selection of techniques to formally test the hypotheses.

4.4.3.1 Screening of data

Analysis of differences of means and graphs of LCC probabilities vs. time were performed for different classifications: owner type, property type, municipio (AP, NU, RM) and farmer type. Analysis of differences of means was conducted among the 70-farmer average probabilities for each LCC transition on the 10 years of the period of study. Specific details and discretionary decisions will be presented in the discussion of results.

4.4.3.2 Testing the assumption of homogeneity of subjects with respect to trends of transitional probabilities of LCC.

Panel data analysis was conducted to formally test the effects of groups (subjects) and time (years) using the fixed effects technique. The Fixed Group Effects Model (FGEM) approach was used to test the assumption of homogeneous probabilities among subjects, commonly used in stochastic predictive models applied to small farmers' LCC decisions. Probabilities calculated for the first five years of the 10-year study period were used in the FGEM .Empirical transition probabilities of the last five years were used for model validation in the Markov chain step of the methodology. The objective is to use the FGEM as the linear function in the modified Markov chain model described in the literature review in equation 3, and developed in equations 5 to13. A poolability test by group (subject) was performed to decide if the model pooling all subjects (Ordinary Least Square model or OLS model) in a single sample is better than exploring clustering of the subjects in a typology that better describes their specific patterns of LCC probabilities (farmer type, owner type, property type). A Least Square Dummy Variable Regression (LSDV) fixed group effects model was run adding dummy variables for each farmer and excluding one variable of the model. Thus, the model turns into a Without -a -Dummy Least

Square Dummy Variable Regression for the test of fixed effects of subjects (in this case the individuals interviewed) and it will be called LSDV1–farmer in the rest of the dissertation . Specific details and challenges encountered during the analysis of data will be discussed in the results chapter.

The equations for the seven LSDV1 models are presented below in equations 14 to 20.. As explained in section 4.4.2, no equations for the non-events $P_{SF\text{-}PF}$ and $P_{NF\text{-}PF}$ were calculated, and only six independent variables were included to avoid multicollinearity. The code for the linear regressions was developed in SAS software and it is presented in Appendix A.

$$P_{PF\text{-}PF} = \beta_0 + P_{PFSF}\,\beta_1 + P_{PFNF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + \ldots + d_{69} \quad \text{(Eq 14)}$$

$$P_{PF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFNF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + \ldots + d_{69} \quad \text{(Eq 15)}$$

$$P_{SF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + \ldots + d_{69} \quad \text{(Eq 16)}$$

$$P_{NF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{SFNF}\,\beta_5 + P_{NFNF}\,\beta_6 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + \ldots + d_{69} \quad \text{(Eq 17)}$$

$$P_{PF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + \ldots + d_{69} \quad \text{(Eq 18)}$$

$$P_{SF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + \ldots + d_{69} \quad \text{(Eq 19)}$$

$$P_{NF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{SFNF}\,\beta_5 + P_{NFSF}\,\beta_6 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + \ldots + d_{69} \quad \text{(Eq 20)}$$

Where $d_1 \ldots d_{69}$ are the dummy variables added to represent the "fixed subject effect" of each of the 70 farmers for which pathway analysis was conducted. As explained above, in the description of the LSDV1-farmer model, one dummy variable has to be dropped out to avoid the "dummy variable trap"[20]. The dummy variable d_{70} was dropped from all the above equations.

4.4.3.3 Testing the assumption of stationary probabilities of LCC.

The Fixed Group Effects Model (FGEM) approach was used to test the assumption of stationary probabilities through time, in the same way that it was used to test homogeneity of subjects. A poolability test for time (year) was performed to decide if the model pooling all probabilities from all years in a single sample is better than exploring temporal changes of the probabilities across time. A Least Square Dummy Variable Regression (LSDV1) fixed time effects model was run adding dummy variables for each year and excluding one variable of the

[20] The "dummy variable trap" is a common way to refer to the presence of multicollinearity due to the addition of an excessive number of dummy variables to the linear regression model. If a categorical variable can take one of n possible values, then only n-1 dummy variables can be added to the model in order to avoid multicollinearity.

model. This model follows the same design rationale as explained for the LSDV1-farmer model in section 4.4.3.1.

The equations for the seven LSDV1 models are presented below in equations 21 to 27.. As explained in section 4.4.2, no equations for the non-events $P_{SF\text{-}PF}$ and $P_{NF\text{-}PF}$ were calculated, and only six independent variables were included to avoid multicollinearity. The code for the linear regressions was developed in SAS software and it is presented in Appendix A.

$$P_{PF\text{-}PF} = \beta_0 + P_{PFSF}\,\beta_1 + P_{PFNF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + y_2 + y_3 + y_4 + y_5 \qquad \text{(Eq 21)}$$

$$P_{PF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFNF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + y_2 + y_3 + y_4 + y_5 \qquad \text{(Eq 22)}$$

$$P_{SF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + y_2 + y_3 + y_4 + y_5 \qquad \text{(Eq 23)}$$

$$P_{NF\text{-}SF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{SFNF}\,\beta_5 + P_{NFNF}\,\beta_6 + y_2 + y_3 + y_4 + y_5 \qquad \text{(Eq 24)}$$

$$P_{PF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + y_2 + y_3 + y_4 + y_5 \qquad \text{(Eq 25)}$$

$$P_{SF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + y_2 + y_3 + y_4 + y_5 \quad \text{(Eq 26)}$$

$$P_{NF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{PFNF}\,\beta_3 + P_{SFSF}\,\beta_4 + P_{SFNF}\,\beta_5 + P_{NFSF}\,\beta_6 + y_2 + y_3 + y_4 + y_5 \quad \text{(Eq 27)}$$

Where $y_2 \ldots y_5$ are the dummy variables added to represent the "fixed time effect" of the probabilities changing over the first five years of the period of study (1992-1993, 1993-1994, 1994-1995, 1995-1996, 1996-1997). As explained above, in the description of the LSDV1-farmer model, one dummy variable has to be dropped to avoid the "dummy variable trap". The dummy variable y_1 , representing the year of study June 1992- June 1993 was dropped from all the above equations.

<u>4.4.4 Comparing explanatory and predictive accuracy of the two modeling approaches.</u>

The "LCC-only variables" approach applies the results from the panel data analysis (LSDV1 for farmers, LSDV1 for years, OLS pooled) to the Markov chain model through matrix multiplication. The "survey-data variables" approach uses the survey data in a multinomial logit regression model to predict the LC class of each pixel. It is important to note that the first modeling approach generates predicted transition probabilities for each of the 9 LCC possible events. In contrast, the multinomial logit regression does not produce values for the 9 feasible LCC events, instead it predicts and assigns one of the three possible final land covers to each pixel.

Hypothesis six stated that land cover changes among small farmers are better explained and predicted when classifying farmers according to how their total landholdings area changes through time (owner typology). In order to test this hypothesis, dummy variables for the owner type and property type will be included in the multinomial logit model. Their contributions to the explanation of the variability of the future LCC event will be compared to decide whether the property size or the total landholdings affect more the LCC decisions in the farm.

To test hypothesis five, regarding the explanatory and predictive capacity, LCC predictions at the pixel level will be compared using the percentage of cells for which future LC was predicted correctly. Methodological details, relevant results and possible explanations for testing these two hypotheses will be explained in detail in the following subsections

4.4.4.1 Modeling approach 1: explaining and predicting LC using only past LCC probabilities.

Modeling approach 1 uses the results from the panel data analysis (sections 4.4.3.2 and 4.4.3.3) in a modified Markov chain. Basically, the predicted probabilities for the 9 LCC transitions are used in a matrix multiplication fashion to obtain a new matrix of 3 x3 with the LCC probabilities of LC in the next stage in time (see figure 5). In sections 4.4.3.2 and 4.4.3.3. 14 LSDV1 models were run for the 7 feasible LCC: 7 models were run using the group effects model and 7 using the time effects model. A pooled model (Ordinary Least Square –pooled, OLS-pooled) ,disregarding differences among subjects and across time, was run as well. Predicted probabilities from the LSDV1-group, LSDV1-time and OLS-pooled models were separately used in the Markov matrix multiplication.

Figure 5 Matrix multiplication of the matrix of LCC transition probabilities

$$A \times A = A^2$$

$$\begin{vmatrix} \mathbf{P_{PF\text{-}PF}} & \mathbf{P_{PF\text{-}SF}} & \mathbf{P_{PF\text{-}NF}} \\ \mathbf{P_{SF\text{-}PF}} & \mathbf{P_{SF\text{-}SF}} & \mathbf{P_{SF\text{-}NF}} \\ \mathbf{P_{NF\text{-}PF}} & \mathbf{P_{NF\text{-}SF}} & \mathbf{P_{NF\text{-}NF}} \end{vmatrix} \mathbf{X} \begin{vmatrix} \mathbf{P_{PF\text{-}PF}} & \mathbf{P_{PF\text{-}SF}} & \mathbf{P_{PF\text{-}NF}} \\ \mathbf{P_{SF\text{-}PF}} & \mathbf{P_{SF\text{-}SF}} & \mathbf{P_{SF\text{-}NF}} \\ \mathbf{P_{NF\text{-}PF}} & \mathbf{P_{NF\text{-}SF}} & \mathbf{P_{NF\text{-}NF}} \end{vmatrix} = \begin{vmatrix} \mathbf{P^*_{PF\text{-}PF}} & \mathbf{P^*_{PF\text{-}SF}} & \mathbf{P^*_{PF\text{-}NF}} \\ \mathbf{P^*_{SF\text{-}PF}} & \mathbf{P^*_{SF\text{-}SF}} & \mathbf{P^*_{SF\text{-}NF}} \\ \mathbf{P^*_{NF\text{-}PF}} & \mathbf{P^*_{NF\text{-}SF}} & \mathbf{P^*_{NF\text{-}NF}} \end{vmatrix}$$

Note: Matrix A is the matrix of LCC transition probabilities resulting from the panel data analysis of the first five years of the study period. When this matrix is multiplied by itself it produces the predicted LCC transition probabilities for year 2002, at the end of the ten-year study period. Three matrix multiplications were performed for each farmer under study, using the predicted probabilities from the LSDV1-farmer, LSDV1-year and OLS-pooled models.

4.4.4.2 Modeling approach 2: explaining and predicting LC using demographic, socioeconomic and land cover variables at the farm level.

A multinomial logit regression model was run in SPSS using selected variables from the survey data to predict future LC class for each pixel in the farm, when we know their current LC classification. This step was followed by a comparison of the contribution of each variable to explaining the variability of the LC classification of every pixel. The percentage of correct LC predictions was compared against the alternative model using only LCC variables.

Variables from the long list of variables in the survey questionnaire were selected based on previous research suggesting their relevance to deforestation, the land –cover change event of main interest for this dissertation. The list of the relevant variables used in the multinomial logit regression model was presented in Table 4.3.

CHAPTER 5: RESULTS AND DISCUSSION OF FINDINGS

This chapter presents the results of the analysis performed for each of the hypotheses being tested in this dissertation. Results are presented in the same order as that for which steps of the methodology were shown in Table 4.1 in the methodology chapter. Relevant findings are discussed with respect to the hypotheses, research objectives, study site and scope of this research. Outputs of statistical software, programming codes, graphs are other analyses are contained in the Appendixes.

5.1 ESTIMATION OF THE EMPIRICAL TRANSITION PROBABILITIES USING PATHWAY ANALYSIS OF PAST LAND-COVER CHANGES.

Empirical transition probabilities were estimated for the 70 pathways of past land cover change (Alto Paraiso, n=27; Nova Uniao, n=27; Rolim de Moura, n=16) following the procedure explained in section 4.4.1. Farm owners were classified according to the change in their total landholdings from 1992 to 2002 in three categories: stable (owner type 1, OT1, n=35), subdividing (owner type 2, OT2, n=17), and expanding (owner type 3, OT3, n=18), regardless of where the land was located. Properties were classified according to their change in lot size from 1992 to 2002 into: stable property (property type 1, PT1), subdividing property (property type 2, PT2), and enlarging property (property type 3, PT3).

Some important methodological notes deserve to be mentioned. First, farms classified in the three property type categories were included in the study. However, only the portion of the farm that is present in both surveys (1992 and 2002) is studied in the pathways of past land cover change and in the pathway analysis to calculate LCC probabilities. Second, property and owner typologies are included in further steps of the analysis (in the analysis of differences of means and in the multinomial logit regression) through the use of dummy variables (OT =1, 2, 3 and PT=1, 2, 3). Initially the "gone category" was included in the pathway analysis, to represent land that was sold or lost and thus was not part of the property in the 2002 survey. I decided to exclude the "gone category" and use only the land of the farms that was "present" during the whole period of study in order to make a consistent pathway analysis of what happens with land through time. Computation of probabilities of a portion of land estimated to be "gone" proved to be computationally difficult given the small sample size. Moreover, this decision was taken also in order to be able to perform panel data analysis without imbalanced data (no data for certain subjects' variables for some years).

In section 4.4.1, an example of a pathway diagram was presented, the one corresponding to farm 9 in line 36 of the municipio Nova Uniao. Table 5.1 shows the average transitional probabilities during the ten-year period of study for the same pathway example (farm 9 in NU line 36). In the same fashion, empirical probabilities were calculated for each of the 9 land cover changes, for each of the ten years and for all seventy farmers in the sample. Similar tables could be shown for each of the 70 farmers in the sample, but instead a summary of the whole sample is presented in Table 5.2.

Table 5.1 Example of average transitional probabilities during the ten-year period of study for farm 9 in NU line 36

LC transition	**Average probability during the 10-year study period**
PF-PF	0.92
PF-SF	0.00
PF-NF	0.08
SF-PF	0.00
SF-SF	0.96
SF-NF	0.04
NF-PF	0.00
NF-SF	0.04
NF-NF	0.96

Table 5.2. Average transitional probabilities during the ten-year period of study for the whole sample of farmers (n=70)

LC transition	**Average probability during the 10-year study period**	**Std dev**
PF-PF	0.8909	0.179
PF-SF	0.0022	0.009
PF-NF	0.0655	0.058
SF-PF	0.0	0.0
SF-SF	0.6555	0.323
SF-NF	0.1131	0.084
NF-PF	0.0	0.0
NF-SF	0.0432	0.049
NF-NF	0.950	0.096

At first glance it stands out that the "persistence" events (PF-PF, SF-SF, NF-NF) are the dominant ones in terms of their high probabilities and it may seem that the "deforestation" events (PF-NF, SF-NF) are unlikely. However a closer look at the numbers and at their interpretation in the context of Markov chains shows a very different conclusion. These probabilities represent a not-irreducible Markov chain since Primary Forest (PF) is a transient state and Secondary Forest (SF) and Non-Forest (NF) are recurrent states[21]. PF is a transient state, which means it can only take place a finite number of times in the Markov process. This is the case given that the events SF-PF and NF-PF have probability of zero and thus no LCC event will produce a PF cover. Once a part of PF converts to other land cover (LC), then that primary forest is gone forever -- with all its qualities of a pristine forest—and the only possible conversions will result in SF and NF. On the other hand SF and NF are recurrent states since they can happen an infinite number of times, involving transitions back and forth between these two land covers (SF-NF, NF-SF) or land use transitions that produce the same land cover (SF-SF, NF-NF).

Given these conditions (PF transient, SF, NF recurrent) Markov theory predicts that with certitude at some future time all primary forest will disappear and that this will be an irreversible process, leaving only hope for re-growth and conversions to secondary forest. This is a very powerful conclusion that sets a high and imminent risk deforestation. We can get a sense of this process if we look at the cumulative effect of the figures in table 5.2. These are average yearly

[21] Non-irreducible Markov chain: a Markov chain where there are more than two classes, so not all states communicate with each other. Two states that communicate are said to be in the same class.
Transient states: States that are not certain to be returned to, even if the process starts in the state. If starting in this state, the expected number of time periods that the process is in this state is finite. A state that is visited only a finite number of times.
Recurrent state: If starting in this state, the expected number of time periods that the process is in this state is infinite. There is always a probability to reenter this state. The state is visited infinitely number of times.

probabilities calculated from the sample of 70 farmers over the ten-year period. Every year approximately 10% of the primary forest is lost into SF or NF (0.0022 plus 0.0655), and it is never replaced. Every year, 35% of the land in SF becomes NF (SF-NF probability=0.6555); and 95% of the land in NF will remain as NF for the next year (NF-NF probability= 0.950). In a year, only 5% of the land in NF will allow for re-growth into SF. If we look at the cumulative process over a number of years it is easy to see why PF is disappearing due to the deforestation process.

In the long run the best feasible scenario is the preservation of SF. If all PF will certainly disappear, then efforts should focus on increasing the probabilities of the events SF-SF and NF-SF. This takes us to a very important part of the deforestation Markov chain: NF as an absorbing state. If NF is an absorbing state, it means that once land converts to NF, this land will never leave this state. Most of previous predictive models assume NF is an absorbing state and carry out this assumption without further proof. In the present dissertation, the capability to estimate empirical transition probabilities allowed us to see that although small, there is a probability for NF-SF conversions. Although NF is not an absorbing state as it has been assumed in previous models of LCC, the probability NF-SF is so low that NF may be considered an absorbing state in some regions of the Amazon. This probability can be increased if degraded land is allowed time to recover and if human interventions and technological improvements are used to regenerate forest.

NF may become an absorbing state in the future if vegetation is not allowed to regegerate. In the case where NF is absorbing, PF is transient, and SF is recurrent, then the Markov chain is not irreducible. Furthermore, according to Markov chain theory, if a finite

Markov chain starts in a transient state, then the chain is certain to enter some closed communication class, then the system is unable to leave certain states. This would mean that, it is certain that all land will end up in the absorbing state NF and it will remain NF. Once more, it is clear that efforts should target the NF-SF transition in order to allow for re-growth and regeneration.

Some key conclusions can be drawn from this Markov theory analysis. First, conservation efforts have to focus on protecting primary forest because once it is gone it will be gone forever. Second, if current practices continue in the future, PF will certainly be eradicated. Third, increased efforts should target the NF-SF event, encouraging the re-growth process. The policy implications of these conclusions will be discussed in more detail in chapter 6.

5.2 SCREENING OF DATA

Results of the descriptive statistical analysis and probability trends among farmers and across years are presented and discussed below. Analysis of differences of means and graphs of LCC probabilities vs. time were performed for different typology classifications: owner type, property type, municipio (AP, NU, RM) and farmer type. The property typology is based on changes on the size of the lot under study. The owner typology is built on changes on total landholdings. Farmer typology is based on subdividing, expanding and stable size of total landholdings and the nature of the owner as same owner, new owner part of the kin family and new owner unrelated. These typologies were described in more detail in section 4.4.3.1.

<u>5.2.1 Exploring classification typologies to describe differences among farmer's LCC probabilities.</u>

This section focuses on the first part of research objective 1: Describe patterns in the distribution of the transitional probabilities of land cover change among small farmers in Rondônia, Brazil, especially different patterns of subdividing, expanding and stable owner types. Table 5.3 shows the results for the screening analysis of relevant LC survey variables in 1992 and 2002, for the whole sample of pathways (n=70). The results of the screening analysis of relevant LC survey variables in 1992 and 2002 using the owner typology are shown in Tables 5.8 and 5.9.

Table 5.3 shows that there are significant statistical differences between several LCC variables in 1992 and 2002 among the 70 farmers sampled (pathway sample.) First, the percentage of farmers with multiple rural properties increased from 13.4% to 31.3%, showing a significance difference at the 0.050 level. Although the average number of owned rural properties decreased from 2.0 to 1.37 properties, there is still a significant difference at the 0.050 level. Thus, we can conclude that more farmers have multiple properties but they do not own a large number of properties. Moreover, given that the average total area owned and the area of the survey plot (primary lot) do not show significant differences at the 0.050 level, we can infer that small farmers are acquiring small parcels of land – not adjacent to the primary lot -- . This makes sense since farmers in Rondônia are predominantly small farmers –with small capital as well -- compared with farmers in other states, for example *Mato Grosso*. These figures also tell us that the land aggregation process is occurring in small incremental steps, where some farmers are subdividing and selling small portions of land and where other farmers are increasing their landholdings also in small steps. If we can extrapolate this land aggregation process from the pathway sample to the large Rondônia sample (n=192 in 1992 and n=228 in 2002) and even to the whole state of Rondônia, then policies should focus on discouraging farmers from fragmenting their land. This recommendation will be more evident in the further discussion of differences among owner types (stable, subdividing, and expanding).

On other hand, the pathway sample does not show statistical significance in the cattle herd size and urban property ownership. As for the cattle variable, these figures reflect only cattle owned by the farmer but do not include third party cattle on the farmers' property. Unfortunately, the 1992 survey questionnaire did not ask for differences among owned and third-party cattle.

Second, there is statistical significant difference at the 0.001 level in the areas and percentages of forest and non-forest land covers with GIS-based figures. The percentage of area in perennials, pasture, and forest --according to survey data—show also significant statistical differences. The forest and non-forest variables are the ones of interest to this dissertation's research question regarding differences in the deforestation probabilities among small farmers in Rõndonia.

Table 5.3 Relevant LCC variables in 1992 and 2002 for the pathway sample (n=70)

Variable label in SPSS	**1992**	**2002**	**Sig.**
sample size (n)	N =70	N =70	
Area of the survey plot (has) (according to survey data)	87.65 (27.90)	77.69 (33.09)	0.060
Area of the survey plot (has) (GIS-based figures)	69.47 (35.24)	81.58 (31.63)	0.054
Owns multiple rural properties (% yes)	13.4 %	31.3 %	0.011*
♦ Number of rural properties owned	2.00 (1.31)	1.37 (0.65)	0.026*
♦ Area of all rural properties owned (alq) (1 alq=2.4 ha)	61.67 (49.57)	41.05 (27.19)	0.060
Heads of cattle the owner owns now (#)	107.58 (82.60)	83.85 (110.023)	0.159
Owns urban properties (% yes)	28.4 %	17.6 %	0.101
Non Forest (ha) (GIS-based figures)	35.30 (27.49)	58.85 (32.51)	0.000***
Non Forest (% of lot area) (GIS-based figures)	0.5020 (0.2067)	0.7179 (0.2166)	0.000***
Forest (ha) (GIS-based figures)	34.17 (20.85)	22.73 (20.82)	0.004***
Forest (% of lot area) (GIS-based figures)	0.4980 (0.2067)	0.2780 (0.2129)	0.000***
Area deforested (ha) (according to survey data)	40.15 (51.42)	54.14 (31.37)	0.063
Annuals (A) (% of lot area) (according to survey data)	6.028 (5.33)	4.52 (5.75)	0.122
Perennials (PE) (% of lot area) (according to survey data)	7.62 (7.67)	12.32 (14.00)	0.018*
Pasture (PA) (% of lot area) (according to survey data)	19.00 (16.44)	44.23 (30.74)	0.000***
Forest (FO) (% of lot area) (according to survey data)	63.34 (17.66)	30.00 (23.07)	0.000***
Percentage of deforestation (% of lot area) (according to survey data)	42.57 (23.40)	70.49 (23.09)	0.000***

* Significant at the 0.050 level, ** Significant at the 0.010 level, *** Significant at the 0.001 level.

The main conclusion from table 7 is that there are statistically significant differences in the area and percentages of forest and non-forest between 1992 and 2002 data for the same set of farmers. The next step is to determine if these differences are only temporal or if there are significant differences in the composition of the small farmers' sample. Tables 5.4, 5.5. and 5.6 will provide the answer to this question.

Table 5.4 shows results of the ANOVA test looking for significant differences among owner types 1, 2, and 3 for some relevant LCC variables in the 1992 survey. Post-Hoc tests were performed to look for significant differences between pairs of owner types. As described in section 4.4.3.1, the owner typology is based in the change in total landholdings between 1992 and 2002. Owner type 1 (OT1) had stable landholdings, while owner type 2 (OT2) had decreased landholdings, and owner type 3 (OT3) had increased landholdings. The ANOVA test did not find a significant difference, even at the 0.050 level, when looking at the three owner types together in the 1992 survey. From this we can conclude that, at the beginning of the study period, farmers were relatively homogeneous in terms of the lot area, percentage of area in forest, annuals, perennials and pasture. From the Post-Hoc test though, we can see a significant difference (at the 0.050 level) between stable and expanding farmers. Differentiation among farmers, with respect to deforestation and land uses, will develop as their total landholdings change due to the land fragmentation/aggregation process as Table 5.5 shows.

Table 5.4. Comparison of relevant LCC variables among farmers in the 1992 survey using the Owner typology.

Variable label in SPSS	OT 1	OT 2	OT 3	ANOVA test Sig	Post Hoc test. Sig
sample size (n)	N =35	N =16	N =19	N/A	
Area of the survey plot (has) (according to survey data)	81.35 (24.49)	87.69 (22.07)	99.87 (34.79)	0.071	1 & 3 (0.056)
Area of the survey plot (alq) (according to survey data)	33.47 (10.11)	36.37 (8.77)	41.53 (14.10)	0.050	1 & 3 (0.038)*
Heads of cattle the owner owns now (#)	96.71 (62.38)	101.71 (61.48)	133.28 (121.84)	0.303	no diff
Annuals (A) (% of lot area) (according to survey data)	5.80 (3.56)	6.07 (7.59)	6.44 (6.30)	0.92	no diff
Perennials (PE) (% of lot area) (according to survey data)	6.92 (6.16)	6.07 (6.49)	10.16 (10.44)	0.246	no diff
Pasture (PA) (% of lot area) (according to survey data)	17.21 (15.50)	16.24 (10.01)	24.55 (21.10)	0.244	no diff
Forest (FO) (% of lot area) (according to survey data)	70.07 (15.29)	71.63 (13.35)	58.85 (22.27)	0.053	1& 3 (0.071) 2& 3 (0.099)

Table 5.5 shows that in 2002, statistically significant differentiation exists among the three owner types, and between pairs or owner types, for most of the LCC variables under study. From the ANOVA test we find that when looking at the three groups together, they are significant different at the 0.001 level for the variables: area of the lot, number of rural properties owned, total area owned, area in non-forest and in annuals. Linking these results to the ones in table 8 it is easy to see that differentiation developed in the ten-year period among owners with different total landholdings. The owner typology helped to unveil these differences among the small farmers, which were not evident differences when comparing the whole sample of farmers in 1992 and 2002. The difference in number of rural properties and total landholdings was expected from the nature of the owner typology. The significant difference in the non-forested area reinforces the idea that differences among owner types affect deforestation practices as well. The percentage area in annuals is a striking result as well, showing that from a group of small farmers that started with similar percentage area in annuals, the expanding owners are moving away from annual cropping. Expanding owners (OT3) decreased their percentage area in annuals (in the primary lot) from 6.44% in 1992 to 1.28% in 2002, while stable farmers (OT1) remained about the same with 5.8% and 4.23%, and subdividing farmers (OT2) increased their land in annuals from 6.07% to 9.16%. OT3 are moving into pasture leasing activities as we can infer from the Post-Hoc test for total cattle (own cattle plus third party cattle). OT2 are not only subdividing and selling their land to cope with farm and living costs; but they are also intensifying their annual cropping activities.

Table 5.5. Comparison of relevant LCC variables among farmers in the survey Ro2002 using the Owner typology.

Variable label in SPSS	OT 1	OT 2	OT 3	ANOVA test Sig	Post Hoc test. Sig
sample size (n)	N =35	N =17	N =18	N/A	
Area of the survey plot (has) (according to survey data)	78.68 (24.89)	49.99 (22.98)	103.10 (34.52)	0.000***	1&2 (0.003), 1 & 3 (0.008), 3 & 2 (0.000)
Area of the survey plot (alq) (according to survey data)	32.78 (10.37)	20.81 (9.54)	42.96 (14.38)	0.000***	1&2 (0.002), 1 & 3 (0.008), 3 & 2 (0.000)
Heads of cattle the owner owns now (#)	75.24 (99.64)	46.87 (45.48)	134.78 (149.51)	0.055	no diff
Total cattle now =own +third party	83.38 (100.27)	63.20 (45.74)	165.56 (151.63)	0.014	1 & 3 (0.031), 2 & 3 (0.024)
Heads of cattle the farmer owned when he arrived to the lot	4.38 (11.97)	3.36 (6.01)	1.94 (7.07)	0.695	no diff
Total cattle when arrived = own +third party	4.38 (11.97)	3.36 (6.01)	1.94 (7.07)	0.399	no diff
Rural properties owned (#)	1.09 (0.29)	1.06 (0.25)	2.25 (0.68)	0.000***	1 & 3 (0.000), 2& 3 (0.000)
Area of all rural properties in (alq) 1 alq=2.4 ha	34.83 (12.80)	21.38 (9.47)	75.45 (31.65)	0.000***	1 & 3 (0.000), 2& 3 (0.000)
Non Forest (ha) (GIS-based figures)	58.71 (25.64)	35.34 (12.06)	80.54 (41.96)	0.000***	1 &2 (0.035), 1 & 3 (0.035), 2 &3 (0.000)
Non Forest (% of lot area) (GIS-based figures)	71.68 (21.42)	67.80 (22.99)	75.36 (22.15)	0.608	no diff
Annuals (A) (% of lot area) (according to survey data)	4.23 (3.68)	9.16 (9.35)	1.28 (1.73)	0.000***	1 & 2 (0.010), 2 & 3 (0.000)
Perennials (PE) (% of lot area) (according to survey data)	14.24 (16.03)	11.60 (13.79)	9.08 (8.96)	0.459	no diff
Pasture (PA) (% of lot area) (according to survey data)	40.30 (29.01)	44.18 (29.56)	52.11 (35.15)	0.44	no diff
Secondary Forest (SF) (% of lot area) (according to survey data)	6.53 (8.89)	10.72 (13.42)	4.50 (8.53)	0.219	no diff
Forest (FO) (% of lot area) (according to survey data)	31.91 (23.44)	28.02 (21.03)	27.80 (24.90)	0.787	no diff

Results from the Post-Hoc test show significant differences between pairs of owner types, particularly between expanding farmers(OT3) and any other type of farmer(OT1, OT2). Comparing the Post-Hoc results from table 5.5 against table 5.4 we can see that differences among pairs of owner types accentuated. In the Post-hoc test for 2002 there are many more variables showing significant difference and now the differences are at the 0.010 and 0.001 levels, compared with the 0.050 level in 1992. As expected, there is a significant difference in the number of rural properties and total area of rural properties, being OT3 the ones that show the largest differentiation. Still, these expanding owners do not have many multiple properties, since the average is 2.25. Since the size of the primary lot for OT3 did not experience a drastic change (41.53 alq to 42.96 alq), we can infer that they tend to retain their original lots and acquire small parcels of land, not-adjacent to that primary lot. The average total landholdings for OT3 is 75.45 alq, which is about the size of 2 original lots. Comparing this with the total landholdings of OT1 (34.83 alq, a bit less than the size of a lot) and the total landholdings of OT2 (21.38 alq, about half the size of a lot), we can see how farmers have evolved due to the fragmentation/aggregation land processes.

Regarding the cattle variables, there is no difference among the owner types when looking at the total cattle owned, neither in 1992 nor in 2002 data. However, in 2002 more cattle variables were collected providing information about the own cattle and third party cattle. In this regard, there are significant differences at the 0.050 level among OT1 and OT3, OT2 and OT3. Basically, expanding owners are the ones showing the largest cattle herds and largest third-party cattle. All farmers, OT1, OT2, OT3, are engaging in renting pasture for third-party cattle. An interesting question to ask in future research is: who owns that cattle?

The Post-Hoc test also shows significant differences among owner types, for the area in non forest and the percentage area in annuals. Comparing the percentage area in annuals in 1992 and 2002 for the OT1, OT2, OT3 we can see that they started with about the same percentage of area in annuals (5.8%, 6.07%, and 6.44% respectively). In 2002 though, OT3 have basically moved away from annual crops in their primary lot (1.28%), while OT2 still give priority to annual crops even in their smaller primary lot (9.16%). OT1 decreased slightly their focus in annual crops from 5.8% to 4.23%. Since annual crops are usually produced for consumption in the farmer' family, we can infer that all small farmers are generating some income from other economic activities (i.e. pasture rental, cattle, off-farm work) so they can provide food to the family without all the labor that annual crops require.

These differences among owner types can be explained by the land aggregation and subdividing patterns experienced in the agricultural frontier of Rondônia. When farmers arrived on the frontier, land was allocated in plots of 42 alq (100 ha). In 1992 there were slight differences in the size of the primary lot. However, as farmers evolved through time the differentiation process became more evident, showing some evidence of the success/failure of their land use and economic practices in the frontier. This differentiation process is relevant to this dissertation because it proves that small farmers do not share all the same characteristics, the same success on the frontier, land use/ land-cover change practices and therefore the same probabilities for deforestation. If expanding and subdividing patterns are developing among small farmers, then the general deforestation and cattle ranching trends that have been studied among small farmers (owning less than 100 ha of land) and large farmers (owning more than 100 ha of land) are likely being replicated at a small scale inside the pool of small farmers (small farmers with expanding, subdividing and stable landholdings). Therefore, predictive land-cover models that assume homogeneity among small farmers are missing details in the description of the system and accuracy in the predictions of the model.

A similar analysis was performed for the same relevant LC variables using the property type, farmer type and municipio classification typologies. Outputs are skipped due to their length and the limited or none significant differences found. Data by property type in 1992 shows significant difference only for the variable percentage of forest. In 2002 no variable shows significant difference.

Data by farmer type shows no significant difference either in 1992 or 2002 among the 9 farmer classes. The farmer typology had very small sub-samples given the use of nine typology classes; in some instances it was not possible to perform the ANOVA and Post-Hoc tests. Data by municipio shows significant differences in 1992 and in 2002, with a stronger difference in 2002. The general difference trends are among Alto Paraiso and Nova Uniao and Alto Paraiso and Rolim de Moura. Differences among municipios may require further analysis of the local planning and forest conservation policies and will be left for future research.

From the several typologies tested, the owner typology was the only one that showed significant differences in the LCC variables under study. We can conclude that changes in total landholdings (owner typology) are more relevant, in defining difference among small farmers in Rondônia, than changes in the size of the plot under study (property typology) or primary lot. This finding is very relevant not only for simulation purposes, but also for policy changes. In both instances, small farmers have been regarded and treated as a homogeneous group based solely on the size of their primary lot (in the state of Rondônia), overlooking the effect that other properties elsewhere can impose in the LCC decisions on the primary lot.

The second part of this section focuses on the following research question: Are there significant differences in the transitional probabilities of LCC among small farmers in Rondônia, especially among subdividing, stable and expanding farmers? For this question, analysis of difference of means tests for each of the 9 LCC classes, were conducted using the different classification typologies.

From the analysis of differences of means using the owner typology, property typology and farmer typology, the only LCC category that showed significant difference among classes was the SF-SF transition. The t-stats were 0.005, 0.037, and 0.011, respectively. Thus, the owner typology was the classification with the most significant difference. The classification by municipio showed significant difference for the SF-SF as well, with a t-stat of 0.018. This last classification showed also significant difference for the PF-NF transition, with a t-stat of 0.002. Thus, the municipio typology was the only classification that showed statistically significant differences in the probability of deforestation for the transition PF-NF. Further exploration of the differences at the municipio level (i.e. land use policies, land occupation and titling procedures) is suggested in the future research section. In order to see if owner types were distributed in a special pattern among municipios, a Chi-square test was performed for the percentage of surveys OT1, OT2, and OT3 in each of the municipios. This test showed no significant difference either in 1992 or 2002 that could indicate certain municipio was predominantly composed by a certain owner type. Such finding would have suggested that specific differences at the municipio level encouraged the expanding, subdividing or stable landholdings process.

The analysis of differences of means was presumably not the best test choice since it misses the information derived from time-series patterns. These tests were conducted using the average probability for each land cover change from the set of 10 –yr probabilities for each farmer. However, these tests were part of an exploratory analysis of data, looking for patterns of differences among small farmers. Table 5.6 shows the average probabilities for each of the 9 LCC events for the three classes in the owner typology and in the whole pathway sample. Using the owner typology, the only LCC that showed significant difference was the SF-SF (t-stat =0.005). From the figures in table 5.6 we could conclude high stability for the "persistence " LC and thus no point to be concerned about the deforestation process. However, the discussion of these probabilities under the scope of the Markov theory shows very different conclusions. To avoid repetition, please refer to table 5.6 in section 5.2 and the discussion of these probabilities from a Markov chain perspective.

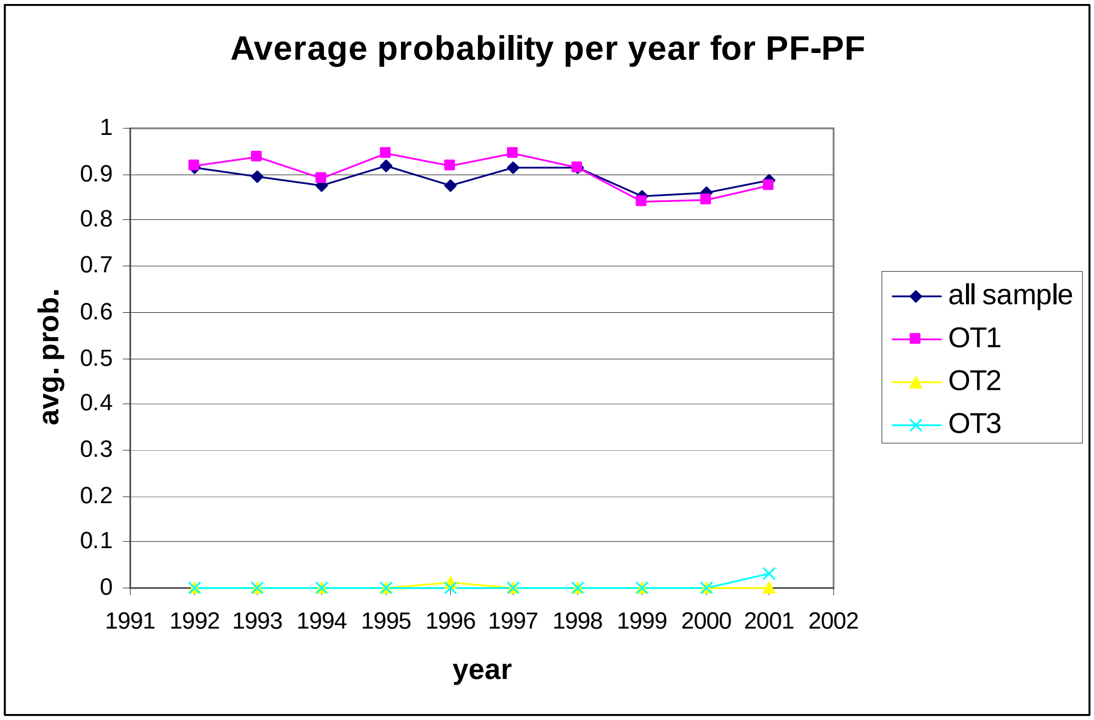
Table 5.6. Average transitional probabilities during the ten-year period of study for the classes in the owner typology and in the whole pathway sample

LC transition	OT1	OT2	OT3	Whole Pathway sample
PF-PF	0.9033	0.0013	0.0030	0.8909
PF-SF	0.0023	0.0013	0.0030	0.0022
PF-NF	0.0687	0.0503	0.0724	0.0655
SF-PF	0.0	0.0	0.0	0.0
SF-SF	0.7452	0.7229	0.4576	0.6555
SF-NF	0.1205	0.1334	0.0869	0.1131
NF-PF	0.0	0.0	0.0	0.0
NF-SF	0.0515	0.0460	0.0270	0.0432
NF-NF	0.930	0.940	0.970	0.950

Several graphs were produced using the transitional LCC probabilities in order to find temporal patterns and patterns in the owner typology that may have escaped previous screening of data. The following graphs were produced for each of the 9 LCC, for the whole pathway sample, and for each owner type (OT1, OT2, OT3): probabilities by year, average probability per year by year, average 10-yr probability by farmer ID. The only LCC change that showed a definite pattern was the persistence on forest, PF-PF, and thus this is the only graph discussed in this section. Other graphs are skipped here and presented in Appendix B.

Figure 6 shows a clear pattern in the probability of forest persistence (PF-PF) among owner types. The probability of forest persistence -- for land not deforested in the initial colonization years – among owner types 1 is very high and close to 1 (avg. probability of 0.9033), compared with the probability for the same LCC among owner types 2 and 3 (avg. probabilities of 0.0013, and 0.0030 respectively). This trend suggests that farmers whose landholdings are enlarging or subdividing through time have a lower incentive to preserve primary forest. This trend could be explained since farmers that acquire more land do so often with the intent to pursue cattle ranching, an activity that requires large extensions of pasture and thus involves forest clearing. On other hand, farmers whose landholding are reducing through time are often farmers that are not being successful in the agricultural frontier and have to sell land to pay for household and farming expenses. Such farmers may not have a big incentive to preserve primary forest since they may sell the rest of their land in a near future and move further into the frontier pristine forested area.

Figure 6. Differentiation of the average probability for forest persistence among the whole sample and the owner typology.



Returning to the premise of homogeneous subjects that stated that probabilities of land cover change are not constant among small farmers in Rondônia, this implies that they should not be treated all as one homogeneous group of farmers. So far, results of the screening analysis suggest that there are slight differences among farmers. This screening using the whole pathway sample (n=70) was likely missing differences among small farmers. Such differences started to unveil using different classifications typologies. Municipio and owner typologies were the only classifications that showed significant differences in the LCC probabilities. A clear pattern was finally uncovered by the owner typology showing that owners with stable properties tend to preserve forest with a much higher probability (0.9033) than owner with subdividing or expanding properties (probs. of 0.0013 and 0.0030).

Initial screening of data used a pooled sample (the whole pathway sample, n=70). The analysis of differences of means and such poolability missed the fact that we have panel data. We were missing the richness of information we can obtain from considering simultaneously cross-sectional and time-series data subsets of transitional probabilities. After this screening of data, it was evident that the nature of the panel data set had to be addressed. Thus, I performed panel data analysis using the fixed group effects model to take into account differences among farmers in their corresponding time series data and cross-sectional data. Results of the panel data analysis are presented and discussed in sections 5.3.1 and 5.3.2

5.2.1.1 The human face of the owner type 1

So far the research results suggest that changes in total area of landholdings are more relevant than changes in the area of the primary lot to land-cover decisions. Thus, the owner typology depicts the differences among small farmers better than the property typology. Let's recall that owners type 1 have stable total landholdings, while owners type 2 have subdividing or decreasing landholdings, and owners type 3 have expanding landholdings. Relevant land cover variables and LCC transition probabilities were presented and discussed in tables 5.4, 5.5 and 5.6. An important conclusion from section 5.1 and especially from figure 6 and table 5.6 is that type 1 owners are more likely to preserve primary forest (probability of PF-PF = 0.9033) than type 2 and 3 owners (probs. of 0.0013 and 0.0030, respectively).

Besides these land-cover variables, we may ask: what is the human face or social characterization of type 1 owners for planning and policy purposes? In order to answer this question a descriptive statistical analysis was conducted, covering socio-economic and demographic household variables among owner types in 2002. As mentioned before in the screening of data for land-cover change variables, it is at the end of the study period that the differentiation among owner types becomes more evident and that is why the statistical analysis was conducted with the 2002 data. Results of the analysis of these household variables are presented in table 5.7 and will be discussed in this section.

Table 5.7. Comparison of relevant socioeconomic and demographic household variables among owner types in 2002

Variable label in SPSS	OT 1	OT 2	OT 3	ANOVA test Sig
sample size (n)	N =35	N =17	N =18	
Age of head of household	56.47 (12.47)	48.66 (12.64)	50.89 (10.14)	0.001**
Years the owner attended school	3.08 (3.178)	3.69 (3.696)	3.39 (3.509)	0.613
How many other owners has this lot had?	1.91 (0.988)	2.12 (0.847)	2.60 (1.621)	0.008**
Wealth index[+] - initial	2.16 (2.165)	2.44 (2.454)	2.81 (3.197)	0.340
Wealth index – now (2002)	6.02 (2.579)	5.16 (2.613)	6.39 (3.601)	0.059
Number of families living in the lot	1.79 (1.009)	1.11 (0.868)	1.27 (0.906)	0.000***
Number of people living in the lot –Fn[++]	7.54 (5.161)	5.26 (3.349)	5.81 (3.988)	0.008**
Level of social participation[+++]	1.23 (0.841)	0.97 (0.882)	1.34 (0.861)	0.066
Rural properties owned (#)	1.09 (0.29)	1.06 (0.25)	2.25 (0.68)	0.000***
Number of men –F1[++++]	1.65 (1.30)	1.53 (1.40)	1.58 (1.01)	0.879
Number of women –F1	1.41 (0.946)	1.28 (0.797)	1.19 (0.732)	0.362
Number of children –F1	1.23 (1.692)	1.28 (1.333)	0.88 (1.219)	0.385
Number of elderly – F1	0.31 (0.572)	0.09 (0.366)	0.16 (0.485)	0.060
F1 total people	4.59 (2.461)	4.19 (2.185)	3.81 (1.816)	0.189
Dependency ratio –F1[+++++]	0.615 (0.681)	0.716 (0.864)	0.445 (0.551)	0.211
Dependency ratio --Fn	0.624 (0.561)	0.711 (0.777)	0.511 (0.560)	0.294

NOTE: [+]The wealth index was calculated based on a list of 18 household items and goods.

[++] Fn represents the sum of all families living in the lot, including the primary family.

[+++] The level of social participation has a scale 1 to 3 based on the participation in syndicate, cooperative or mutual help group.

[++++] F1 is the primary family or the owner's family.

[+++++] Dependency ratio = (children + elderly)/ (women + men)

Table 5.7. Comparison of relevant socioeconomic and demographic household variables among owner types in 2002 Continuation …

Variable label in SPSS	OT 1	OT 2	OT 3	ANOVA test Sig
sample size (n)	N =35	N =17	N =18	
Paid farm workers (daily salaries paid/yr) –F1	47.18 (74.689)	42.31 (104.62)	84.58 (176.06)	0.212
Number of people working off-farm last year --F1	0.50 (0.954)	0.65 (1.231)	0.37 (0.926)	0.468
Number of relatives living in a rural lot – F1	13.64 (23.56)	12.76 (15.46)	12.79 (17.39)	0.964
Number of relative living in an urban lot – F1	6.59 (9.92)	11.33 (19.73)	7.93 (14.39)	0.226
Paid farm workers (daily salaries paid/yr) –Fn	50.47 (81.176)	46.70 (108.50)	85.27 (170.52)	0.210
Number of people working off-farm last year --Fn	0.57 (0.880)	0.87 (1.331)	0.41 (0.734)	0.063
Number of relatives living in a rural lot – Fn	14.24 (24.07)	10.90 (13.16)	11.55 (16.81)	0.582
Number of relative living in an urban lot – Fn	7.34 (10.42)	11.15 (18.64)	7.84 (14.05)	0.297
Owner lives in the lot (%yes)	85.4 %	64.6 %	61.0 %	0.002**
Owner has been the only owner of the plot	52.4%	31.3%	27.1%	0.003**
Definitive land title	75.6%	46.9%	60.3%	0.002**
Multifamily	49.4%	27.7%	32.2%	0.016*
Earned money from renting pasture	15.5%	10.9%	10.8%	0.707
Rented pasture from others	30.0%	23.1%	15.4%	0.625
Syndicate	68.8%	53.3%	57.8%	0.157
Mutual help group	49.4%	36.7%	66.7%	0.010*
Cooperative	7.9%	6.6%	11.4%	0.669

Table 5.7. Comparison of relevant socioeconomic and demographic household variables among owner types in 2002 Continuation …

Variable label in SPSS	OT 1	OT 2	OT 3	ANOVA test Sig
sample size (n)	N =35	N =17	N =18	
Credit	21.0%	21.3%	26.4%	0.734
Savings account	10.3%	8.3%	6.8%	0.803
Checking account	32.9%	29.5%	38.6%	0.617
Owner pays daily laborers	55.9%	41.0%	69.4%	0.047*
Members of F1 work off-farm	26.0%	41.9%	16.3%	0.027*

The best way to characterize owner type 1 is by comparing the particular household variables of this group with owner types 2 and 3. Thus, a discussion of the statistically significant variables from table 5.7 will be discussed below.

First, owners type 1 tend to be older than OT2 and OT3. Years of schooling do not show significant difference among owner types and in fact all small farmers have limited level of education. OT1 tend to be the only owner that has possessed the property under study. Ahigher percentage of Owners type 1 have definitive land title (75.6%), compared with 46.90% and 60.3% of OT2, and OT3, respectively.

Second, the family composition of OT1 was not significantly different from OT2 and OT3's families, with respect to the number of men, women, elderly, and children in the primary family. There was no significant difference in the size and the dependency ratio of the primary family. When looking at all the families living in the lot (Fn), there was a significant difference in the number of families. And a larger number of families --and a corresponding larger total number of people, Fn -- live in farms owned by OT1. However, there was no significant difference in the Fn dependency ratio.

Third, there were no significant differences between owner types regarding the average number of daily wages paid, number of people working off-farm, number of relatives living in urban or rural areas. This statement applies for both, the primary family (F1) and all families together (Fn). However, when we look at the percentage of owners that pay daily laborers and the percentage of F1 families that pursue off-farm work, there were significant differences among owner types (0.050 level of significance). As expected a larger percentage of OT3 were able to pay daily laborers.

Fourth, although the level of social participation (i.e. participation in syndicate and cooperative) showed no significant differences among owner types, there is a significant difference in the percentage of owner types participating in mutual help groups. About 50% of OT1 participate in mutual help groups, compared with 66.7% of OT3, who have the largest percentage of participation With respect to credit, savings and checking accounts, there were no significant differences among owner types.

Fifth, regarding wealth and possessions, there were no significant differences among owner types, either in the initial or the final wealth index that includes some basic household items and small durable goods. As expected, there was a significant difference in the number of rural properties owned by different owner types, and owners type 1 in general only own one property –their primary lot. There was no significant difference in the percentage of owner types earning money from renting pasture, or in the percentage of farmers renting pasture from others.

In conclusion, if policy makers and planners want to encourage OT1 to continue their pattern of high forest conservation rates, efforts should focus on securing land titling, providing health care and alternative sources of income for the OT1's family members and elderly owners to remain in the lot. Moreover, the larger number of families and corresponding larger total number of people living on the lot represent labor force that should be encouraged to stay on the farm pursuing environmentally sustainable annual crops, perennials and small cattle ranching carried out in a balanced way. This balanced way includes pasture rotation, rotation of annual crops, agro-forestry projects and small and large cattle. All activities could be performed in designated areas of the farm allowing other areas for rotation and regeneration. The general purpose is to discourage new clearing while still providing sources of food and income for the large number of people living in the lot.

5.2.2 Descriptive statistics and graphs by owner typology to describe temporal differences in farmer's LCC probabilities.

This section focuses on exploring temporal trends in the LCC transitional probabilities in order to test if the assumption of stationary probabilities through time holds for the LCC process among small farmers in Rondônia.

Table 5.2,already discussed in section 5.1, showed the 10-year average transition probabilities for each of the 9 LCCclasses for the whole pathway sample (n=70). Appendix B presents graphs of the yearly average probability for each LCC plotted against time. From the figures in table 5.6 we could conclude high stability for the "persistence " LC and thus no point to be concerned about the deforestation process. However, the discussion of these probabilities under the scope of the Markov theory shows very different conclusions. To avoid repetition, please refer to table 5.2 in section 5.1 and the discussion of these probabilities from a Markov chain perspective.

An interesting result is the feasibility of non-negative probabilities of NF to SF suggesting that if land is abandoned when it is not so damaged, it can recover. This also leaves space for future technological improvements that may be used to guide forest regeneration.

Table 5.2 Average transitional probabilities during the ten-year period of study for the whole sample of farmers (n=70)

LC transition	Average probability during the 10-year study period	Std dev
PF-PF	0.8909	0.179
PF-SF	0.0022	0.009
PF-NF	0.0655	0.058
SF-PF	0.0	0.0
SF-SF	0.6555	0.323
SF-NF	0.1131	0.084
NF-PF	0.0	0.0
NF-SF	0.0432	0.049
NF-NF	0.950	0.096

Analysis of differences of means was conducted among the 70-farmer average probabilities for each LCC transition on the 10 years of the period of study. Thus each average analyzed in this test was calculated using 70 measures in each year (one for each farmer for each LCC in each year). Outputs are skipped here due to their length and limited significant results. The only LCC that showed significant difference across years was PF-NF. The post hoc test shows the main differences being between the years: 1992 and 2001, 1993 and 2001, 1996 and 2001, 1994 and 2001, 1997 and 2001. This suggests that a major change occurred in year 2001 causing farmers to deforest more. This difference is however very localized (reflecting changes in deforestation mostly in 2001). The general conclusion is that transition probabilities of LCC do not significantly change over time, as opposed to the conclusion of previous section where probabilities change significantly among subjects.

A possible explanation for these two trends is that all small farmers are subject to the same exogenous variables that may change yearly, such as interest rate, demand for agricultural cash products, government subsidies, etc. Such exogenous variables may not have changed drastically over the 10-year period of study and therefore changes over time on the LCC probabilities are minimal. This explanation relies on the premise that small farmers are more responsive to endogenous variables, as was mentioned in the literature review.

These results suggest that hypothesis 2 cannot be rejected. Hypothesis 2 proposes that probabilities could be assumed constant through time for a short study time period for a homogeneous group of farmers, specifically owner types (stable, subdividing, expanding). Since the yearly average probability of each LCC is about the same in each of the ten years studied and the standard deviation is about the same as well, this suggests that probabilities can be considered stationary through time but not constant among farmers. This could be explained since small farmers show more differences in their land aggregation behavior, demographic variables and other household variables and such variability affects their probabilities of LCC. Probabilities thus, change from farmer to farmer, but each farmer tends to keep the same land cover and land use strategies that have worked for him in the past.

Furthermore, comparing results from sections 5.3.1 and 5.3.2 we can say that differences among farmers are more significant than trends across time when analyzing the probabilities of LCC. Further analysis using panel data and the fixed group effects and fixed time effects were conducted to explore these trends among farmers and across time and in order to formally test two common assumptions used in stochastic predictive models –stationary probabilities and homogeneity of subjects.

5.3 TESTING MODELING ASSUMPTIONS

In this section I test two commonly used assumptions about farmers and their probabilities in stochastic predictive models–stationary probabilities and homogeneity of subjects. Results from the screening analysis suggest significant differences among farmers but not across time. Panel data analysis was conducted to formally test the effects of groups (farmers) and time (years) using the fixed effects technique.

5.3.1 Testing the assumption of homogeneity of subjects with respect to trends of transitional probabilities of LCC.

The Fixed Group Effects Model (FGEM) approach was used to test the assumption of homogeneous probabilities among subjects, a commonly used assumption in stochastic predictive models applied to small farmers' LCC decisions. Estimated empirical transition probabilities for the first five years of the 10-year study period were used in the FGEM. The empirical transition probabilities of the last five years of the study period were used for model validation in the Markov chain step of the methodology. The objective is to use the FGEM as the linear function in the modified Markov chain model described in section 4.4.2 The Fixed Group Effects Model (FGEM) approach and the LSDV1-farmer model were described in detail in section 4.4.3.2 and in equations 14-20.

A poolability test by group (farmer) was performed to choose between a model using all subjects in a single sample (OLS-pooled model) and a model clustering farmers in subgroups or typologies (LSDV1-farmer). The model with better goodness of fit measures (adjusted R square and Sum of Standard Errors) thus not only explained greater percentage of the variability of the probabilities of LCC, but also provide empirical data to test if it is worthwhile to pursue an analysis using a farmer classification.

A large number of regressions were performed (70 individual –OLSs, OLS-pooled model and LSDV1-model) using each of the 9 LCC as the independent variable. The software outputs, description and analysis of the models would take hundreds of pages. Therefore, I will only discuss the models developed for the PF-NF transition, which serves both to describe the deforestation and forest-persistence processes.

5.3.1.1 Fixed Group Effects Model

In order to find differences among farmers' LCC decisions, we could in principle run 70 cross-sectional linear regression models (one for each farmer) using the Ordinary Least Square Regression (OLS). Then we could compare the corresponding R squares among themselves and against the R square of the pooled OLS using all farmers in the same sample (OLS-pooled). Such procedure sounds reasonable but requires the more formal approach of FGEM to account for all farmers' differences at the same time. In fact the 70 OLS regressions were run in order to perform the poolability test by group that will be discussed in section 5.3.1.2.

Before choosing a Fixed Effect model to test the group and time effects, some exploratory linear regressions were conducted. These regressions were run using the probability of the PF-NF transition as the dependent variable. The objective was to find out how the probability for that LCC was affected by previous probabilities for the LCC: PF-PF, SF-NF, NF-NF, PF-SF, SF-SF and NF-SF. Equation 11 is presented again to show these relationships among probabilities.

$$P_{PF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 11)}$$

As explained in detail in section 4.4.3.1 the rationale behind the set of panel data equations is that the probability of a given land cover change depends on probabilities of other land-cover changes occurring in the present or previous years, in the specific farm under study and in the other farms as well. Basically these functions depict the interdependence among the LCC probabilities, and the temporal patterns in the probabilities series.

The “non-event” transitions were eliminated from the model due to their zero probability and in order to avoid redundancy. These exploratory regressions reflect in fact the nature of the Markov chains, where future LC can depend on present and/or past LC. For the scope of this research we are not yet making distinction among pixels on the same farm. Thus, the PF-NF in the exploratory regressions is in fact the probability of any forested pixel in a given farm becoming non-forested in the next year. This probability is being affected in some fashion by the LCC transitional probabilities of pixels in other land covers to shift to any of the other possible LC.

The exploratory linear regressions included:

- OLS pooled regression for the whole sample of farmers
- LSDV1 (Least Square Dummy Variable with one variable omitted) Fixed Group Effect
- Within Effect Model Fixed Group Effect
- LSDV1 Fixed Time Effect
- Within Time Effect Model

The model chosen was the LSDV1 Fixed Group Effect model because its intercept is the actual parameter estimate of the dropped dummy variable. Thus it allows for a straightforward interpretation, besides its computational advantages when using the SAS and SPSS statistical software packages. This model adds one dummy variable for each farmer and it requires that one of the 70 dummy variables be dropped in order to avoid the "dummy variable trap" (perfect multicollinearity).

OLS models and LSDV1 models were run for each of the 9 LCC, using some of the remaining LCC probabilities as the independent variables. The REG procedure of SAS was used first. Later, the TSCSREG procedure of SAS was used because it provides a test for fixed effects. In total 21 models were run to test differences among farmers. The SAS program for the exploratory models and the selected models can be found in Appendix A. Due to the voluminous amount of results, only the results for the PF-NF transition are shown in Table 5.8. The same LSDV1 models were run in SPSS to obtain additional statistical tests, such as the Durbin Watson statistic, which is used to test for autocorrelation.

The explicit equation for the OLS-pooled model is described by equation 11.

$$P_{PF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 11)}$$

The explicit equation for the LSDV1-farmer model is described by equation 18.

$$P_{PF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + \ldots + d_{69} \qquad \text{(Eq 18)}$$

Where $d_1 \ldots d_{69}$ are the dummy variables added to represent the "fixed subject effect" of each of the 70 farmers for which pathway analysis was conducted. As explained above, in the description of the LSDV1-farmer model, one dummy variable has to be dropped out to avoid the "dummy variable trap". The dummy variable d_{70} was dropped of the above equation. One of the many outputs of the regression models is provided as example in Appendix C.

Table 5.8. Results of the OLS-pooled regression and the LSDV1-farmer regression to test for fixed group effects.

	OLS pooled	**LSDV1 -farmer**
Dependent variable:	PF-NF	PF-NF
Independent variables:	PF-PF, SF-NF, NF-NF, PF-SF, SF-SF, NF-SF	PF-PF, SF-NF, NF-NF, PF-SF, SF-SF, NF-SF, d1-d69
Adjusted R square	0.3424	0.5847
Durbin Watson stats	1.937 (no autocorrelation)	1.952 (no autocorrelation)
Coefficients are significant for the variables:	PF-PF	PF-PF, PF-SF, SF-SF, all farmer dummies d1-d69

In order to reject the assumption of homogeneous probabilities among small farmers in Rondônia we have to prove that there are significant differences among farmers' LCC probabilities. If this is true then "small farmers" cannot be all treated as one group for LCC modeling purposes and exploration of what makes these farmers behave differently is justified.

The OLS pooled regression represents the situation when we consider there is no difference among small farmers with regard to their transition probabilities. The LSDV1 model is equal to the OLS pooled with 69 additional variables, one for each farmer. Both models show no autocorrelation since the Durbin Watson statistic falls in the region that rejects the autocorrelation hypothesis.

The LSDV1 fixed group effects model is a better model than the OLS pooled one. We arrive at this conclusion based not only on the improved R square, but also on the many additional significant coefficients. In fact all dummy variables' coefficients prove to be significant, which means that definitely there are differences among farmers.

In order to formally test which model is better, and if there are fixed group effects in the data, a partial F test is conducted. This test is formally called test for fixed group effects of panel data (Hun Myoung Park, 2005) Relevant results are shown in Table 5.9 and calculations are included in Appendix D.

Table 5.9. Partial F test for fixed group effects

	OLS pooled	**LSDV1 -farmer**
Adjusted R square	0.3424	0.5847
SSE	7.38247	11.68799
d.f. numerator = n-1	69	
d.f. denominator = nT-n-k	624	
F value	5.27	

The null hypothesis in this F-test is that some of the farmers' coefficients are equal to zero. Then, rejecting the null hypothesis means that adding the dummy farmer parameters improves the goodness of fit of the model and thus the LSDV1 model is preferred over the OLS-pooled model. From the large F value we can conclude that the null hypothesis is rejected and that the LSDV1 model is better than the OLS pooled model.

The conclusion from this section is that there are group effects due to differences among the small farmers that prevent a pooled sample analysis from accurately depicting the trends in their probability patterns as a whole group. The test for fixed-subject effects does not give insight about what differences exist among farmers and which variables contribute to these differences in their probabilities of LCC. For that effect, the relevance of the owner type and property typologies was tested in the multinomial logit regression model and results are discussed in section 5.4.

5.3.1.2 Poolability test

The purpose of the previous section was to test for fixed-subject effects in the panel data. To infer if significant differences among subjects existed that should be considered in further analysis and/or classification of the data. In this section, the poolability test is applied to assess the validity of using a pooled regression. The poolabilitity test is another way to answer the question: can all small farmers be treated as a whole sample for purposes of studying their probabilities? Or in other words: if we use a pooled regression, are we significantly decreasing our capability to explain the variability of the independent variable PF-NF? This F test was conducted and detailed calculations are included in Appendix D. The F-stats is equal to 0.4244.

This test asks if the slopes are the same across groups, which is the main assumption of the fixed and random effects models, allowing only differences in intercepts and error variances. The small F statistic obtained does not reject the null hypothesis in favor of poolable data with respect to farmers. Thus, we can conclude that when using a pooled sample of small farmers we are missing relevant information to explain the variability of the LCC transition probabilities. There are significant differences among small farmers and they should not be treated as a homogeneous sample for modeling and policy purposes.

5.3.2 Testing the assumption of stationary probabilities of LCC.

The Fixed Group Effects Model (FGEM) approach was used to test the assumption of stationary probabilities through time, in the same way that it was used to test homogeneity of subjects. A poolability test for time (year) was performed to decide if the model pooling all probabilities from all years in a single sample is better than exploring temporal changes of the probabilities across time.

5.3.2.1 Fixed Time Effects Model

This section tests the assumption of stationary probabilities through time, which is a broadly used assumption in stochastic predictive models of LCC. To find differences in the LCC probabilities across years, we could in principle run one cross-sectional linear regression model for each year using the Ordinary Least Square Regression (OLS). Then we could compare the corresponding R squares among themselves and against the R square of the pooled OLS using all probabilities from all years in the same sample. As mentioned in section 5.3.1, there is a more formal way to test for fixed-time effects.

A LSDV1 fixed time effects model was run adding dummy variables for each year and excluding one dummy variable from the model. The SAS programs for fixed time effects are presented in Appendix A. Due to the voluminous amount of results, only the results for the PF-NF transition are shown in Table 5.10. The same LSDV1 models were run in SPSS to obtain additional statistical tests, such as the Durbin Watson statistic, which is used to test for autocorrelation.

The explicit equation for the OLS-pooled model is described by equation 11.

$$P_{PF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6 \qquad \text{(Eq 11)}$$

The explicit equation for the LSDV1-year model is described by equation 25.

$$P_{PF\text{-}NF} = \beta_0 + P_{PFPF}\,\beta_1 + P_{PFSF}\,\beta_2 + P_{SFSF}\,\beta_3 + P_{SFNF}\,\beta_4 + P_{NFSF}\,\beta_5 + P_{NFNF}\,\beta_6$$
$$+ y_2 + y_3 + y_4 + y_5 \qquad \text{(Eq 25)}$$

Where $y_2 \ldots y_5$ are the dummy variables added to represent the “fixed time effect” of the probabilities changing over the first five years of the period of study (1992-1993, 1993-1994, 1994-1995, 1995-1996, 1996-1997). As explained before, in the description of the LSDV1-farmer model, one dummy variable has to be dropped to avoid the “dummy variable trap”. The dummy variable y_1 , representing the year of study June 1992- June 1993 was dropped of the above equation.

Table 5.10. Results of the OLS pooled regression and the LSDV1-year regression to test for time effects.

	OLS pooled	**LSDV1 -year**
Dependent variable:	PF-NF	PF-NF
Independent variables:	PF-PF, SF-NF, NF-NF, PF-SF, SF-SF, NF-SF	PF-PF, SF-NF, NF-NF, PF-SF, SF-SF, NF-SF, yr1-yr4
Adjusted R square	0.3424	0.371
Durbin Watson stats	1.937 (no autocorrelation)	2.01 (no autocorrelation)
Estimated coefficients are statistically significant for the variables:	PF-PF	PF-PF

In order to reject the assumption of homogeneous probabilities across time in Rondônia we have to prove that there are significant differences in the probabilities across years. If we reject the null hypothesis then there are significant differences in the probabilities across time, thus we cannot find a single value that can be used to model deforestation in different points in time. If that is the case, transition probabilities have to be calculated every year and there is the possibility that no predictable pattern may exist at all.

The OLS pooled regression represents the situation in which we consider we can find a "magic probability number" that represents the probability for a given LCC, regardless of the year. If such number exists, then there is no need for yearly survey data collection or satellite imagery. This looks at first instance as a very ambitious statement. Therefore formal testing for both fixed-time effects and poolability across time is needed. The LSDV1 fixed time effects model is equal to the OLS pooled with additional dummy variables, one for each year. Both models show no autocorrelation since the Durbin Watson stats falls in the region that rejects the autocorrelation hypothesis.

The LSDV1 fixed group effects models is a slightly better model than the OLS pooled one. The improvement in the R square is relatively small (from 0.3424 to 0.371). There is no improvement in the number of statistically significant coefficients. Although the four dummy variables added to the model improved slightly the R square, they proved not to be significant in explaining the variability of the probability of PF-NF.

In order to formally test if there are fixed-time effects in the data, a partial F test is conducted. Relevant results are shown in Table 5.11 and one the regression outputs with the F statistic can be found in Appendix E.

Table 5.11 Partial F test for fixed-time effects

	OLS pooled	**LSDV1 -time**
Adjusted R square	0.3424	0.371
SSE	7.38247	11.68799
d.f. numerator = n-1	9	
d.f. denominator = nT-n-k	684	
F value	3.45	

The null hypothesis in this F-test is that some of the dummy parameters for the years are equal to zero. Rejecting the null hypothesis means that adding the dummy year parameters improves the goodness of fit of the model and thus, the LSDV1-year model is better than the OLS-pooled model.

From the F value of 3.45, and P value of 0.0004, we can conclude that there are no time effects. This would mean that farmers keep their same LCC decision patterns, at least for a period of time like the 10 years used in this research, which is a relatively short period of time for modeling purposes. The F test for the fixed time effects indicates that hypothesis 3 -- Probabilities are not constant through time for the sample of farmers as a whole – can be rejected. However, this is not a conclusive statement and analysis of a larger period of study should be performed before extrapolating conclusions.

To further inquire on this hypothesis and hypothesis 4 (which are very interrelated), a poolability test by year was conducted. So far the fixed-time effects test indicates that there are not significant changes through time in the LCC probabilities that impose temporal patterns and justify the use of separate regression models.

5.3.2.2 Poolability test by year

The conclusion from the previous section is that there are not fixed-time effects that justify further analysis of temporal trends. In this section we will test if the OLS-pooled model should be preferred from individual regression models by year. The poolabilitity test is another way to answer the question: do farmers' LCC probabilities change significantly from year to year? This F test was conducted and detailed calculations are included in Appendix D. The F stats is equal to 2.99.

This test asks if the slopes are the same across years, which is the main assumption of the fixed and random effects models, allowing only differences in intercepts and error variances. The null hypothesis is that the fixed-time effects models is better than the OLS-pooled model. The large F-statistic rejects the null hypothesis in favor of poolable panel data with respect to time.

So far we can conclude that there are fixed-group effects indicating significant differences among the LCC probabilities by farmer. On the other hand, no fixed-time effects were found. In other words, there is more fixed-effects variability in the LCC probabilities from farmer to farmer than across time.

Does this mean that there are only group effects and no time effects? Not necessarily, since there is the possibility of random time effects to exist or the case of both group and time effects. Future extension of this dissertation research will include testing for random-group effects, random-time effects, and two way group and time effects.

5.4 COMPARING MODELING APPROACHES: "LCC-ONLY VARIABLES" VERSUS "SURVEY-DATA VARIABLES"

This section compares the explanatory and predictive capacity of two modeling approaches to predict land cover change. The first one using only pathways of past land cover change and the second one using household survey data (demographic, socioeconomic and land cover data at the farm level).

The "LCC-only variables" approach applies the results from the panel data analysis (LSDV1 for farmers, LSDV1 for years, OLS pooled) to the Markov chain model through matrix multiplication. The "survey-data variables" approach uses the survey data in a multinomial logit regression model to predict the LC class of each pixel. It is important to note that the first modeling approach generates predicted transition probabilities for each of the 9 LCC. In contrast, the multinomial logit regression does not produce values for the 9 feasible LCC, instead it chooses one of the three possible final land covers for each pixel. Therefore, in order to compare the two models, I looked at the percentage of cells with a correct LC prediction. The multinomial logit regression directly produces the predicted LC class for each pixel, thus the only calculation involves the percentage of correct pixel classifications. The probabilities resulting from the Markov multiplication are used to assign the LC to each pixel (assigning the final LC with highest probability to each initial land-cover).

Hypothesis six stated that land cover changes among small farmers are better explained and predicted when classifying farmers according to how their total landholdings area changes through time (owner typology). In order to test this hypothesis, dummy variables for the owner type and property type will be included in the multinomial logit model. Their contributions to the explanation of the variability of the future LCC will be compared to decide whether the property size or the total landholdings affect more the LCC decisions on the farm.

To test hypothesis five, regarding the explanatory and predictive capacity, LCC predictions at the pixel level will be compared using the percentage of cells for which future LC was predicted correctly. Methodological details, relevant results and possible explanations for testing these two hypotheses will be explained in detail in the following subsections

5.4.1 Modeling approach: explaining and predicting LC using demographic, socioeconomic and land cover variables at the farm level.

The starting premise is that demographic, socioeconomic and land cover survey data at the farm level significantly affect our capacity to predict future land cover. In fact the premise assumes that survey data very likely will improve our explanatory and predictive accuracy. In making their LCC decisions farmers may take into consideration other factors besides farm's LC, e.g. family labor, credit, and value of other landholdings. The rationale behind the assumption is that small farmers are more sensitive and responsive to changes in endogenous variables compared with exogenous variables.

To test this premise a multinomial logit regression model was run in SPSS using selected variables of the survey data to predict future LC class for each pixel in the farm, when we know their current LC classification. This step was followed by a comparison of the contribution of each variable to explaining the variability of the LC classification of every pixel. The percentage of correct LC predictions will be compared in section 5.4.2 against the alternative model using only LCC variables.

Variables from the long list of variables in the survey questionnaire (more than 300 variables) were selected based on previous research suggesting their relevance in affecting deforestation, which is the main LCC studied here, given the forested, non-forested LC classes used. The list of variables included in the multinomial regression model is presented in Table 5.12.

Table 5.12. Independent variables used in the multinomial regression model

Independent variables	
Municipio (1,2,3)	Mutual help group (0/1)
Owner Type (1,2,3)	Interest in planting native tree species (0/1)
Multifamily household (0/1)	Area of the lot in alqueires
Land title (0/1)	Percentage area in annuals
Planted native tree species (0/1)	Percentage area in forest
Property type (1,2,3)	Number of people living in the lot
Credit (0/1)	Number of people living in the lot * Total cattle owned now
Savings account (0/1)	Total cattle owned now
Checking account (0/1)	F1 dependency ratio
Urban properties (0/1)	Syndicate (0/1)
Other rural properties (0,1)	Cooperative (0/1)
Paid farm workers (0/1)	Extracts forest products (0/1)
Off-farm work (0/1)	
Dependent variable	
Land cover class of the pixel (PF, SF, NF)	

These variables have been linked to deforestation in previous research. Some of the references that helped to define this list of variables included: Dale et al, 1994; Godoy, et al. 1997; Browder, 1996; Pfaff, 1999; Lambin, 2003; Evans, 2001; Fearnside P.M, 2001; Perz, 2002; Pan et al, 2004; Mertens, 2002.

The way these variables can be expected to be related to deforestation is explained as follows. First, the variables credit, savings account and checking account are likely to positively covary with deforestation, under the assumption that farmers are more likely to use available capital to pursue the economic activity of highest return, in this case cattle. Since cattle-ranching requires large areas of pasture, it is likely that farmers will clear land to pursue or expand this economic activity. Second, ownership of other rural and/or urban properties is also likely to increase deforestation on the primary lot. The assumption is that farmers are more financially secure due to other properties, not only for the available land to pursue other economic activities, but also for the investment value that the properties represent. Third, social participation in syndicates, cooperatives and mutual help groups may indirectly increase deforestation. The rationale is that these forms of social participation may increase access to credit, checking and savings accounts, and organized manual labor. Forth, paid farm workers and income from off-farm work are likely to increase deforestation as well. The existence of paid farm workers can be interpreted as the owner’s large financial capacity to even pay extra labor, which may also be used for clearing practices if needed. Remittances from off-farm work are money that can be used to pay for the farm and living expenses. Thus, as any other form of capital, it may be used for entrepreneurship activities and cattle is likely to be top in the list.

Fifth, number of people living in the lot, multiple families living in the lot may be negatively correlated with deforestation. This is a surprising result is contrary to what we expected based on previous research. However, we will try to explain this phenomenon. The presence of multiple families in the lot or simply a large number of people living in the lot may be related to intense labor activities such as annual crops. Therefore, high values of these variables may be linked to reduced deforestation. Sixth, dependency ratio and definitive title may have positive or negative effects in deforestation, depending of the farmer's particular conditions. The dependency ratio –equal to the number of children and elderly divided by the number of working age men and women—is an indirect measure of the proportion of economically active and inactive people in the farm. If the dependency ratio is closer to zero, then there is more available labor in the farm and the effects will follow the pattern described for number of people and multiple families. If the dependency ratio is closer to 1.0 or higher then there is limited available labor in the farm and we can expect limited deforestation if the family focus on annual crops. However, the family may be an empty nester that is focusing on cattle activities thanks to the remittances received from family members working off-farm. As we can see, there are many intricate ways in which the variables may relate to deforestation in each particular context.

Seventh, land title has produced ambivalent results in its effects on deforestation, as it has already mentioned in the literature review. Farmers may clear land due to the pressure to prove land occupation to obtain a land title. On the other side, clear property rights and titling may reduce aversion to cattle investment and entrepreneurship, encouraging clearing of land. Eighth, forest extraction, either timber or other products, may create an incentive to clear forest to make a living. Ninth, interest in planting native tree species and actually having planted native tree species are variables linked to reduced deforestation. These variables may show a strong interest in forest conservation and possibly technical and financial assistance to pursue alternative activities. Tenth, the larger the amount of cattle owned, the larger the incentive to convert forest into pasture. It is important to clarify that a farmer may have not only his own cattle in his property, but also third-party cattle. Unfortunately, the 1992 survey did not include specific questions in this regard, and only the 2002 disaggregated data is available. Eleventh, the interaction between the number of people living in the lot and the total cattle owned may provide insight about the labor used for cattle purposes and a composite effect on deforestation. I included this variable based on my experience with the Rondônia database and my personal curiosity on the cattle/people interaction. The multinomial logit model allows setting interactions between pairs of variables. However, the coefficient will not tell if there is a direct or indirect relationship.

Twelfth, the effect of the variables area of the lot in alqueires, percentage of area in annuals, percentage area in forest, on deforestation is explained as follows. The area of the lot may correlate with higher probabilities of deforestation, but the actual effect on deforestation will depend on the existence of other rural or urban properties, the owner typology (expanding, enlarging and sub-diving landholdings), among other variables. The percentage of area in annuals tells us how relevant is this economic activity for the farmer’s family in terms of food and income production. Absence or low percentages of land in annuals will indicate that few or no people are living in the plot, or that other economic activities such as cattle are the main economic endeavor. Of course, the percentage of area in forest will affect how much can be cleared in the future and what other land cover decisions can be made with adjacent land.

Finally, the municipio, owner typology and property typology were added to the model to see which typology has significant effect and greater effect when explaining and predicting the final land-cover of the pixel. The significance can be observed from the coefficients and their significance level. The level in the contribution to explain LCC can be related to the contribution of each variable to the -2 log likelihood, which is a measure of the capacity of the model to explain the variability of the dependent variable (final land-cover).

Let's remember that the dependent variable in the multinomial logit regression is the land cover change of the pixel, which can assume three values (NF, SF, PF). Thus, this model predicts the final land-cover for each pixel. The model indirectly predicts deforestation when the initial land cover of a pixel is PF and the final is NF. In general, a logit regression model has a dependent variable a categorical variable that can assume two values –then the model is called logit—or three or more values –then the model is called multinomial logit. The independent variables can be numeric or categorical and interactions between variables are possible as well.

The longer the list of variables, the larger is the probability of correlation among the variables. Thus, a correlation analysis of the variables proposed for this study was performed. The multinomial logit model automatically excludes redundant variables (since a step-wise regression was used), but it does not allow the user to select which one of the correlated variables to delete. Since there is no perfect advice about which variables to eliminate from the model in case they are correlated, several models were run and compared. In total, 78 models were run trying different combinations of the variables.

Table 5.13 presents the results of the model with the largest improvement in the initial -2 Log likelihood measure. and the largest pseudo R square. The coefficients and their corresponding significance as presented as well. Table 5.14 shows the individual likelihood contributions of each independent variable. The SPSS output is included in Appendix F. Results of the variables with significant coefficients and likelihood contributions will be discussed in detail in this section.

Table 5.13. Results from the multinomial logit regression model. B coefficients and significance.

Dependent variable: LC of the pixel (PF, NF, SF) / Independent variables:	**B coefficient**	**Significance**
Municipio = 2	0.912	0.000***
Owner Type = 1	-1.631	0.000***
Owner type =2	-1.368	0.012**
Multifamily household = 0	-0.159	0.724
Land title = 0	0.739	0.015**
Planted native tree species = 0	1.169	0.000***
Property type = 1	-0.127	0.783
Property type = 2	-0.365	0.520
Credit = 0	-3.075	0.004**
Savings account = 0	0.512	0.285
Checking account = 0	-2.959	0.000***
Urban properties = 0	-1.991	0.000***
Other rural properties = 0	-1.156	0.035*
Paid farm workers = 0	-0.039	0.858
Off-farm work = 0	3.418	0.000***
Syndicate = 0	1.729	0.000***
Cooperative = 0	2.019	0.009**
Extracts forest products = 0	-0.985	0.002**
Mutual help group = 0	0.148	0.664
Interest in planting native tree species= 0	-0.085	0.624
Area of the lot in alqueires	0.072	0.000***
Percentage area in annuals	0.011	0.553
Percentage area in forest	0.033	0.005**
Number of people living in the lot	0.238	0.001**
Number of people living in the lot * Total cattle owned now	-0.001	0.022*
Total cattle owned now	0.017	0.010*
F1 dependency ratio	0.106	0.605
Initial -2 log likelihood	998.615	
Final -2 log likelihood	396.166	
Nagelkerke Pseudo R-Square	0.20	

Table 5.14. Results from the multinomial logit regression model.
Likelihood contributions of independent variables.

Dependent variable: LC of the pixel (PF, NF, SF) / Independent variables:	**Likelihood contribution of independent variables**	**Sig. Chi square test Likelihood Ratio Tests**
Municipio (1,2,3)	414.586	.000***
Owner Type (1,2,3)	418.370	.000***
Multifamily household (0/1)	397.691	.466
Land title (0/1)	403.388	.027*
Planted native tree species (0/1)	418.257	.000***
Property type (1,2,3)	399.434	.514
Credit (0/1)	405.274	.011*
Savings account (0/1)	397.748	.453
Checking account (0/1)	431.416	.000***
Urban properties (0/1)	453.027	.000***
Other rural properties (0,1)	402.284	.047*
Paid farm workers (0/1)	397.460	.523
Off-farm work (0/1)	443.390	.000***
Syndicate (0/1)	431.634	.000***
Cooperative (0/1)	407.920	.003**
Extracts forest products (0/1)	410.197	.001**
Mutual help group (0/1)	396.524	.836
Interest in planting native tree species (0/1)	399.295	.209
Area of the lot in alqueires	419.800	.000***
Percentage area in annuals	397.648	.477
Percentage area in forest	408.525	.002**
Number of people living in the lot	410.488	.001**
Number of people living in the lot * Total cattle owned now	403.311	.028*
Total cattle owned now	405.003	.012*
F1 dependency ratio	396.542	.828
Initial -2 log likelihood	998.615	
Final -2 log likelihood	396.166	
Nagelkerke Pseudo R-Square	0.20	

The significance of the B coefficients in table 5.13 is related with the significance of the chi-square statistic of table 5.14. Variables that show significance in one table also show significance in the other table. This happens because the -2-log likelihood is a surrogate measure of the R square.

In multinomial logit models, the -2-Log likelihood is a measure of the models' capacity to explain the variability of the dependent variable. In other words, it is a similar measure to the R square. However, the -2 log likelihood does not follow a scale from 0-1 and cannot be directly compared to an R square. That is why statistical packages provide with a pseudo R square.

The initial -2-log likelihood showed in tables 5.13 and 5.14 is the total variability in the values of the dependent variable. The final -2-log likelihood is the part of the variability not explained by the model after all the independent variables have been added in the model in a step-wise fashion. Therefore, this measure is equivalent to the sum of residuals in a ordinary least square model.

The SPSS program for multinomial logit models works with iterations of "educated guesses" assigning values to the category of the dependent variable. Iteration zero has no independent variables and thus the best possible guess for predicting the dependent variable is made by assigning the actual distribution of categories of the dependent variable. Appendix F shows the SPSS output and the initial distribution of land cover categories in the pixels: 36% are NF, 24% are SF and 40% are PF. The initial -2-log likelihood is obtained from iteration zero.

After this iteration, sets of variables are added (redundant variables are excluded by the model automatically) and a new -2-log likelihood is calculated until the parameter estimates converge.

The multinomial logit model converged in 6 iterations, improving the initial -2 Log likelihood from 998.615 to 396.166. Iteration zero consists of prediction of future LC when we use only the knowledge of present LC of the pixel, as it will be explained later in this section. Therefore, the inclusion of the survey variables helped to explain about two thirds of the -2 log likelihood (602.449), showing that these variables improved our estimation of the future LCC when compared with iteration zero.

In table 5.14, when looking at the Likelihood contribution of each independent variable, and its corresponding chi-square significance, we can infer which variables contribute more to the explanation and prediction of the LC classification of the pixels. The chi-square statistic for a given variable is the difference in -2 log-likelihoods between the final model (the one including all variables listed in the table) and a reduced model (the one omitting the given variable). Therefore the values in the column labeled “likelihood contribution of each independent variable” tell us how much this variable contributes to explain LC classification of the pixels. Then, looking at the chi-square statistic we can tell if this variable provides a statistically significant contribution at the 0.050, 0.010 or 0.001 levels.

In table 5.13, the “significance” column indicates if the coefficient estimated for each parameter or variable is significantly contributing to the model. Thus, this “significance” column is equivalent in meaning to the chi-square statistic of table 5.14. Again, the contribution of a variable to the whole model can be significant at the 0.050, 0.010 or 0.001 levels.

The main difference between tables 5.13 and 5.14 is that the first measures the contribution of each variable to the model's capacity to explain the variability of the dependent variable. These contributions lack meaning unless we know what the initial -2 log-likelihood is. On other hand, in table 5.13 the B coefficients tell directly for each independent variable what is its contribution to the prediction of the pixel's LC. This can be a positive or negative contribution to the probability of the pixel's LC, where the probability value moves between zero and one. For example, the variable, percentage area in annuals has a coefficient of 0.011 which means that an extra percent of area in annuals adds 0.011 to the probability to predict the LCC of the pixel. When the variable "interest in planting native tree species" is equal to "no" the coefficient is -0.085 which means that not having interest in planting native trees decreases to 0.085 the probability to predict the LCC of the pixel. Very large negative or positive numbers mean that the contribution of this variable is extremely important and it may even suffice to provide an accurate guess of the pixel's LCC.

Of course, the sign and value of the B coefficient for a given variable is irrelevant if this coefficient is not significant at least at the 0.050 value (under the column "significance"). Therefore, this section will discuss only the variables that proved to significantly contribute to the model or otherwise to the hypotheses of this dissertation.

First, of the variables credit, checking account and savings account, only the first two proved to be significant in explaining and predicting the pixels' land cover (LC). The absence of credit had a coefficient of -3.075 and the absence of checking account had a coefficient of -2.959, which means both variables highly reduce the probability of land cover change (LCC). Both coefficients were significant at the 0.001 level. This makes sense given that the capital from checking and credit accounts is readily available and can be used for economic entrepreneurship activities –cattle being the most lucrative one. Since cattle-ranching requires vast areas of pasture, it is likely that farmers will clear land to pursue or expand this economic activity. Regarding savings accounts, few farmers have extra money after paying for farm and living expenses and thus, few farmers can have savings accounts. Furthermore, in the event small farmers happen to have a bit of extra money they may prefer to invest it in cattle since then their savings are "visible and growing" and can be cashed at any time by selling the cattle.

Second, for the variables "ownership of other rural properties" and "ownership of other urban properties", both proved to be significant in the multinomial logit model. The absence of other urban properties had a coefficient of -1.991 (significant at the 0.001 level) and the absence of other rural properties had a coefficient of -1.156, which means both variables highly reduce the probability of land cover change (LCC). A possible explanation is that farmers are more financially secure due to other properties, not only for the available land to pursue other economic activities, but also for the monetary value that the properties represent. Thus, farmers can engage in economic activities of high investment and return, such as cattle, which linkage to clearing and deforestation was just explained above.

Third, of the variables social participation in syndicates, cooperatives and mutual help groups, only the first two proved to be significant in explaining and predicting the pixels' land cover (LC). The absence of syndicate participation had a coefficient of 1.729 (significant at the 0.001 level) and the absence of cooperative participation had a coefficient of 2.019 (significant at the 0.010 level), which means both variables highly increase the probability of land cover change (LCC). This could be explained due to the restrictions or group agreements that people in the syndicate may have to accept to be part of the group. If the syndicate or cooperative is focused on agro-forestry, forest conservation or alternative eco-friendly production then, it may be required for all members to follow strategies that pursue land conservation. In order to get more insight about these two variables, further research should be pursued about the nature of each syndicate and cooperative.

Fourth, for the variables "paid farm workers" and "income from off-farm work", only the latter one proved to be significant in the multinomial logit model. The absence of off-farm workers had a coefficient of 3.418 (significant at the 0.001 level).This means that family with no remittances from off-farm work will have a highly increased probability of land cover change (LCC). This phenomenon can be explained in the following way. When people work in off-farm locations some family members may be sending remittances back to the farm and affecting to some degree LCC decisions and investments in the land. Since, remittances from off-farm work are money that can be used to pay for the farm and living expenses, absence of remittances forces the farmer to pursue any economic activity in the farm to ensure a living.

Fifth, of the variables “number of people living in the lot”, “multiple families living in the lot”, and “dependency ratio” only the first one proved to be significant in explaining and predicting the pixels’ land cover (LC). The variable “number of people living in the lot” had a coefficient of 0.238 and was significant at the 0.010 level. This means that one extra person living in the lot will increase the probability of LCC by 0.238. The presence of multiple families in the lot was not a significant variable, which tells that what really matters is the number of people in the lot and not the kin relationships among them.

Sixth, the variable “definitive title” proved to be significant in explaining and predicting the pixels’ land cover (LC). The absence of definitive land title had a coefficient of 0.739 (significant at the 0.050 level), which means that absence of title increases the probability of land cover change (LCC). As mentioned before in the literature review, research shows that land title has produced ambivalent results in its effects on deforestation. In this case, farmers may be clearing land due to the pressure to prove land occupation to obtain a land title or they are just trying to get the most benefit of the land before they move to another place.

Eighth, for the variables “extraction of forest products”, “planted native tree species”, and “interest in planting native-tree species”, only the first two variables proved to be significant in explaining and predicting the pixels’ land cover (LC). The absence of forest extraction had a coefficient of 0.985 (significant at the 0.010 level). This means that when farmers are not engaged in forest extraction activities, then the probability for land cover change decreases. This In other words, if a farmer is extracting either timber or other forest products, he is more likely to make land cover change decisions. This can be explained given that the income from forest extraction may create an incentive to clear forest to make a living. The variable “non-planting native tree species” had a coefficient of 1.169 (significant at the 0.001 level) while the variable “interest in planting native tree species showed no significance even at the 0.050 level. The contrast between these two variables corroborates the difference between individuals’ intentions and their actual decisions. In this case, interest in forest conservation is better expressed by actual actions taken by planting native tree species. If farmers plant native tree species they are more likely to shoe their commitment to forest conservation in other land –cover decisions they make.

Ninth, the variables “total cattle owned” and the interaction between variables “number of people living in the lot * total cattle owned” proved to be significant in explaining and predicting the pixels’ land cover (LC). An extra head of cattle will increase the probability of LCC by 0.017 (significant at the 0.050 level) while a decimal increment in the ratio “number of people living in the lot* total cattle owned” will reduce the probability of LCC by 0.001 (significant at the 0.050 level). Such patterns can be explained as follows. The larger the amount of cattle owned, the larger the incentive to convert forest into pasture. The use of land for cattle ranching affects also the LCC decisions as we can infer from the significant value of the variable total cattle. This result agrees with literature, suggesting that even small farmers have increased their herds and cleared land and are moving in the direction of large cattle ranching activities (Walker, Moran, and Anselin, 2000).The interaction between the number of people living in the lot and the total cattle owned provides insight about the labor used for cattle or annual purposes. Thus, this explains why an increment in the number of people with respect to the number of cattle may produce a small decrement in land-cover change. That extra people may be working on keeping the area on annual crops productive –keeping in mind that cattle ranching requires a lower amount of labor. It is important to clarify that a farmer may have not only his own cattle in his property, but also third-party cattle. Unfortunately, the 1992 survey did not include specific questions in this regard, and only the 2002 disaggregated data is available.

Tenth, of the variables "area of the lot in alqueires", "percentage area in forest", "percentage of area in annuals", only the first two variables proved to be significant in explaining and predicting the pixels' land cover (LC). An extra alqueire in the area of the lot increases the probability of LCC by 0.072 (significant at the 0.001 level), while an increment in the percentage area in forest increases the probability of LCC by 0.033 (significant at the 0.010 level). Increment in the area of the lot correlate with higher probabilities of LCC (including deforestation) because more land availability can be translated into more possibilities for economic endeavors. Farmers can use part of the land for annuals, another part for perennials and other for cattle; or the whole land can be used for cattle ranching. The main idea is that vast amount of land provides security and more income opportunities for farmers. Given a large amount of land available, the main limiting conditions on the amount of clearing that can be done are: available labor for forest clearing and available land for cattle investment or other economic investment. On other hand, the percentage of area in annuals tells us how relevant is this economic activity for the farmer's family in terms of food and income production. The coefficient of this variable was not significant meaning that the percentage area in annuals and in general annuals cropping are not a key factor when making land-cover change decisions. Finally, the more forest available in the land the higher the probability for land-cover change. This could be explained by the fact that forest is still seen by farmers as a source of income of economic activities, either directly related to the forest products or indirectly related to other activities that can be done after forest clearing.

Finally, the municipio, owner typology and property typology were added to the model to see which typology has significant effect and greater effect when explaining and predicting the final land-cover of the pixel. The municipio and owner typology proved to be significant at the 0.050 level, while the property typology was not significant. Owners type 1 had the coefficient with the highest significance (0.001 level).

Municipio and owner type proved to significantly contribute to the explanation and prediction of LC, while property type does not. Let's recall that the owner typology is based in the area of total landholdings, while the property typology is based in the area of the primary lot under study. These figures indicate that total landholdings are more relevant than size of the studied property in explaining LCC in such property. Total landholdings may affect more significantly LCC in the primary lot because land in other rural or urban areas is not only available land for other economic activities (e.g. cattle ranching, agriculture, timber extraction), but it is also an economic asset by itself. Land speculation encourages small farmers with limited economic success to sell their properties, while successful farmers expand their profitable economic activities and landholdings. Table 5.13 indicates that possession of both rural and urban properties significantly affect LCC in the primary lot.

In a broader context, the implication of this result affects how small farmers are classified and judged as a whole homogeneous group by many social researchers and modelers. The "poolability" of small farmers is a common generalization and assumption in many models. The results suggest that small farmers' LCC cannot be accurately depicted and their behavior and clearing practices cannot be judged and modeled when farmers are judged solely on the size of the farm under study. This finding corroborates conclusions from the LSDV1 models that show differences among farmers' LCC probabilities when panel data probabilities were used in the calculation.

We also see a significant interaction between the size of the cattle herd and the number of people living in the lot. This interaction can be explained by the low demand of labor imposed by cattle ranching activities. This is favorable for aging households or for those seeking to free labor for off-farm work. (Walker, Moran, and Anselin, 2000). The significance of cattle ranching activities in an originally agricultural frontier reflects also the evolving nature of small farmers toward cattle ranchers in Rondônia.

From the case processing summary in the SPSS output (see Table 5.15) we find that the distribution of cells among the three possible land cover classes is: NF (36%), SF (24.1%), PF (40%). These values come from Katie Budreski's CART pixel classification and reflect the actual LC in year 1992, which is also the year of the survey data used here. Therefore, by looking at this descriptive data of LCC we have an idea of the probability of any pixel in the sample to be in the PF, SF, NF classes. If I take one pixel randomly from this sample and try to guess its cover class, it will most likely be PF, according to the figures from Table 5.13. If I assign the PF classification as my best guess for all pixels, I will have a success rate of 40%. Thus, my goal while using a predictive model is to improve this success rate.

Table 5.15 Actual distribution of cells among land covers in year 1992

		N	Marginal Percentage
1= NF , 2=SF 3=PF 0= nothing	1.00	1115	36.0%
	2.00	746	24.1%
	3.00	1239	40.0%

Although the pseudo R square from Table 5.13 is quite low (0.20) we find that the percentage of correct LC classification is 52.4% (Table 5.16.) when using the proposed multinomial logit regression model. This means an improvement of 12.4% in the percentage of correct cells, showing that the survey variables indeed improve our capability to predict LCC when compared to the iteration zero scenario of the multinomial regression. This percentage of accuracy will be compared with the modeling approach using LCC probabilities panel data and Markov chains.

Table 5.16. Percentage of correct predictions for the multinomial logit model

Classification				
	Predicted			
Observed	1.00	2.00	3.00	Percent Correct
1.00	637	74	404	57.1%
2.00	234	152	360	20.4%
3.00	330	74	835	67.4%
Overall Percentage	38.7%	9.7%	51.6%	52.4%

5.4.2 Modeling approach: explaining and predicting LC using only past LCC.

This section uses the results from the panel data analysis (sections 5.3.1 and 5.3.2) in a modified Markov chain. Basically, the predicted probabilities for the 9 LCC transitions are used in a matrix multiplication fashion to obtain a new matrix of 3 x3 with the LCC probabilities of LC in the next stage in time. In section TTTT 14 LSDV1 models were run for the 7 feasible LCC, 7 using the group effects model and 7 using the time effects model. A pooled model disregarding differences among subjects and across time was run as well. Predicted probabilities from the LSDV1-group, LSDV1-farmer and OLS-pooled models were separately used in the Markov matrix multiplication.

Going back to the hypothesis for research objective 3, we should compare results from sections 5.3.1 and 5.3.2. Hypothesis 3 states that the model using household survey data has more explanatory and predictive power than the model using only historical land cover changes. Thus, predictive accuracy of the 3 models, measured by the percentage of cells for which LC was predicted correctly, is compared and discussed in this section.

Since the LSDV1 models and the OLS pooled model were explained in detail in previous sections (sections 5.3.1 and 5.3.2), this section focuses only on the Markov matrix multiplication and on the accuracy of the predictions. Table 5.17 summarize the results from the LSDV1-farmer, LSDV1-year and OLS pooled models.

Table 5.17 Summary of results for the LSDV1-farmer, LSDV1-year and the OLS-pooled models.

	OLS pooled	**LSDV1-farmer**	**LSDV1-year**
Dependent variable:	PF-NF	PF-NF	PF-NF
Independent variables:	PF-PF, SF-NF, NF-NF, PF-SF, SF-SF, NF-SF	PF-PF, SF-NF, NF-NF, PF-SF, SF-SF, NF-SF, d1-d69	PF-PF, SF-NF, NF-NF, PF-SF, SF-SF, NF-SF, yr1-yr4
Adjusted R square	0.3424	0.5847	0.371
Durbin Watson stats	1.937 (no autocorrelation)	1.952 (no autocorrelation)	2.01 (no autocorrelation)
Coefficients are significant for the variables:	PF-PF	PF-PF, PF-SF, SF-SF, all farmer dummies d1-d69	PF-PF

Regressions were run using probabilities of past LCC for the years 1992 to 1995 and this time period was then assumed as the new time interval or stage.

Each of these models generated predicted probabilities for the 9 LCC transitions in Table 5.18.

Table 5.18 Matrix of LCC transitions.

		Final land cover		
		PF	SF	NF
Initial land cover	PF	**PF →PF**	**PF →SF**	**PF →NF**
	SF	**SF →PF**	**SF →SF**	**SF →NF**
	NF	**NF →PF**	**NF →SF**	**NF →NF**

NOTE:Since there are three feasible land cover classifications studied in this research, the initial land cover and the final land cover have to fall in one of the three categories PF (Primary Forest), SF (Secondary Forest or re-growth) and NF (non forest). The nine possible combinations or land cover change transitions are highlighted.

A pixel can be classified only in one land cover class, following the quality of the Markov chains where we can classify the system at any given time (stage) in only one of the feasible states. Thus, a final land cover class was selected for every pixel, choosing the one with the highest probability in each of the rows of Table 5.18. This predicted LC class was then compared against the CART pixel classification developed by Katie Budreski. The accuracy of each model, represented by the percentage of correct cells is shown in Table 5.19.

Table 5.19. Accuracy of the LSDV1 and OLS models used in conjunction with a Markov chain model

	OLS pooled + Markov	**LSDV1-farmer + Markov**	**LSDV1-year + Markov**
% cells correct	57.1811	59.2531	57.5435
Adjusted R square from LSDV1 alone	0.3424	0.5847	0.371

This table shows the % accuracy in the prediction of the pixel LC in 2002, when the results of the matrix multiplication were compared against actual LC class in that year.

From the figures in this table we can conclude that the model that accounts for farmers differences provides a slightly higher accuracy than the model that account for time difference and the model that simply used all the data in a single pool. Although the difference in the R square of the models is relatively big (0.5847 is a good R square in social science research, and very high when compared with 0.37 and 0.34), the percentage of correct predicted land cover at the pixel level shows a small difference. This could be explained by the fact that in Markov chain theory, the LCC transition probabilities approach a stable value (limiting probabilities) when the system is approaching equilibrium.

If we take into account time or farmer differences we arrive to a predictive accuracy that is about the same; and very similar in value to the accuracy of the case when all differences are disregarded . This could be explained by the fact that the frontier is reaching an equilibrium state, forced by the 50% rule that required farmers to keep 50% of their land as primary forest. This rule has been updated requiring now to keep 80% of the land as primary forest. The Brazilian government has limited resources to enforce this requirement and some farms are almost completely deforested. However, the 50% rule may still put some pressure on farmers as we see from Table 5.15, the actual percentage of land in primary forest is 40%. Such percentage may also reflect the tendency of farmers to overestimate their amount of forest and this percentage may have been intended to be 50%. We should also consider the area contribution of secondary forest, which is a hard class to identify not only by farmers but also by satellite imagery.

CHAPTER 6: CONCLUSIONS AND FUTURE RESEARCH

The main goal of this dissertation was to predict land cover change (LCC) firstly using only knowledge of past LCC and secondly using demographic, socioeconomic and land cover data at the farm level. The purpose was to compare the explanatory and predictive accuracy of both approaches, while testing two commonly used assumptions in models of LCC. The first assumption considers all small farmers homogeneous regarding their probabilities of LCC. Thus, small farmers are commonly regarded as a single group for modeling purposes. The second assumption refers to temporal changes in the probabilities of LCC. Relevant results and conclusions related to these research goals are discussed below.

6.1 MAIN RESULTS AND CONCLUSIONS WITH RESPECT TO THE RESEARCH GOALS

6.1.1 Homogeneity of subjects.

Regarding the homogeneity of subjects -- in this case small farmers --, several analyses were performed: descriptive statistical analysis, panel data analysis (composed of fixed time and subject effects models) and a multinomial logit regression. All of them arrive to the same general conclusion that small farmers cannot be regarded as a homogenous group with regard to their probabilities for LCC transitions and their characteristic variables. The owner typology explains better the farmer's different patterns in forest preservation. A brief summary of the results obtained from the statistical analysis performed, and the conclusions derived are discussed in this section.

Firstly, the descriptive statistical analysis of the estimated empirical transition probabilities can be discussed in the context of the Markov chain theory. At first glance, the 10-year average transition probabilities for the pathway sample (n=70) show high probabilities for the persistence events (PF-PF, SF-SF, NF-NF). However, these probabilities show that PF is a transient state and SF and NF are recurrent states. Then, Markov theory predicts that with certitude at some future time all primary forest will disappear and that this will be an irreversible process, leaving only hope for re-growth and conversions to secondary forest. We can get a sense of this process if we look at the cumulative effect of the probabilities. Every year approximately 10% of the primary forest is lost (into SF or NF) and it is never replaced. Every year, 35% of the land in SF becomes NF; and 95% of the land in NF will remain as NF for the next year. In a year, only 5% of the land in NF will allow for re-growth into SF. If we look at the cumulative process over a number of years it is easy to see why PF is disappearing due to the deforestation process. Most of previous predictive models assume NF is an absorbing state and carry out this assumption without further proof. In the present dissertation, the capability to estimate empirical transition probabilities allowed us to see that although small, there is a probability for NF-SF conversions. This probability can be increased if degraded land is allowed time to recover and if human interventions and technological improvements are used to regenerate forest. Some key conclusions can be drawn from this Markov theory analysis. First, conservation efforts have to focus on protecting primary forest because once it is gone it will be gone forever. Second, if current practices continue in the future, PF will certainly be eradicated. Third, increased efforts should target the NF-SF event, encouraging the re-growth process.

Tests of differences of means for the years 1992 and 2002 show that there are not statistically significant differences (even at the 0.050 level) among owner types at the beginning of the period of study when looking at the list of relevant land cover variables (table 5.3). From this we can conclude that, at the beginning of the study period, farmers were relatively homogeneous in terms of the lot area, percentage of area in forest, annuals, perennials and pasture. Differentiation among farmers, with respect to deforestation and land uses developed as their total landholdings change due to the land fragmentation/aggregation process as the analysis of differences of means in 2002 shows. The percentage area in annuals is a striking result, showing that from a group of small farmers that started with similar percentage area in annuals, the expanding owners are moving away from annual cropping. Expanding owners (OT3) decreased their percentage area in annuals (in the primary lot) from 6.44% in 1992 to 1.28% in 2002, while stable farmers (OT1) remained about the same with 5.8% and 4.23%, and subdividing farmers (OT2) increased their land in annuals from 6.07% to 9.16%. The Post-Hoc test also shows significant differences among owner types, for the area in non forest and the percentage area in annuals.

These differences among owner types can be explained by the land aggregation and subdividing patterns experienced in the agricultural frontier of Rondônia. This differentiation process is relevant to this dissertation because it proves that small farmers do not share all the same characteristics, success in the frontier, land use/ land-cover change practices and therefore probabilities for deforestation. If expanding and subdividing patterns are developing among small farmers, then the general trends small farmers/ large farmers are likely being replicated at a small scale inside the pool of small farmers. Therefore, predictive land-cover models that assume homogeneity among small farmers are missing details in the description of the system and accuracy in the predictions of the model.

Similar analyses of differences of means were performed for the same relevant LC variables using the property type, farmer type and municipio classification typologies. From the several typologies tested, the owner typology was the only one that showed significant differences in the LCC variables under study. We can conclude that changes in total landholdings (owner typology) are more relevant, in defining difference among small farmers in Rondônia, than changes in the size of the plot under study (property typology) or primary lot. This finding is very relevant not only for simulation purposes, but also for policy changes. In both instances, small farmers have being regarded and treated as a homogeneous group based solely in the size of their primary lot (in the state of Rondônia), overlooking the effect that other properties elsewhere can impose in the LCC decisions in the primary lot.

Several graphs were produced using the transitional LCC probabilities in order to find temporal patterns and patterns in the owner typology that may have escaped previous screening of data. A clear pattern was finally uncovered by the owner typology showing that owners with stable properties tend to preserve forest with a much higher probability (0.9033) than owner with subdividing or expanding properties (probs. of 0.0013 and 0.0030). This trend could be explained since farmers that acquire more land do so often with the intent to pursue cattle ranching, an activity that requires large extensions of pasture and thus involves forest clearing. On other hand, farmers whose landholding are reducing through time are usually farmers that are not being successful in the agricultural frontier and have to sell land to pay for household and farming expenses. Such farmers may not have a big incentive to preserve primary forest since they may sell the rest of their land in a near future and move further into the frontier pristine forested area.

Secondly, panel data analysis was conducted to test for fixed-group effects. The Fixed Group Effects Model (FGEM) approach (specifically the LSDV1 technique) was used to test the assumption of homogeneous probabilities among subjects, a commonly used assumption in stochastic predictive models applied to small farmers' LCC decisions. Future probabilities of land –cover change were calculated as a function of a panel data set of past LCC probabilities. The rationale behind the set of panel data equations is that the probability of a given land cover change depends on probabilities of other land-cover changes occurring in the present or previous years, in the specific farm under study and in the other farms as well. Basically these functions depict the interdependence among the LCC probabilities, and the temporal patterns in the probabilities series.

The partial F test for fixed-group effects shows that the LSDV1 fixed group effects model is a better model than the OLS pooled one. We arrive at this conclusion based not only on the improved R square (from R^2 = 0.3424 of the OLS-pooled model to R^2 = 0.5847 of the LSDV1-farmer model), but also on the many additional significant coefficients. In fact all dummy variables' coefficients prove to be significant, which means that definitely there are differences among farmers. Furthermore, the poolability test by farmer (F-stats= 5.27) concludes as well that the LSDV1 model is preferred over the OLS-pooled model.

Thirdly, a multinomial logit regression model (or the "survey-data variables" approach as called in this dissertation) was performed to get insight about how survey variables at the farm level affect probabilities of land-cover change. The multinomial logit model converged in 6 iterations, improving the initial -2 Log likelihood from 998.615 to 396.166. Therefore, the inclusion of the survey variables helped to explain about two thirds of the -2 log likelihood (602.449), which is a measure of the models' capacity to explain the variability of the dependent variable. The Nagelkerke Pseudo R-square was equal to 0.20.

The following variables had statistically significant coefficients and thus with a significant contribution in explaining and predicting the pixels' land cover category: credit, checking account, ownership of other rural properties, ownership of other urban properties, social participation in syndicates, cooperatives, income from off-farm work, number of people living in the lot, multiple families living in the lot, definitive title, extraction of forest products, planted native tree species, total cattle owned, the interaction variable "number of people living in the lot * total cattle owned" (a ratio variable that proved inverse relationship with the probability of LCC), area of the lot in alqueires, percentage area in forest.

Several classifications were explored using the multinomial logit regression model where independent variables were composed of survey data at the farm level. The owner type, property type and municipio provided statistically significant contributions in explaining land cover class (Forest, Non-Forest, Secondary Forest). The municipio and owner typology proved to be significant at the 0.050 level, while the property typology was not significant even at the 0.050 level. Owners type 1 had the coefficient with the highest significance (0.001 level). Changes in the total area of landholdings proved a stronger influence in farmer's LCC decisions in their main property (primary lot) when compared to changes in the area of the primary lot. These findings reflect the land aggregation and subdivision processes experienced in the agricultural frontier of Rondônia, Brazil. When farmers have other rural or urban properties, they have additional area for agricultural and cattle ranching purposes, among other land demanding economic activities. Moreover, that additional land is an economic asset by itself that may facilitate farmers' access to credit in order to change their productive and clearing strategies in the primary lot. The presence of differences among municipios encourages further research regarding local land use and forest conservation plans and policies. In a broader context, the implication of this result affects how small farmers are classified and judged as a whole homogeneous group by many social researchers and modelers. The "poolability" of small farmers is a common generalization and assumption in many models. The results suggest that small farmers' LCC cannot be accurately depicted and their behavior and clearing practices cannot be judged and modeled when farmers are judged solely on the size of the farm under study. This finding corroborates conclusions from the LSDV1 models that show differences among farmers' LCC probabilities when panel data probabilities were used in the calculation.

6.1.2 Stationary probabilities

Regarding the assumption of stationary probabilities of land-cover change, several analysis were performed: descriptive statistical analysis and graphics, panel data analysis (composed of fixed time effects models) and a poolability test. All of them arrive to the same general conclusion that not fixed-time effects were found. In other words, we can analyze the data as a pooled sample of probabilities, disregarding temporal trends in the LCC probabilities since not significant temporal changes are observed. The LCC probabilities of a given farmer can be considered nearly stationary for that given farmer.

Analysis of differences of means was conducted among the 70-farmer average probabilities for each LCC transition on the 10 years of the period of study. The general conclusion is that transition probabilities of LCC do not significantly change over time, as opposed to the conclusion of previous section where probabilities change significantly among subjects.

There are no significant differences in the LCC probabilities across time when using a fixed time effects model. Panel data analysis of the LCC empirical transition probabilities (LSDV1 fixed time effects model) concludes that a Ordinary Least Square (OLS) pooled version of the probabilities can be chosen over a LSDV1-time model. The poolability test also indicates that a pooled model can be used without risk of missing temporal changes in the LCC probabilities, given that not trends were found.

After testing of the two modeling assumptions –homogeneous subjects and stationary probability—we can conclude that small framer's LCC probabilities show differences among subjects but not differences across time. This would mean that farmers' household and farm variables affect their LCC decisions, but they keep on the long run the LCC strategy that works for them.

In conclusion, there are significant differences among small farmers and they should not be regarded as a single group for modeling, planning and policy purposes related with land cover change. Results from the panel data analysis of the empirical LCC transition probabilities indicate that farmers are not homogeneous with respect to their probabilities of LCC. The presence of differences among farmers in the LSDV1-fixed group effect by farmer suggest that further differentiation or classification of farmers into homogeneous subgroups will depict better their LCC decisions.

6.1.3 Comparing the explanatory and predictive capacity of the "LCC only-variables" and the "survey-data variables" models.

When applying the results of the panel data analysis to a modified markov chain model the LSDV1-farmer model provides a slightly better accuracy than the LSDV1-time and the OLS-pooled models. This suggests again that taking into account farmers' differences will provide a better predictive model. However, the difference in the percentage of correct predicted land cover is so small (59.25%, compared with 57.54% and 57.18%, respectively) that we could still use any of the three models for LCC prediction based solely in past LCC probabilities. Thus, I would suggest using the OLS-pooled model when the objective of the planner is to identify areas at high risk of deforestation, without regard to the underlying causes. The OLS-pooled model will provide results very similar in accuracy to the panel data analysis of past LCC with reduced calculations.

However, if the objective of the planner is to identify areas at high risk of a specific LCC, and then to identify the variables that may be causing this change, then the choice will be the multinomial logit model, which has an accuracy of 52%. The multinomial logit model could be run using cross-sectional survey data and if panel survey data is available for all the years when we have past LCC data, then the a multinomial logit panel data will provide even better results.

6.2 Policy and planning implications

As a planner, I see simulation models as tools for planers' decision making and not as the "planning process" itself. Thus, the purpose of this research was not to predict with 100% accuracy what the real system does, which will be the same as reproducing the system and which requires a level of detail that goes beyond the simplifications made in models. Instead, the purpose is to gain insight about general trends, causes or processes in order to increase our knowledge of the process, which might better inform policies to manage deforestation processes.

The main finding for policy and planning purposes is that owners type 1 –with stable landholdings-- tend to preserve forest with a much higher probability (0.9033) than owners with subdividing or expanding properties (probs. of 0.0013 and 0.0030 respectively).

Some relevant findings that describe owner type 1 include: owners type 1 tend to be older than OT2 and OT3; OT1 tend to be the only owner that has possessed the property under study. More type 1 owners have definitive land title (75.6%), compared with 46.90% and 60.3% of OT2, and OT3 respectively. A larger number of families --and a corresponding larger total number of people -- live in farms owned by OT1. However, there is not significant difference in the composition of the primary family (F1) or in its dependency ratio.

In conclusion, if policy makers and planners want to encourage OT1 to continue their pattern of high forest conservation rates, efforts should focus on securing land titling, providing health care and alternative sources of income for the OT1's family members and elderly owners to remain in the lot. Moreover, the larger number of families and corresponding larger total number of people living in the lot represent labor force that should be encouraged to stay in the farm pursuing environmentally sustainable annual crops, perennials and small cattle ranching carried out in a balanced way. This balanced way includes pasture rotation, rotation of annual crops, agro-forestry projects and small and large cattle. All activities could be performed in designated areas of the farm allowing other areas for rotation and regeneration. The general purpose is to discourage new clearing while still providing sources of food and income for the large number of people living in the lot.

6.3 Contribution to research base

The contribution of this research is both, methodological and theoretical. Methodologically, the importance of this research lies in the estimation and analysis of empirical transition probabilities. To the best of my knowledge, probabilities per farmer, per year, per LCC class and their temporal trends in a panel sample have not been formally tested in previous research. Another contribution to simulation methology for deforestation is the testing of underlying assumptions about the transition probabilities through the comparison of predictive accuracy when probabilities calculated under different assumptions are used in a Markov chain model.

Theoretically, this work tests variables at the farm and household level that have been suggested by the literature as causes of LCC and deforestation in the Amazon. The rationale supporting the selection of these variables relies on frontier and deforestation theories. This dissertation focused on exploring the variables in the context of their contribution in a LCC predictive model. Testing of the demographics, household dynamics and other frontier theories have been amply studied by Amazon forest theorists and it is out of the scope of this research.

6.4. Future research

The findings in this dissertation research answered the research questions proposed initially at a conclusive or partial level. Results strongly support differences among farmers and encourage further research to explore random time effects. Both, partial and conclusive findings raised new questions that I want to leave for my future academic research or for new generations of modelers to come.

I propose the following questions for further exploration:

- How do local policies of land use and forest conservation efforts at the municipio level affect small farmers' LCC probabilities?
- What is the role of the exogenous variables in small farmer's LCC decision making once that they engage in large scale cattle ranching projects?
- How do pixel-related variables such as distance to road, water bodies and closest non-forest patch affect small farmer's LCC decision-making?
- How do neighboring farmers' decisions affect the LCC decisions of a given farmer?

- Is spatial autocorrelation linked in any fashion to temporal autocorrelation?
- How long does it take for each farm to become all non forest? Could we group farms that started with the same percentage of forest and ended up with zero percentage of NF, into clusters with same rates of deforestation?
- How can this model approach and its assumptions work for deforestation in urban settings?
- Would a second or third order Markov chain better predict LCC (futureland covers to be a function of not only the present land cover but also past land covers)?

BIBLIOGRAPHY

Aaviksoo, K., 1995. Simulating vegetation dynamics and land use in a mire landscape using a Markov model. Landscape and Urban Planning , 31, 129-142.

Abramovitz, J., 1998. Taking a stand: cultivating a new relationship with the world's forests. World Watch Institute, Washington, D.C.,

Ahn, T., Ostrom E., Gibson, C., 1998. Scaling issues in the social sciences. IHDP Working Paper No. 1, Bonn, 85 pp.

Almeida, A.L. Campari, J.S., 1995. Sustainable settlement in the Brazilian Amazon. Oxford University Press, New York.

Almeida, Oriana Trinidade de, Adalberto Veríssimo, Angélica Toniolo, Christopher Uhl, Marli Maria Mattos, Paulo Barreto, and Ricardo Tarifa. 1996. *A evolução da Fronteira* Amazônica (The evolution of the Amazon Frontier). Belém, Brazil: Instituto do Homem e Meio Ambiente da Amazonia.

Alves, D., 1999. An analysis of the geographical patterns of deforestation in Brazilian Amazon in the 1991-1996 period. In: Wood, C., Porro, R. (Eds), Patterns and Processes of Land Use Change and Forest Change in the Amazon. University of Florida Press, Gainesville.

Andersen, L.E., 1996. The causes od deforestation in the Brazilian Amazon. J. Environ. Dev. 5, 309-328.

Andersen, L.E., Reis E.J., 1997. Deforestation, development and government policy in the Brazilian Amazon: An Econometric Analysis. Working Paper 513. Instituto de Pesquisa Econômica Applicada (IPEA), Rio de Janeiro.

Anderson, M.C., 1966. Ecological groupings of plants. *Nature* 212, 54-56

Angelsen A., Kaimowitz D., 1999. Rethinking the causes of deforestation: lessons from economic models. *The World Bank Research Observer*, 14 (1) 73–98.

Angelsen, A., 1994. Shifting cultivation expansion and intensity of production: the open economy case. Working paper 3. Christian Michelsen Institute, Fantoft, Bergen.

Annon, 1998b. Brazil's rainforest plan under fire. *Associated Press*, 25 November 1998.

Anon.: 1997c, Malaysian loggers in the Amazon. Action Alert, Rainforest Action Network, 24 February 1997.

Anon.: 1998a, NGOs press for more forest reserves. *Inter-Press Service,* 23 December 1998.

Aspinall, R., 2004. Modelling land use change with generalized linear models – a multi-model analysis of change between 1860 and 2000 in Gallatin Valley, Montana. Journal of Environmental Management 72, 91-103.

Babbie, E., 2001. The practice of social research, 9 edition

Baker, W.L., 1989. A review of models of landscape change. *Landscape Ecology* 2(2) 111-133

Barbier, E.B. 2000. Links between economic liberalization and rural resource degradation in the developing regions. *Agric. Econ.* 23(3): 299-310

Barbier, E.B., 1997. The economic determinants of land degradation in developing countries. Philos. Trans. R. Soc. Ser. B. 352, 891-899

Barbier, E.B., 2000. Links between economic liberalization and rural resource degradation in the developing regions. Agricultural Economics 23, 229-310.

Barraclough, S.L., Ghimire, K.B., 2000. Agricultural expansion and tropical deforestation: Poverty, international trade and Land use. Earthscan Publications, London.

Berry, M., Hazen, B., MacIntyre, R.L., Flamm, R., 1996. LUCAS: a system for modeling land-use change. IEEE Comput. Sci. Eng. 3, 24-35.

Bhat, U.N., 1984. Elements of Applied Stochastic Processes., Wiley, New York.

Binswanger H.P., 1991. Brazilian Policies that encourage deforestation in the Amazon. World Development 19 (7), 821-829

Blaikie, P., Brookfield, H. 1987. *Land degradation and society*. London: Methuen. 296 pp.

Bockstael, N.E., 1996. Modeling economics and ecology: the importance of a spatial perspective. Am. J. Agr. Econ. 78, 1168-1180.

BRA/87/037 1989 Cooperação Técnica na Implementação do Programa Integrado de Desenvolvimento POLONOROESTE. O Problema da Ecologia Regional no Desenvolvimento Rural e forestall dos Estados de Mato Grosso e Rondônia. Documento de Trabalho No. 5 Brasília, SUDEO, MINTER, FAO, PNUD.

Browder, 1996. Reading colonists landscapes: social interpretations of tropical forest patches in an Amazonian agriculture frontier. In: Schelhas J., Greenberg, R. (Eds.) Forest Patches in Tropical Landscapes. Island Press, Washington, D.C., pp. 285-299.

Browder, J. 1988. Public Policy and deforestation in the Brazilian Amazon. In: Repetto R. and Gillis. (eds), Public policies and the misuse of forest resources. Cambridge University Press, Cambridge, UK, pp. 247-297

Browder, J. O. 1988. Public policy and deforestation in the Brazilian Amazon. Pp. 247-298 in *Public Policies and the Misuse of Forest Resources*, edited by R. R. a. M. Gillis. Cambridge: Cambridge University Press.

Brown J.C., 2001. Responding to Deforestation: Productive Conservation, the World Bank, and Beekeeping in Rondônia, Brazil. *The Professional Geographer*, 53(1) 106-118(13)

Brown J.C., 2001. Responding to Deforestation: Productive Conservation, the World Bank, and Beekeeping in Rondônia, Brazil. *The Professional Geographer*, 53(1) 106-118(13)

Brown, Katrina, and David Pearce, eds. 1994. *The Causes of Tropical Deforestation.* Vancouver: University of British Columbia Press.

Brown, L.A., 1970. On the use of Markov chains in movement research. *Econ. Geogr.* 46, 393-403.

Bunker, Stephen G. 1984. "Modes of extraction, unequal exchange, and the progressive underdevelopment of an extreme periphery: The Brazilian Amazon, 1600-1980. *American journal of Sociology* 89:1017-64

Buschbacher, R., Uhl, C., Serrao, E.A.S., 1988. Abandoned pastures in eastern Amazonia. II. Nutrient stocks in the soil and vegetation. *Journal of Ecology*, 76 , 682-699

Buschbacher, R., Uhl, C., Serrao, E.A.S., 1988. Abandoned pastures in eastern Amazonia. II. Nutrient stocks in the soil and vegetation. *Journal of Ecology*, 76 , 682-699

Campbell, D. T., Stanley J. C., 1963. Experimental and quasi-experimental design for research. Houghton Mifflin Co: Boston.

Chomitz, K, Gray, D., 1996. Roads, land use and deforestation: a spatial model applied to Belize. World Bank Econ. Rev. 10, 487-512

Clarke, A. B., Disney, R., 1970. Probability and random processes for engineers and scientists. Wiley, New York.

Collins, L., 1975. An introduction to Markov chain analysis. Concepts and techniques in Modern Geography. Geo Abstracts, University of East Anglia, Norwich.

Coy, M. 1986. Realidade atropela os planos de POLONOROESTE. Pau Brazil (11, Ano II): 67-73.

Coy, M. 1987 Rondônia: Frente pioneira e programa Polonoroeste. O processo de diferenciação sócio-econômica na periferia e os limites do planejamento público. Tubinguen Geogrsphhische Studien 95:253-270.

Coy, M., 1987. Rondônia: Frente Pioneira e Programa POLONOROESTE: O Processo de Diferenciação Sócio-Econômica na Periferia e os Limites do Planejamento Publico. In:

Kohlepp, G., Schraeder, A. (eds). Homem e Natureza na Amazônia. Simpósio Internacional e Interdisciplinar, Blauberen, 1986. ADLAF, Tübingen Geographische Studien, 95, 253-270.

Cultural Survival, Inc. 1981 *In the path of POLONOROESTE: Endangered peoples of Western Brazil.* Cambridge, M.A.: Cultural Survival Inc.

Dale V.H., O'Neill, R.V., Southworth, F., Pedlowski, M., 1994. Modeling effects of land management in the Brazilian Amazonian settlement of Rôndonia. *Conservation biology* , 8 (1), 196-206

Dale, V.H., Houghton, R.A., Hall, C.A.S., 1991. Estimating the effects of land use change on global atmospheric CO_2 concentrations. *Canadian Journal of Forestry Research*, 21, 87-90.

Dale, V.H., O'Neill, R.V., Southworth, F., Pedlowski, M, A. 1994. Modeling effects of land management in the Brazilian settlement of Rondônia. Conservation Biology 8 (1) 196-206.

Dale, V.H., Pearson S.M. 1997, In: W.F. Laurance and R.O. Bierregaard, Jr., Eds., *Tropical Forest Remnants: Ecology, Management, and Conservation of Fragmented Communities.* University of Chicago Press, Chicago, pp. 400-409.

DeVaus, David, 2002. "Constructing questionnaires" in surveys in Social Research. Fifth Edition. Routledge: Taylor and Francis. Pp.94-118.

Duffy, S.B., Corson, M.S., Grant, W.E., 2001. Simulating land-use decisions in the La Amistad Biosphere Reserve buffer zone in Costa Rica and Panama. *Ecological Modelling*, 140 (1-2) 9-29

Duffy, S.B., Corson, M.S., Grant, W.E., 2001. Simulating land-use decisions in the La Amistad Biosphere Reserve buffer zone in Costa Rica and Panama. *Ecological Modelling*, 140 (1-2) 9-29

Duncan, E.M.J., Mcgregor, D.F.M., Vieira N.A.Q. 1990 Pasture development in the cleared forest land in northern Amazonia. The Geographical Journal 15:283-296.

Evans T. P., Manire, A., De Castro, F., Brondizio, E., McCracken, S., 2001. A dynamic model of household decision-making and parcel level land cover change in the eastern Amazon. *Ecological Modeling*, 143 (1-2), 95-11

FAO-CP. 1987. Brazil: Northwest I, II and III technical review: final report. Rome: Food and Agriculture Organization Cooperative Program.

Fearnside P.M, 2001. Land tenure issues as factors in environmental destruction in Brazilian Amazonia: the case of southern Para. World Development Vol. 29 N. 8 pp.1361-1372

Fearnside P.M. 1999 Forests and global warming mitigation in Brazil: opportunities in the Brazilian forest sector for responses to global warming under the clean development mechanism. *Biomass and Bioenergy*, 16(3) 171-189

Fearnside, P. M. 1996. Amazonian deforestation and global warming: carbon stocks in vegetation replacing Brazil's Amazon forest. Forest Ecology and Management 80, 21-34

Fearnside, P.M. 1987, *Conserv. Biol*. 1, 214-221

Fearnside, P.M. 1989. A Ocupação Humana de Rondônia: Impactos, Limites, e Planejamento. Programa do Tropico Unido POLONOROESTE. Relatório de Pesquisa. Brasília, CNPq.

Fearnside, P.M. 1990 In: A.B. Anderson, ed. *Alternatives to Deforestation: Towards Sustainable Use of the Amazonian Rain Forest.* Columbia University Press, New York, pp. 235-251

Fearnside, P.M. 1993, *Ciencia Hoje* 19, 6-8

Fearnside, P.M. 1995 *Environ. Conserv*. 22, 7-19

Fearnside, P.M., 1996. Amazonian deforestation and global warming: carbon stocks in vegetation replacing Brazil's Amazon forest. *Forest Ecology and Management* 80, 21-34.

Fearnside, P.M., 2001. Land-tenure issues as factors in environmental destruction in Brazilian Amazonia: the case of southern Pará. World Dev. 29 (8) 1361-1372

Fearnside, Philip M. 2000. Deforestation impacts, environmental services and the international community. Pp. 11-24 in Anthony Hall, ed. *Amazonia at the crossroads.* London: University of London Press.

Feller, W., 1968. An introduction to Probability Theory and its Applications. Vol. 1 3rd Ed. Wiley, New York.

Finn, J.T., 1985. Analysis of land use change statistics through the use of Markov chains. *In*: Proceedings of the 1985 Summer Computer Simulation Conference. p 514-519. The Society for Computer Simulation.

Flamm, R., Turner, M., 1994. Alternative model formulations for a stochastic simulation of landscape change. Landscape Ecol. 9, 37-46.

Food and Agricultural Organization (FAO). 1995 *Forest Resources Assessment 1990: Global Synthesis*, Rome: FAO.

Foresta, R. 1991. *The limits of providence: Amazon conservation in the age of development.* University of Florida Press, Gainesville, Florida.

Fox, J., Krummel, J., Yarnasarn, S., Ekasingh, M., Podger N. 1995. Land use and landscape dynamics in northern Thailand: assessing change in three upland watersheds. *Ambio* 24(6): 328-34

Frohn, R.C., Dale, V.H., Jimenez, B.D. 1990 Colonization, road development and deforestation in the Brazilian Amazon Basin of Rondônia. ORNL/TM-11470. Oak Ridge National Laboratory, Oak Ridge, Tennnessee.

Geoghegan et al, 2001. Modeling tropical deforestation in the Southern Yucatan peninsular region: comparing survey and satellite data. Agriculture, Ecosystems and Environment 85. 25-46.

Gerwing, J.J., 2002. Degradation of forest through logging and fire in the eastern Brazilian Amazon. Forest Ecology and Management 157, 131-141.

Ginsberg, R.B., 1972. Incorporating causal structure and exogenous information with probabilistic models: With special reference to gravity, migration and Markov chains. *J. Math. Soc.* 2, 83-103

Ginsberg, R.B., 1973. Stochastic models of residential and geographic mobility for heterogeneous populations. Envir. And Planning A. 5: 113-124

Godoy, R. et al. 1997.Household Determinants of deforestation by Amerindians in Honduras. World Development, vol 25, No. 6, pp. 977-987.

Gujarati, Damonar N. (2003) Basic Econometrics, 4th Edition. Boston: McGraw Hill.

Hageman, H. 1996. Bancos, Incendiários, e Florestas Tropicais (Banks, Fires and Tropical Forests). Rio de Janeiro: FASE/IBASE/Instituto Socioambiental.

Hall, Anthony. 1989. *Developing Amazonia: Deforestation and social conflicts in Brazil's Carajás programme*. Manchester, England; Manchester University Press.

Hall, Anthony. 1997. *Sustaining Amazonia: Grassroots Actions for Productive Conservation. Manchester*, England: Manchester University Press.

Harary, F., Lipstein, B., Styan, G.P.H., 1970. A matrix approach to nonstationary chains. *Operations Res.* 18, 1168-1181

Hazen, B.C., Berry, M.W., 1997. The simulation of land-cover change using a distributed computing environment. Simulat. Pract. Theory 5, 489-514.

Hecht S. and Cockburn A. 1989. The fate of the forest –developers, destroyers and defenders of the Amazon. Verso, London, UK, 357 pp.

Hecht S.B. 1989. The sacred cow in the Green Hell: livestock and forest conversion in the Brazilian Amazon. The Ecologist 19:229-234

Hecht, S.B. 1981. Deforestation in the Amazon Basin: Magnitude, dynamics and soil resource effects. Studies in Third World Societies, 13, 61-110

Hecht, S.B., and Cockburn, A. 1990. *Fate of the forest: Developers, destroyers and defenders of the Amazon.* New York: Harper Collins.

Holdsworth, A.R., Uhl, C., 1997. Fire in Amazonian selectively logged rain forest and the potential for fire reduction. *Ecol. Appl.* 7, 713-725.

Horn, H.S., 1975. Markovian processes of forest succession. In: Ecology and Evolution of Communities. pp 196-211. Edited by M.L. Cody and J.M. Diamond. Belknap Press, Cambridge, Mass.

Houghton, R.A., Hobbie, J.E., Melillo, J. M., Moore, B., Peterson, B.J., Shaver, G.R., Woodwell, G.M., 1983. Changes in the carbon content of terrestrial biota and soil between 1860 and 1980: Net release of CO_2 to the atmosphere. Ecological monograms, 53, 235-262.

Houghton, R.A., Hobbie, J.E., Melillo, J. M., Moore, B., Peterson, B.J., Shaver, G.R., Woodwell, G.M., 1983. Changes in the carbon content of terrestrial biota and soil between 1860 and 1980: Net release of CO_2 to the atmosphere. Ecological monograms, 53, 235-262.

Hsiao, C. 2003. Analysis of Panel Data. Cambridge University Press, second edition.

Hulst, R. van 1979. On the dynamics of vegetation: Markov chains as models of succession. *Vegetatio* 40, 3-14.

Hurrell, A. 1991. The Politics of Amazonian Deforestation. *Journal of Latin American Studies* 23: 197-215

INPE, 2001. Monitoramento de floresta Amazônica Brasileira por satélite 1999-2000. Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, Brazil.

INPE: 1996, Deforestation estimates in the Brazilian Amazon, 1991-1994. Instituto Nacional de Pesquisas Espaciais, San Jose dos Campos, Brazil.

INPE: 1998, Deforestation estimates in the Brazilian Amazon, 1995-1997. Instituto Nacional de Pesquisas Espaciais, San Jose dos Campos, Brazil.

Instituto Brasileiro de Geografia e Estatística (IBGE). 1979. Censo Agropecuário 1975 (Agricultural Census of Brazil 1975) Rio de Janeiro: IBGE

Instituto Brasileiro de Geografia e Estatística (IBGE). 1990. Censo Agropecuário 1985 (Agricultural Census of Brazil 1985) Rio de Janeiro: IBGE

Instituto Brasileiro de Geografia e Estatística (IBGE). 1990. Censo Agropecuário 1985 (Agricultural Census of Brazil 1985) Rio de Janeiro: IBGE

Instituto Brasileiro de Geografia e Estatística (IBGE). 1998a Censo Agropecuário 1995/96 (Agricultural Census of Brazil 1995/96). Available on-line at IBGE website: http://www.ibge.org

Instituto Nacional de Pesquisas Espaciais (INPE). 2000 *Desmatamento da Amazonia Brasileira,* 1998-1999 (Deforestation of the Brazilian Amazon, 1998-1999) Available on-line at INPE website: http://www.inpe.br/Informacoes_Eventos/amz1998_1999/pagina7.htm

Irwin, e. G., Geoghegan, J., 2001. Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems and the environment* (85) 7-23.

Jahan, S., 1986. The determination of stability and similarity of Markovian land use change processes. A theoretical and empirical analysis. Socio-Econ. Plan. Sci. 20, 243-251

Kaimowitz, David, and Arild Angelson. 1998. E*conomic Models of Tropical Deforestation: A Review.* Jakarta, Indonesia: Center for International Forestry Research.

Kainmowitz, D., Angelsen, A., 1998. economic models of tropical Deforestation: A review. Centre for International Forestry Research, Jakarta, 139 pp.

Kates, R.W., Haarmann V. 1992. Where the poors live: Are the assumptions correct? *Environment* 34(4): 4-28

Keck, M. and k. Sikkink. 1998. *Activists Beyond Borders: Advocacy Networks in International Politics.* Ithaca, N.Y. and London: Cornell University Press.

Kemeny, J.G., Snell, J.L., 1960. Finite Markov Chains. D. Van Nostrand, Princeton, New Jersey.

Kolk A., 1998. From Conflict to Cooperation: International Policies to Protect the Brazilian Amazon *World Development*, 26(8) 1481-1493(13)

Kolk, A. 1996. *Forests in International Environmental Politics: International Organisations, NGOs and the Brazilian Amazon.* Utrecht: international Books.

Kolk, Ans. 1996. *Forests in International Environmental Politics.* Utrecht, Netherlands: International.

LaGro, J.A., DeGloria, S.D., 1992. Land use dynamics within an urbanizing non-metropolitan county in New York state (USA). Landscape Ecol. 7, 275-289.

Lambin E.F., Geist, H. J., Lepers, E., 2003. Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources, (*28) 205-241.

Lambin, E.F., 1997. Modeling and monitoring land cover change processes in tropical regions. *Progress in Physical Geography* 21 (3) 375-393.

Lambin, E.F., Geist, H.J., Lepers, E, 2003. Dynamics of land-use and land-cover change in tropical regions. Annu. Rev. Environ. Resour. 28, 205-241

Lambin, E.F., Geist, H.J., N.D. Global land-use and cover change: What have we learned so far? *Global changes Newsletter* No. 46 27-30

Lambin, E.F., Rounsevell, M.D.A., Geist, H.J., 2000. Are agricultural land-use models able to predict changes in land-use intensity? *Agriculture, Ecosystems & Environment* 82 (1-3), 321-331.

Landis, J., 1995. Imagining land use futures: applying the California urban futures model. J. Am. Plann. Assoc. 61, 438-457

Landis, J., Zhang, M., 1998a. The second generation of the California urban futures model. Part 1: model logic and theory. Environ. Plann. A 25, 657-666.

Landis, J., Zhang, M., 1998b. The second generation of the California urban futures model. Part 2: specification and calibration results of the land-use change submodel. Environ. Plann. A 25, 795-824.

Laurance W.F., 2000. Mega-Development Trends in the Amazon: Implications for Global Change *Environmental Monitoring and Assessment*, 61(1), pp. 113-122(10)

Leonard H.J., 1989. Environment and the poor: development strategies for a common agenda. U.S.-Third World Policy Perspect. 11:3-45.

Lopez, E., Bocco, G., Mendoza, M., Duhau, E., 2001. Predicting land-cover and land-use change in the urban fringe. A case in Morelia city, Mexico. Landscape and Urban Planning, 55, 271-285.

Lovejoy, Thomas e. 2000. Amazonian forest degradation and fragmentation: Implications for biodiversity conservation. Pp. 41-57 in Anthony Hall, ed., *Amazonia at the Crossroads.* London: University of London Press.

Ludeke, A. K., Maggio, R.C., Reid, L.M., 1990. An analysis of anthropogenic deforestation using logistic regression and GIS. J. Environ. Manage. 32, 247-259

Machado, L. 1991. Frontera agricola en la Amazonia Brasilena. Pages 395-427 in V. I. Coloquio de Geografia Geral. Madrid, Spain.

Mahar, D.H., 1989. Government policies and deforestation in Brazil's Amazon Region. The World Bank, Washington, D.C., USA, 56 pp.

Majar, D.J. 2000. Agro-Ecological zoning in Rondônia, Brazil: what are the lessons? Pp. 115-28 in Anthony Hall ed., *Amazonia at the Crossroads.* London, University of London Press.

Massy, W.F. Montgomery, D.B. and Morrison, D.G. 1970. Stochastic models of buying behavior. M.I.T. Press, Cambridge, Mass.

Mertens B.; Poccard-Chapuis R.; Piketty M.-G.; Lacques A.-E.; Venturieri A. 2002. Crossing spatial analyses and livestock economics to understand deforestation processes in the Brazilian Amazon: the case of Sao Felix do Xingu in South Para. *Agricultural Economics*, 27(3) 269-294

Mertens B.; Poccard-Chapuis R.; Piketty M.-G.; Lacques A.-E.; Venturieri A. 2002. Crossing spatial analyses and livestock economics to understand deforestation processes in the Brazilian Amazon: the case of Sao Felix do Xingu in South Para. *Agricultural Economics*, 27(3) 269-294
Mertens, B., Lambin, E., 1997. Spatial modeling of deforestation in southern Cameroon. Appl. Geogr. 17, 143-162.

Miller, Thomas I. 1994. “Designing and conducting surveys” in Handbook of Practical Program Evaluation. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer, eds San Francisco: Jossey-Bass. Pp. 271-292.

Millikan, B. H. 1988. The dialects of devastation: Tropical deforestation, land degradation, and society in Rondônia, Brazil. M.A. thesis. University of California, Berkeley, California.

Millikan, Brent. 1992. Tropical deforestation, land degradation, and society: Lessons from Rondônia Brazil. *Latin American Perspectives*19: 45-72

Molofosky, J., Hall, C.A.S., Myers, N., 1985. A comparison of tropical forest surveys. U.S. Department of Energy, Carbon Dioxide Research Program TR032, Washington, D.C.

Moran, Emilio F. 1981. *Developing the Amazon.* Bloomington: Indiana University Press.

Moran, Emilio F. 1989 “Adaptation and Maladaptation in Newly Settled Areas.” Pp 20-41 in Debra A. Schumann and William L. Partridge, eds., *The Human Ecology of Tropical Land Settlement in Latin America.* Boulder, Colo.: Westview.

Myers, N., 1980. The present status and future prospects of the tropical moist forest. *Environmental conservation*, 7, 101-114.

Nelson, G., Hellerstein, D., 1997. Do roads cause deforestation? Using satellite images in econometric analysis of land use. Am. J. Agr. Econ. 79, 80-88

Nepstad D.; Carvalho G.; Cristina Barros A.; Alencar A.; Paulo Capobianco J.; Bishop J.; Moutinho P.; Lefebvre P.; Lopes Silva U.; Prins E. 2001. Road paving, fire regime feedbacks, and the future of Amazon forests. *Forest Ecology and Management*, 154(3) 395-407(13)

Nepstad D.; Carvalho G.; Cristina Barros A.; Alencar A.; Paulo Capobianco J.; Bishop J.; Moutinho P.; Lefebvre P.; Lopes Silva U.; Prins E. 2001. Road paving, fire regime feedbacks, and the future of Amazon forests. *Forest Ecology and Management*, 154(3) 395-407(13)

Nepstad, D.C., Adalberto Veríssimo, Ana Alencar, Carlos Nobre, Eirivelthon Lima, Paul Lefebvre, Peter Schlesinger, Christopher Potter, Paulo Moutinho, Elza Mendoza, Mark Cochrane and Vanessa Brooks. 1999. Large-scale impoverishment of Amazonian Forests by logging and fire. *Nature* 398: 505-08

Nepstad, D.C., et al, 1992. Biotic impoverishment of Amazonian forests by rubber tappers, loggers and cattle ranchers. *Adv. In Economic Botany* 9, 1-14.

Nepstad, D.C., Klink, C.A., Uhl, C., et al 1997 *Ciencia e Cultura* 49, 73-86

Nielsen, T.L., Zöbisch, M.A. 2001. Multifactorial causes of land-use change: land-use dynamics in the agropastoral village of Im Mial, northwestern Syria. *Land Degrad. Dev.* 12(2): 143-61

Organization for Economic Cooperation and Development (OECD). 1997. Brazilian Agriculture: Recent policy changes and trade prospects." Working Paper No. 55. Paris, OECD.

Pan, W.K. Y., Walsh, S.J., Bilsborrow, R. E., Frizzelle, B., G., Erlien, C.M., Baquero, F., 2004. Farm-level models of spatial patterns of land use and land cover dynamics in the Ecuadorian Amazon**.** *Agriculture, Ecosystems & Environment,* 101,(2-3) 117-134

Pedlowski M.A.; Dale V.H 1992. Land use practices in Ouro Preto do Oeste, Rondônia, Brazil. ORNL/TM 12062, Oak Ridge, TN.

Pedlowski M.A.; Dale V.H.; Matricardi E.A.T.; da Silva Filho E.P. 1997 Patterns and impacts of deforestation in Rôndonia, Brazil. *Landscape and Urban Planning*, 38(3) 149-157(9)

Pedlowski, M.A. and V. H. Dale. 1992 Land use practices in Ouro Preto Do Oeste, Rondônia, Brazil. ORNL-TM 3850. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Perz S.G. 2002 The Changing Social Contexts of Deforestation in the Brazilian Amazon. *Social Science Quarterly*, 83(1) 35-52

Perz S.G. 2002 The Changing Social Contexts of Deforestation in the Brazilian Amazon. *Social Science Quarterly*, 83(1) 35-52

Perz, S.G., 2001. Population growth and Net Migration in the Brazilian Legal Amazon 1970-1996". Forthcoming in Charles H. Wood and Roberto Porro, eds., *Deforestation and Land Use in the Amazon.* Gainesville: University of Florida Press.

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

Pichón, Francisco J. 1996. Settler Agriculture and the Dynamics of Resource Allocation in Frontier Environments" *Human Ecology* 24:341-71

Pickles, A., 1980. Models of movement: A review of alternative methods. *Envir. and Planning A* 12, 1383-1404

Portela, R; Rademacher, I., 2001. A dynamic model of patterns of deforestation and their effect on the ability of the Brazilian Amazonia to provide ecosystem services. *Ecological modeling* (143) 115-146

Post W. M., Peng, T.H., Emanuel, W., King, A.W., Dale, V.H., DeAngelis, D.L., 1990. The global carbon cycle. *American scientist*, 78 (4) 310-326

Post W. M., Peng, T.H., Emanuel, W., King, A.W., Dale, V.H., DeAngelis, D.L., 1990. The global carbon cycle. *American scientist*, 78 (4) 310-326

Rejmanek, M., Sasser, C.E., Gosselink, J.G., 1987. Modeling of vegetation dynamics in the Mississippi River deltaic plain. *Vegetatio* 69, 133-140

Rich, B. 1994 *Mortaging the Earth: The World Bank, Environment, Impoverishment, and the Crisis of Development.* Boston, M.A.: Earthscan.

Riebsame, W.E., Meyer, W.B., Turner, B.L., 1994. Modeling land use and cover as part of global environmental change. *Climate Change* 28 (1-2) 45-64

Roberts, M.R., and Hruska, A.J., 1986. Predicting diameter distributions: A test of the stationary Markov model. *Canadian Journal of Forest Research*. 16, 130-135

Ross, S. M., 1997. Introduction to probability models. Sixth edition. Academic Press.

Rounsevel, M.D.A., Evans, S.P., Mayr, T.R., Audsley,E., 1998. Integrating biophysical and socio-economic models for land use studies. In: Proceedings of the ITC-ISSS Conference on Geo-Information for Sustainable Land Management, Enschede, 17-21 August 1997, 368 pp.

Rudel, T. and Horowitz, B. 1993. *Tropical Deforestation: Small Farmers and Land Clearing in the Ecuatorian Amazon.* Boulder, Colo: Westview

Saiko, T.A., Zonn IS. 2000. Irrigation expansion and dynamics of desertification in the Circum-Aral region of Central Asia. *Appl. Geogr.* 20(4): 349-67

Sanchez, T. W., 2004. Land use and growth impacts from highway capacity increases. Journal of Urban Planning and Development. ASCE. P 75-82.

Schmink, M. and Wood, C.H. 1987. The "Political Ecology" of Amazonia. Pp.38-57 in Peter D. Little and Michael M. Horowitz, eds., *Lands at risk in the Third World: Local level perspectives. Boulder, Colo.: Westview.*

Schmink, M. and Wood, C.H. 1992. *Contested Frontiers in Amazonia.* New York: Columbia University Press.

Schomberg, W., 1999. Brazil suspends issuing of Amazon clearing permits. Reuters News Service, 12 February 1999.

Schomberg, W.: 1998b, "Brazil introduces new law to protect environment" Reuters News Service, 13 February 1998.

Serrão, E. Adilson S., and Alfredo K. O. Homma. 1993. "Brazil" Pp. 265-351 in *Sustainable Agriculture and the Environment in the Humid Tropics.* Washington, D.C.: National Academy Press.

Shukla, J., Nobre C., Sellers, P., 1990. Amazon deforestation and climate change. *Science* 247, 1322-1325.

Sierra, R., Stallings J. 1998. The dynamics and social organization of tropical deforestation in Northwest Ecuador, 1983-1995. *Hum Ecol.* 26(1) 135-61

Simon, Herbert A. 1997. *Models of Bounded Rationality: Empirically Grounded Economic Reason.* Cambridge: MIT Press.

Skole, D.L., Chomentowski, W.H., Sala, W.A., Nobre, A.D., 1994. Physical and human dimensions of deforestation in Amazônia. Bioscience 44, 314-321

Summers, P.M., Browder, J. O., Pedlowski, M.A., 2004. Tropical forest management and silvicultural practices by small farmers in the Brazilian Amazon: recent farm-level evidence from Rôndonia. Forest *Ecology and management* (192) 161-177.

Thorner, Daniel, Basil Kerblay and R.E.F. Smith, eds. 1986. *A.V. Chayanov on the Theory of Peasant Economy.* Madison: University of Wisconsin Press.

Thornton, P.K., Jones, P.G., 1998. A conceptual approach to dynamic agricultural land-use modeling. Agric. Syst. 57 (4), 505-521.

Toni, Fabiano. 1999. State –Society Relations on the Agricultural Frontier. Doctoral dissertation, University of Florida. Available online at http://etd.fcla.edu/etd/uf/1999/amg2051/toni_f.pdf

Tucker, C.J., Holben, B.N., and Goff, T.E. 1984. Intensive forest clearing in Rondônia, Brazil, as detected by satellite remote sensing. Remote Sensing Environment 15: 255-261

Tucker, C.J., Townsend, J.R.G., Goff, T.E., and Holben, B.N. 1986. Continental and global scale remote sensing of land cover. Pp. 221-241 *In:* J.R. Trabalka and D.E. Reichle ed. The changing carbon cycle: A global analysis. Springer-Verlag, New York.

Turner, Billie Lee II, David Sloke, Steven Sanderson, Günther Fischer, Louise Fresco and Rik Leemans. 1995. *Land Use and Land Cover Change: Science/Research Plan*. Stockholm and Geneva: International Geosphere-Biosphere Program/ HDP.

Turner, M.G. 1987. Spatial simulation of landscape changes in Georgia: A comparison of three transition models. Landscape Ecology 1:29-36

Uhl, C., Buschbacher, R., 1985. A disturbing synergism between cattle ranch practices and selective tree harvesting in the eastern Amazon. Biotropica 17, 265-268

Uhl, C., Kauffman, J.B., 1990. Deforestation fire susceptibility and potential tree responses to fire in the eastern Amazon. *Ecology* 71, 437-449

UN Food Agric. Organ. 2001. Global forest resources assessment 2000 (FRA 2000): main report, *FAO For.* Pap. 140, FAO, Rome.

Veiga, J.B., Alves, A.M., Poccard-Chapuis, R., Thales, M.C., Da Costa, P.A., Grijalva, J.O., Chamba, T.V., Costa, R.M., Piketty, M-G., Tourrand, J-F., 2001. Cattle ranching, land use and deforestation in Brazil, Peru and Ecuador. Annual report for the Inter-American Institute, Gainesville, USA.

Veldkamp, A., Fresco, L., 1996. CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. Ecol. Model. 91, 231-248.

Veldkamp, A., Fresco, L., 1997b. Reconstructing land use drivers and their spatial scale dependence for Costa Rica. (1973 and 1984). Agric. Syst. 55, 19-43.

Verissimo, A., Barreto, P., Mattos M., Tarifa, R., Uhl, C., 1992. Logging impacts and prospects for sustainable forest management in an old Amazon frontier: the case of Paragominas. Forest Ecology and Management 55, 169-199.

Walker, R., 2004. Theorizing land-cover and land-use change: the case of tropical deforestation. International Regional Science Review, 27 (3), 247-270.

Walker, R., Drzyzga, S.A., Li, Y., Qi, J., Cladas, M., Arima, E., Vergara, D., shortcoming in Ecological Applications. A behavioral model of landscape change in the Amazon Basin: the colonist case.

Walker, R., Moran, E.F., Anselin, L., 2000. Deforestation and cattle ranching in the Brazilian Amazon: external capital and household processes. *World Dev.* 28 (4), 683-699.

Walker, R.T., Moran E.F., and Anselin L., 2000. "Deforestation and cattle ranching in the Brazilian Amazon: External capital and household processes." *World Development* 28:683-99.

Walker, Robert T., and Alfredo K.O. Homma. 1996 Land Use and Land Cover Dynamics in the Brazilian Amazon: Overview." *Ecological Economics* 18:67-80.

Wassenaar, T., Lagacherie, P., Legros, J.P., Rounsevel, M.D.A., 1999. Modeling wheat yield responses to soil and climate variability at the regional scale. *Clim. Res.* 11, 209-220

Watts, M, Bosshardt, W., How instructors make a difference: panel data estimates from principles of Economics courses. The Review of Economics and Statistics, 73 (2), 1991, 336-340.

Weng, Q., 2002. Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modeling. Journal of Environmental Management, 64, 273-284.

White, R., Engelen, D., Uljee, I., 1997. The use of constrained cellular automata for high resolution modeling of urban land-use dynamics. Environmental Planning B 24 (3), 323-343

Whitmore, T.C. 1997, In: W.F. Laurance and R.O. Bierregaard, Jr., Eds. *Tropical Forest Remnants: Ecology, Management, and Conservation of Fragmented Communities.* University of Chicago Press, Chicago, pp. 3-12

Wilson, E.O. 1989. Scientific American, 261(3) p. 108

Wood, Charles H., David Skole, Stephen Perz, and André Caetano. 1996 "Population and Deforestation in the Brazilian Amazon." Paper presented at the Population Association of America meetings, Miami, Fla., April 6-8.

Wooldridge, J.M., 2000. Introductory econometrics. A modern approach. South-Western College Publishing.

Yamada M.; Gholz H.L. 2002. An evaluation of agroforestry systems as a rural development option for the Brazilian Amazon. *Agroforestry Systems*, 55(2), pp. 81-87(7)

Zhouri A., 2004. Global–Local Amazon Politics: Conflicting Paradigms in the Rainforest Campaign *Theory, Culture & Society*, 21(2) 69-89

Glossary of terms

Absorbing state. If the system ever enters this state, it can never leave.

Endogenous variables. Socioeconomic and demographic variables at the household level.

Exogenous variables. Macroeconomic variables.

First order property, Markov property, first order dependency, first order behavior. One of the two main assumptions of a Markov chain, this property says that The future (state) of a process depends just on the present (state) and not on the previous states of the process. Thus, the probability of the distribution among land covers in time t+1 depends only on the current distribution at time t and on the transition probabilities. In terms of conditional probabilities this is usually expressed as:

$$\Pr[X_n = i_n \mid X_1=i_1; \ldots ; X_{n-1} = i_{n-1}] = \Pr[X_n = i_n \mid X_{n-1} = i_{n-1}]$$

Forest. Extensive debate exists about the lack of a standard definition of forest and deforestation and the diverging information provided by national and international agencies. For example FAO's Tropical Forest Resource Assessment, 2000 changed the definition of forest from 20% to 10% crown cover in a specific area. Due to the ambiguity in the definitions of primary and secondary forest and the divergence of data provided by governmental agencies, the dependent variable in this research will not be obtained through the survey instrument or governmental sources. Instead, Satellite images of the area of study will be used to estimate the percentage of area deforested after overlaying property boundaries.

Higher order effects. A first-order process was defined in Assumption 1 or Markov property. In a Markov chain with higher order effects the future state (land cover at time t+1) of a process depends not only on the present state (land cover at time t) and transition probabilities, but also on past states (land covers at time t-1, t-2, t-3, etc). It is possible to model higher order effects, changing the definition of the state to include present and previous land covers. For example in a second-order Markovian chain the definition of the state would include the land cover in the present and previous time period. Additional data from at least two time intervals following the initial observation is needed in that case.

Irreducible Markov chain. When eventual transitions from any state (land cover) to any other state (land cover) are possible, although they may not be possible in one step. A Markov chain in which all states can communicate with each other, even when the probability may be very low or the transition may take more than one time period.

Non irreducible Markov chain. It is not possible to move from any state to any other state in the Markov chain. A Markov chain where there are more than two classes, so not all states communicate with each other. Two states that communicate are said to be in the same class.

Land cover. Land cover refers to the "attributes of a part of the Earth's land surface and immediate subsurface, including biota, soil, topography, surface and groundwater, and human structures." One type of land cover is forest. (Turner et al, 1993 in Lambin et al, 2000)

Land use. "Land use refers to the purposes for which humans exploit the land cover." For example: recreation, timber production, wildlife conservation are land uses for the land cover forest. (Turner et al, 1993 in Lambin et al, 2000)

Land cover conversion. Land cover conversion is the complete replacement of one land-cover type by another. (Turner et al, 1993 in Lambin et al, 2000)

Land cover modification. Land cover modification refers to "more subtle changes that affect the character of the land cover without changing its overall classification" (Turner et al, 1993 in Lambin et al, 2000)

Legal Amazon. The Legal Amazon is made up of all of the North region (the states Acre, Amapa, Amazonas, Para, Rondônia, Roraima, and Tocatins) plus parts of the states of Maranhao, Mato Grosso, and Goias. The southern edge is the 16th parallel, and the eastern edge is the 44th meridian.

Markov chain process or Markov chain model. A process in which the future state (future land cover) can be predicted knowing the present state (present land cover) and the transition probability matrix.

Pathways. Pathways of land cover change or cronosequences are graphic reconstructions or flow charts of land cover changes in each farm over the 10-yr period. This reconstruction is an approximation of main land cover transitions based on satellite images, survey data, and interviews (in 2003) performed in a sub-sample of 80 farms. Such data will be used to calculate proximate empirical probabilities for land cover change.

Recurrent state. If starting in this state, the expected number of time periods that the process is in this state is infinite. There is always a probability to reenter this state. The state is visited infinitely number of times.

Stationary probabilities, homogeneous or stationary property. One of the main assumptions of the Markov chain theory, it states that transition probabilities are stationary or constant through time. If the assumption of non-stationary probability transitions is not satisfied, then Markov models can be applied to provide answers to "what if" questions.

Stochastic model. Stochastic models are used to simulate stochastic processes, which are phenomena that vary to some degree unpredictably as time goes on. If we observe the process in several occasions under presumable "identical" conditions, the resulting observation would be different. The main characteristic of stochastic processes is the use of probabilities.

Transient states. States that are not certain to be returned to, even if the process starts in the state. If starting in this state, the expected number of time periods that the process is in this state is finite. A state that is visited only a finite number of times.

Transition probabilities, transitional probabilities. Transition probabilities of land cover change represent the probability of each pixel of changing from one land cover to another land cover.

Transitional probability matrix, matrix of transition probabilities. It is a n x n matrix that contains all the transition probabilities among the n states (land covers).

Unit of analysis. Unit at which analysis is conducted and conclusions formulated.

Unit of observation. Unit at which the land cover classification is obtained or observed.

APPENDIX A

SAS programs for exploratory models and for the fixed group and time effects models.

EXPLORATORY MODELS

```
proc reg data= nancy.newpanel;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf;
run;

proc reg data= nancy.newpanel;
model pfnf =  d1-d69 pfpf sfnf nfnf pfsf sfsf nfsf;
test
d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15=d16=d17=d18=d19=d20=d21=d2
2=d23=d24=d25=d26=
d27=d28=d29=d30=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45=d
46=d47=d48=d49=d50=d51=
d52=d53=d54=d55=d56=d57=d58=d59=d60=d61=d62=d63=d64=d65=d66=d67=d68=d69=0;
run;
proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf/FIXONE;
run;
proc reg data= nancy.newpanel;
model pfnf =  y2-y10 pfpf sfnf nfnf pfsf sfsf nfsf;
test y2=y3=y4=y5=y6=y7=y8=y9=y10=0;
run;
proc sort data= nancy.newpanel;
by year farmr;
proc TSCSREG data=nancy.newpanel;
id year farmr;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf/FIXONE;
run;
proc reg data= nancy.newpanel;
model pfnf =  d1-d69  y2-y10 pfpf sfnf nfnf pfsf sfsf nfsf;
test
d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15=d16=d17=d18=d19=d20=d21=d2
2=d23=d24=d25=d26=
d27=d28=d29=d30=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45=d
46=d47=d48=d49=d50=d51=
d52=d53=d54=d55=d56=d57=d58=d59=d60=d61=d62=d63=d64=d65=d66=d67=d68=d69=y2=y3
=y4=y5=y6=y7=y8=y9=y10=0;
run;
proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf/FIXTWO;
run;
```

```
proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf/RANONE;
run;
proc sort data= nancy.newpanel;
by  year farmr;
proc TSCSREG data=nancy.newpanel;
id  year farmr;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf/RANONE;
run;
proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf/RANTWO;
run;
proc sort data= nancy.newpanel;
by farmr;
proc REG data=nancy.newpanel;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf;
by farmr;
run;
proc sort data= nancy.newpanel;
by year;
proc REG data=nancy.newpanel;
model pfnf = pfpf sfnf nfnf pfsf sfsf nfsf;
by year;
run;
```

SAS PROGRAMS FOR FIXED GROUP AND TIME EFFECTS

```
proc reg data= nancy.newpanel;
model pfpf = pfsf pfnf sfsf sfnf nfsf nfnf;
run;


proc reg data= nancy.newpanel;
model pfsf = pfpf pfnf sfsf sfnf nfsf nfnf;
run;


proc reg data= nancy.newpanel;
model sfsf = pfpf pfsf pfnf sfnf nfsf nfnf;
run;


proc reg data= nancy.newpanel;
model nfsf = pfpf pfsf pfnf sfsf sfnf nfnf;
run;


proc reg data= nancy.newpanel;
model pfnf = pfpf pfsf sfsf sfnf nfsf nfnf;
run;


proc reg data= nancy.newpanel;
model sfnf = pfpf pfsf pfnf sfsf nfsf nfnf;
run;


proc reg data= nancy.newpanel;
model nfnf = pfpf pfsf pfnf sfsf sfnf nfsf;
run;


proc reg data= nancy.newpanel;
model pfpf =  d1-d69 pfsf pfnf sfsf sfnf nfsf nfnf;
test d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15
=d16=d17=d18=d19=d20=d21=d22=d23=d24=d25=d26=d27=d28=d29=d30
=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45
=d46=d47=d48=d49=d50=d51=d52=d53=d54=d55=d56=d57=d58=d59=d60
=d61=d62=d63=d64=d65=d66=d67=d68=d69=0;
run;



proc reg data= nancy.newpanel;
model pfsf =  d1-d69 pfpf pfnf sfsf sfnf nfsf nfnf;
test d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15
=d16=d17=d18=d19=d20=d21=d22=d23=d24=d25=d26=d27=d28=d29=d30
=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45
=d46=d47=d48=d49=d50=d51=d52=d53=d54=d55=d56=d57=d58=d59=d60
=d61=d62=d63=d64=d65=d66=d67=d68=d69=0;
run;


proc reg data= nancy.newpanel;
model sfsf =  d1-d69 pfpf pfsf pfnf sfnf nfsf nfnf;
test d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15
```

```
=d16=d17=d18=d19=d20=d21=d22=d23=d24=d25=d26=d27=d28=d29=d30
=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45
=d46=d47=d48=d49=d50=d51=d52=d53=d54=d55=d56=d57=d58=d59=d60
=d61=d62=d63=d64=d65=d66=d67=d68=d69=0;
run;


proc reg data= nancy.newpanel;
model nfsf =  d1-d69 pfpf pfsf pfnf sfsf sfnf nfnf;
test d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15
=d16=d17=d18=d19=d20=d21=d22=d23=d24=d25=d26=d27=d28=d29=d30
=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45
=d46=d47=d48=d49=d50=d51=d52=d53=d54=d55=d56=d57=d58=d59=d60
=d61=d62=d63=d64=d65=d66=d67=d68=d69=0;
run;


proc reg data= nancy.newpanel;
model pfnf =  d1-d69 pfpf pfsf sfsf sfnf nfsf nfnf;
test d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15
=d16=d17=d18=d19=d20=d21=d22=d23=d24=d25=d26=d27=d28=d29=d30
=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45
=d46=d47=d48=d49=d50=d51=d52=d53=d54=d55=d56=d57=d58=d59=d60
=d61=d62=d63=d64=d65=d66=d67=d68=d69=0;
run;


proc reg data= nancy.newpanel;
model sfnf =  d1-d69 pfpf pfsf pfnf sfsf nfsf nfnf;
test d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15
=d16=d17=d18=d19=d20=d21=d22=d23=d24=d25=d26=d27=d28=d29=d30
=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45
=d46=d47=d48=d49=d50=d51=d52=d53=d54=d55=d56=d57=d58=d59=d60
=d61=d62=d63=d64=d65=d66=d67=d68=d69=0;
run;


proc reg data= nancy.newpanel;
model nfnf =  d1-d69 pfpf pfsf pfnf sfsf sfnf nfsf;
test d1=d2=d3=d4=d5=d6=d7=d8=d9=d10=d11=d12=d13=d14=d15
=d16=d17=d18=d19=d20=d21=d22=d23=d24=d25=d26=d27=d28=d29=d30
=d31=d32=d33=d34=d35=d36=d37=d38=d39=d40=d41=d42=d43=d44=d45
=d46=d47=d48=d49=d50=d51=d52=d53=d54=d55=d56=d57=d58=d59=d60
=d61=d62=d63=d64=d65=d66=d67=d68=d69=0;
run;


proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model pfpf = pfsf pfnf sfsf sfnf nfsf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model  pfsf = pfpf pfnf sfsf sfnf nfsf nfnf/FIXONE;
run;
```

```
proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model  sfsf = pfpf pfsf pfnf sfnf nfsf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model nfsf = pfpf pfsf pfnf sfsf sfnf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model  pfnf = pfpf pfsf sfsf sfnf nfsf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model  sfnf = pfpf pfsf pfnf sfsf nfsf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by farmr year;
proc TSCSREG data=nancy.newpanel;
id farmr year;
model  nfnf = pfpf pfsf pfnf sfsf sfnf nfsf/FIXONE;
run;


proc reg data= nancy.newpanel;
model pfpf = y2-y5 pfsf pfnf sfsf sfnf nfsf nfnf;
test y2=y3=y4=y5=0;
run;


proc reg data= nancy.newpanel;
model pfsf = y2-y5 pfpf pfnf sfsf sfnf nfsf nfnf;
test y2=y3=y4=y5=0;
run;


proc reg data= nancy.newpanel;
model sfsf = y2-y5 pfpf pfsf pfnf sfnf nfsf nfnf;
test y2=y3=y4=y5=0;
run;


proc reg data= nancy.newpanel;
model nfsf = y2-y5 pfpf pfsf pfnf sfsf sfnf nfnf;
test y2=y3=y4=y5=0;
run;
```

```
proc reg data= nancy.newpanel;
model pfnf = y2-y5 pfpf pfsf sfsf sfnf nfsf nfnf;
test y2=y3=y4=y5=0;
run;


proc reg data= nancy.newpanel;
model sfnf = y2-y5 pfpf pfsf pfnf sfsf nfsf nfnf;
test y2=y3=y4=y5=0;
run;


proc reg data= nancy.newpanel;
model nfnf = y2-y5 pfpf pfsf pfnf sfsf sfnf nfsf;
test y2=y3=y4=y5=0;
run;


proc sort data= nancy.newpanel;
by year farmr;
proc TSCSREG data=nancy.newpanel;
id year farmr;
model pfpf = pfsf pfnf sfsf sfnf nfsf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by year farmr;
proc TSCSREG data=nancy.newpanel;
id year farmr;
model pfsf = pfpf pfnf sfsf sfnf nfsf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by year farmr;
proc TSCSREG data=nancy.newpanel;
id year farmr;
model sfsf = pfpf pfsf pfnf sfnf nfsf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by year farmr;
proc TSCSREG data=nancy.newpanel;
id year farmr;
model nfsf = pfpf pfsf pfnf sfsf sfnf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by year farmr;
proc TSCSREG data=nancy.newpanel;
id year farmr;
model pfnf = pfpf pfsf sfsf sfnf nfsf nfnf /FIXONE;
run;


proc sort data= nancy.newpanel;
by year farmr;
proc TSCSREG data=nancy.newpanel;
id year farmr;
```

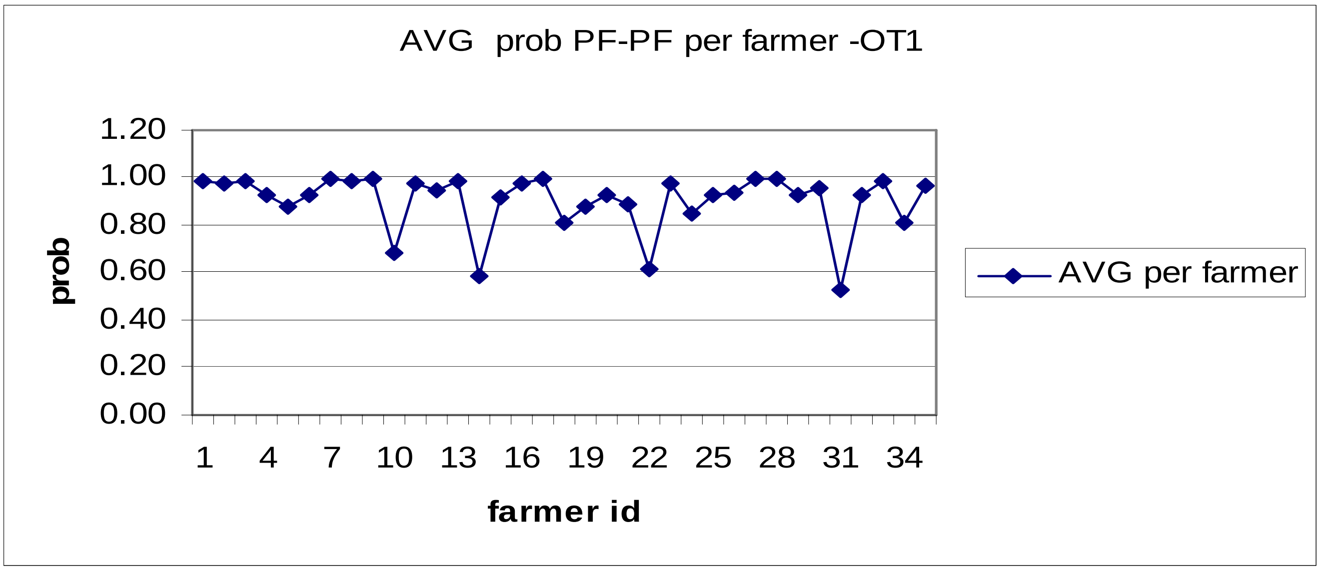
```
model sfnf = pfpf pfsf pfnf sfsf nfsf nfnf/FIXONE;
run;


proc sort data= nancy.newpanel;
by year farmr;
proc TSCSREG data=nancy.newpanel;
id year farmr;
model nfnf = pfpf pfsf pfnf sfsf sfnf nfsf/FIXONE;
run;


proc sort data= nancy.newpanel;
by farmr;
proc REG data=nancy.newpanel;
model pfpf = pfsf pfnf sfsf sfnf nfsf nfnf;
by farmr;
run;


proc sort data= nancy.newpanel;
by farmr;
proc REG data=nancy.newpanel;
model pfsf = pfpf pfnf sfsf sfnf nfsf nfnf;
by farmr;
run;


proc sort data= nancy.newpanel;
by farmr;
proc REG data=nancy.newpanel;
model sfsf = pfpf pfsf pfnf sfnf nfsf nfnf;
by farmr;
run;


proc sort data= nancy.newpanel;
by farmr;
proc REG data=nancy.newpanel;
model nfsf = pfpf pfsf pfnf sfsf sfnf nfnf;
by farmr;
run;


proc sort data= nancy.newpanel;
by farmr;
proc REG data=nancy.newpanel;
model pfnf = pfpf pfsf sfsf sfnf nfsf nfnf;
by farmr;
run;


proc sort data= nancy.newpanel;
by farmr;
proc REG data=nancy.newpanel;
model sfnf = pfpf pfsf pfnf sfsf nfsf nfnf;
by farmr;
run;


proc sort data= nancy.newpanel;
by farmr;
proc REG data=nancy.newpanel;
```

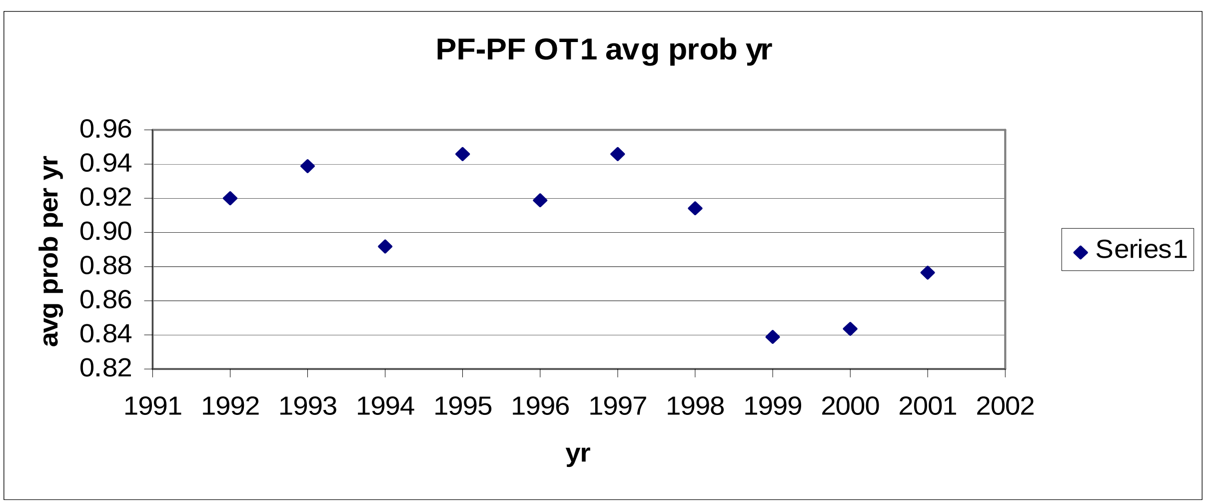
```
model nfnf = pfpf pfsf pfnf sfsf sfnf nfsf;
by farmr;
run;


proc sort data= nancy.newpanel;
by year;
proc REG data=nancy.newpanel;
model pfpf = pfsf pfnf sfsf sfnf nfsf nfnf;
by year;
run;


proc sort data= nancy.newpanel;
by year;
proc REG data=nancy.newpanel;
model pfsf = pfpf pfnf sfsf sfnf nfsf nfnf;
by year;
run;


proc sort data= nancy.newpanel;
by year;
proc REG data=nancy.newpanel;
model sfsf = pfpf pfsf pfnf sfnf nfsf nfnf;
by year;
run;


proc sort data= nancy.newpanel;
by year;
proc REG data=nancy.newpanel;
model nfsf = pfpf pfsf pfnf sfsf sfnf nfnf;
by year;
run;


proc sort data= nancy.newpanel;
by year;
proc REG data=nancy.newpanel;
model pfnf = pfpf pfsf sfsf sfnf nfsf nfnf;
by year;
run;


proc sort data= nancy.newpanel;
by year;
proc REG data=nancy.newpanel;
model sfnf = pfpf pfsf pfnf sfsf nfsf nfnf;
by year;
run;


proc sort data= nancy.newpanel;
by year;
proc REG data=nancy.newpanel;
model nfnf = pfpf pfsf pfnf sfsf sfnf nfsf;
by year;
run;
```

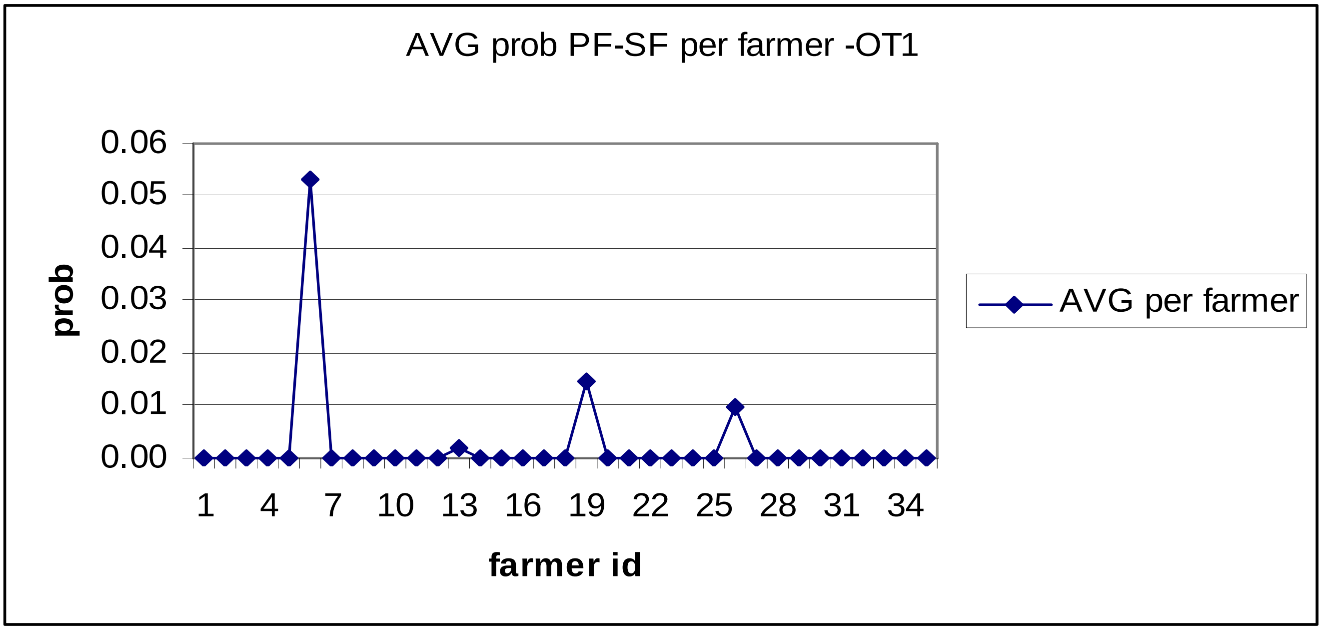
APPENDIX B

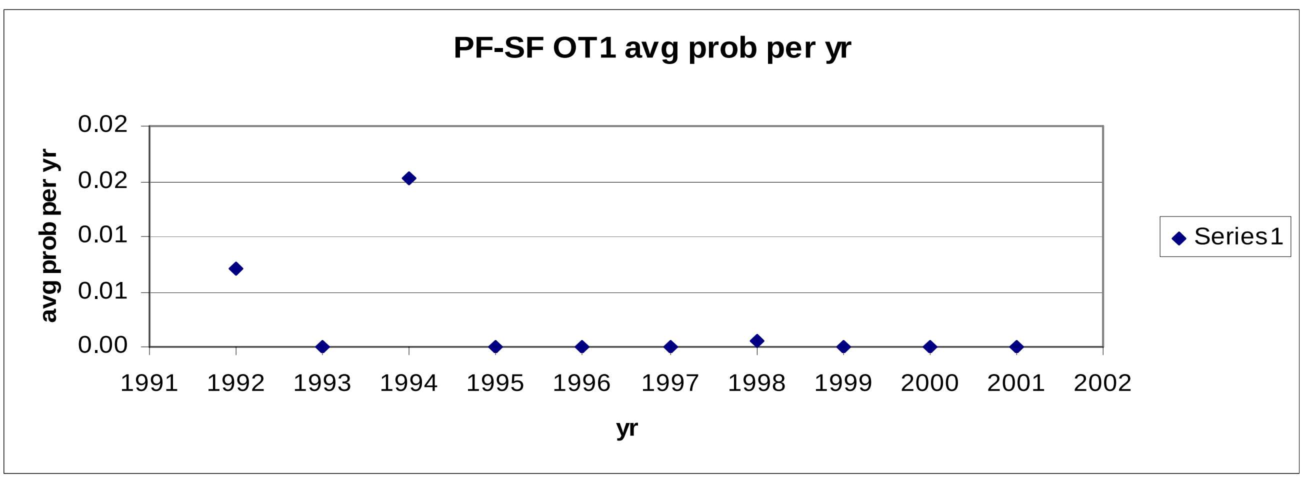
Graphs of average probabilities per LCC to explore trends by farmer and by year

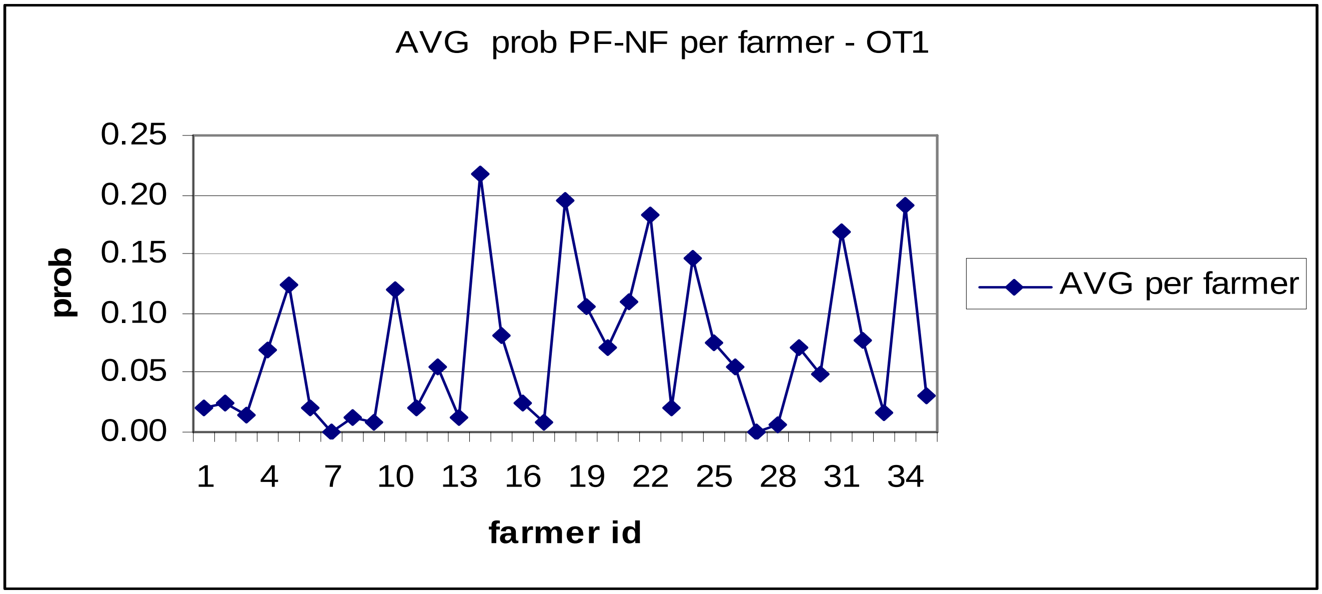
ANALYSIS OF PROBABILITIES – OWNER TYPE 1

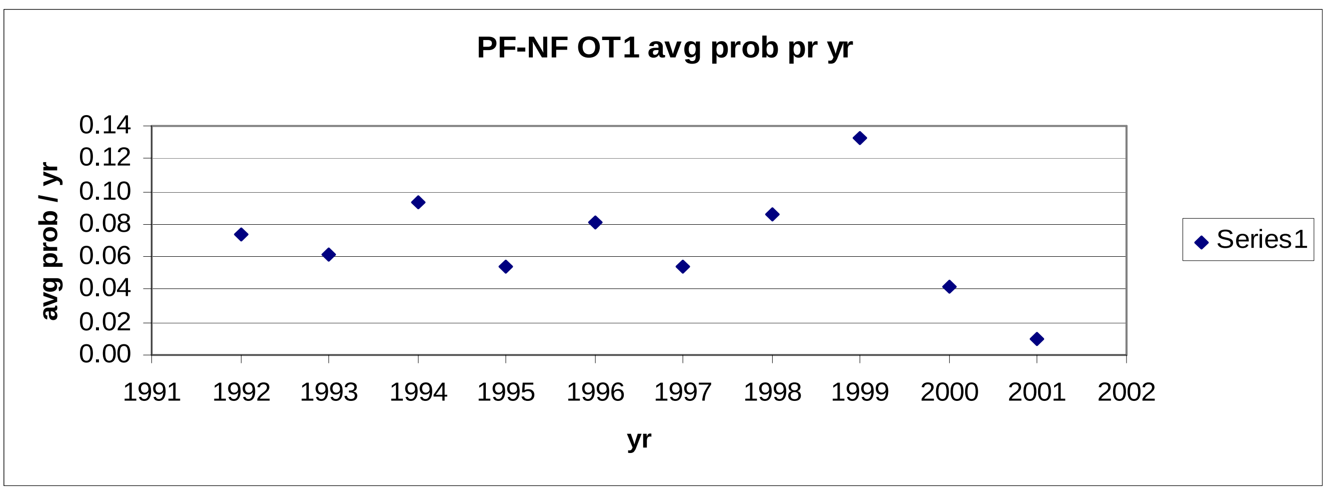


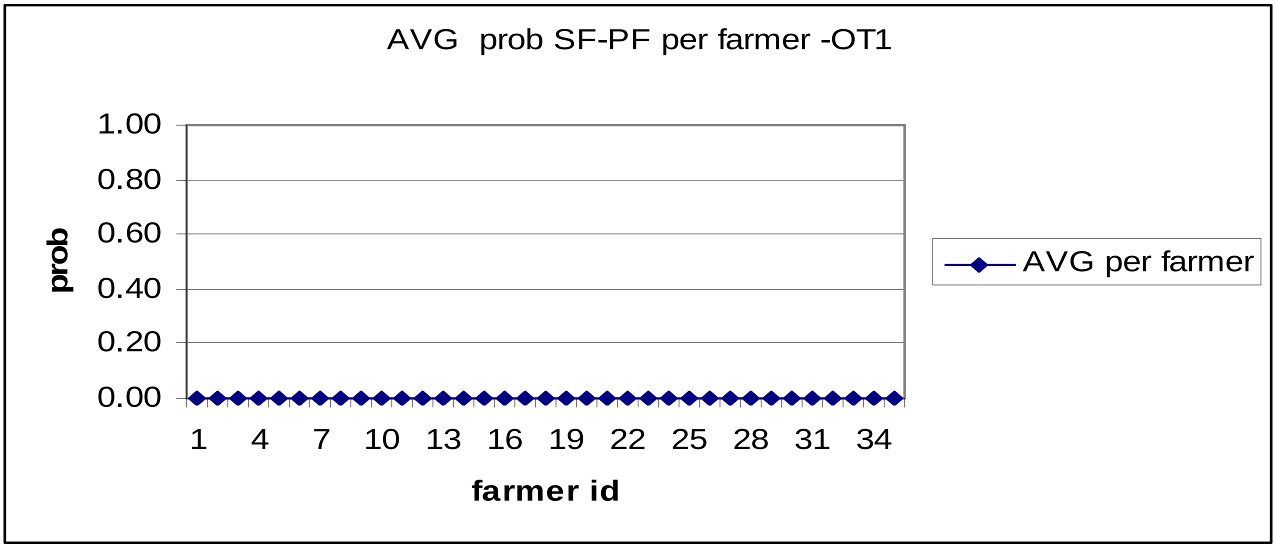


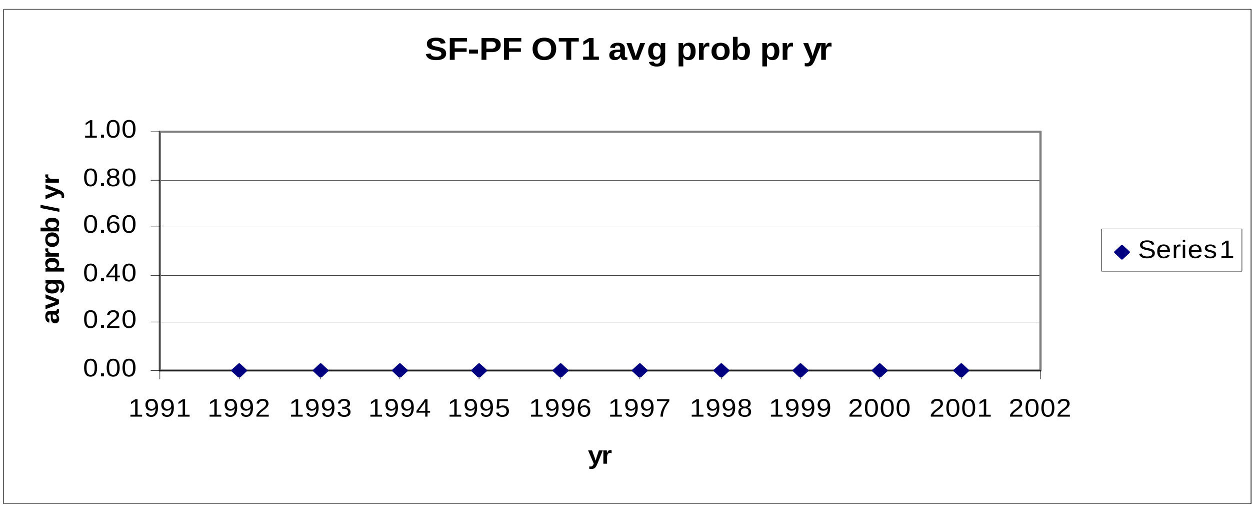

AVG prob PF-SF per farmer -OT1
0.06
0.05
0.04
0.03
0.02
0.01
0.00
prob
1
4
7
10
13
16
19
22
25
28
31
34
farmer id
AVG per farmer

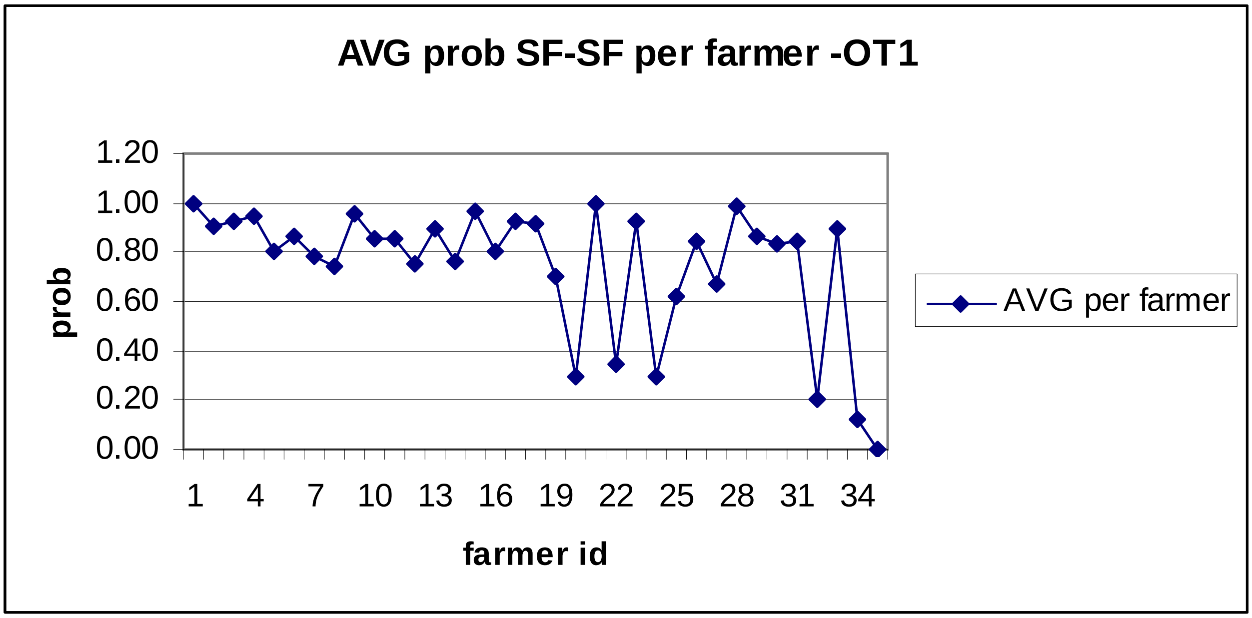

PF-SF OT1 avg prob per yr
0.02
0.02
0.01
0.01
0.00
avg prob per yr
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

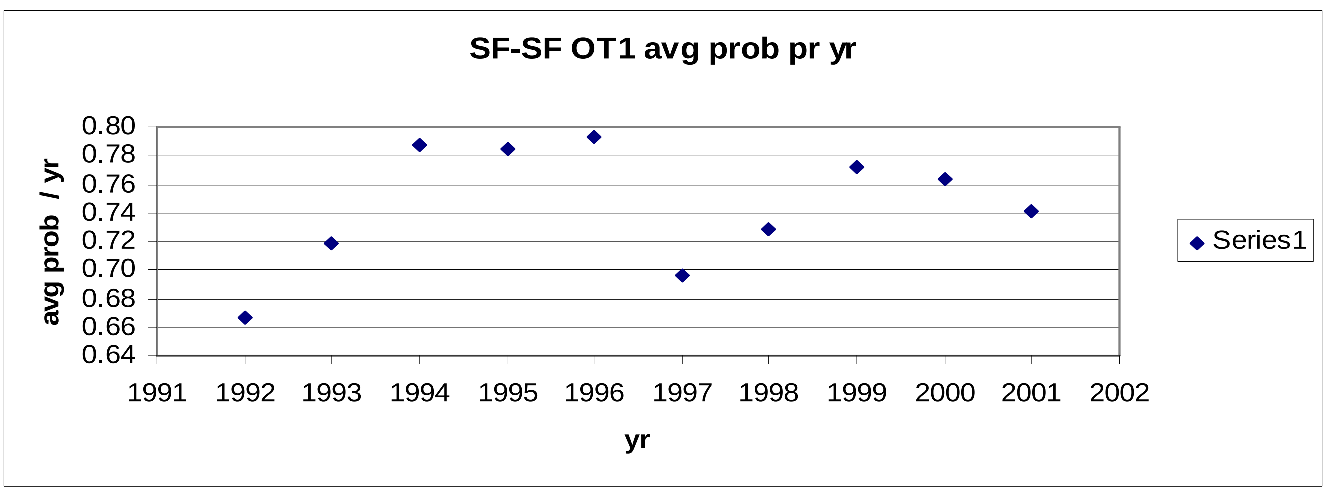

AVG prob PF-NF per farmer - OT1
prob
0.25
0.20
0.15
0.10
0.05
0.00
1
4
7
10
13
16
19
22
25
28
31
34
farmer id
AVG per farmer

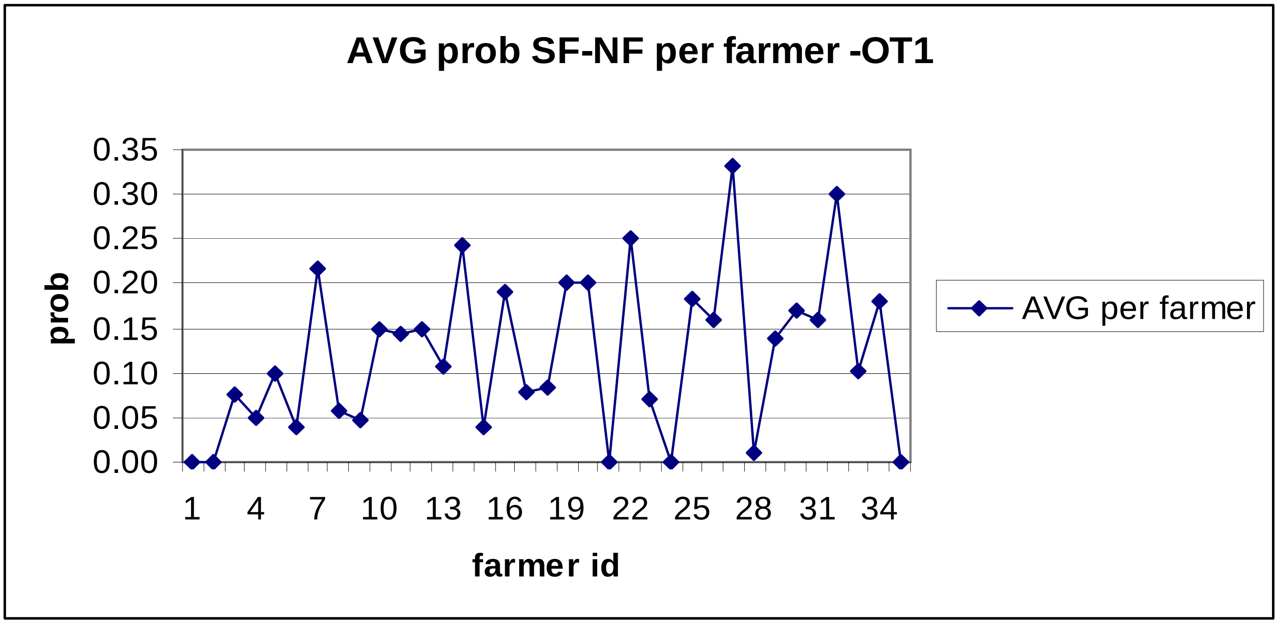

PF-NF OT1 avg prob pr yr
avg prob / yr
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

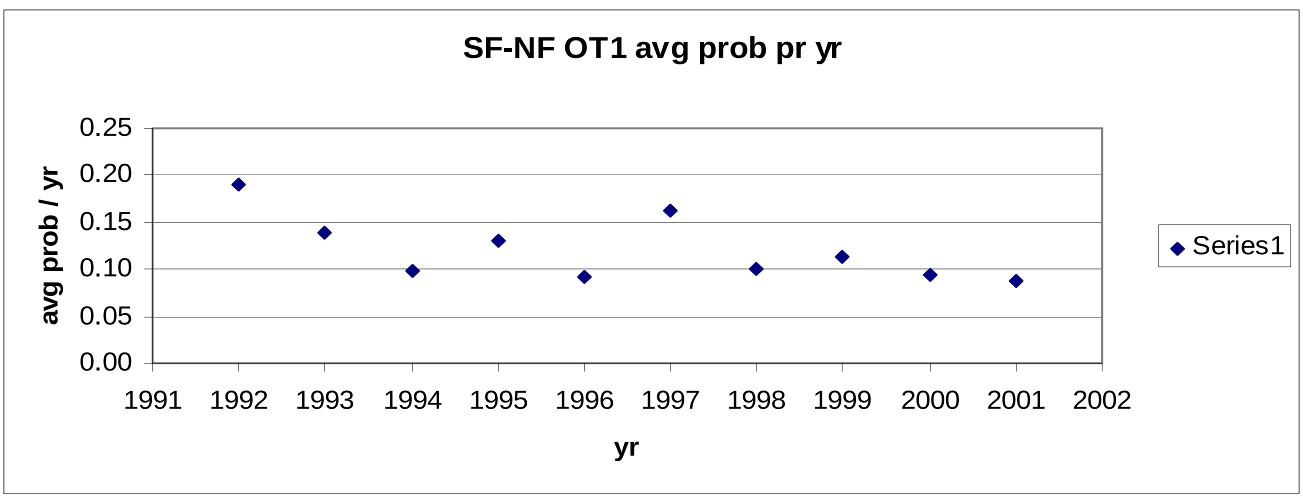

AVG prob SF-PF per farmer -OT1
prob
1.00
0.80
0.60
0.40
0.20
0.00
1
4
7
10
13
16
19
22
25
28
31
34
farmer id
AVG per farmer

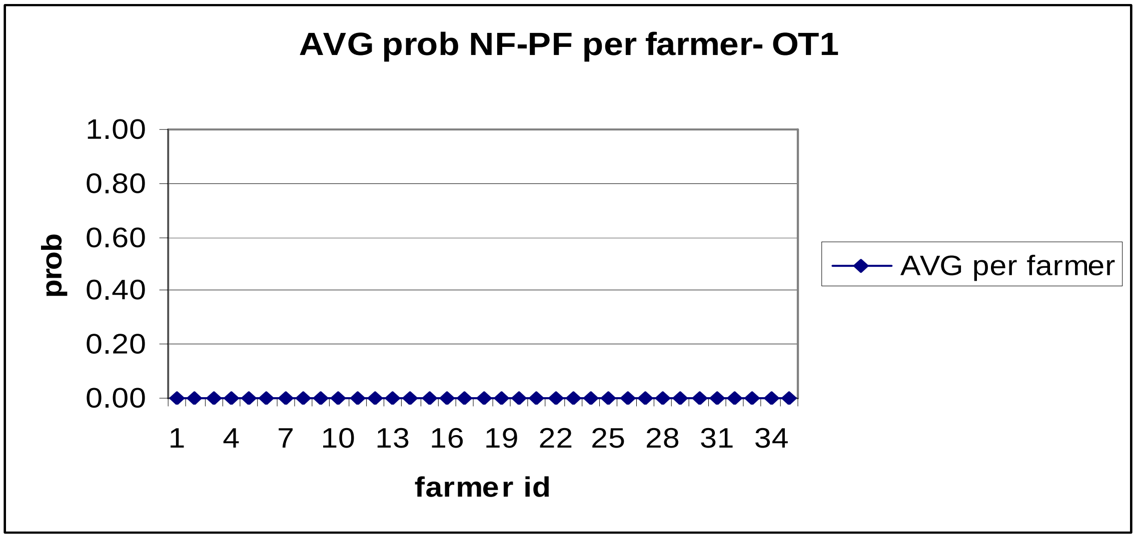

SF-PF OT1 avg prob pr yr
avg prob / yr
1.00
0.80
0.60
0.40
0.20
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1


AVG prob SF-SF per farmer -OT1
1.20
1.00
0.80
0.60
0.40
0.20
0.00
prob
1
4
7
10
13
16
19
22
25
28
31
34
farmer id
AVG per farmer

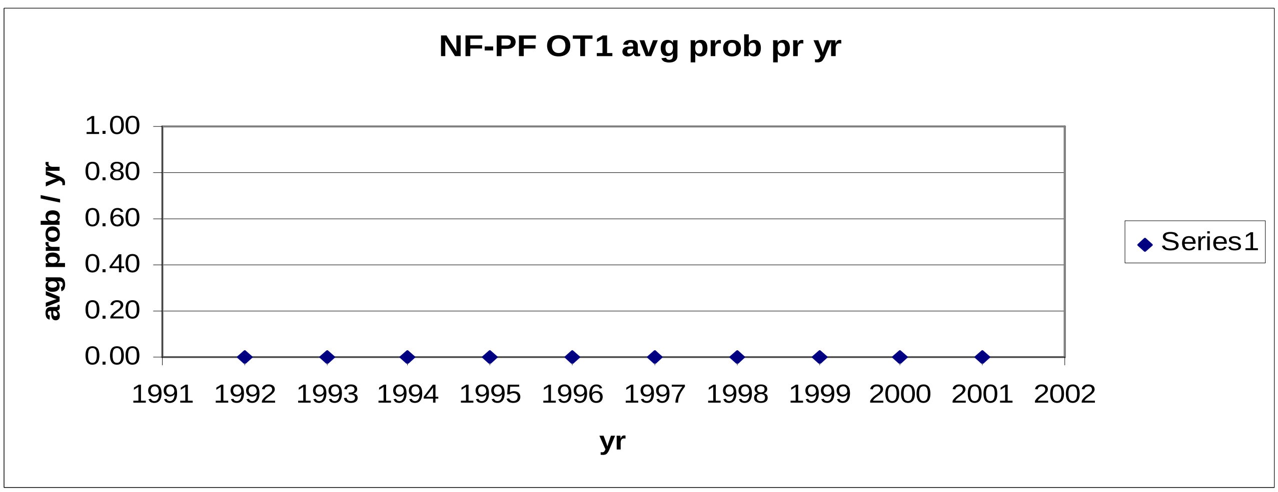

SF-SF OT1 avg prob pr yr
0.80
0.78
0.76
0.74
0.72
0.70
0.68
0.66
0.64
avg prob / yr
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1


AVG prob SF-NF per farmer -OT1
prob
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1
4
7
10
13
16
19
22
25
28
31
34
farmer id
AVG per farmer

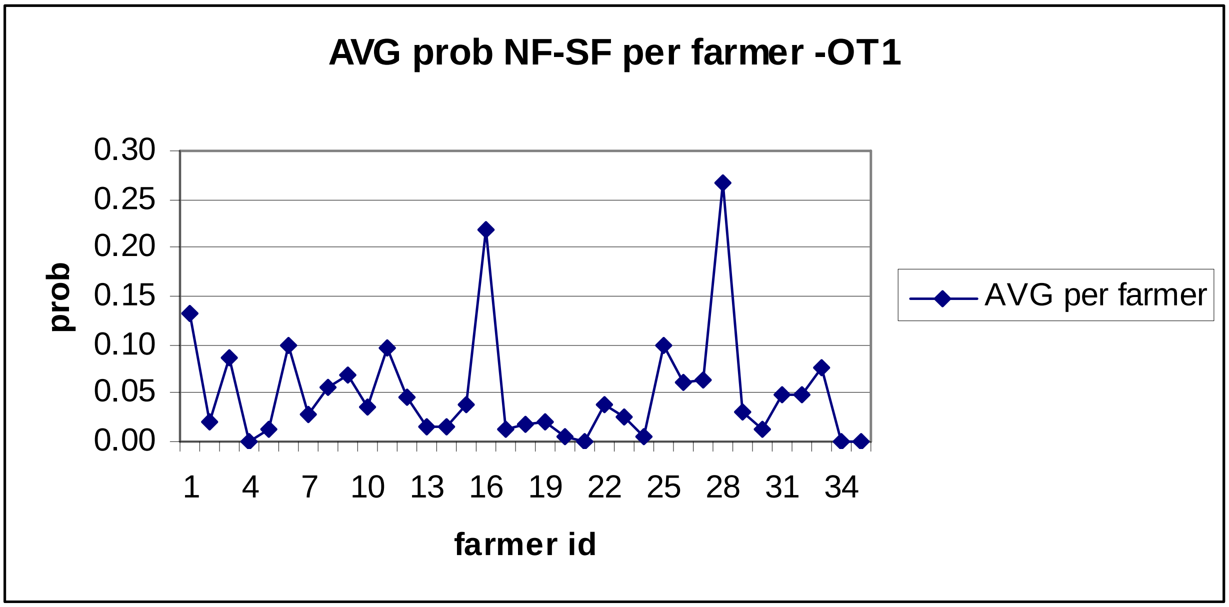

SF-NF OT1 avg prob pr yr
avg prob / yr
0.25
0.20
0.15
0.10
0.05
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

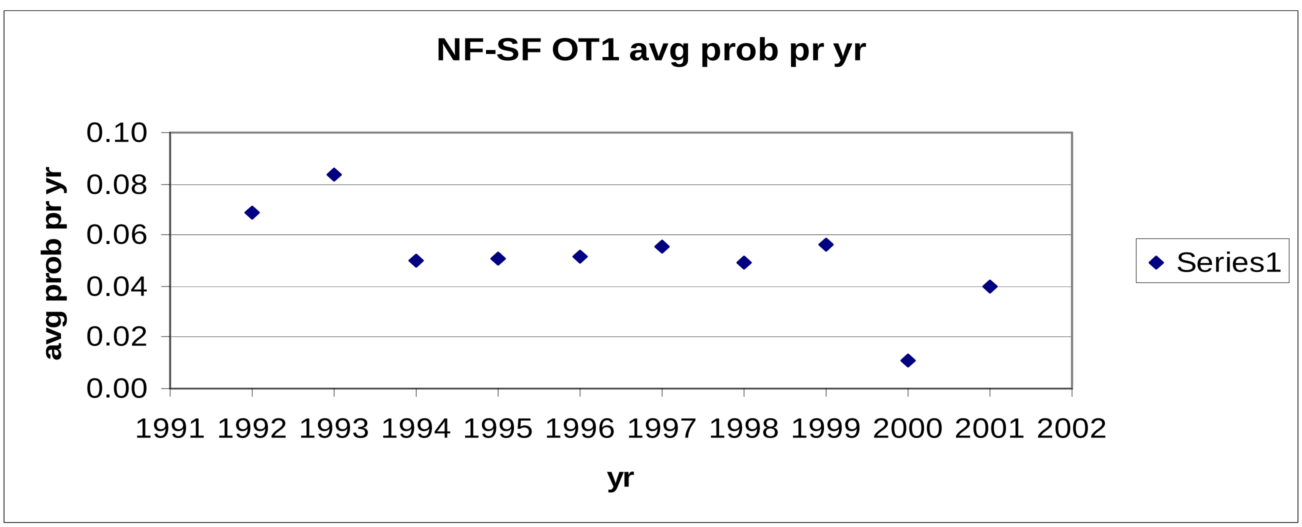


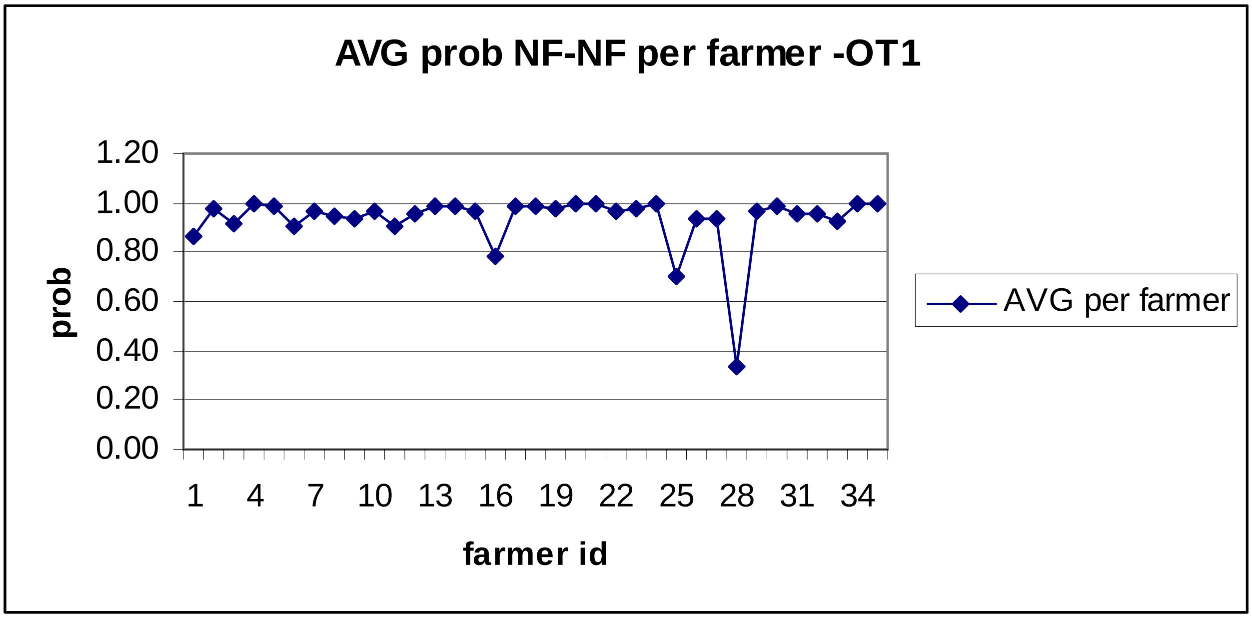
AVG prob NF-PF per farmer- OT1
1.00
0.80
0.60
0.40
0.20
0.00
prob
1
4
7
10
13
16
19
22
25
28
31
34
farmer id
AVG per farmer

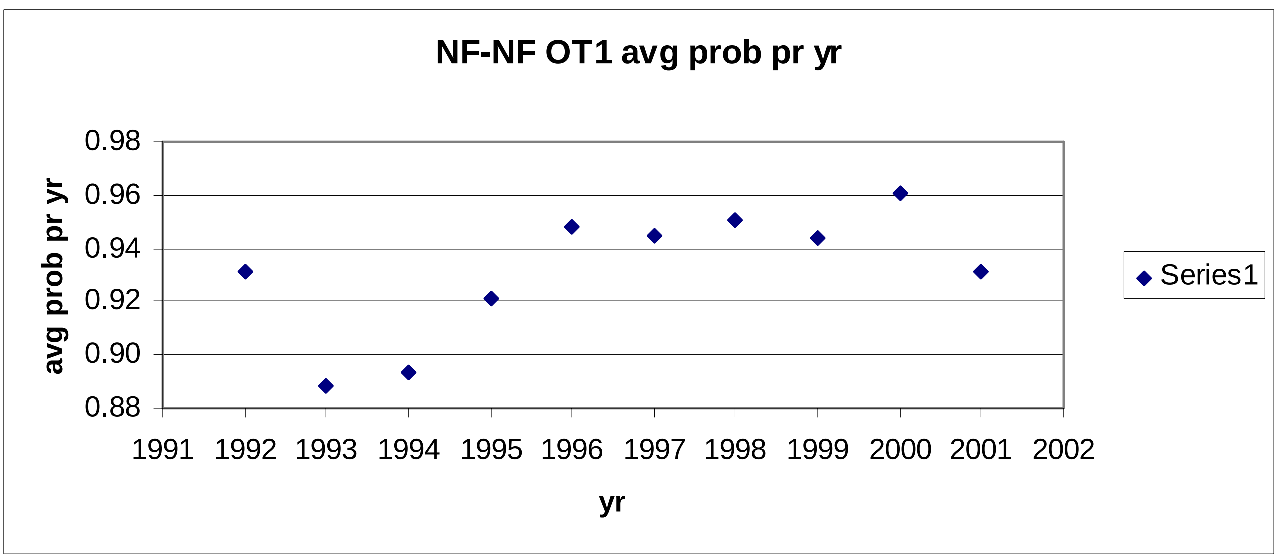


NF-PF OT1 avg prob pr yr
1.00
0.80
0.60
0.40
0.20
0.00
avg prob / yr
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

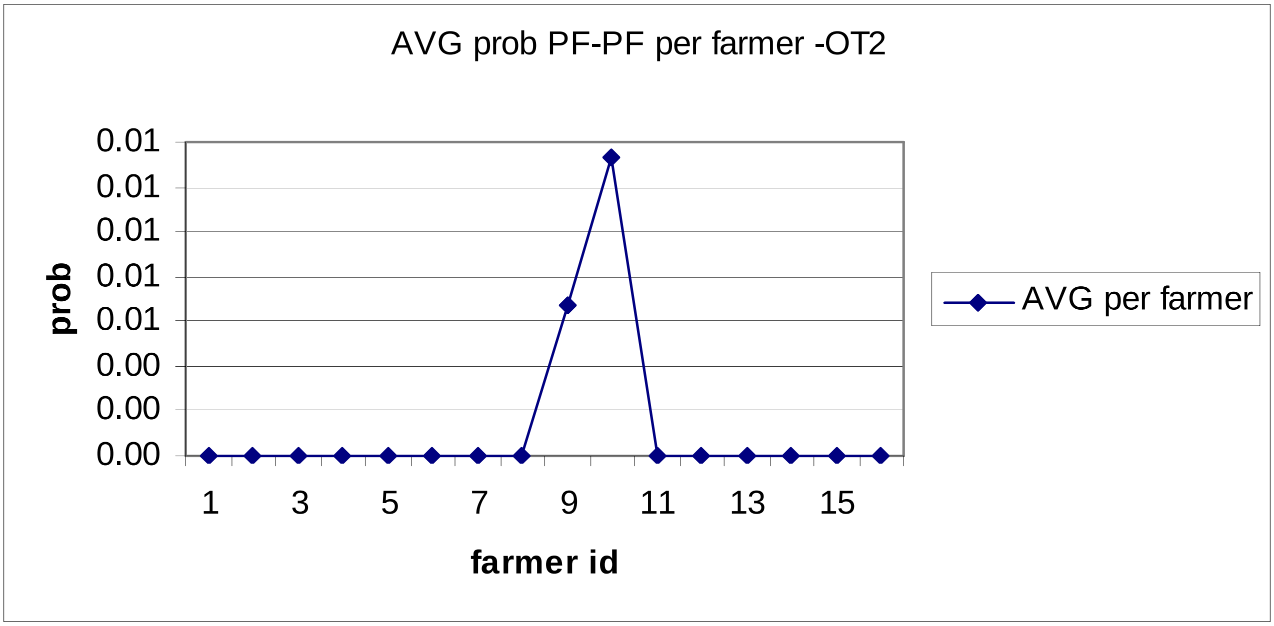

AVG prob NF-SF per farmer -OT1
prob
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1
4
7
10
13
16
19
22
25
28
31
34
farmer id
AVG per farmer

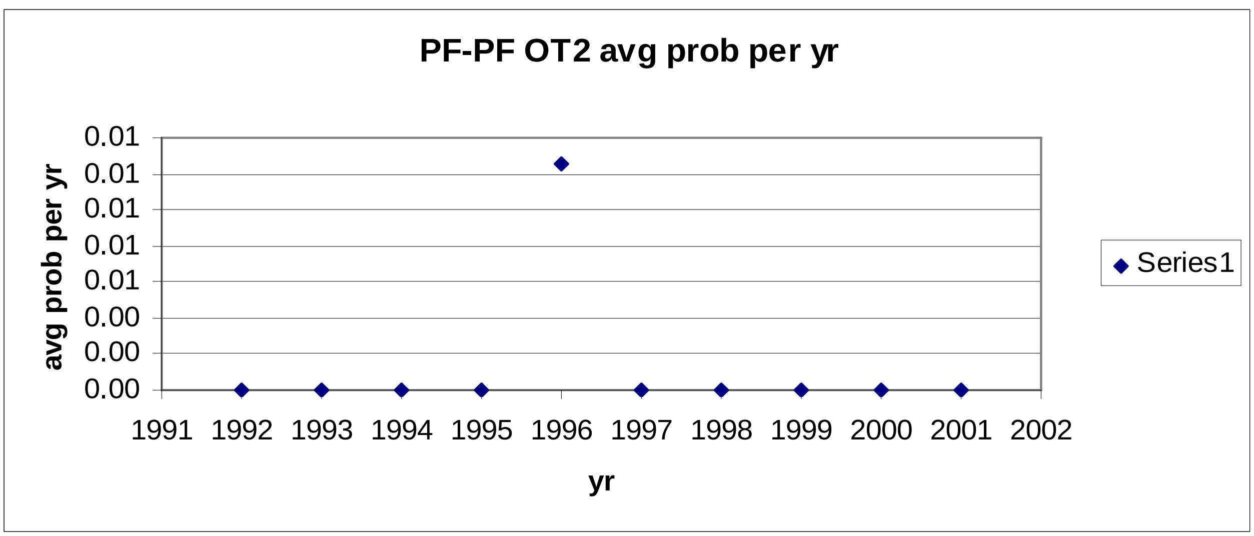

NF-SF OT1 avg prob pr yr
avg prob pr yr
0.10
0.08
0.06
0.04
0.02
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

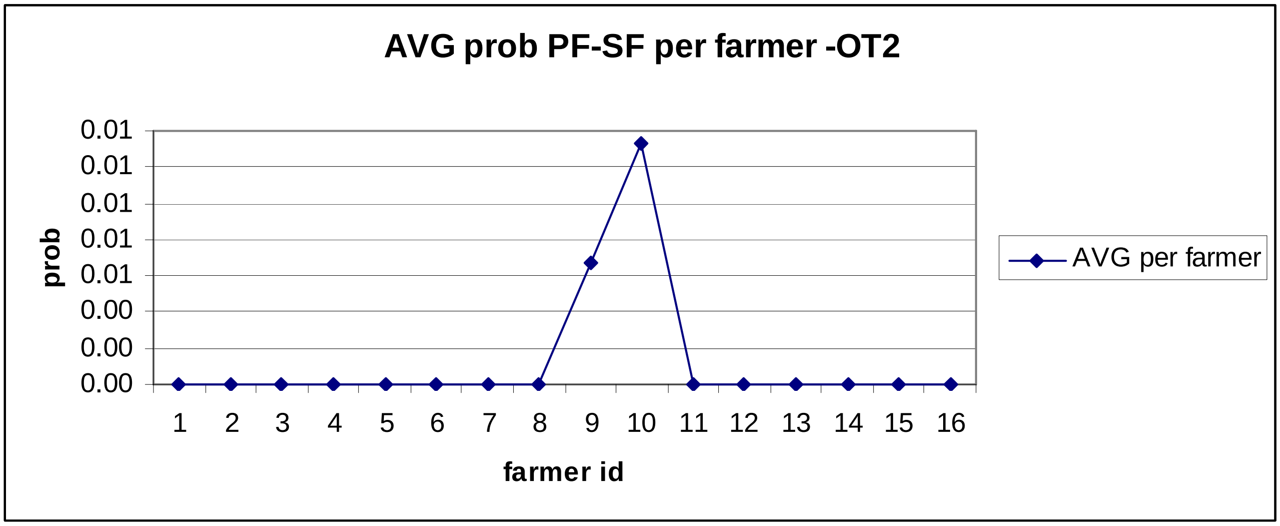

AVG prob NF-NF per farmer -OT1
prob
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1
4
7
10
13
16
19
22
25
28
31
34
farmer id
AVG per farmer

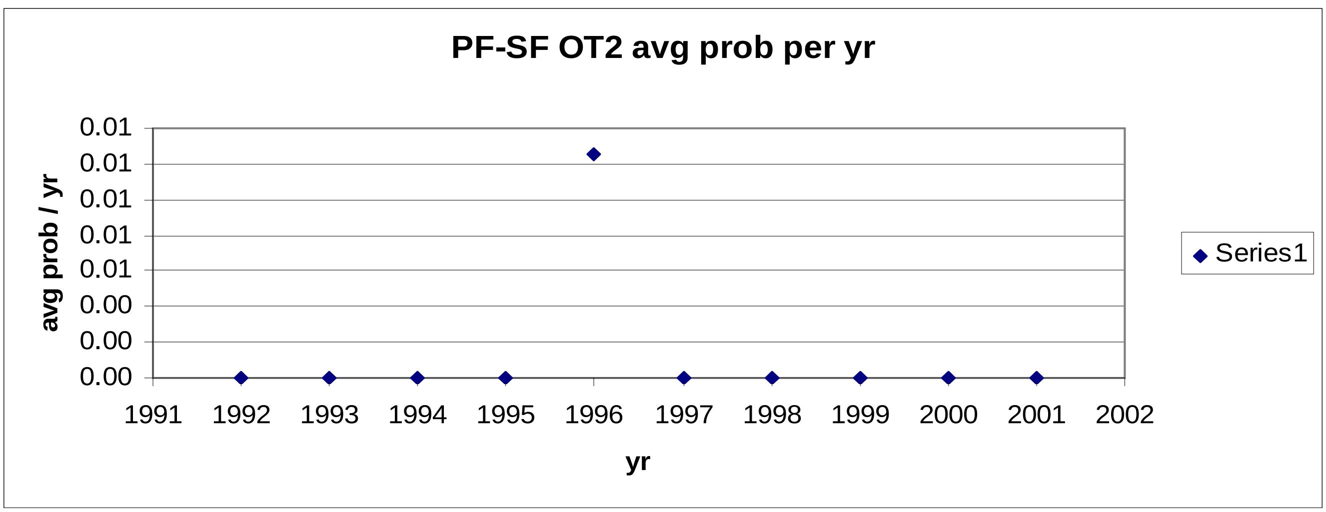

NF-NF OT1 avg prob pr yr
avg prob pr yr
0.98
0.96
0.94
0.92
0.90
0.88
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

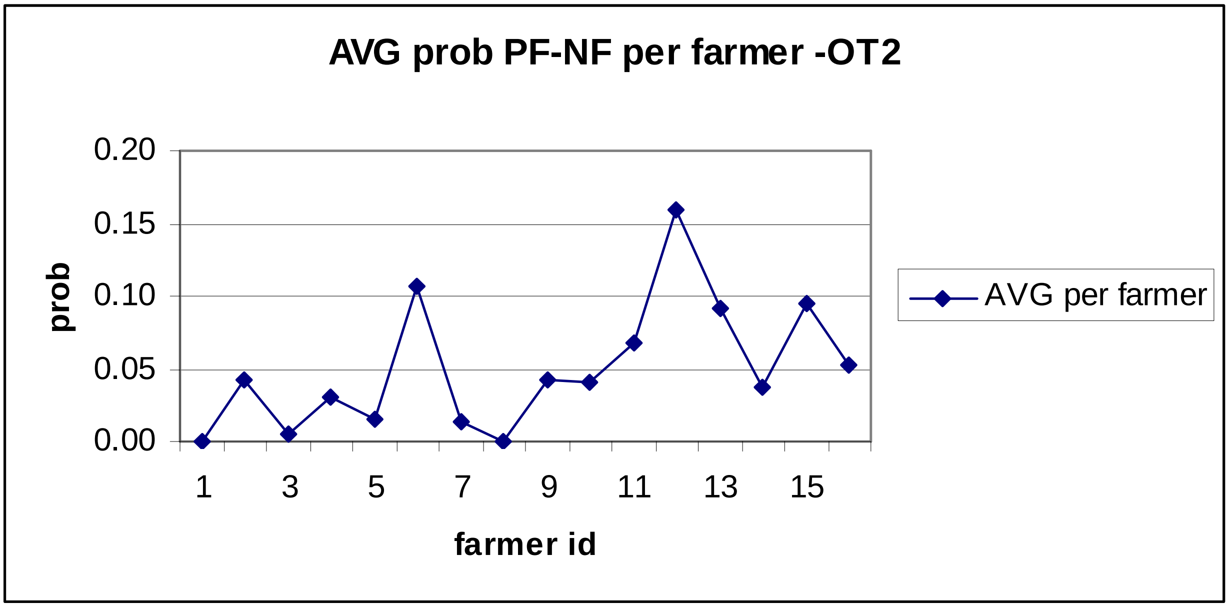
ANALYSIS OF PROBABILITIES BY OT 2

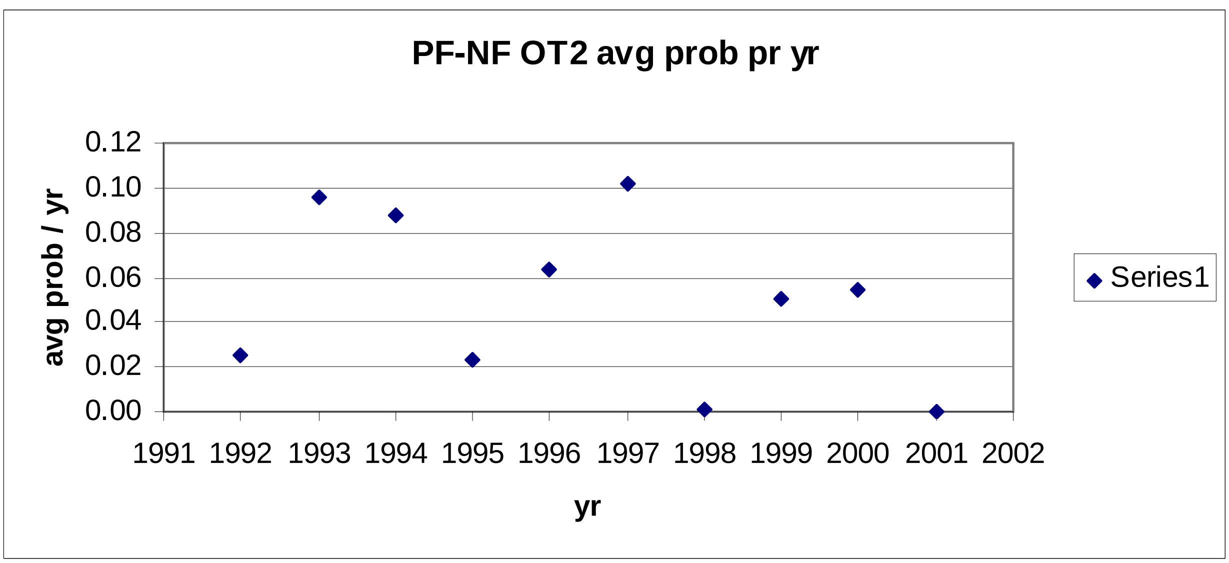


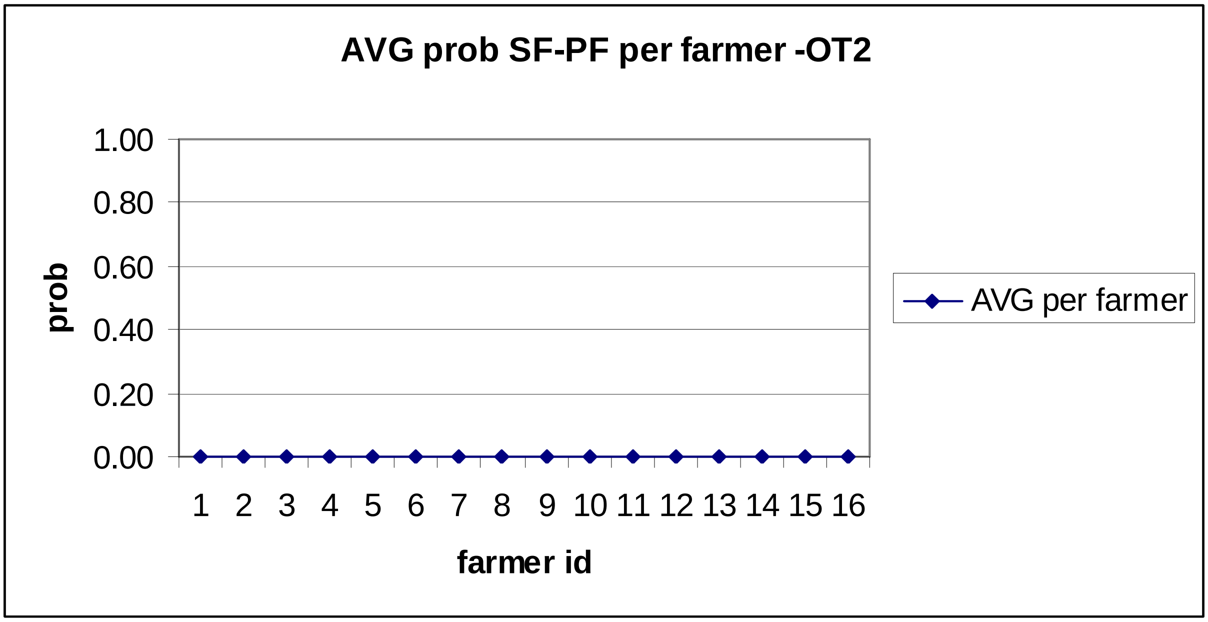



AVG prob PF-SF per farmer -OT2
prob
0.01
0.01
0.01
0.01
0.01
0.00
0.00
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
farmer id
AVG per farmer

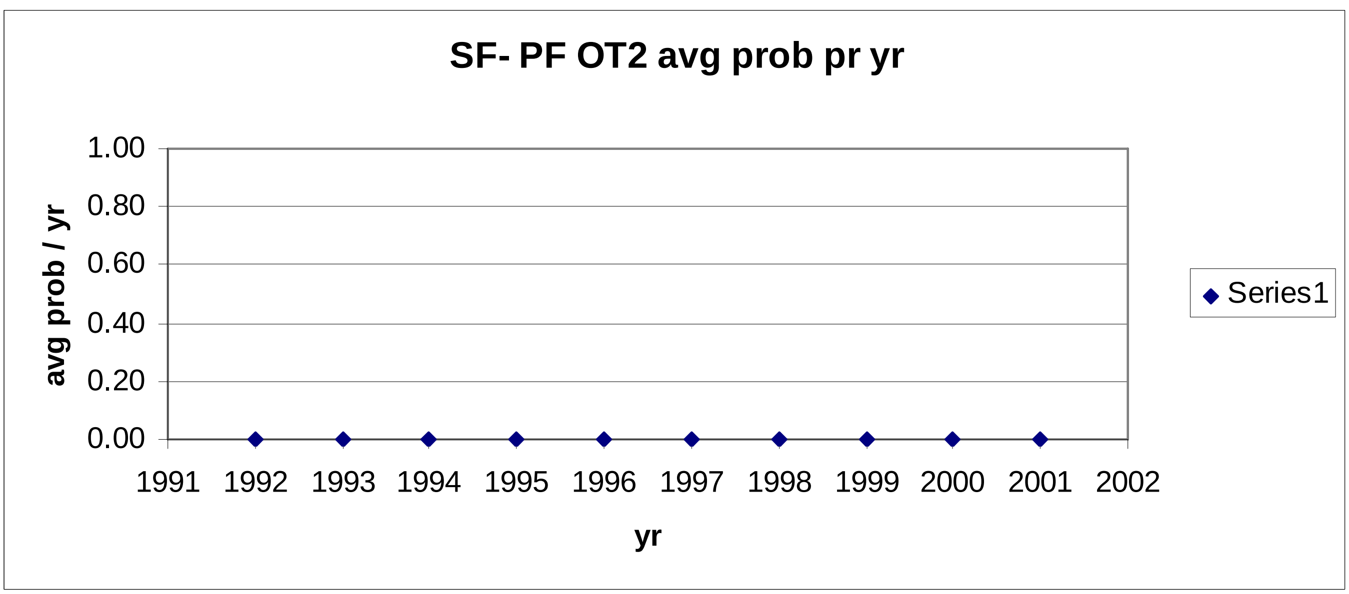

PF-SF OT2 avg prob per yr
avg prob / yr
0.01
0.01
0.01
0.01
0.01
0.00
0.00
0.00
1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002
yr
Series1


AVG prob PF-NF per farmer -OT2
prob
0.20
0.15
0.10
0.05
0.00
1
3
5
7
9
11
13
15
farmer id
AVG per farmer

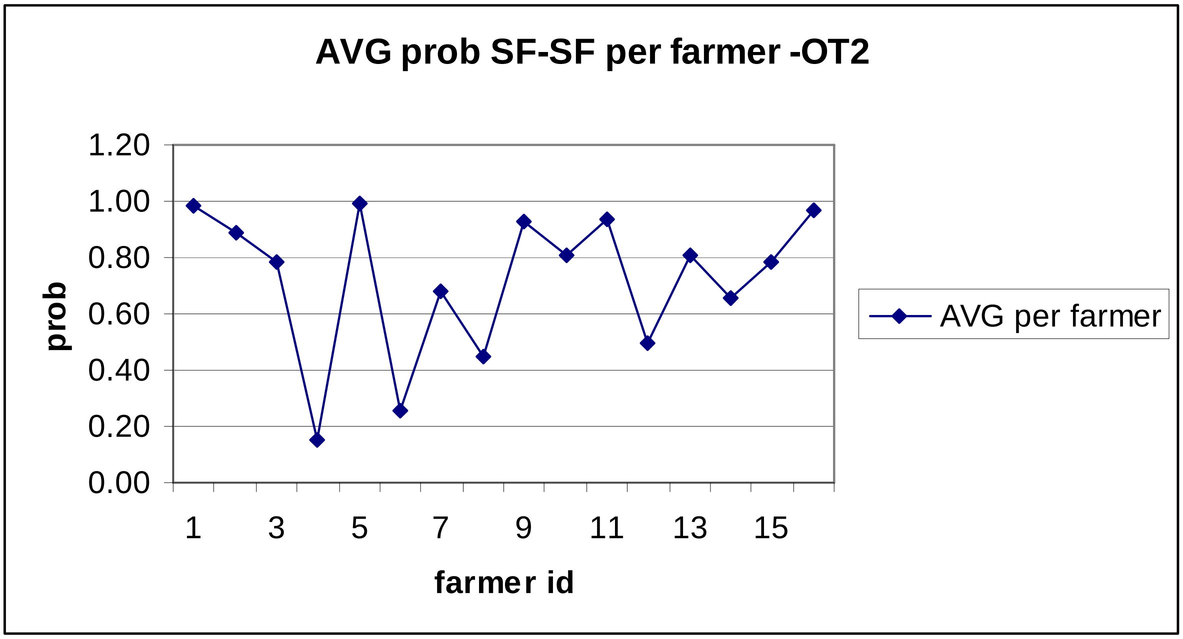

PF-NF OT2 avg prob pr yr
avg prob / yr
0.12
0.10
0.08
0.06
0.04
0.02
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

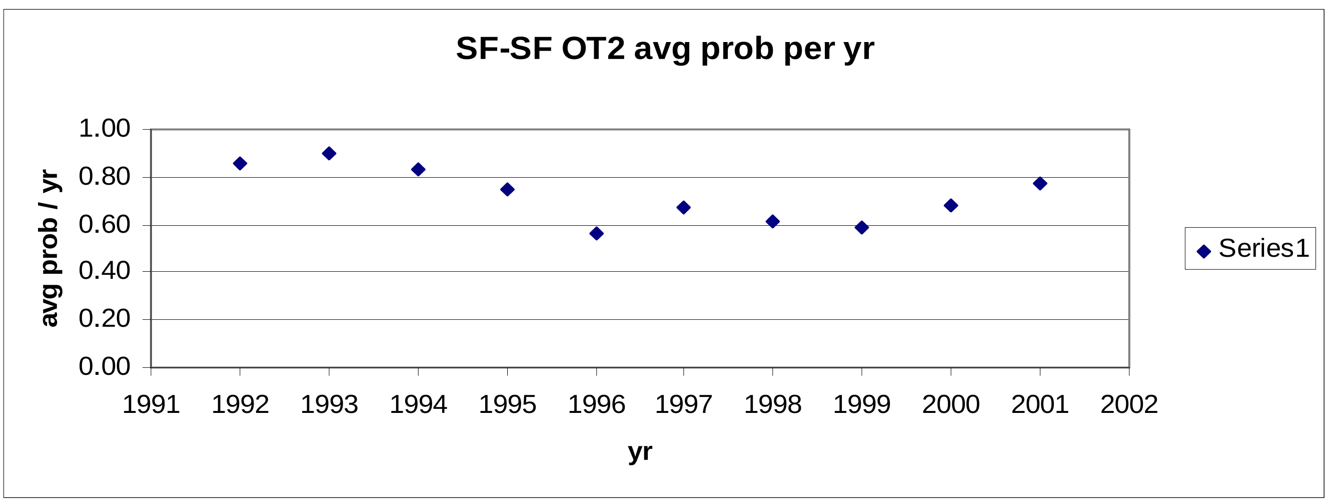


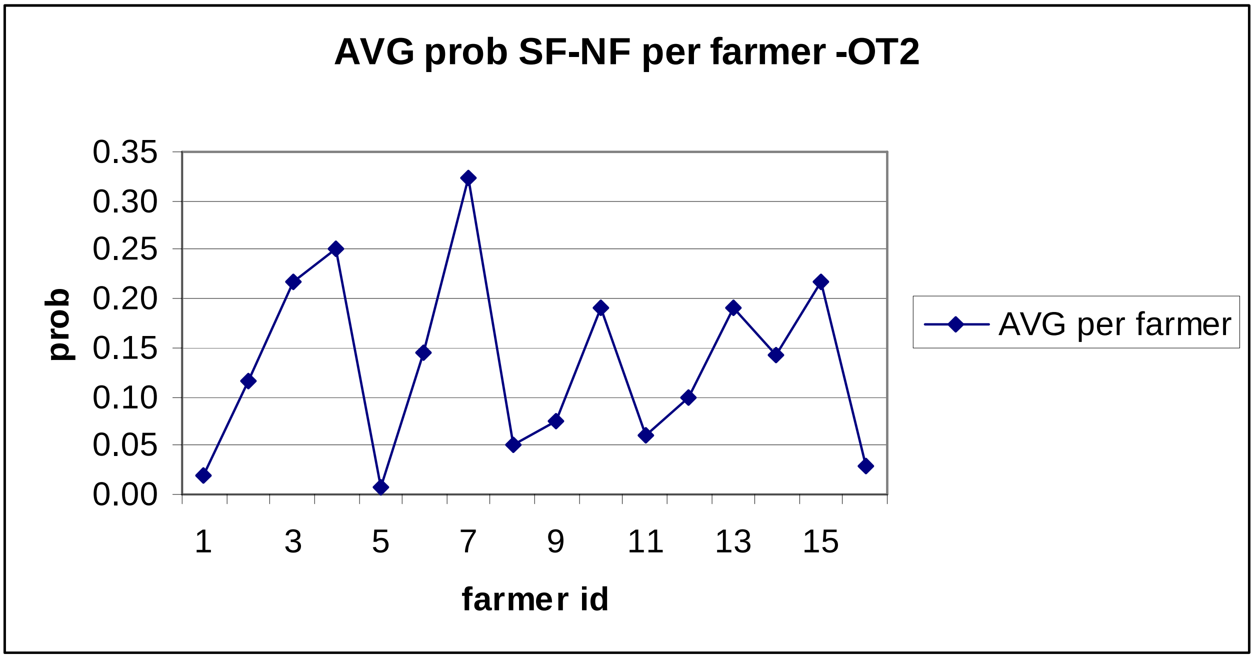
AVG prob SF-PF per farmer -OT2
1.00
0.80
0.60
0.40
0.20
0.00
prob
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
farmer id
AVG per farmer

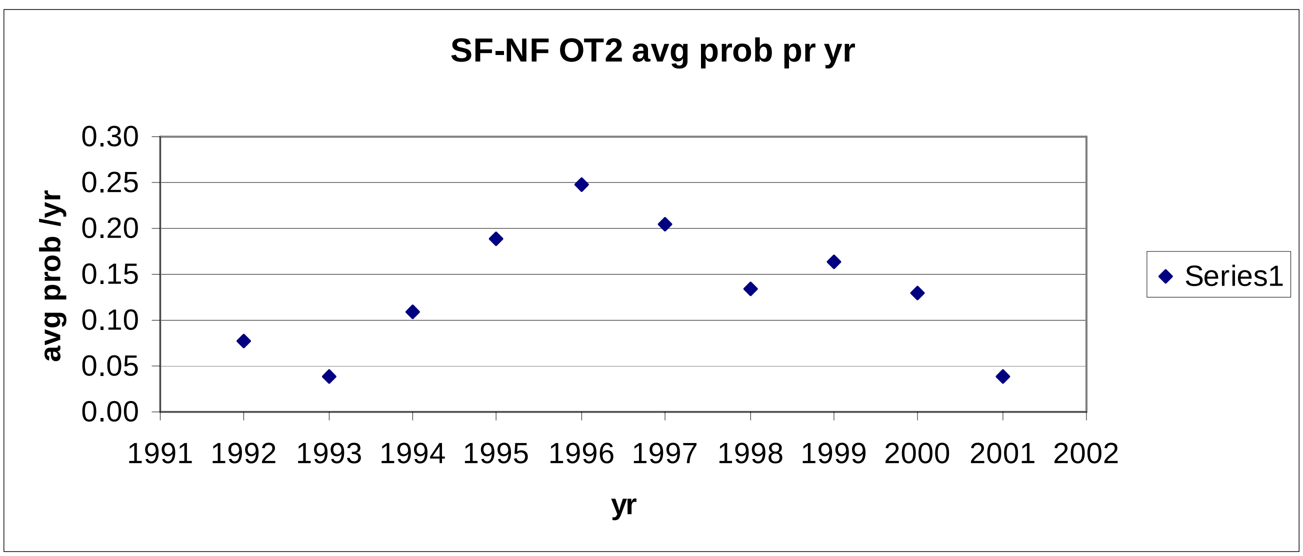


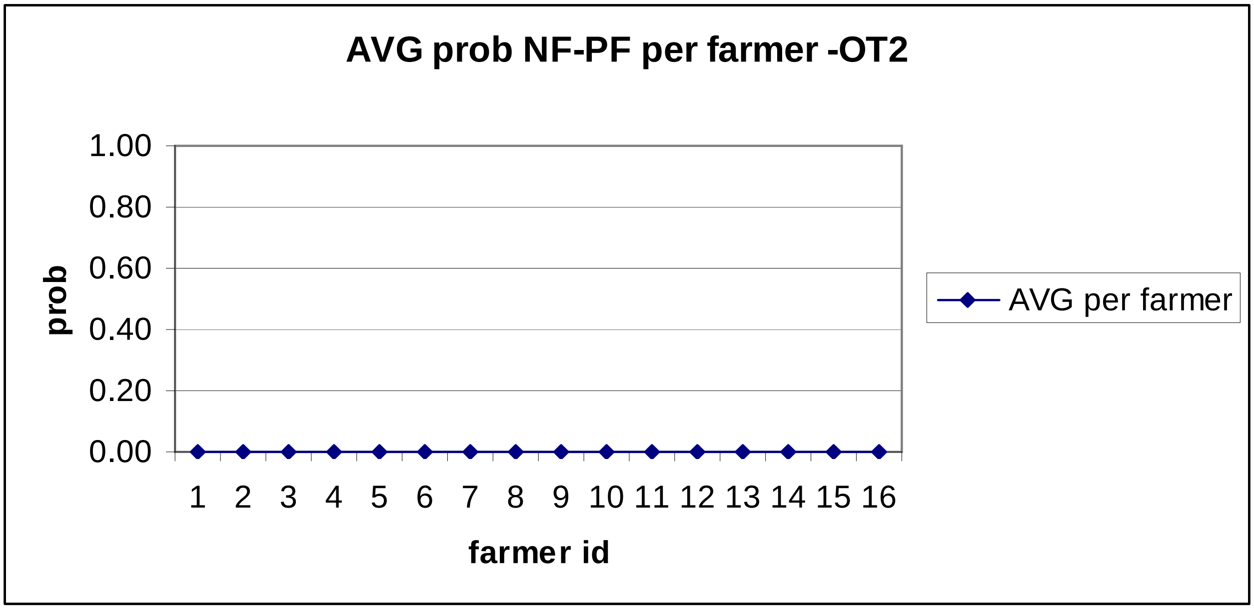
SF- PF OT2 avg prob pr yr
1.00
0.80
0.60
0.40
0.20
0.00
avg prob / yr
1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002
yr
Series1

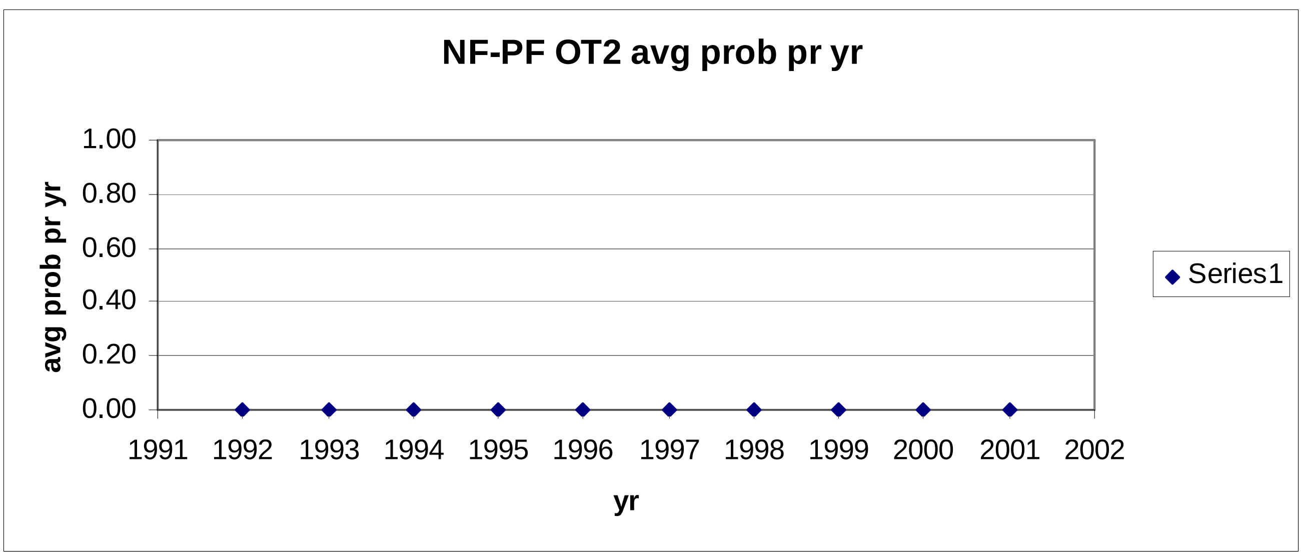

AVG prob SF-SF per farmer -OT2
prob
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1
3
5
7
9
11
13
15
farmer id
AVG per farmer

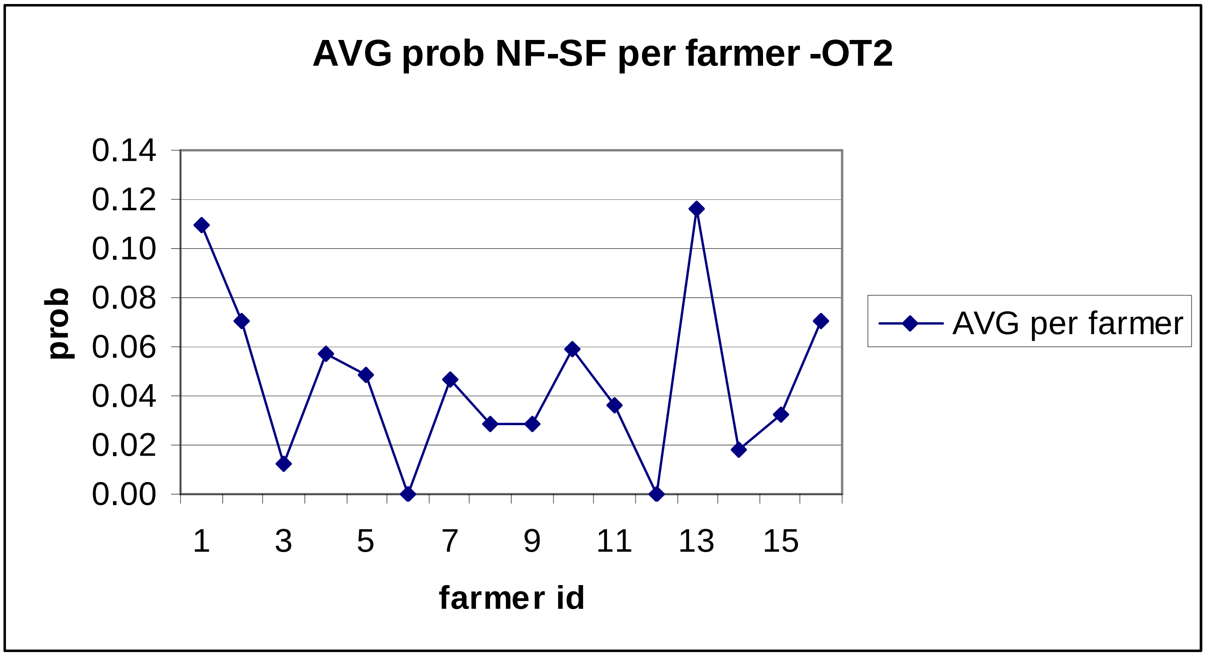

SF-SF OT2 avg prob per yr
avg prob / yr
1.00
0.80
0.60
0.40
0.20
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

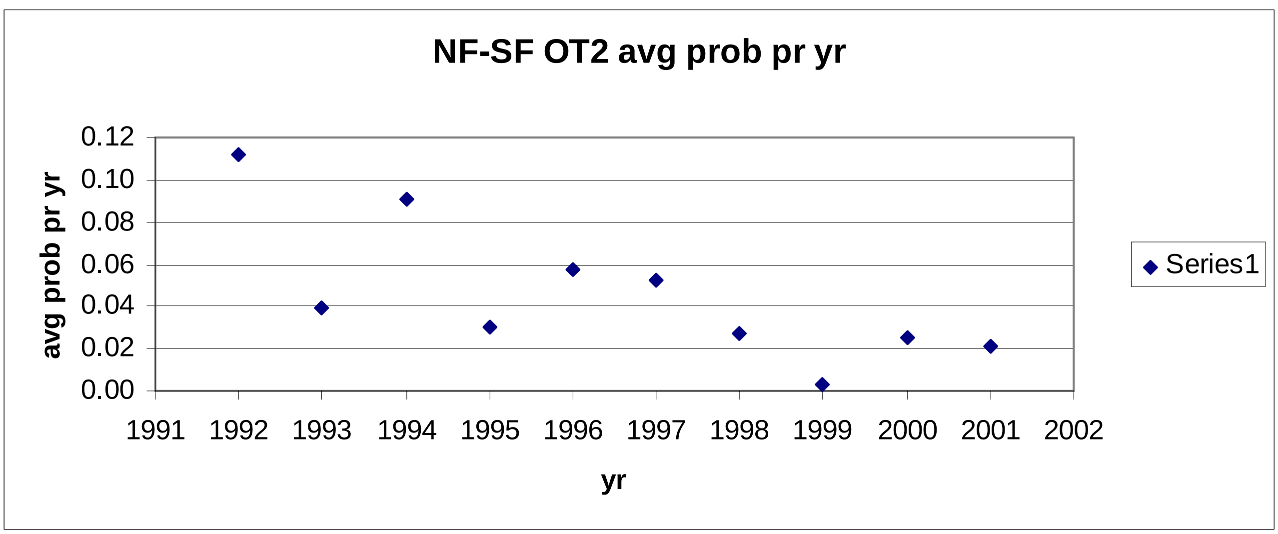

AVG prob SF-NF per farmer -OT2
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
prob
1
3
5
7
9
11
13
15
farmer id
AVG per farmer

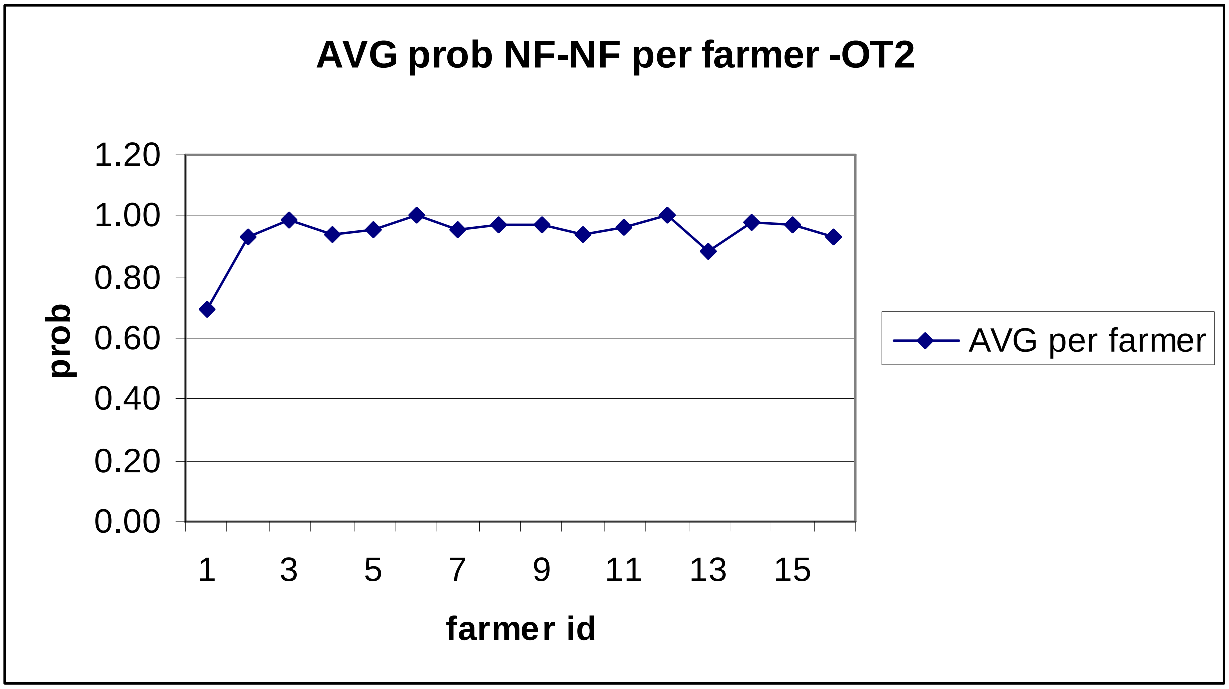

SF-NF OT2 avg prob pr yr
0.30
0.25
0.20
0.15
0.10
0.05
0.00
avg prob /yr
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

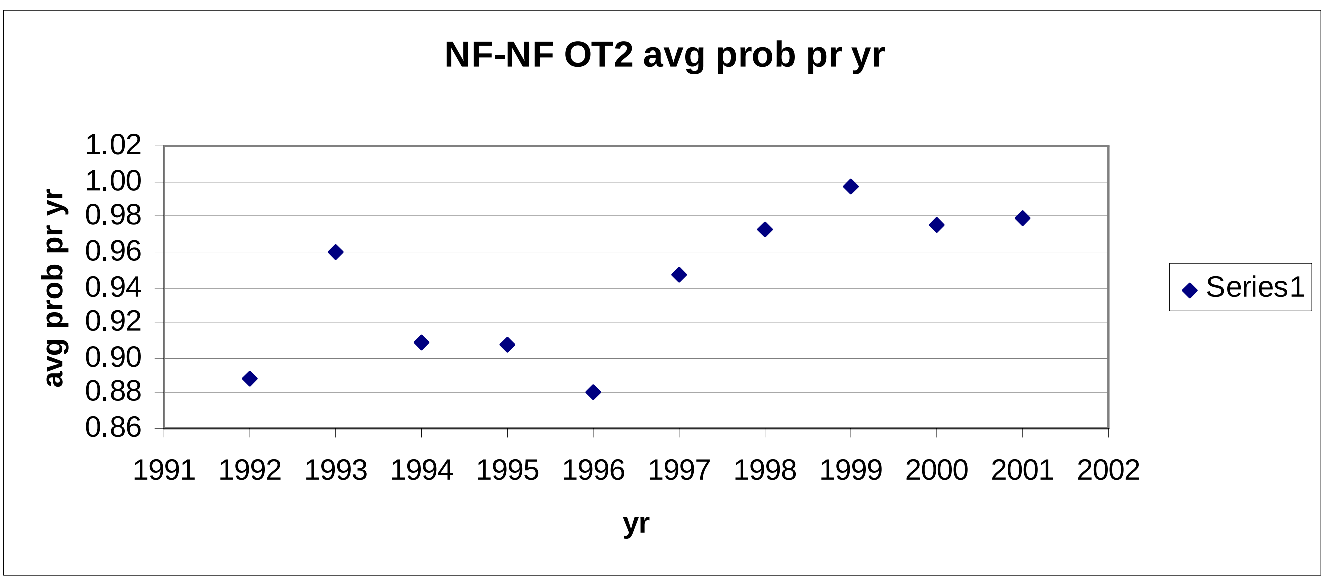

AVG prob NF-PF per farmer -OT2
prob
1.00
0.80
0.60
0.40
0.20
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
farmer id
AVG per farmer


NF-PF OT2 avg prob pr yr
avg prob pr yr
1.00
0.80
0.60
0.40
0.20
0.00
1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002
yr
Series1

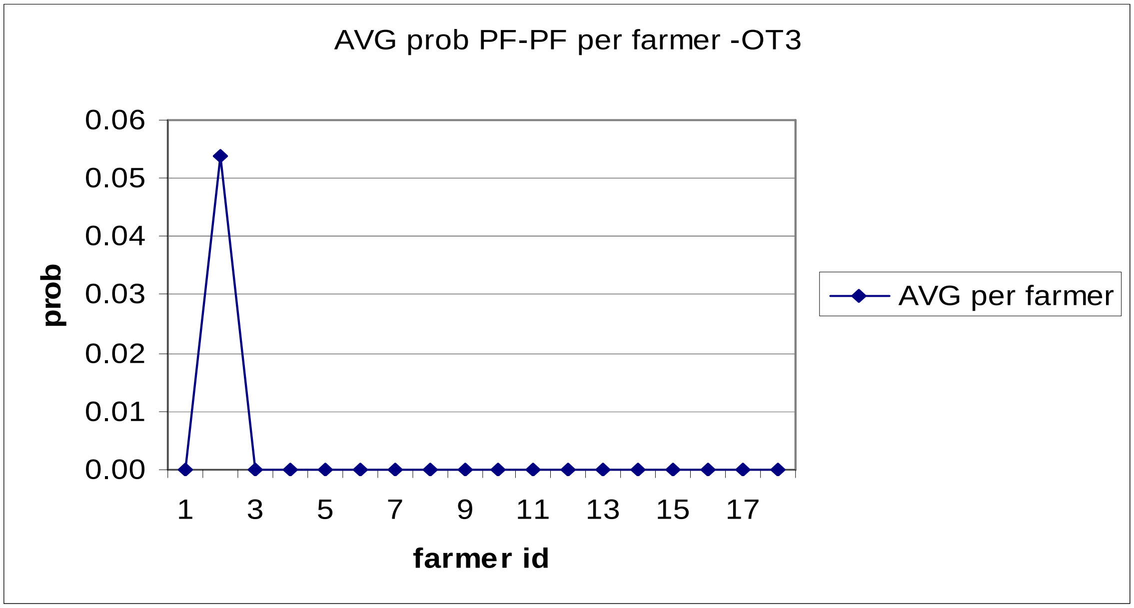

AVG prob NF-SF per farmer -OT2
prob
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
1
3
5
7
9
11
13
15
farmer id
AVG per farmer

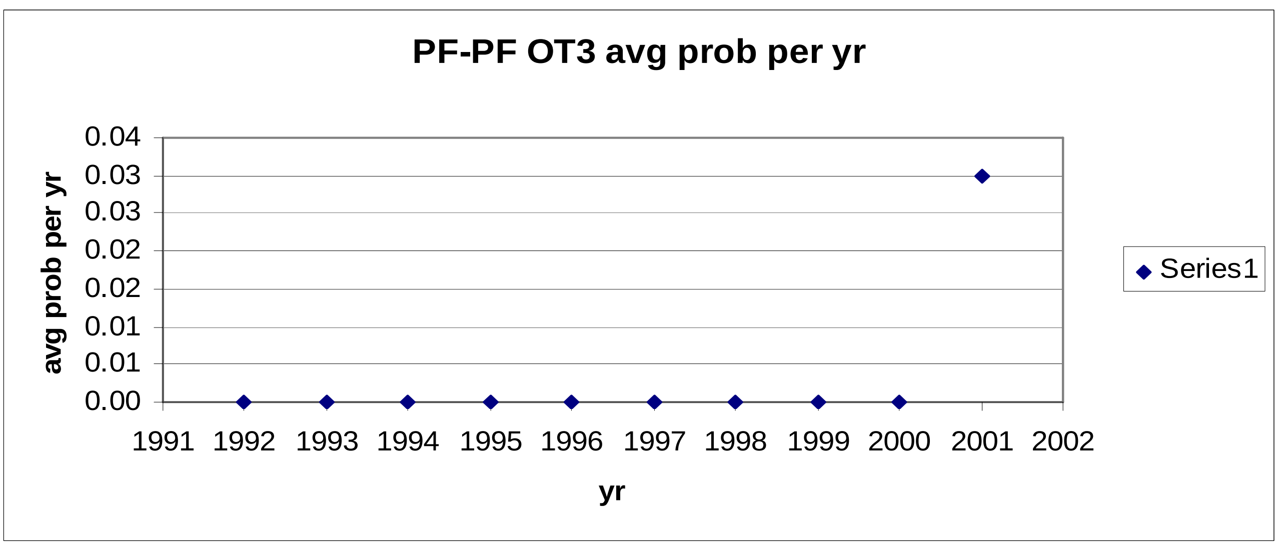

NF-SF OT2 avg prob pr yr
avg prob pr yr
0.12
0.10
0.08
0.06
0.04
0.02
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

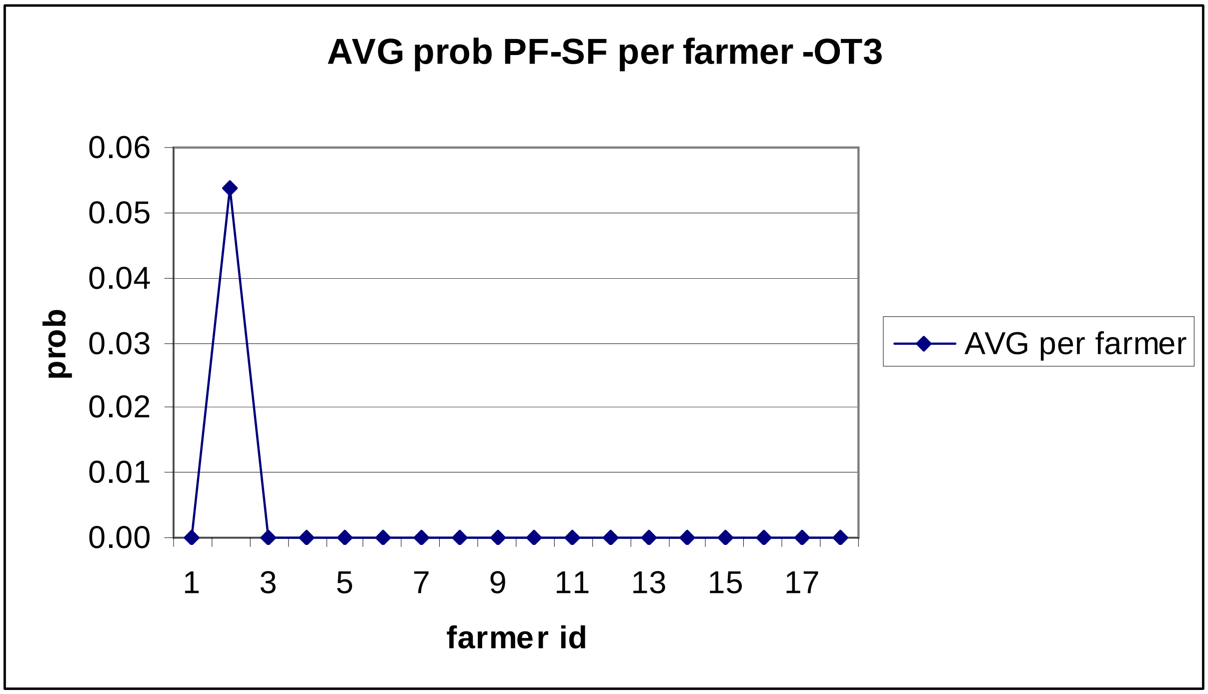

AVG prob NF-NF per farmer -OT2
prob
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1
3
5
7
9
11
13
15
farmer id
AVG per farmer

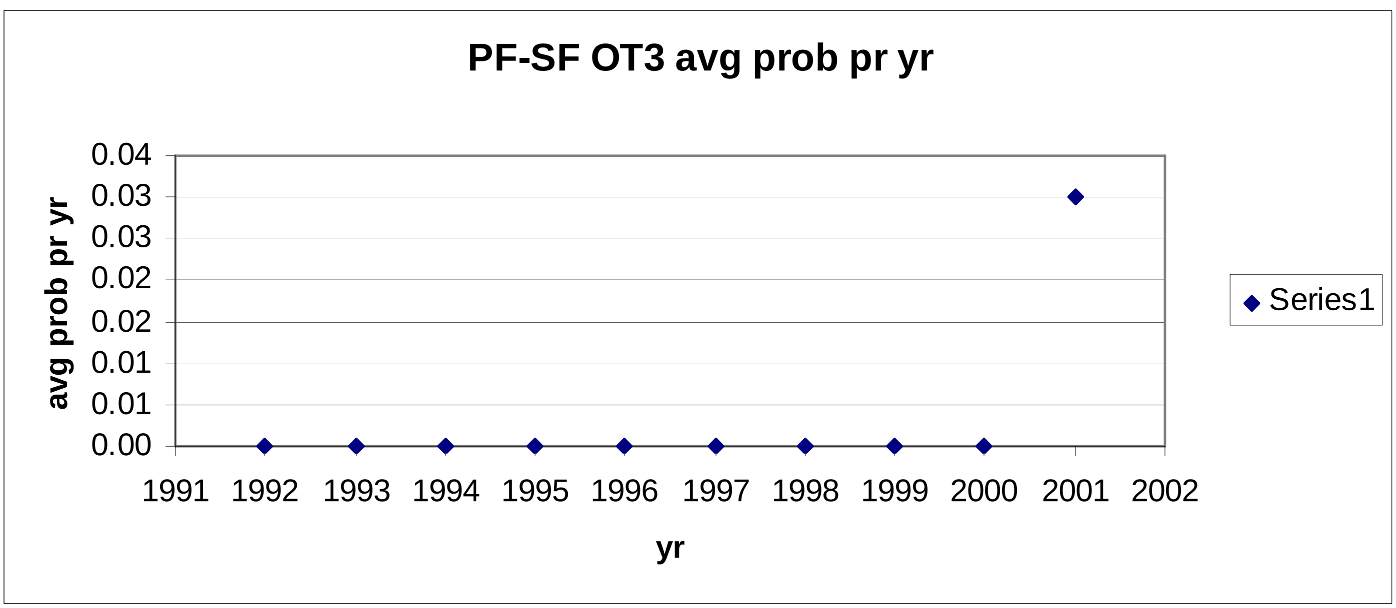

NF-NF OT2 avg prob pr yr
avg prob pr yr
1.02
1.00
0.98
0.96
0.94
0.92
0.90
0.88
0.86
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

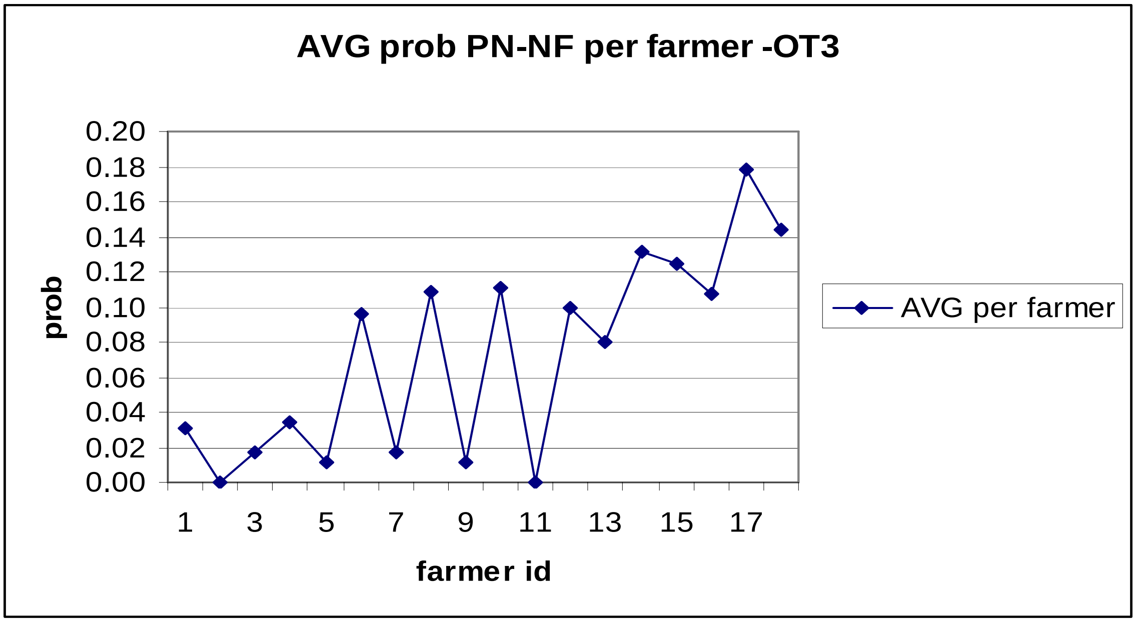
ANALYSIS PROBABILITIES – OT3

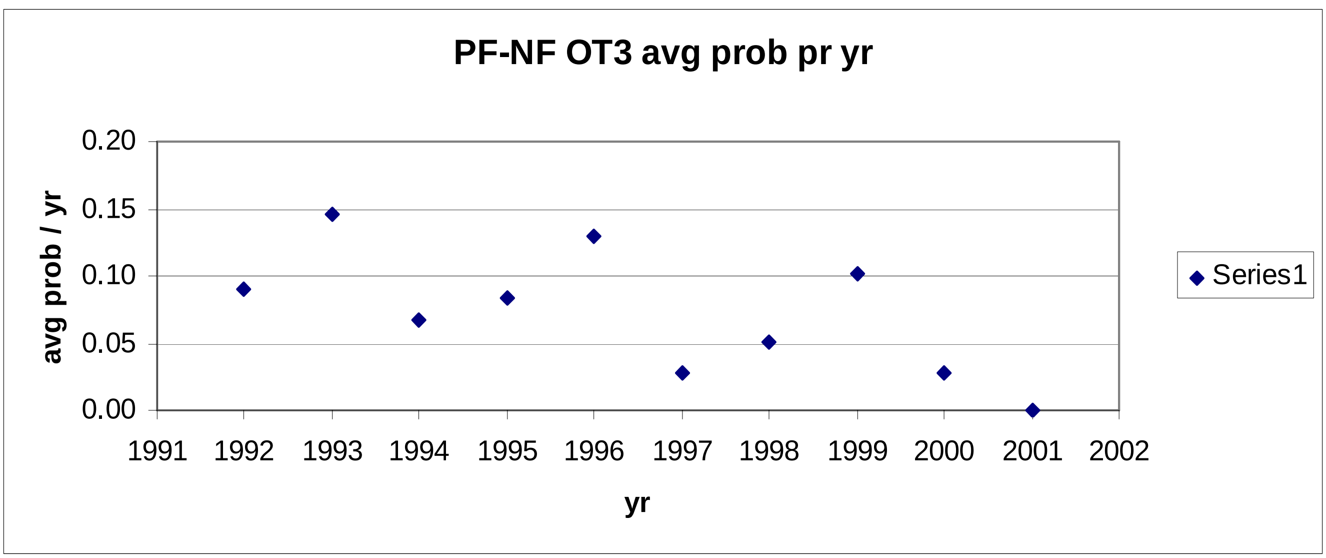


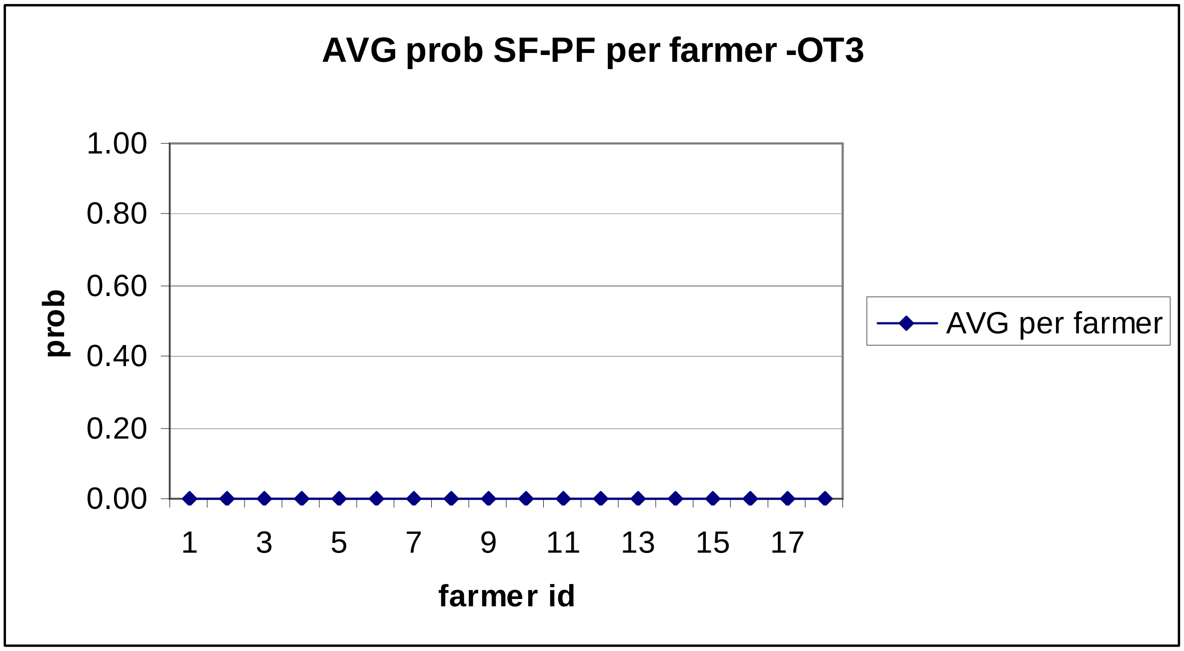


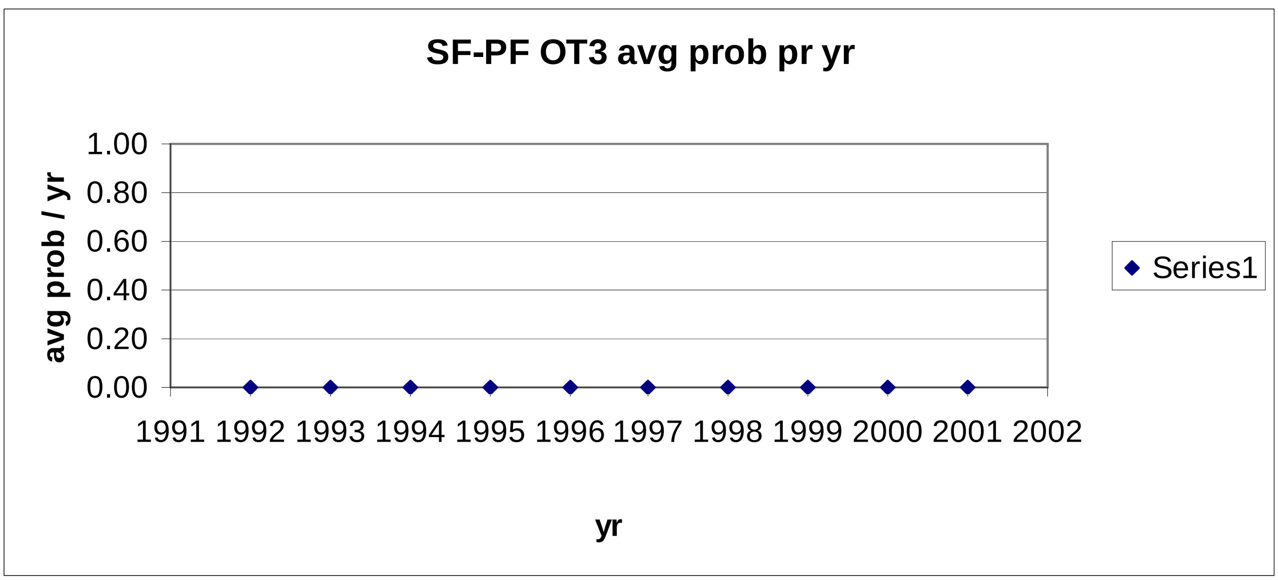

AVG prob PF-SF per farmer -OT3
prob
0.06
0.05
0.04
0.03
0.02
0.01
0.00
1
3
5
7
9
11
13
15
17
farmer id
AVG per farmer

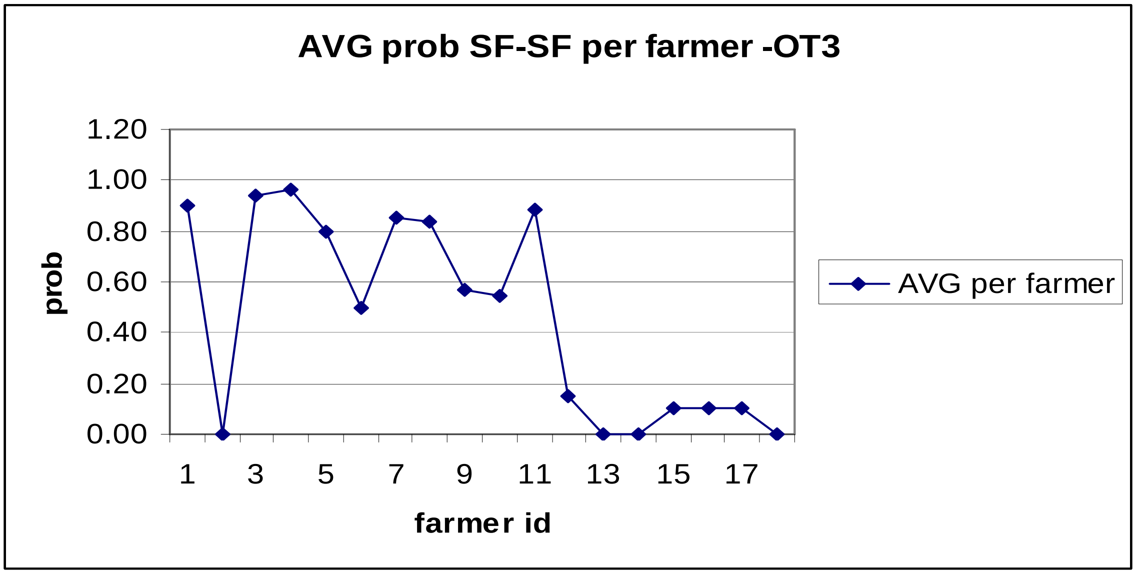

PF-SF OT3 avg prob pr yr
avg prob pr yr
0.04
0.03
0.03
0.02
0.02
0.01
0.01
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

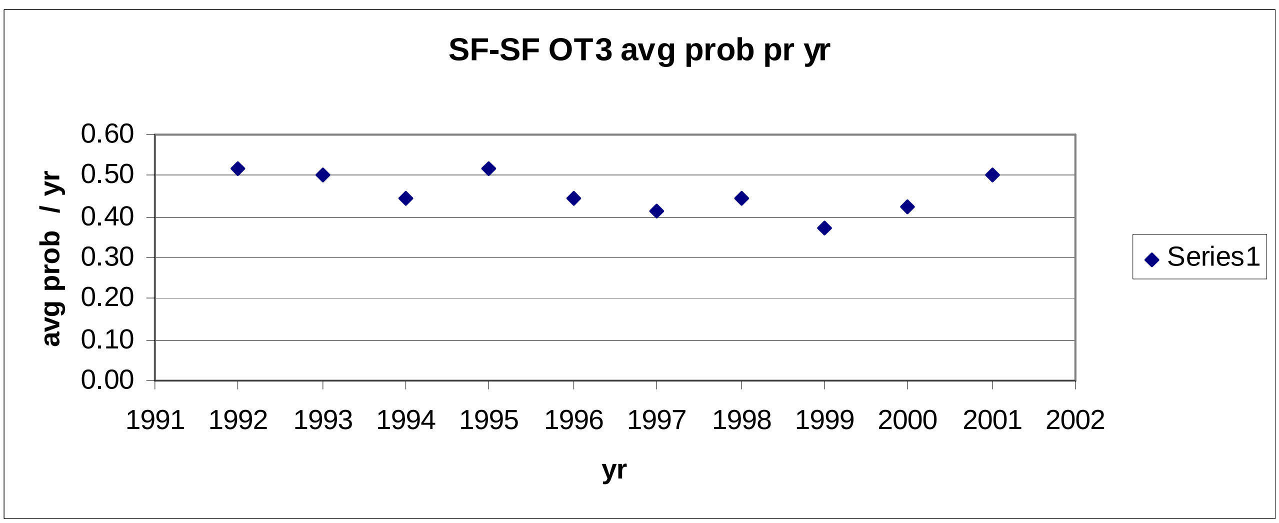

AVG prob PN-NF per farmer -OT3
prob
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
1
3
5
7
9
11
13
15
17
farmer id
AVG per farmer

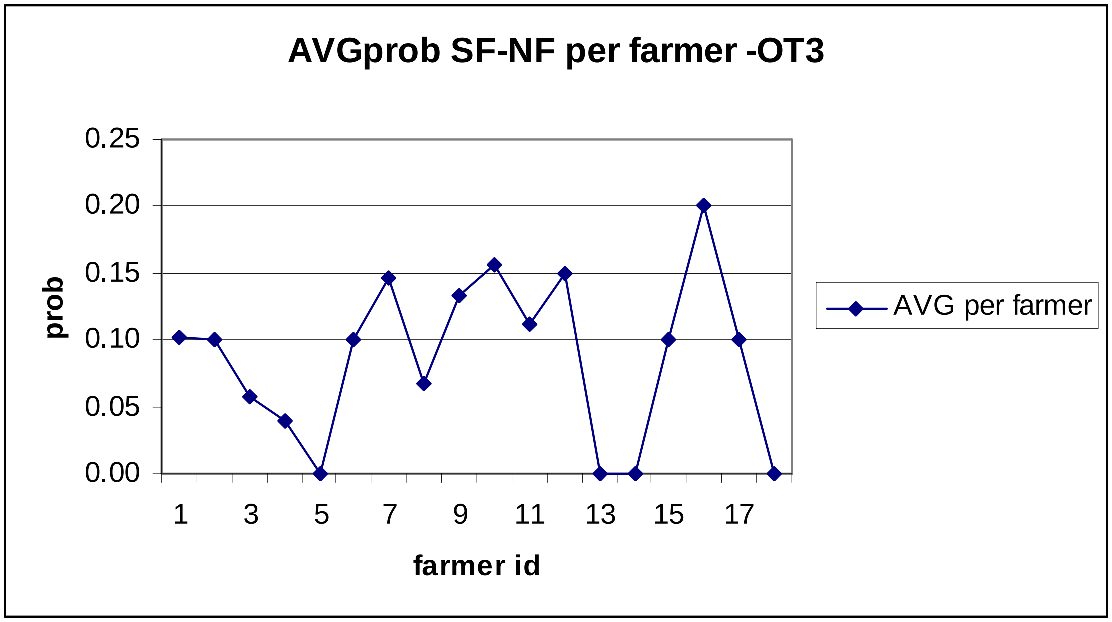

PF-NF OT3 avg prob pr yr
avg prob / yr
0.20
0.15
0.10
0.05
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

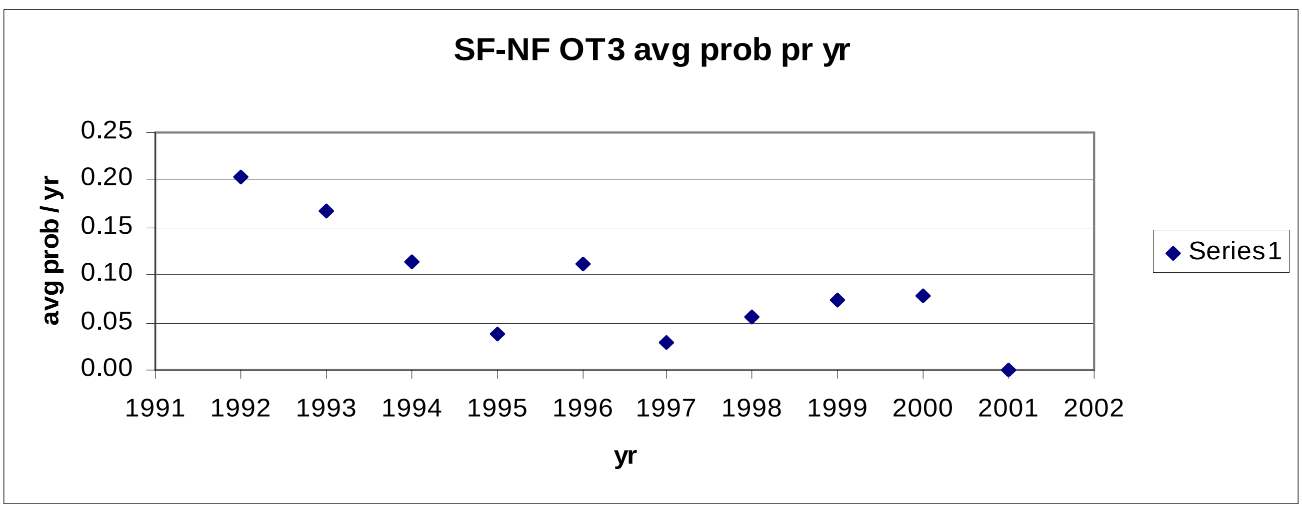

AVG prob SF-PF per farmer -OT3
prob
1.00
0.80
0.60
0.40
0.20
0.00
1
3
5
7
9
11
13
15
17
farmer id
AVG per farmer

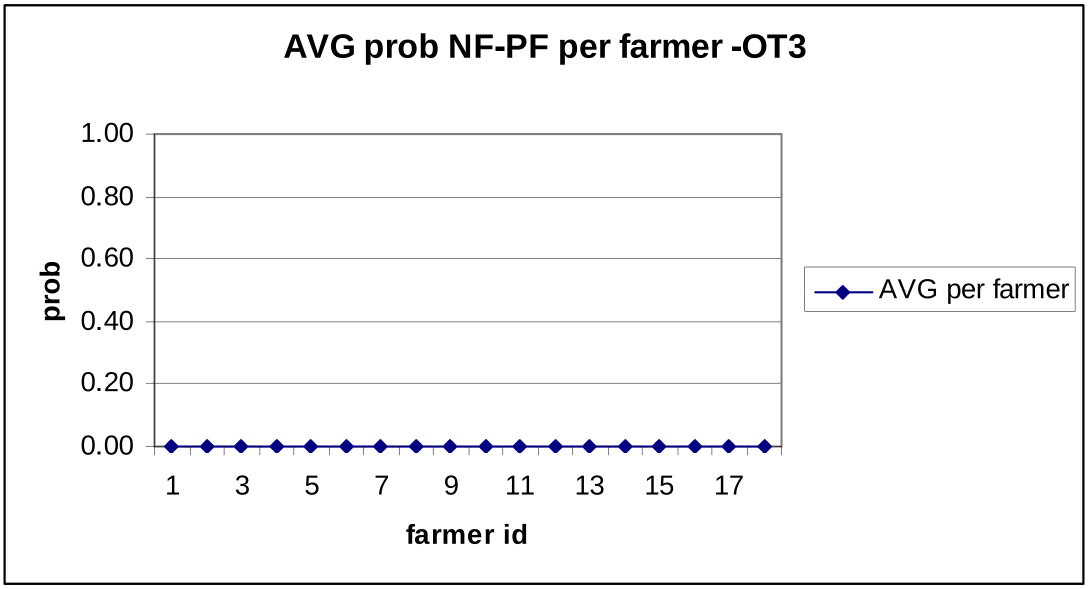

SF-PF OT3 avg prob pr yr
avg prob / yr
1.00
0.80
0.60
0.40
0.20
0.00
1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002
yr
Series1

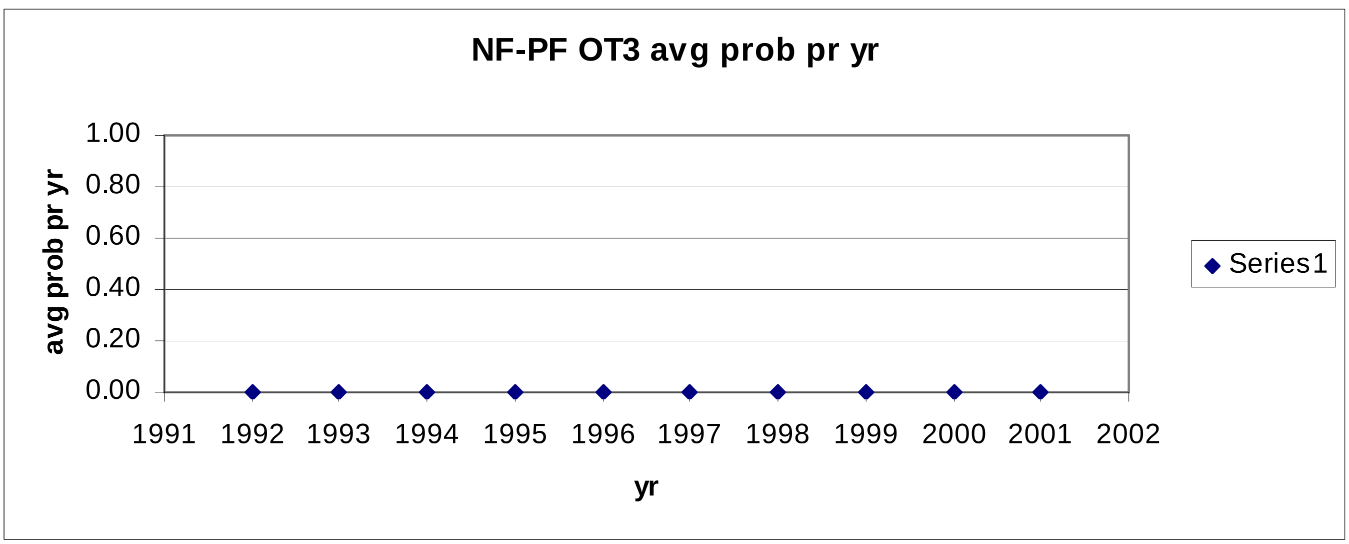

AVG prob SF-SF per farmer -OT3
prob
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1
3
5
7
9
11
13
15
17
farmer id
AVG per farmer

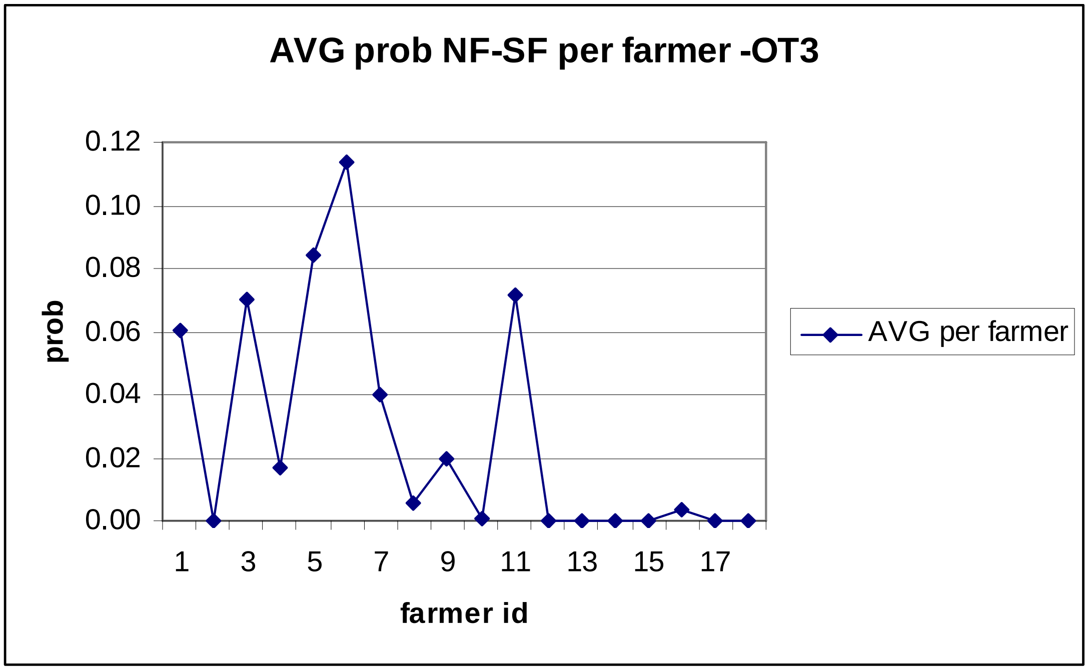

SF-SF OT3 avg prob pr yr
avg prob / yr
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1


AVGprob SF-NF per farmer -OT3
prob
0.25
0.20
0.15
0.10
0.05
0.00
1
3
5
7
9
11
13
15
17
farmer id
AVG per farmer

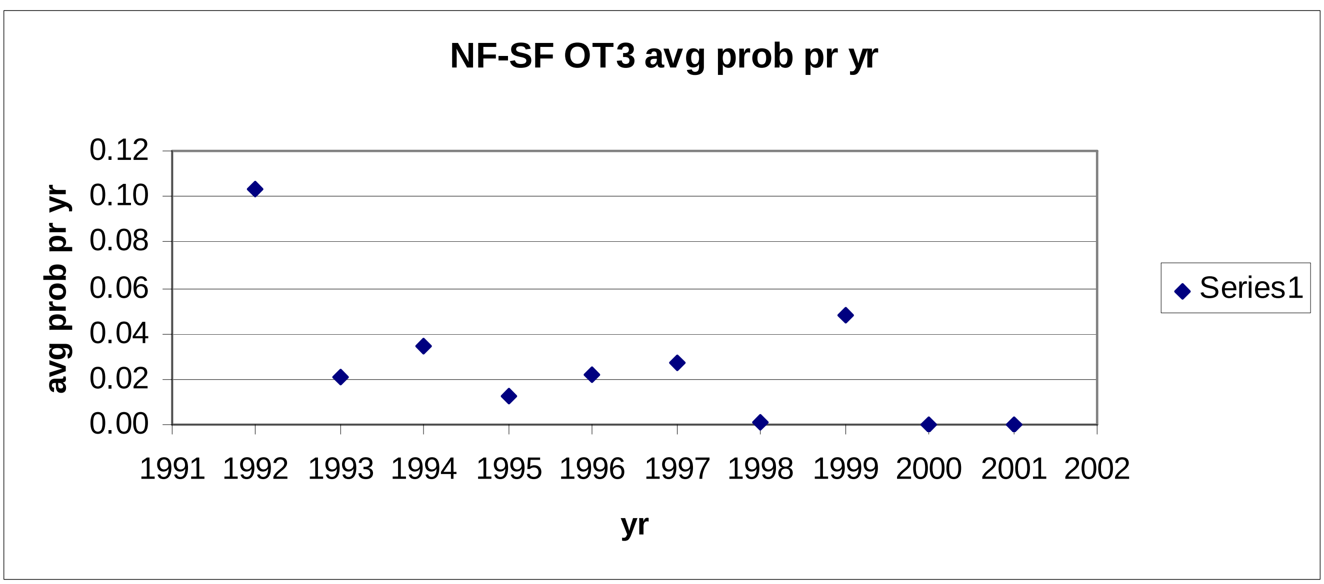

SF-NF OT3 avg prob pr yr
avg prob / yr
0.25
0.20
0.15
0.10
0.05
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1


AVG prob NF-PF per farmer -OT3
prob
1.00
0.80
0.60
0.40
0.20
0.00
1
3
5
7
9
11
13
15
17
farmer id
AVG per farmer

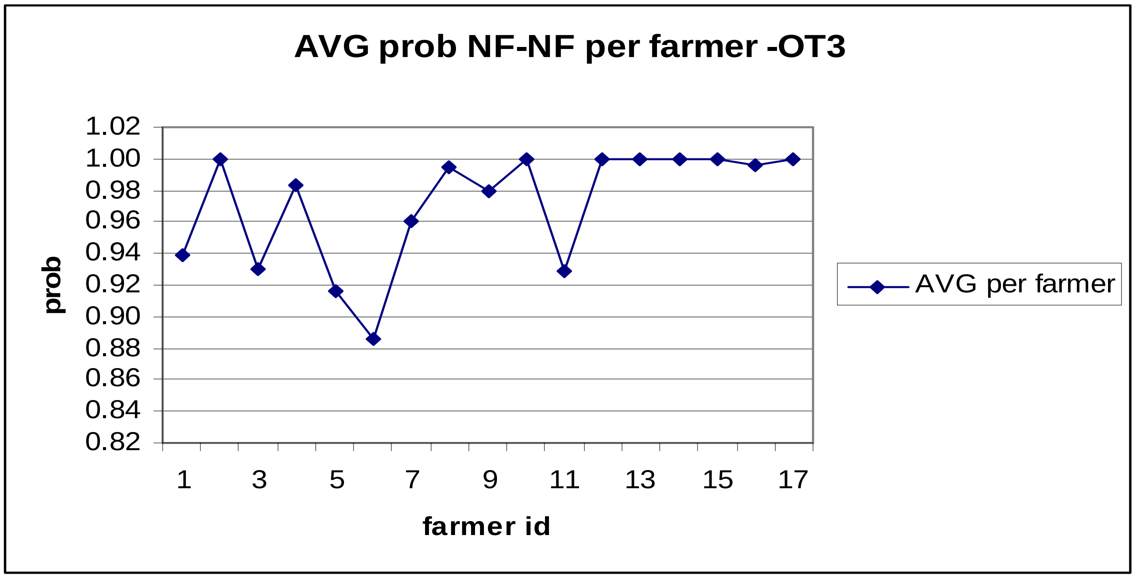

NF-PF OT3 avg prob pr yr
avg prob pr yr
1.00
0.80
0.60
0.40
0.20
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1



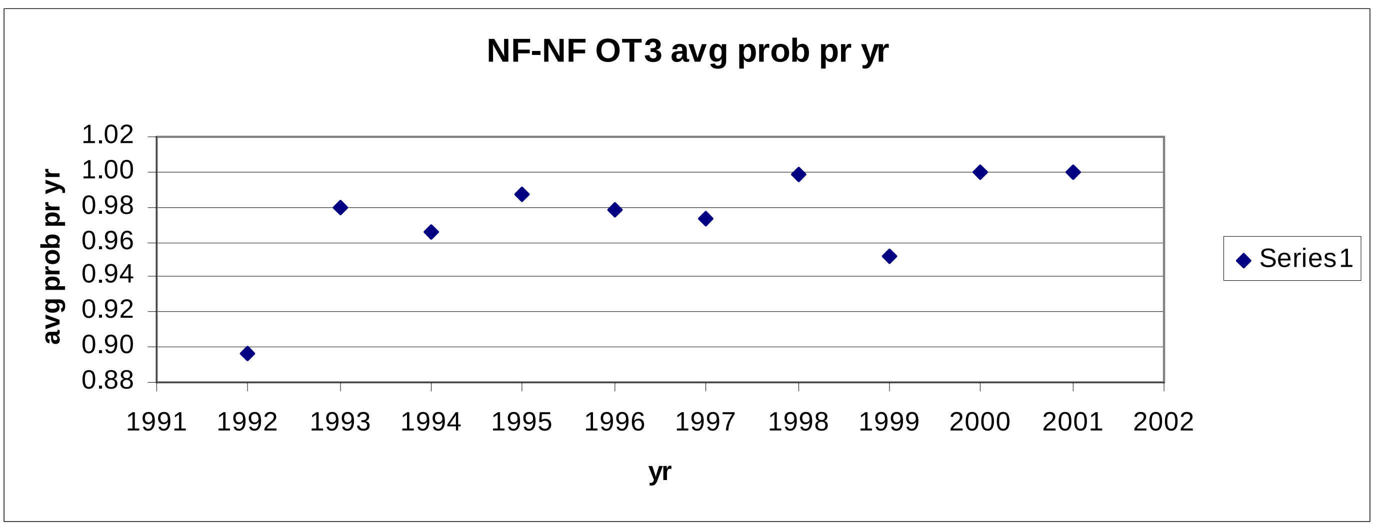
AVG prob NF-SF per farmer -OT3
prob
0.12
0.10
0.08
0.06
0.04
0.02
0.00
1
3
5
7
9
11
13
15
17
farmer id
AVG per farmer



NF-SF OT3 avg prob pr yr
avg prob pr yr
0.12
0.10
0.08
0.06
0.04
0.02
0.00
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1



AVG prob NF-NF per farmer -OT3
prob
1.02
1.00
0.98
0.96
0.94
0.92
0.90
0.88
0.86
0.84
0.82
1
3
5
7
9
11
13
15
17
farmer id
AVG per farmer



NF-NF OT3 avg prob pr yr
avg prob pr yr
1.02
1.00
0.98
0.96
0.94
0.92
0.90
0.88
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
yr
Series1

APPENDIX C

Example of a regression output for the fixed group effects model

The SAS System   04:12 Wednesday, April 2, 2008   9

The REG Procedure
Model: MODEL1
Dependent Variable: pfpf pfpf

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
d25	d25	1	-0.00712	0.05521	-0.13	0.8974
d26	d26	1	-0.01146	0.05589	-0.21	0.8376
d27	d27	1	-0.01497	0.05520	-0.27	0.7864
d28	d28	1	-0.00649	0.05515	-0.12	0.9063
d29	d29	1	-0.00367	0.05523	-0.07	0.9471
d30	d30	1	-0.00466	0.05579	-0.08	0.9334
d31	d31	1	0.00693	0.05537	0.13	0.9005
d32	d32	1	0.00691	0.05649	0.12	0.9026
d33	d33	1	-0.01168	0.05518	-0.21	0.8324
d34	d34	1	-0.03397	0.05543	-0.61	0.5402
d35	d35	1	-0.00028405	0.05681	-0.00	0.9960
d36	d36	1	-0.23382	0.05558	-4.21	<.0001
d37	d37	1	-0.00006966	0.05514	-0.00	0.9990
d38	d38	1	-0.01069	0.05535	-0.19	0.8468
d39	d39	1	0.01124	0.05615	0.20	0.8413
d40	d40	1	0.01234	0.05545	0.22	0.8240
d41	d41	1	0.03803	0.05832	0.65	0.5146
d42	d42	1	-0.00489	0.05559	-0.09	0.9299
d43	d43	1	0.00117	0.05519	0.02	0.9831
d44	d44	1	-0.00235	0.05525	-0.04	0.9661
d45	d45	1	-0.18775	0.05692	-3.30	0.0010
d46	d46	1	-0.01156	0.05518	-0.21	0.8342
d47	d47	1	-0.00262	0.05522	-0.05	0.9622
d48	d48	1	-0.01447	0.05525	-0.26	0.7935
d49	d49	1	-0.02001	0.05526	-0.36	0.7174
d50	d50	1	-0.21272	0.05620	-3.79	0.0002
d51	d51	1	-0.01665	0.05529	-0.30	0.7634
d52	d52	1	-0.00665	0.05539	-0.12	0.9044
d53	d53	1	-0.00330	0.05645	-0.06	0.9534
d54	d54	1	0.02460	0.05820	0.42	0.6727
d55	d55	1	0.02245	0.05663	0.40	0.6919
d56	d56	1	-0.00570	0.05516	-0.10	0.9177
d57	d57	1	-0.00040413	0.05515	-0.01	0.9942
d58	d58	1	-0.00351	0.05522	-0.06	0.9494
d59	d59	1	-0.00226	0.05589	-0.04	0.9677
d60	d60	1	-0.00003250	0.05593	-0.00	0.9995
d61	d61	1	0.00480	0.05519	0.09	0.9308

d62	d62	1	-0.00128	0.05519	-0.02	0.9815
d63	d63	1	-0.00206	0.05519	-0.04	0.9703
d64	d64	1	0.00430	0.05518	0.08	0.9379
d65	d65	1	-0.21894	0.05525	-3.96	<.0001
d66	d66	1	0.00349	0.05533	0.06	0.9497
d67	d67	1	-0.00332	0.05519	-0.06	0.9520
d68	d68	1	0.00548	0.05528	0.10	0.9211

The TSCSREG Procedure

Dependent Variable: pfsf pfsf

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Pr > |t|	Label
CS54	1	-0.00194	0.0138	-0.14	0.8881	Cross Sectional Effect 54
CS55	1	0.000911	0.0151	0.06	0.9517	Cross Sectional Effect 55
CS56	1	-0.002	0.0138	-0.15	0.8842	Cross Sectional Effect 56
CS57	1	-0.00129	0.0135	-0.10	0.9236	Cross Sectional Effect 57
CS58	1	-0.0018	0.0137	-0.13	0.8961	Cross Sectional Effect 58
CS59	1	-0.00157	0.0131	-0.12	0.9043	Cross Sectional Effect 59
CS60	1	-0.00187	0.0131	-0.14	0.8862	Cross Sectional Effect 60
CS61	1	-0.00249	0.0137	-0.18	0.8563	Cross Sectional Effect 61
CS62	1	-0.00184	0.0131	-0.14	0.8888	Cross Sectional Effect 62
CS63	1	-0.02065	0.0142	-1.46	0.1451	Cross Sectional Effect 63
CS64	1	-0.00138	0.0138	-0.10	0.9202	Cross Sectional Effect 64
CS65	1	-0.00151	0.0132	-0.11	0.9094	Cross Sectional Effect 65
CS66	1	-0.00272	0.0131	-0.21	0.8361	Cross Sectional Effect 66
CS67	1	-0.04756	0.0151	-3.15	0.0017	Cross Sectional Effect 67
CS68	1	-0.00085	0.0138	-0.06	0.9507	Cross Sectional Effect 68
CS69	1	-0.00315	0.0132	-0.24	0.8113	Cross Sectional Effect 69
Intercept	1	0.055576	0.0181	3.06	0.0023	Intercept
pfpf	1	-0.05722	0.00920	-6.22	<.0001	pfpf
pfnf	1	-0.04145	0.0106	-3.91	0.0001	pfnf
sfsf	1	0.001394	0.00452	0.31	0.7576	sfsf
sfnf	1	0.002567	0.00535	0.48	0.6318	sfnf
nfsf	1	-0.00427	0.0142	-0.30	0.7646	nfsf
nfnf	1	0.001154	0.0126	0.09	0.9271	nfnf

The SAS System 04:12 Wednesday, April 2, 2008 44

The TSCSREG Procedure

Dependent Variable: sfsf sfsf

Model Description

Estimation Method	FixOne
Number of Cross Sections	70
Time Series Length	10

Fit Statistics

SSE	41.7596	DFE	624
MSE	0.0669	Root MSE	0.2587
R-Square	0.6886		

F Test for No Fixed Effects

Num DF	Den DF	F Value	Pr > F
69	624	14.88	<.0001

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Pr > |t|	Label
CS1	1	0.961206	0.1159	8.29	<.0001	Cross Sectional Effect 1
CS2	1	0.999614	0.1163	8.60	<.0001	Cross Sectional Effect 2
CS3	1	0.900366	0.1157	7.78	<.0001	Cross Sectional Effect 3
CS4	1	0.972507	0.1160	8.38	<.0001	Cross Sectional Effect 4
CS5	1	0.977699	0.1157	8.45	<.0001	Cross Sectional Effect 5
CS6	1	0.853777	0.1159	7.36	<.0001	Cross Sectional Effect 6
CS7	1	0.068763	0.1173	0.59	0.5578	Cross Sectional Effect 7
CS8	1	0.888832	0.1175	7.56	<.0001	Cross Sectional Effect 8
CS9	1	0.92193	0.1161	7.94	<.0001	Cross Sectional Effect 9
CS10	1	0.779797	0.1158	6.73	<.0001	Cross Sectional Effect 10
CS11	1	0.984157	0.1159	8.49	<.0001	Cross Sectional Effect 11

The TSCSREG Procedure

Dependent Variable: sfsf sfsf

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Pr > |t|	Label
CS12	1	0.969263	0.1174	8.25	<.0001	Cross Sectional Effect 12
CS13	1	0.976348	0.1186	8.23	<.0001	Cross Sectional Effect 13
CS14	1	0.954823	0.1160	8.23	<.0001	Cross Sectional Effect 14
CS15	1	0.979087	0.1159	8.45	<.0001	Cross Sectional Effect 15
CS16	1	0.921324	0.1160	7.94	<.0001	Cross Sectional Effect 16
CS17	1	0.946153	0.1162	8.15	<.0001	Cross Sectional Effect 17
CS18	1	0.984592	0.1157	8.51	<.0001	Cross Sectional Effect 18
CS19	1	0.80088	0.1159	6.91	<.0001	Cross Sectional Effect 19
CS20	1	0.555321	0.1163	4.78	<.0001	Cross Sectional Effect 20
CS21	1	0.841231	0.1159	7.26	<.0001	Cross Sectional Effect 21
CS22	1	0.946178	0.1159	8.16	<.0001	Cross Sectional Effect 22
CS23	1	0.306092	0.1162	2.63	0.0087	Cross Sectional Effect 23
CS24	1	0.867786	0.1158	7.49	<.0001	Cross Sectional Effect 24
CS25	1	-0.00404	0.1157	-0.03	0.9722	Cross Sectional Effect 25
CS26	1	0.961835	0.1158	8.31	<.0001	Cross Sectional Effect 26
CS27	1	0.651889	0.1158	5.63	<.0001	Cross Sectional Effect 27
CS28	1	0.998021	0.1158	8.62	<.0001	Cross Sectional Effect 28
CS29	1	0.925773	0.1184	7.82	<.0001	Cross Sectional Effect 29
CS30	1	0.338878	0.1160	2.92	0.0036	Cross Sectional Effect 30
CS31	1	0.98056	0.1158	8.47	<.0001	Cross Sectional Effect 31
CS32	1	0.926545	0.1175	7.89	<.0001	Cross Sectional Effect 32

Dependent Variable: sfsf sfsf

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Pr > |t|	Label
CS33	1	0.972962	0.1158	8.41	<.0001	Cross Sectional Effect 33
CS34	1	0.953539	0.1163	8.20	<.0001	Cross Sectional Effect 34
CS35	1	0.880214	0.1165	7.55	<.0001	Cross Sectional Effect 35
CS36	1	0.652318	0.1426	4.58	<.0001	Cross Sectional Effect 36
CS37	1	0.970973	0.1158	8.39	<.0001	Cross Sectional Effect 37
CS38	1	0.927824	0.1162	7.99	<.0001	Cross Sectional Effect 38
CS39	1	0.634409	0.1160	5.47	<.0001	Cross Sectional Effect 39
CS40	1	0.960342	0.1160	8.28	<.0001	Cross Sectional Effect 40
CS41	1	0.818893	0.1162	7.05	<.0001	Cross Sectional Effect 41
CS42	1	0.4216	0.1160	3.63	0.0003	Cross Sectional Effect 42
CS43	1	0.271183	0.1173	2.31	0.0211	Cross Sectional Effect 43
CS44	1	0.992676	0.1158	8.57	<.0001	Cross Sectional Effect 44
CS45	1	0.525574	0.1181	4.45	<.0001	Cross Sectional Effect 45
CS46	1	0.973995	0.1158	8.41	<.0001	Cross Sectional Effect 46
CS47	1	0.289051	0.1160	2.49	0.0129	Cross Sectional Effect 47
CS48	1	-0.00459	0.1157	-0.04	0.9684	Cross Sectional Effect 48
CS49	1	0.708287	0.1188	5.96	<.0001	Cross Sectional Effect 49
CS50	1	0.973737	0.1158	8.41	<.0001	Cross Sectional Effect 50
CS51	1	0.938025	0.1161	8.08	<.0001	Cross Sectional Effect 51
CS52	1	0.878763	0.1166	7.53	<.0001	Cross Sectional Effect 52
CS53	1	0.550533	0.1161	4.74	<.0001	Cross Sectional Effect 53

Dependent Variable: sfsf sfsf

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Pr > |t|	Label
CS54	1	0.922034	0.1165	7.92	<.0001	Cross Sectional Effect 54
CS55	1	0.959998	0.1278	7.51	<.0001	Cross Sectional Effect 55
CS56	1	0.94461	0.1159	8.15	<.0001	Cross Sectional Effect 56
CS57	1	0.745806	0.1158	6.44	<.0001	Cross Sectional Effect 57
CS58	1	0.934664	0.1159	8.06	<.0001	Cross Sectional Effect 58
CS59	1	-0.00939	0.1159	-0.08	0.9355	Cross Sectional Effect 59
CS60	1	0.153828	0.1159	1.33	0.1850	Cross Sectional Effect 60
CS61	1	0.912292	0.1161	7.85	<.0001	Cross Sectional Effect 61
CS62	1	0.218093	0.1161	1.88	0.0607	Cross Sectional Effect 62
CS63	1	0.977272	0.1194	8.18	<.0001	Cross Sectional Effect 63
CS64	1	0.986506	0.1159	8.51	<.0001	Cross Sectional Effect 64
CS65	1	0.383158	0.1164	3.29	0.0011	Cross Sectional Effect 65
CS66	1	0.148811	0.1162	1.28	0.2007	Cross Sectional Effect 66
CS67	1	0.124039	0.1348	0.92	0.3578	Cross Sectional Effect 67
CS68	1	0.962757	0.1160	8.30	<.0001	Cross Sectional Effect 68
CS69	1	0.217761	0.1164	1.87	0.0619	Cross Sectional Effect 69
Intercept	1	-0.08137	0.1620	-0.50	0.6156	Intercept
pfpf	1	0.168369	0.0837	2.01	0.0448	pfpf
pfsf	1	0.109572	0.3548	0.31	0.7576	pfsf
pfnf	1	0.262516	0.0946	2.77	0.0057	pfnf
sfnf	1	-0.62655	0.0403	-15.54	<.0001	sfnf
nfsf	1	-0.07871	0.1262	-0.62	0.5329	nfsf
nfnf	1	-0.08994	0.1117	-0.81	0.4209	nfnf

APPENDIX D

Poolability tests by farmer and by year

POOLABILITY TEST BY YEAR

	F test time effects	Poolability test- time						
Durbin Watson	R2		SSE					
1.937	0.342	SSE - OLS pooled	6.086					
		n				70		
		T				10		
		k	6	vars	plus a constant, equals	7		
Durbin Watson	R2		SSE					
1.95	0.526	SSE - yr 92	0.5395					
1.879	0.675	SSE - yr 93	0.70123					
1.979	0.489	SSE - yr 94	0.98783					
2.095	0.336	SSE - yr 95	0.77147					
1.842	0.534	SSE - yr 96	1.01533					
2.079	0.334	SSE - yr 97	0.71946					
2.078	0.452	SSE - yr 98	0.91387					
2.005	0.589	SSE - yr 99	1.88722					
2.049	0.129	SSE - yr00	1.04346					
2.056	0.023	SSE - yr01	0.10701					
		Sum e'e	8.68638					
	numerator	SSE OLS pooled - sum	-2.60038	divided by	(T-1)K	63	equals	-0.04128
	denominator	sum SSE	8.68638	divided by	T (n-K)	630	equals	0.013788
		F=	-2.993629107					
	poolable trough time							

POOLABILITY TEST BY FARMER

F test group effects	Poolability test by farmer							
R2		SSE						
0.342	SSE - OLS pooled	6.086						
	n				70			
	T				10			
	k	6	vars	plus a constant, equals	7			
R2		SSE						
1	SSE - farmer 1							
1	SSE - farmer 2	8.75669E-16						
1	SSE - farmer 3	0						
1	SSE - farmer 4	1.01679E-15						
1	SSE - farmer 5	0						
1	SSE - farmer 6	0						
nc	SSE - farmer 7	0						
1	SSE - farmer 8	0						
nc	SSE - farmer 9	0						
1	SSE - farmer 10							
1	SSE - farmer 11	1.16762E-15						
0.386	SSE - farmer 12	0.54987						
nc	SSE - farmer 13	0						
1	SSE - farmer 14	0						
1	SSE - farmer 15	1.82458E-15						
1	SSE - farmer 16							
1	SSE - farmer 17	4.06023E-15						
1	SSE - farmer 18	0						
1	SSE - farmer 19							
1	SSE - farmer 20	0						
1	SSE - farmer 21	0						
1	SSE - farmer 22	5.72782E-15						
1	SSE - farmer 23	1.54702E-15						
1	SSE - farmer 24	0						
1	SSE - farmer 25	0						
1	SSE - farmer 26							
1	SSE - farmer 27							
1	SSE - farmer 28	7.11846E-16						
0.54	SSE - farmer 29	0.56319						

1	SSE - farmer 30	0						
1	SSE - farmer 31	0						
1	SSE - farmer 32	0						
1	SSE - farmer 33							
1	SSE - farmer 34	0						
1	SSE - farmer 35							
nc	SSE - farmer 36	0						
1	SSE - farmer 37	0						
1	SSE - farmer 38	0						
1	SSE - farmer 39	0						
nc	SSE - farmer 40	0						
1	SSE - farmer 41	0						
1	SSE - farmer 42							
0.259	SSE - farmer 43	0.66667						
1	SSE - farmer 44	0						
0.317	SSE - farmer 45	0.69967						
1	SSE - farmer 46	2.26261E-15						
1	SSE - farmer 47	0						
1	SSE - farmer 48	0						
1	SSE - farmer 49	0						
1	SSE - farmer 50	0						
1	SSE - farmer 51	7.67899E-16						
nc	SSE - farmer 52	0						
1	SSE - farmer 53	0						
1	SSE - farmer 54	0						
1	SSE - farmer 55							
1	SSE - farmer 56							
1	SSE - farmer 57							
1	SSE - farmer 58	0						
1	SSE - farmer 59	0						
1	SSE - farmer 60	0						
1	SSE - farmer 61	0						
1	SSE - farmer 62	0						
0.734	SSE - farmer 63	0.26104						
1	SSE - farmer 64							
1	SSE - farmer 65	0						
1	SSE - farmer 66	0						
0.101	SSE - farmer 67	0.88889						
1	SSE - farmer 68							
1	SSE - farmer 69	0						
1	SSE - farmer 70	0						
	Sum e'e	3.07946						
numerator	SSE OLS pooled - sum	3.00654	divided by	(n-1)K	483	equals	0.00622	
denominator	sum SSE	3.07946	divided by	n (T-K)	210	equals	0.01466	

	F=	0.424487184						
The small F statistic does not reject the null hypothesis	Non poolable by farmer.							
in favor of poolable panel data with respect to farmers.								

APPENDIX E

Example of a regression output for the fixed time effects model

```
                          The SAS System       11:54 Saturday, March 29, 2008  13

                         The REG Procedure
                           Model: MODEL1

              Test 1 Results for Dependent Variable pfnf

                                       Mean
       Source             DF         Square    F Value    Pr > F

       Numerator          69        0.06240       5.27    <.0001
       Denominator       624        0.01183

                          The SAS System       11:54 Saturday, March 29, 2008  14

                         The REG Procedure
                           Model: MODEL1
                    Dependent Variable: pfnf pfnf

                Number of Observations Read         700
                Number of Observations Used         700


                        Analysis of Variance

                                 Sum of           Mean
Source                  DF      Squares         Square    F Value    Pr > F

Model                   15      6.59382        0.43959      26.89    <.0001
Error                  684     11.18036        0.01635
Corrected Total        699     17.77418


             Root MSE              0.12785    R-Square     0.3710
             Dependent Mean        0.06551    Adj R-Sq     0.3572
             Coeff Var           195.16680


                         Parameter Estimates

                                  Parameter       Standard
Variable     Label        DF       Estimate          Error    t Value    Pr > |t|

Intercept    Intercept     1        0.38387        0.05301       7.24      <.0001
y2           y2            1        0.01349        0.02168       0.62      0.5340
y3           y3            1        0.00714        0.02174       0.33      0.7427
y4           y4            1       -0.01482        0.02177      -0.68      0.4963
y5           y5            1        0.00959        0.02170       0.44      0.6589
y6           y6            1       -0.01118        0.02170      -0.52      0.6065
```

y7	y7	1	-0.01240	0.02179	-0.57	0.5693
y8	y8	1	0.01169	0.02176	0.54	0.5913
y9	y9	1	-0.05055	0.02189	-2.31	0.0213
y10	y10	1	-0.07280	0.02188	-3.33	0.0009
pfpf	pfpf	1	-0.38410	0.02026	-18.96	<.0001
sfnf	sfnf	1	0.01059	0.02029	0.52	0.6018
nfnf	nfnf	1	0.02480	0.04589	0.54	0.5890
pfsf	pfsf	1	-0.29336	0.16364	-1.79	0.0735
sfsf	sfsf	1	0.01899	0.01217	1.56	0.1191
nfsf	nfsf	1	-0.01486	0.05892	-0.25	0.8010

The REG Procedure
Model: MODEL1
Dependent Variable: pfnf pfnf

Number of Observations Read	700
Number of Observations Used	700

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	6.59382	0.43959	26.89	<.0001
Error	684	11.18036	0.01635		
Corrected Total	699	17.77418			

Root MSE	0.12785	R-Square	0.3710
Dependent Mean	0.06551	Adj R-Sq	0.3572
Coeff Var	195.16680		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	0.38387	0.05301	7.24	<.0001
y2	y2	1	0.01349	0.02168	0.62	0.5340
y3	y3	1	0.00714	0.02174	0.33	0.7427
y4	y4	1	-0.01482	0.02177	-0.68	0.4963
y5	y5	1	0.00959	0.02170	0.44	0.6589
y6	y6	1	-0.01118	0.02170	-0.52	0.6065
y7	y7	1	-0.01240	0.02179	-0.57	0.5693
y8	y8	1	0.01169	0.02176	0.54	0.5913
y9	y9	1	-0.05055	0.02189	-2.31	0.0213
y10	y10	1	-0.07280	0.02188	-3.33	0.0009
pfpf	pfpf	1	-0.38410	0.02026	-18.96	<.0001
sfnf	sfnf	1	0.01059	0.02029	0.52	0.6018
nfnf	nfnf	1	0.02480	0.04589	0.54	0.5890
pfsf	pfsf	1	-0.29336	0.16364	-1.79	0.0735
sfsf	sfsf	1	0.01899	0.01217	1.56	0.1191

nfsf	nfsf	1	-0.01486	0.05892	-0.25	0.8010

The REG Procedure
Model: MODEL1

Test 1 Results for Dependent Variable pfnf

Source	DF	Mean Square	F Value	Pr > F
Numerator	9	0.05640	3.45	0.0004
Denominator	684	0.01635		

APPENDIX F

SPSS output for the Multinomial logit regression model

Regression
Multinomial logistic regression.
Sample: pathway subsample NU & RM, survey Rondônia 1992

Warnings

The number of dimensions requested for the Observed and Predicted Frequencies table has exceeded the limit (20) of the pivot table subsystem. Therefore, this table will not be produced.
There are 56 (31.6%) cells (i.e., dependent variable levels by subpopulations) with zero frequencies.

Case Processing Summary

		N	Marginal Percentage
1= NF , 2=SF 3=PF 0= nothing	1.00	1115	36.0%
	2.00	746	24.1%
	3.00	1239	40.0%
R11.1-MUN	2	1900	61.3%
	3	1200	38.7%
R-OwnerType	1	1500	48.4%
	2	800	25.8%
	3	800	25.8%
R30.1.1 multifamily	0	1400	45.2%
	1	1700	54.8%
R115.1 definite title	0	800	25.8%
	1	2300	74.2%
R520-Has the owner planted any native tree species?	0	2400	77.4%
	1	700	22.6%
R-Property Type	1	2000	64.5%
	2	900	29.0%
	3	200	6.5%
R616-During last year did someone linked to the lot received credit ?	0	3000	96.8%
	1	100	3.2%
R618-Does the owner have a savings account?	0	2400	77.4%
	1	700	22.6%
R619-Does the owner have a checking account?	0	2300	74.2%
	1	800	25.8%
R635-Does the owner have any urban properties in Rondonia?	0	2500	80.6%
	1	600	19.4%
R701-Does the owner have other rural properties in Rondonia?	0	2800	90.3%
	1	300	9.7%

Q3-R31.7.1 Diarias 0/1 for F1	0	1400	45.2%
	1	1700	54.8%
Q3-R 31.8.1 Off farm work for F1 0/1	0	2600	83.9%
	1	500	16.1%
R601-Does the owner participate in a workers sindicate?	0	2500	80.6%
	1	600	19.4%
R610-Does the owner participate in a cooperative?	0	2900	93.5%
	1	200	6.5%
R501-Has the owner extracted any forest products during 2001?	0	1700	54.8%
	1	1400	45.2%
R605-Does the owner participate in a mutual help group?	0	2200	71.0%
	1	900	29.0%
R31.1 mora neste lote mesmo	1	3100	100.0%
R522-Has the owner any interest in planting native tree species?	0	1300	41.9%
	1	1800	58.1%
Valid		3100	100.0%
Missing		900	
Total		4000	
Subpopulation		59	

a The dependent variable has only one value observed in 28 (47.5%) subpopulations.

Iteration History

Iteration	N Step-halving	-2 Log Likelihood	1= NF , 2=SF 3=PF 0= nothing 1.00 Intercept	[MUN=2]	[OWNR_TYP=1]	[OWNR_TYP=2]	[MULTFAMI=0]
0	0	998.615	-.1054502	.0000000	.0000000	.0000000	.0000000

1	0	419.590	-3.3542494	-.5305426	-1.6353057	-1.2691915	-.1033841
2	0	396.542	-2.1285946	-.8552044	-1.6239397	-1.3329369	-.1411954
3	0	396.166	-1.9019691	-.9101906	-1.6301086	-1.3656474	-.1578428
4	0	396.166	-1.8913445	-.9117506	-1.6306077	-1.3676407	-.1590002

5	0	396.166	- 1.8913289	-.9117521	- 1.6306088	- 1.3676447	-.1590026
6	1	396.166	- 1.8913289	-.9117521	- 1.6306088	- 1.3676447	-.1590026

Redundant parameters are not displayed. Their values are always zero in all iterations.
a The parameter estimates converge. Last absolute change in -2 Log Likelihood is .000, and last maximum absolute change in parameters is 1.645290E-11.

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	998.615			
Final	396.166	602.449	54	.000

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	131.168	62	.000
Deviance	112.075	62	.000

Pseudo R-Square

Cox and Snell	.177
Nagelkerke	.200
McFadden	.090

Likelihood Ratio Tests

Effect	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	396.166	.000	0	.
MUN	414.586	18.420	2	.000
OWNR_TYP	418.370	22.204	4	.000
MULTFAMI	397.691	1.526	2	.466
D_TITLE	403.388	7.223	2	.027
NATRE_01	418.257	22.092	2	.000
PROP_TYP	399.434	3.268	4	.514
CREDIT	405.274	9.108	2	.011
SAVINGS	397.748	1.583	2	.453
CHECKING	431.416	35.250	2	.000
URB_PROP	453.027	56.862	2	.000
RURLOT01	402.284	6.119	2	.047
F1_DIA01	397.460	1.295	2	.523
F1OFFW01	443.390	47.225	2	.000
SINDICAT	431.634	35.468	2	.000
COOPERAT	407.920	11.754	2	.003
XTRACT_F	410.197	14.032	2	.001
MUTUAL	396.524	.358	2	.836
LIVE_LOT	396.166	.000	0	.
INTNATRE	399.295	3.130	2	.209
LOT_A_AL	419.800	23.634	2	.000
AN_PERLO	397.648	1.482	2	.477
FO_PERLO	408.525	12.359	2	.002
NPERLSUR	410.488	14.322	2	.001
NPERLSUR * OWNCAT_N	403.311	7.145	2	.028
OWNCAT_N	405.003	8.837	2	.012
F1DEPRAT	396.542	.376	2	.828

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Parameter Estimates

1= NF , 2=SF 3=PF 0= nothing	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound

1.00	Intercept	-1.891	2.489	.577	1	.447			
	[MUN=2]	-.912	.255	12.745	1	.000	.402	.244	.663
	[MUN=3]	0	.	.	0	.	.	.	.
	[OWNR_TYP=1]	-1.631	.356	20.975	1	.000	.196	9.745E-02	.393
	[OWNR_TYP=2]	-1.368	.545	6.291	1	.012	.255	8.748E-02	.742
	[OWNR_TYP=3]	0	.	.	0	.	.	.	.
	[MULTFAMI=0]	-.159	.451	.124	1	.724	.853	.352	2.064
	[MULTFAMI=1]	0	.	.	0	.	.	.	.
	[D_TITLE=0]	.739	.303	5.929	1	.015	2.093	1.155	3.792
	[D_TITLE=1]	0	.	.	0	.	.	.	.
	[NATRE_01=0]	1.169	.298	15.366	1	.000	3.220	1.794	5.778
	[NATRE_01=1]	0	.	.	0	.	.	.	.
	[PROP_TYP=1]	-.127	.460	.076	1	.783	.881	.358	2.169
	[PROP_TYP=2]	-.365	.568	.414	1	.520	.694	.228	2.113
	[PROP_TYP=3]	0	.	.	0	.	.	.	.
	[CREDIT=0]	-1.610	1.058	2.315	1	.128	.200	2.510E-02	1.591
	[CREDIT=1]	0	.	.	0	.	.	.	.
	[SAVINGS=0]	.512	.479	1.142	1	.285	1.669	.652	4.270
	[SAVINGS=1]	0	.	.	0	.	.	.	.
	[CHECKING=0]	-2.959	.555	28.449	1	.000	5.186E-02	1.748E-02	.154
	[CHECKING=1]	0	.	.	0	.	.	.	.
	[URB_PROP=0]	-1.991	.269	54.609	1	.000	.137	8.050E-02	.231
	[URB_PROP=1]	0	.	.	0	.	.	.	.
	[RURLOT01=0]	-1.156	.549	4.440	1	.035	.315	.107	.922
	[RURLOT01=1]	0	.	.	0	.	.	.	.
	[F1_DIA01=0]	-.039	.217	.032	1	.858	.962	.628	1.472
	[F1_DIA01=1]	0	.	.	0	.	.	.	.
	[F1OFFW01=0]	3.418	.516	43.952	1	.000	30.497	11.103	83.763
	[F1OFFW01=1]	0	.	.	0	.	.	.	.
	[SINDICAT	1.729	.306	31.965	1	.000	5.633	3.094	10.256

	=0]								
	[SINDICAT=1]	0	.	.	0	.	.	.	.
	[COOPERAT=0]	2.019	.775	6.789	1	.009	7.528	1.649	34.371
	[COOPERAT=1]	0	.	.	0	.	.	.	.
	[XTRACT_F=0]	-.985	.314	9.803	1	.002	.374	.202	.692
	[XTRACT_F=1]	0	.	.	0	.	.	.	.
	[MUTUAL=0]	.148	.341	.188	1	.664	1.160	.594	2.263
	[MUTUAL=1]	0	.	.	0	.	.	.	.
	[LIVE_LOT=1]	0	.	.	0	.	.	.	.
	[INTNATRE=0]	-.085	.174	.240	1	.624	.918	.652	1.292
	[INTNATRE=1]	0	.	.	0	.	.	.	.
	LOT_A_AL	-.072	.016	19.620	1	.000	.931	.902	.961
	AN_PERLO	.011	.018	.352	1	.553	1.011	.976	1.047
	FO_PERLO	.033	.012	7.951	1	.005	1.033	1.010	1.057
	NPERLSUR	.238	.071	11.236	1	.001	1.269	1.104	1.459
	NPERLSUR * OWNCAT_N	-.001	.000	5.221	1	.022	.999	.998	1.000
	OWNCAT_N	.017	.007	6.599	1	.010	1.017	1.004	1.030
	F1DEPRAT	.106	.205	.268	1	.605	1.112	.744	1.661
2.00	Intercept	-3.879	2.695	2.072	1	.150			
	[MUN=2]	.231	.217	1.133	1	.287	1.260	.823	1.929
	[MUN=3]	0	.	.	0	.	.	.	.
	[OWNR_TYP=1]	-.814	.402	4.091	1	.043	.443	.201	.975
	[OWNR_TYP=2]	-.315	.562	.314	1	.575	.730	.242	2.196
	[OWNR_TYP=3]	0	.	.	0	.	.	.	.
	[MULTFAMI=0]	.406	.380	1.142	1	.285	1.500	.713	3.159
	[MULTFAMI=1]	0	.	.	0	.	.	.	.
	[D_TITLE=0]	-.001	.384	.000	1	.999	.999	.471	2.121
	[D_TITLE=1]	0	.	.	0	.	.	.	.
	[NATRE_01=0]	1.174	.310	14.344	1	.000	3.234	1.762	5.935
	[NATRE_0	0	.	.	0	.	.	.	.

1=1]								
[PROP_TYP=1]	.682	.496	1.893	1	.169	1.978	.749	5.223
[PROP_TYP=2]	.599	.636	.888	1	.346	1.820	.524	6.329
[PROP_TYP=3]	0	.	.	0	.	.	.	.
[CREDIT=0]	-3.075	1.055	8.488	1	.004	4.619E-02	5.837E-03	.366
[CREDIT=1]	0	.	.	0	.	.	.	.
[SAVINGS=0]	-.139	.550	.064	1	.800	.870	.296	2.556
[SAVINGS=1]	0	.	.	0	.	.	.	.
[CHECKING=0]	-2.524	.590	18.272	1	.000	8.015E-02	2.519E-02	.255
[CHECKING=1]	0	.	.	0	.	.	.	.
[URB_PROP=0]	-1.056	.300	12.429	1	.000	.348	.193	.626
[URB_PROP=1]	0	.	.	0	.	.	.	.
[RURLOT01=0]	.256	.597	.184	1	.668	1.292	.401	4.158
[RURLOT01=1]	0	.	.	0	.	.	.	.
[F1_DIA01=0]	-.252	.230	1.206	1	.272	.777	.495	1.219
[F1_DIA01=1]	0	.	.	0	.	.	.	.
[F1OFFW01=0]	1.959	.545	12.903	1	.000	7.090	2.435	20.644
[F1OFFW01=1]	0	.	.	0	.	.	.	.
[SINDICAT=0]	1.165	.308	14.341	1	.000	3.206	1.754	5.859
[SINDICAT=1]	0	.	.	0	.	.	.	.
[COOPERAT=0]	2.648	.887	8.916	1	.003	14.131	2.484	80.375
[COOPERAT=1]	0	.	.	0	.	.	.	.
[XTRACT_F=0]	-.983	.325	9.144	1	.002	.374	.198	.708
[XTRACT_F=1]	0	.	.	0	.	.	.	.
[MUTUAL=0]	-.078	.319	.060	1	.807	.925	.495	1.728
[MUTUAL=1]	0	.	.	0	.	.	.	.
[LIVE_LOT=1]	0	.	.	0	.	.	.	.
[INTNATRE=0]	.272	.201	1.841	1	.175	1.313	.886	1.946
[INTNATR	0	.	.	0	.	.	.	.

E=1]								
LOT_A_AL	-.051	.015	12.018	1	.001	.950	.923	.978
AN_PERLO	-.014	.018	.586	1	.444	.986	.952	1.022
FO_PERLO	.040	.014	8.478	1	.004	1.041	1.013	1.069
NPERLSUR	.175	.066	7.088	1	.008	1.191	1.047	1.355
NPERLSUR * OWNCAT_N	-.001	.000	4.370	1	.037	.999	.999	1.000
OWNCAT_N	.015	.006	5.248	1	.022	1.015	1.002	1.028
F1DEPRAT	-.025	.208	.014	1	.904	.975	.648	1.467

a This parameter is set to zero because it is redundant.

Classification

	Predicted			
Observed	1.00	2.00	3.00	Percent Correct
1.00	637	74	404	57.1%
2.00	234	152	360	20.4%
3.00	330	74	835	67.4%
Overall Percentage	38.7%	9.7%	51.6%	52.4%