



Published in final edited form as:

*Appl Geogr.* 2011 January ; 31(1): 210–222. doi:10.1016/j.apgeog.2010.04.005.

## Land Use Change on Household Farms in the Ecuadorian Amazon: Design and Implementation of an Agent-Based Model

Carlos F. Mena<sup>1</sup>, Stephen J. Walsh<sup>2,3</sup>, Brian G. Frizzelle<sup>3</sup>, Yao Xiaozheng<sup>3</sup>, and George P. Malanson<sup>4</sup>

<sup>1</sup>Dept. of Life & Environmental Sciences, Universidad San Francisco de Quito, Ecuador

<sup>2</sup>Dept. of Geography, University of North Carolina at Chapel Hill, USA

<sup>3</sup>Carolina Population Center, University of North Carolina at Chapel Hill, USA

<sup>4</sup>Dept. of Geography, University of Iowa, USA

### Abstract

This paper describes the design and implementation of an Agent-Based Model (ABM) used to simulate land use change on household farms in the Northern Ecuadorian Amazon (NEA). The ABM simulates decision-making processes at the household level that is examined through a longitudinal, socio-economic and demographic survey that was conducted in 1990 and 1999. Geographic Information Systems (GIS) are used to establish spatial relationships between farms and their environment, while classified Landsat Thematic Mapper (TM) imagery is used to set initial land use/land cover conditions for the spatial simulation, assess from-to land use/land cover change patterns, and describe trajectories of land use change at the farm and landscape levels. Results from prior studies in the NEA provide insights into the key social and ecological variables, describe human behavioral functions, and examine population-environment interactions that are linked to deforestation and agricultural extensification, population migration, and demographic change. Within the architecture of the model, agents are classified as active or passive. The model comprises four modules, i.e., initialization, demography, agriculture, and migration that operate individually, but are linked through key household processes. The main outputs of the model include a spatially-explicit representation of the land use/land cover on survey and non-survey farms and at the landscape level for each annual time-step, as well as simulated socio-economic and demographic characteristics of households and communities. The work describes the design and implementation of the model and how population-environment interactions can be addressed in a frontier setting. The paper contributes to land change science by examining important pattern-process relations, advocating a spatial modeling approach that is capable of synthesizing fundamental relationships at the farm level, and links people and environment in complex ways.

### Keywords

Agent Based Model; Household Decision-Making Model; Northern Ecuadorian Amazon

---

© 2010 Elsevier Ltd. All rights reserved.

Corresponding Author: Carlos F. Mena, cmena@usfq.edu.ec, Phone: 593-9-449-7506.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## 1.0 Introduction

The tropical rainforest of the Northern Ecuadorian Amazon (NEA) is an area where complex interactions occur among a number of important and diverse stakeholders: (a) spontaneous colonists, who have migrated from other parts of the country and established farms in the rainforest; (b) newly emerging communities and market centers that are expanding services, providing a source for off-farm employment of colonists, and affecting Land use/land cover (LULC) in direct and indirect ways; (c) indigenous peoples, who continue to follow traditional practices of living mostly off the forest, but are increasingly involved in commercial agriculture, wage employment, and other contacts with oil companies, colonists, and other external agents in a transition to a more market-oriented economy; (d) oil companies, who built the original roads in the region to lay pipelines and continue to find, and extract, more petroleum in colonist and indigenous areas, as well as in national parks; and (e) government agencies such as the Ministry of Environment, influenced by both Ecuadorian and international NGOs, to set aside and monitor large areas of the Amazon for conservation and protection of the extraordinarily rich ecological and cultural diversity of the region. Relationships among the five types of stakeholders are complex, in part, because of feedbacks between spatial patterns and rates of LULC change on household farms. In general, LULC change associated with tropical deforestation is complex – links and feedbacks between population and environment create dynamics with emergent properties, especially at the advancing fronts of frontier development. These feedbacks in turn constrain future changes in LULC and the interactions between people and environment.

This paper reports the design, implementation, and results of a spatial simulation model that was developed to assess the drivers of land use/land cover change in the NEA. The model is based on a longitudinal, socio-economic and demographic survey of colonist' households conducted in 1990 and 1999, an assembled satellite image time-series of LULC change for 1973–2007, and GIS coverages of infrastructure, geographic access, and resource endowments of farms. This work is based on previous models and simulations developed using Cellular Automata (CA) models (Messina and Walsh 2001, 2005; Walsh et al. 2008) and statistical models of the drivers of land use/land cover dynamics (Murphy et al. 1997; Pan et al. 2004; Barbieri and Carr 2005; Pan and Bilsborrow 2005; Mena et al. 2006b). Our spatial simulation model is an Agent-Based Model (ABM) of household behavior, landscape dynamics, and LULC change. The ABM incorporates the land use decision-making behavior of the key stakeholders. At present, the design of the ABM addresses household decision-making of agricultural colonists, who are responsible for the greatest changes that have occurred to date in the NEA. Colonists gained access to the region via the roads built by oil companies that made isolated areas accessible for settlement, and encouraged the development and expansion of communities as service centers and employment nodes.

### 1.1 Study Area

The Northern Ecuadorian Amazon (NEA -- see Figure 1) is a region undergoing dramatic change. Complex social, economic, and demographic processes are developing at multiple temporal and spatial scales (ECORAE 1996; Bilsborrow 2003; Bilsborrow et al. 2004; Mena et al. 2006b), which make the region an optimal laboratory to study feedbacks, adaptation, and non-linear relationships between coupled natural-human systems. The region covers approximately 20,000 km<sup>2</sup> and encompasses exceptional biological and cultural diversity. It has been characterized as a hotspot of biodiversity (Myers 1990) with very high levels of alpha biodiversity (Pitman et al. 2003). Cultural diversity is also high, as the NEA is home to many indigenous communities from different ethnic groups that have adapted to the Amazonian environment over hundreds of years. Many of these communities are currently undergoing significant change due to their contact with other groups and their socio-economic systems.

Further, the agricultural frontier of the NEA is characterized by rapid deforestation, exploitation of natural resources, in-migration of colonists, socio-cultural change of native populations, and integration to regional and global markets, making the region an ideal social-ecological laboratory to study complex and dynamic population-environment interactions. The NEA has seen these processes explode since 1964, when the discovery of petroleum triggered infrastructure development and spontaneous agricultural colonization. Population increased rapidly: 8-percent growth from 1974–1982, 6-percent growth from 1982–1990, and almost 4-percent from 1990–2001, mostly attributed to in-migration from the Andes, and in all cases, more than double the national average (Bilsborrow et al. 2004). The NEA has one of the highest rural population densities of the entire Amazon Basin (ECORAE 1996). Linked with this colonization and resource extraction is ecological degradation. Over the past 20-years, Ecuador has had the highest annual deforestation rate in South America (Food and Agriculture Organization 2001) at 2.5-percent (1986–1996) and 1.8-percent (1996–2002) (Mena et al. 2006b).

The large influx of migrant families to the NEA since the 1970s has resulted in extensive spontaneous colonization of the region and household settlement on agricultural plots of land. As families acquired knowledge and skills over time to produce agricultural products and survive in the very difficult environment, and as household demographic behavior changed, additional land clearing has occurred. Farmers followed the initial planting of subsistence crops with cash crops, and little by little, increased the area of land in pasture for cattle (Uquillas 1984). Further in-migration has continued, contributing to a dramatic process of subdivision or fragmentation of landholdings (Bilsborrow et al. 2004), which in turn has had feedbacks on LULC through deforestation and an initial emphasis on subsistence agriculture, followed by commercial agriculture, thereby delaying the evolution towards ever more pasture and leading to more intensive forms of land use in the cultivation of crops. But, in other cases, LULC changes take the trajectory to fallow and forest succession (Sirén and Brondizio 2009). Meanwhile, migration of children of the original settlers as well as continuing in-migration from elsewhere in the country has fed the growth of local towns, increasing markets for farm products as well as opportunities for off-farm employment. Both of these factors have feedbacks on farm LULC. Meanwhile, two large national parks were established in the region in the 1970s (the Cuyabeno Wildlife Reserve and the Yasuni National Park), and indigenous populations in the Amazon received legal, communal land titles to huge areas of forest, where they had lived for centuries, mainly in the early 1990s (though adjudication continues), both circumscribing further expansion of colonist farms at the extensive margin. Petroleum companies have recently found substantial new deposits of oil, mainly in national parks and indigenous lands, and are negotiating to expand oil production in the region. This is likely to contribute to further the growth of local towns and infrastructure, and to encourage changes in indigenous communities. Thus, patterns of LULC of colonists will likely continue to evolve and change, with further in-migration and fragmentation of plots, increased off-farm employment, and additional wage labor, while changes in LULC of indigenous populations are also likely as a result of growing market contact, increased petroleum production, and adaptive behaviors related to interactions with other stakeholders in the region and exogenous factors, such as social and environmental policies, and their mediation according to ethnicity, socio-economic and demographic conditions, and individual, household, and group responses to changing social and ecological dynamics.

## 1.2 Background & Context

In a complex system like the NEA, deforestation has complex patterns (Malanson et al. 2006a) and multiple components interact in ways that link patterns and processes across scales (Malanson et al. 2006b). Complex systems focus on irreducible complexity arising

out of apparent simplicity, emerging from nonlinearities due to large numbers of relationships involving feedbacks at one or more levels in the system (Cilliers 1998; Malanson 1999; Crawford et al. 2005; Walsh and McGinnis 2008). Complex systems are generally far from equilibrium (Bak 1998) with a constant set of interactions that maintain the organization of the system through negative feedbacks or that alter subsequent alternatives in state space through positive feedbacks. Thus, complexity theory holds that systems cannot be suitably understood without a focus on the feedbacks and nonlinearities that lead to emergent multi-scale phenomena (Matthews et al. 1999; Manson 2001). Moreover, biocomplexity encompasses the many linkages within and among ecological systems, physical systems, and human systems (Michener et al. 2001). These linkages are particularly dynamic in frontier settings. Feedbacks between people and environment can constrain or even reverse initial changes in land use through system dynamics and non-linear feedbacks (Matthews et al. 1999; Manson 2001; Wolfram 1984; Blackman 2000; Walsh 2007). Critical points in the spatial structure of land use patterns and feedbacks can produce a system with identifiable future alternative states in which instabilities can “flip” a system into another regime of behavior by changing the patterns/processes that control land use change (Parker et al. 2003).

In general, this research is motivated by questions that seek understanding in broad areas of human-environment interactions and complex systems, including: How does a complex systems approach help explore the internal mechanisms and provide plausible explanations of dynamic systems? How do results derived from applying complexity theory help in understanding decision-making across levels of social organization ranging from individual households to national governments? How do fundamental characteristics of complex dynamics of coupled natural-human systems and the limits of predictability pertain to sustainable development? How will complexity theory help us understand LULC change in the frontier of human-environment interactions? Understanding the dynamics of a system given a set of rules and initial conditions, as opposed to determining the rules that produce an end state, is at the heart of this new epistemology.

## 2.0 Data & Methods

This paper describes the design and implementation of an ABM that integrates important elements of household decision-making in the NEA, as represented through empirical analyses of farm data collected through the longitudinal and cross-sectional analyses, landscape dynamics in the NEA, as represented through a satellite image time-series, and spatial connections among geographic features and survey farms derived within a GIS.

The design of the ABM and the representation of socio-economic and demographic characteristics of farm households are the corner-stone of a model that includes the interactions of 2,048 household farms in our spatial simulation of LULC dynamics in the NEA. Of that total, approximately 100 are survey farms that were assessed in 1990 and 1999 through a longitudinal socio-economic and demographic survey. Note that the longitudinal survey consists of approximately 450 household farms, but the model described here was developed for a subset of the study area, hence a fewer number of survey farms are included.

We used the results of prior data collection activities and statistical and spatial analyses to leverage the design and development of our ABM. In 1990, the first household survey in the NEA was carried out, on migrant colonist farms. A two-stage sampling design was used to select a sample of farm plots or *fincas*, settled by spontaneous migrant families. The sample frame was a list of farm “sectors” provided by the Ecuadorian land-titling agency, IERAC. In the first stage, sectors were sampled with probabilities of selection proportional to estimated size. Thus, 64 sectors were selected from the total of 297 sectors settled by that

date. In the second stage, clusters of farm plots were randomly selected from each sector in proportion to sector size. The final sample comprised 408 farm plots on which 418 families were living, representing 5.9-percent of the colonist plots in the main colonization area of the NEA (Pichon and Billsborrow 1999). In 1999, the same plots were re-visited, and all farms and new subdivisions in the same geographic space were (re-) interviewed, resulting in a sample of 767 farms plus 111 additional house lots under 0.5 ha (called *solares*) created by subdivision and parcelization, especially along roads and near major towns. As in 1990, detailed questionnaires were administered separately to the head of household and spouse. Each farm in 1999 was also geo-located via GPS technology.

While all surveyed farms are spatially-referenced to Earth coordinates through the use of a GPS technology and farm attributes are associated through the 1990 and 1999 socio-economic and demographic surveys, the farms that surround the survey farms were located and attributed, i.e., LULC and socio-demographic characteristics, in various ways. Using Ecuadorian government maps of development sectors that were provided by INDA, the number of farms in each sector were specified, but their spatial location, orientation, and size were not indicated. To geographically locate the non-survey farms within each development sector, a Landsat TM image time-series was used to delineate farm boundaries of the 2,048 farms that occurred within the study area and a surrounding 3-km spatial buffer (See Figure 2). Images were classified using a hybrid approach. Pre-processing steps included atmospheric correction, geometric correction, and extensive fieldwork for geodetic control and land-use validation. For geometric correction, the 1996 image was selected as the original, or master, image and was rectified using a set of geodetic control points (GCP) collected using global positioning system (GPS) receivers and topographic maps. The remaining images were rectified using the master image as the reference (i.e., relative registration) (Frizzelle, 2004). A hybrid classification approach uses a combination of unsupervised and supervised approaches applied to defined spectral features, such as an extended set of layers that includes Landsat spectral bands, principal components, and fractional cover and vegetation indices (Messina et al., 2001; Walsh et al., 2003) to extract the land-use/land-cover classes. Normally, the farms measures approximately 50-hectare in size and are spatially organized as  $250 \times 2000$ -m units. Socio-economic and demographic characteristics were randomly assigned to the non-survey farms, given the characteristics of the approximately 100 survey farms that occurred within the study area. To incorporate spatial connections between farms in our model (e.g., labor sharing of adjacent and nearby farms), we delineated survey and non-survey farms in a 3-km spatial buffer around the intensive study area.

The ABM uses our extensive longitudinal household survey data to represent heterogeneous agents. Further, we develop the ABM on a spatially-explicit landscape grid, where the actions of agents affect the conditions of cells. Feedbacks between the value of a cell and its spatial conditions are represented, i.e., the location of a cell relative to other changes and features determine the choices made by agents. The movement and interaction of agents depend on their individual and household characteristics. Agents respond to the landscape that they are embedded – this alters the probabilities of change on which their dynamics depend. Importantly, the socio-economic and demographic structure of each household also changes through time in response to the changing landscape. For instance, the out-migration and off-farm employment of household members are affected by LULC change patterns, adaptive and alternative household livelihood strategies, farm access to roads and other farms, and the evolving characteristics of towns made more accessible through road building and enhancement, as well as the expanding geographic “reach” of towns via buses, taxis, and other modes of local and regional transportation. Our approach assumes discrete space and time that has consequences for the model and the processes being considered (e.g.,

Durrett and Levin 1994), and the design approximates human behavior in an agricultural setting reasonably well.

### 3.0 Results

#### 3.1 Model Overview

Our ABM for the NEA is written in Java and designed to operate within the *Repast Symphony* modeling environment (<http://repast.sourceforge.net/>). It is comprised of four primary modules – **Initialization**, **Demography**, **Migration**, and **Agriculture** – and more than 20 sub-modules that operate interdependently and utilize data in a number of different formats. The **Initialization Module** is called first and executes one time, initializing the landscape and social agents at time  $t_0$ . The model then proceeds on a set iteration, one year in the case of the current model, and runs on a farm-by-farm basis, calling the other three modules in order based on certain criteria. The first module to be called after initialization is always the **Demography Module** that contains algorithms to calculate aging, fertility, mortality, marriage, and asset gain/loss. Then, if the farm has positive assets, the model calls the **Agriculture Module** that uses many of the socio-economic and demographic (SED) variables from the social agents to calculate land use change on the farm. However, if the farm has zero or negative assets, the model calls the **Migration Module** to calculate how many people on the farm out-migrate to find other employment and return remittances to the household. In this situation, no land use change occurs on the farm.

The model contains agents for both the social actors and the physical landscape. Each person living on one of the farms is represented as an agent, and the household in which the agents live are also an agent. The landscape is composed of an array of cells, each of which contains information related to ownership, land use, and other physical environmental characteristics. The social agents operate within the confines of their farm, and interact with the landscape cells on their property. The social agents can also interact with other agents outside of their household.

The model is simulating LULC change for over 2,000 farms. The LULC and parcels data on the farms are based on our 1986 LULC classification of Landsat TM data. The LULC data for the farms have been generalized so that all parcel sizes at initialization (i.e. time  $t_0$ ) are integers (in hectares). There are six LULC types in the model: primary forest, successional vegetation, pasture, subsistence agriculture, commercial agriculture, and barren/urban. Interaction between nearby farmers related to land use change practices are included in the model, but farms are still modeled independently, meaning that they are not simulated within the context of a community or region – they are insulated from outside influences other than the input agricultural cost data and what farmers may learn from their neighbors in the model. The key assumptions made in the design of the ABM are as follows: (1) farmers make decisions based on immediate, current information; no future predictions are taken into account, but prior knowledge is used to make more informed decisions; (2) a farmer will always convert as much land as possible based upon labor and available resources. We use a simple definition of assets that includes any income and other material wealth; (3) off-farm labor pool always has sufficient labor to satisfy the needs of the farmer; (4) all farms have high geographic accessibility to the NEA; farms are referenced relative to their distance to all-weather paved roads, (5) motivation for changing LULC is driven by income (wealth) maximization, so the farmer will attempt to implement land change for profit maximization; (6) a farmer will satisfy the consumption strategy before choosing to implement any other type of land use change; and (7) a farmer makes a land change decision each year; some conversions may take multiple years to be realized.

Initialization of the model comes from two sources of data: a longitudinal, socio-economic and demographic, household survey and a rich database of spatial datasets, both vector and raster, representing environmental, topographic, and built features. The social agents are generated based on the household data gathered in the survey, using a selected set of variables to attribute the agents' characteristics. The landscape agents are built from several derived raster datasets, including land use, age of land use per cell, distance of each cell to the nearest road, and distance of each cell to the location of the dwelling unit of the farm. These datasets create the initial conditions of each *Cell* agent, and those agent characteristics are aggregated as necessary to set the characteristics of the *Parcels* and the *Farms*.

### 3.2 Model Agents

Following initialization, the farmer has access to landscape information such as the number of parcels on the farm, the type of land cover in each parcel, and the size of each parcel. The farmer is able to implement changes in land use type on an individual cell, an entire land parcel, or a subset of a parcel. Farmers, cells, and parcels are agents in our model.

This model, however, separates agents into landscape, social, active and passive agents. On the social-side of the model, there is one household resident on each farm, and it is represented by an active *Household* agent. The household is active (opposed to passive), because it has the capability to make decision and execute actions. Within each *Household* are one or more people, represented by passive *Individual* agents. At every time-step, each *Individual* grows older, has a chance of dying, has a chance of getting married (if single), and possibly moving away, and has a chance of giving birth (if female and of reproductive age). Depending upon the needs of the farm, socio-economic status, and the needs of the household, *Individuals* can out-migrate to engage in off-farm employment and can send remittances back to the household as part of an alternative household livelihood strategy. All decisions made on the farm are made by the head of the household (e.g., the farmer), so for ease of functionality, the *Household* agent acts as the proxy for the head of the household, and thus it is the *Household* and not one particular *Individual* on the farm that is the active agent.

The landscape agents, all of which are passive, are the *Cell* (or *LulcCell*), the *Parcel*, and the *Farm*. The *LulcCell* is an object that represents each of the raster cells in the landscape. Contiguous *LulcCells* with the same land use type on the same property form a *Parcel*, and all *Parcels* on the same property form a *Farm*. The farmer *Individual*, via the *Household* agent, makes changes to land use on the farm by converting all or a portion of one or more *Parcels* to a different land use type, which functionally changes the land use attribute of all *LulcCells* within the *Parcels* selected for change. Currently, the farm is set at the typical 50-hectare size commonly seen in this region, and it cannot be subdivided into smaller properties, although in subsequent models land subdivision will be explicitly addressed.

### 3.3 Model Operation

Figure 3 is a generalized flowchart showing the operation of the model and its various modules that govern the actions and reactions that are guided through computations and defined relationships. The raster datasets used in the current model include: (1) land use classes, (2) parcel IDs, (3) farm IDs, (4) age, in years, of land use by cell, (5) slope angle, (6) Euclidean distance from each cell to the nearest main road, and (7) Euclidean distance of each cell to its farm's dwelling unit. Each layer is read into the **Initialization Module** and converted to an array. The values in the array are used to initialize the cell's attributes. One layer initializes one kind of attribute of the cells. The slope angle and distance arrays are static, in that they do not change over the course of the model run. The other arrays, however, are dynamic and are changed through various model processes and for various

time-steps. Any change to the attribute's value of the cells occurs at the end of each iteration or time-step.

At the same time, the model initializes the social agents via the socio-economic and demographic data from the longitudinal household survey (see Figure 4). Agents of the same type are stored in lists that allow for searches and facilitate iterative procedures. Agents in a list can be accessed by all modules and their attributes can be updated as needed. The initial conditions for these agents are set from the longitudinal household survey data for the farms. Data from the household survey is reported at the level of the household, and exact values for the various individuals living in the household are not available. To resolve this problem, we have developed an algorithm that takes the aggregated household information and creates all of the *Individuals* for the *Household* (e.g., 1 male adult, 1 female adult, 3 children), setting their individual characteristics through a constrained random attribute generator. Any changes to the *Individual* and *Household* agents (e.g., births, deaths, marriage) occur at the end of each time-step in the **Demography Module** that is described below and in Figure 4.

Each model time-step begins by determining changes to the household demographics. Four tables containing birth and death rates for single and married individuals are read into the **Life Table Module** and converted to arrays. The birth and death rate data are derived from a regional life table model (Coale et al. 1983) that is representative of the NEA during this period of study (Bilsborrow, personal communication, 2007). The life table arrays are used in the **Demography Module** to determine, for each *Individual*, whether or not a birth or death occurs in that respective year (see Figure 6).

The mortality probabilities are selected from the tables based on the *Individual's* age and sex, and the current year. A random number is then generated for each *Individual*, and if that number is less than the mortality probability, the *Individual* dies and is removed from the household roster list. The fertility function is applied to all female *Individuals* of child-bearing age (15–50), both single and married. The structure of the fertility tables is the same as the life tables. An *Individual's* marriage status, age group, and current year are used to select the appropriate fertility percentage. As with mortality, a random number is generated for each female *Individual*, and if it is less than the fertility percentage, the female gives birth. The gender of the birth is randomly determined by the current NEA ratio of 1:1.05 (i.e., female/male). A new *Individual* agent is created for the birth, and it is added to the *Individual* list.

The marriage function is applied to all single *Individuals* of both sexes. A probability of marriage is taken from the marriage table. Once again, a random number is generated for each single *Individual*, and if it is less than the probability of marriage, the *Individual* gets married. There is also a probability that the new couple will leave the farm and move elsewhere. If this occurs, the newly married *Individual* is removed from the household roster list. If the couple stays on the farm, a new *Individual* is added into the household. These arrays are available for the **Demography Module**, following the completion of the mortality, fertility, and marriage components of the module for all *Individuals*, the *Household* agents are updated according to the new rosters.

Following completion of the **Demography Module**, the model tabulates the assets of each *Household* to determine whether to access the **Migration Module** or the **Agriculture Module**. If a *Household's* assets are zero or lower, no land use change is possible so the model moves to the **Migration Module** (see Figure 5) and each member of the *Household* has a probability of moving away for the year. This probability is calculated from the Barbieri (2005) out-migration model, then a random number is generated for each



*Individual*, and if it is less than the *Individual's* probability of out-migration, the *Individual* will leave the farm. All *Individuals* that leave the farm are removed from the *Individual* list. At the end of the year (i.e., 1 time-step), all the migrants move back to the farm or send remittances to the *Household*. If the entire *Household* leaves the *Farm*, no land use change occurs and all existing land use types on the farm, with the exception of Primary Forest and Barren, convert to fallow in the form of the Succession. If an abandoned *Household* earns a sufficient amount of money through off-farm employment, it returns to the farm and a new agriculture cycle begins the next year.

If the *Household* has assets available, the model instead moves to the **Agriculture Module**. This module simulates the household determining what, if any, land use changes will occur on the farm (see Figure 6). The first option is to convert land into subsistence agriculture to satisfy the consumptive needs of the household, the default is at 1-hectare of subsistence agriculture per person and can be modified according to different scenarios. The **Consumption Module** calculates the amount of subsistence agriculture on the *Farm* and the number of *Individuals* in the household. We define the consumption rate as 0.5 hectare per individual. If the subsistence agriculture needed (i.e., number of individuals  $\times$  0.5) is greater than the number of hectares of subsistence agriculture, (i.e., the consumption strategy is not satisfied), then a land use change is made to increase the subsistence agriculture on the *Farm* to the level necessary to satisfy the consumption strategy. A land use type (other than subsistence agriculture) is selected as the **From Land Use** (FLU) type and subsistence agriculture is set as the **To Land Use** (TLU) type. The difference between *Individuals* and hectares is set as the amount of land to change. However, if the consumption strategy is met, then the model moves to the **Decision-Maker Module** in which a different type of land use change may occur.

The **Decision-Maker Module** is a sub-module within the larger **Agriculture Module**. In this module, the farmer agent selects one of three differing land use change strategies to change cells (*Change Pairs*), based upon pre-determined probabilities of occurrence (see Figure 7). The three strategies are (1) Best Return, which calculates the change pair based upon the conversion that provides the greatest increase in household income; (2) Social Interaction, in which the farmer copies the land use change from a neighbor; and (3) Random Change that randomly chooses the FLU and TLU to represent the farmer making a less than optimum decision. The probabilities can be set by the user via the Repast user interface, so they are not hard-coded into the model. The default values in the current model 85-percent for Best Return, 10-percent for Social Interaction, and 5-percent for Random Change. These values can be changed to explore different scenarios.

If the Best Return strategy is chosen, the model will access a dataset of market prices for cattle (the Pasture land use type is used as a proxy for cattle), subsistence crops, and commercial crops. The datasets are static arrays, created from external text files of real market price data for the region through 2006, with extrapolated values used for future years. A dataset of maintenance costs is available for all land use types; the data are based on regional sources for 2007, and extrapolated for all other years. The market price and maintenance cost data are used with the set of land use types on the *Farm* to determine the change pair that will result in the maximum possible profit increase for the farmer. Profit is calculated as the difference in market price for land use type *i* and maintenance cost for land use type *j*. Future versions of the model will incorporate price trends to test model sensitivity to various changes.

If the Social Interaction strategy is selected, the model will access a neighborhood array that is created during the model initialization, and is based on farm IDs and the topological relationships of farms. The farmer will look to his neighbors, randomly choose the one of

them, and select the FLU and TLU land use types for change on his farm. Currently, a farmer's neighbors are only adjacent farms. In future versions, we will implement a social network for connecting farmers to others, based on a hierarchical series of network ties that include local neighbors, community neighbors, and kinship ties.

If the Random Change strategy is selected, the model will randomly select two land use types, one for the FLU and one for the TLU, and will implement this land use change. If the land use selected for the FLU is not present on the farm, then no land use change occurs during this iteration.

Following the land use change strategy, the model calculates the amount of land that the farmer is able to change. This is known as the *Change Size*, and it is determined by a combination of the amount of assets that the farm is allowed to use for land use change (currently set at 50-percent, but the threshold value can be set by the analyst in the user interface) and maintenance costs. Household assets are an integer value that is continuously updated in the **Agriculture Module** through the calculation of assets based on a regression (OLS) model of household farm assets by Murphy (2001). A dataset of maintenance costs is included and represents the number of males needed per hectare to maintain each type of land use in the model. These values are used in an algorithm that calculates the maximum changeable area, in whole hectares, for each farm.

The land use Change Pair, the integer Change Size, the list of parcels on the farm, and an array of suitability scores (calculated for each cell during the model initialization using the slope and distance arrays) are sent to the **Find Parcel Module** (see Figure 7) to implement the land use change on the farm. The module first searches all of the *Parcels* on the *Farm* in an attempt to identify a *Parcel* with the FLU type and a size equal to the Change Size. If such a *Parcel* is found, its land use (i.e., the land use variable in all of the *LulcCell* agents within the *Parcel*) is changed to the TLU type. If no FLU *Parcel* of the Change Size exists on the *Farm*, the module will then search the parcel list again looking for a FLU *Parcel* with a size greater than the Change Size. If such a *Parcel* is found, the module will use the suitability scores to split the *Parcel* into two smaller *Parcels*, so that one of the two new *Parcels* has a size equal to the Change Size. It is that *Parcel* that will have its land use converted to the TLU type. If no FLU *Parcel* larger than the Change Size exists, then the module will begin a third search of the parcel list looking for a FLU *Parcel* smaller than the Change Size. The first such *Parcel* that is found ( $P_1$ ) will be converted to the TLU type, and its size ( $A_1$ ) will be stored. Then the search continues until another FLU *Parcel* is found. If the size ( $A_2$ ) of that second *Parcel* ( $P_2$ ) is equal to the Change Size minus  $A_1$ , then  $P_2$  will be converted to the TLU and the search ends. If  $A_2$  is greater than the Change Size minus  $A_1$ , then  $P_2$  will be split using the suitability score to form a new *Parcel* with the required remaining size. If  $A_2$  is less than the Change Size minus  $A_1$ , then  $P_2$  is wholly converted to the TLU type, and the search continues as described above until either the Change Size amount of land is converted or the module runs out of *Parcels* of the FLU type. Finally, if no *Parcel* on the *Farm* has the appropriate FLU type, then no change is implemented during this iteration.

All of the above processes run for each *Household* (and thereby, each *Farm*) in the model during each time-step. After all *Households* have been handled, the model advances to the next time-step and continues the processing. During the course of the model run, there is an option to activate the **Data Out Module** and output a text file containing variables of interest, reported by time-step. A second module, **Process Output**, creates a text file containing procedural information, also reported by time-step per *Household*. Such information includes the change strategy, the Change Pairs, and the Change Size. Also included are the number of migrants, births, deaths, hired laborers, total assets, and

remittances. These files are used to assess model performance, relying upon a series of output metrics as diagnostics. Figure 8 shows an example of LULC change on a cluster of farms for years 1, 8, 16, and 25 of the 1990–2015 simulation period, the main output of the model. Figure 9 shows household demographics for the 25-year period, specifically, the mean number of male adults, female adults, children on household farms, as well as the mean number of individuals per household and the mean number of migrants. Figure 10 shows the mean household assets of farmers in development sectors for the 1990–2015 simulation period, indicating considerable variability in farmer's ability to accrue assets early and late in the simulation period. Finally, Figure 11 shows the patterns of land use/land cover change during the 25-year study period.

#### 4.0 Discussion & Conclusions

This paper reports on the design and implementation of an ABM for the simulation of land use/land cover change in the Northern Ecuadorian Amazon (NEA). The ABM simulates the decision-making processes at the household level and it is based on a longitudinal socio-economic and demographic, household survey, a satellite image time-series, and GIS coverages. The model is empirically grounded, based on all the rich information that the Ecuador Project has collected since 1990 in the NEA and the scientific publications generated over the last decade. The model uses active and passive, and social and landscape agents to represent the household as the unit of decision-making. The household responds in different and particular ways to endogenous (demographic and economic) conditions and to exogenous changes (e.g., labor markets) and, also interacts with neighbors as part of their decision-making. Four main modules, **Initialization**, **Demography**, **Agriculture**, and **Migration**, operate individually, but they are linked through key characteristics of the system (e.g., assets of the household) to provide spatial patterns of the farm and landscape or non-spatial demographic or socioeconomic conditions of the household and community.

While the interaction of neighbors through imitation and its incorporation within ABMs has been a challenge for particular study areas (e.g., Schmit and Rounsevell 2006), in the NEA, the direct diffusion of agricultural practices has not been well studied, but empirical evidence supports that this process of decision-making, particularly, in remote, frontier settings, commonly occurs. This ABM has the capability to allow farmers to imitate agricultural patterns of neighbors, but the user also can interactively modify decision-making parameters of farmers according to different conditions and geographic settings.

This paper reports on the conceptualization, structure, and functions of the model as a possible design template for other investigators and study areas in which the behavior of farmers is explicitly linked to the proximate changes in land use patterns in the Amazon region and elsewhere. In this stage, the model synthesizes findings for other studies carried out in the NEA, therefore, the model is process oriented, in the sense that it captures the main demographic, social, and economic variables and relationships at work in the study area. But the model is also spatially-explicit, which suggest a further emphasis on pattern-process relations, specifically related to landscape change that is created by social actions of agents. Subsequent versions of the model will simulate the spatial configuration of patches within a farm that will facilitate the study of land use/land cover structure of neighboring forests remnants and their implication to ecosystem goods and services, including biodiversity. There is empirical evidence that the linkage and synergism between process and pattern oriented models benefit can substantially benefit studies of land use policy (Castella and Verburg 2007).

An important process that has not yet been incorporated into the model is the subdivision of farms. Subdivision is important in that it creates new deforestation fronts and the expansion

of the extensification of agriculture, generates greater population densities, and further fragments land use patterns on household farms. A subdivision module is under development that will accommodate changes in ownership caused by the reallocation of land through kinship ties or land sale. The existence of multiple households on a single farm has considerable implications for LULC dynamics, as well as for influencing LULC on adjacent farms through labor sharing, replicating nearby LULC patterns, and sharing information through social networks on crop prices, farm management, and farm-community interactions.

The process of design and implementation of the model has served as a synthesis tool where findings of past team efforts related to socioeconomic, demographic, and geographical processes are examined in a holistic and integrative way that has led to the enhancement of knowledge of pattern-process relations. The challenges of information integration in the context of an ABM are many. For instance, the architecture of the model grew rapidly as our empirical understanding increased, and so too our concern that we not over-specifying the model and render it deterministic, thereby, diminishing the inherent complexity of the system under study. The dispute between simplification and complexification of the model (Boero et al. 2008) was present during the design and implementation stages, although the models succeeded in retaining fundamental issues such as emergence and self-organization through the behavior of household agents operating within a dynamic environment.

## Acknowledgments

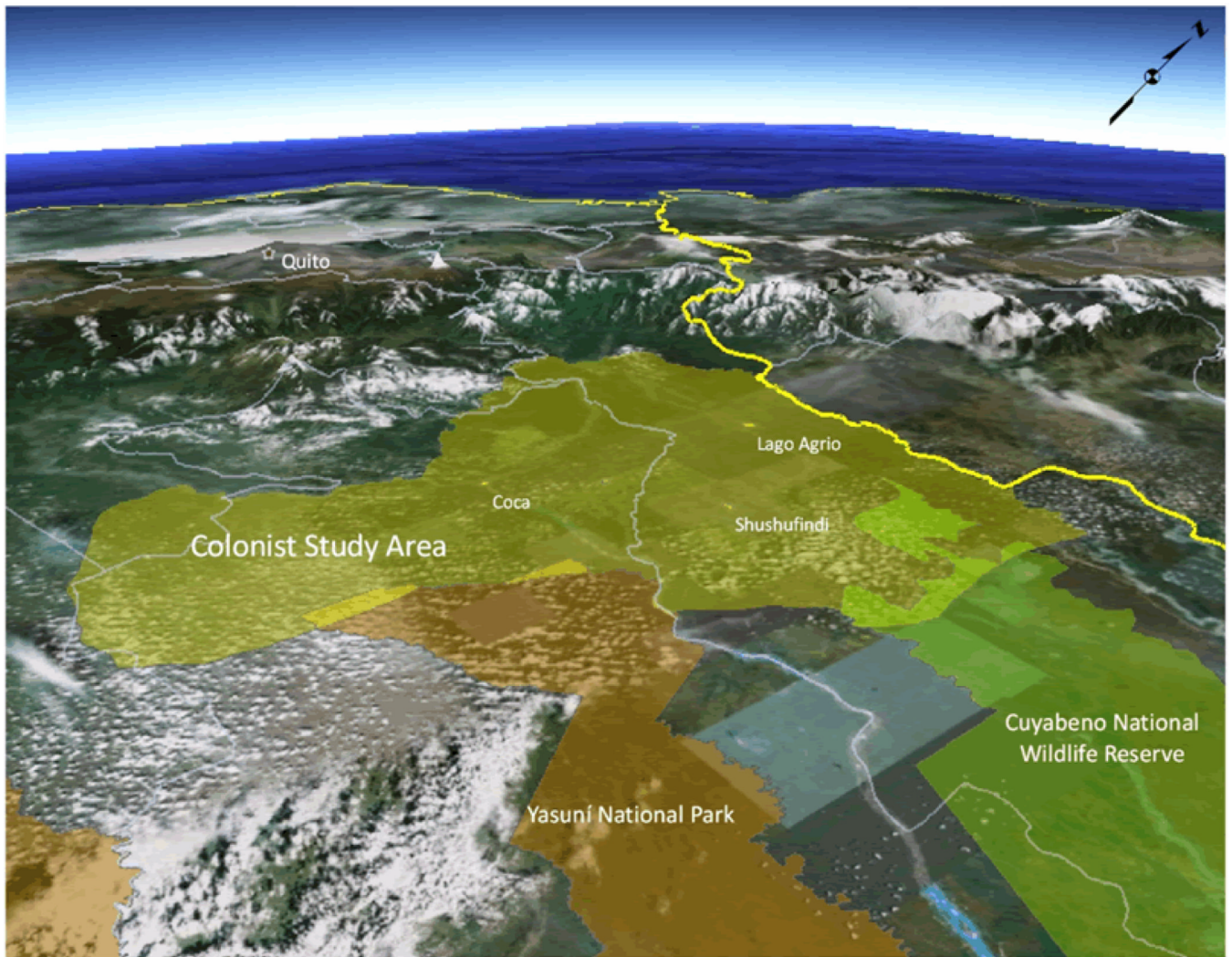
This work was supported by a grants from the National Science Foundation (BE/CNH 0410048), US National Aeronautics and Space Administration (NASA) (NCC5-295), the National Institutes of Health (NIH) (R01-HD38777-01). Carlos Mena would like to thank the NASA Earth System Fellowship (NNG04GR12H).

## Cited References

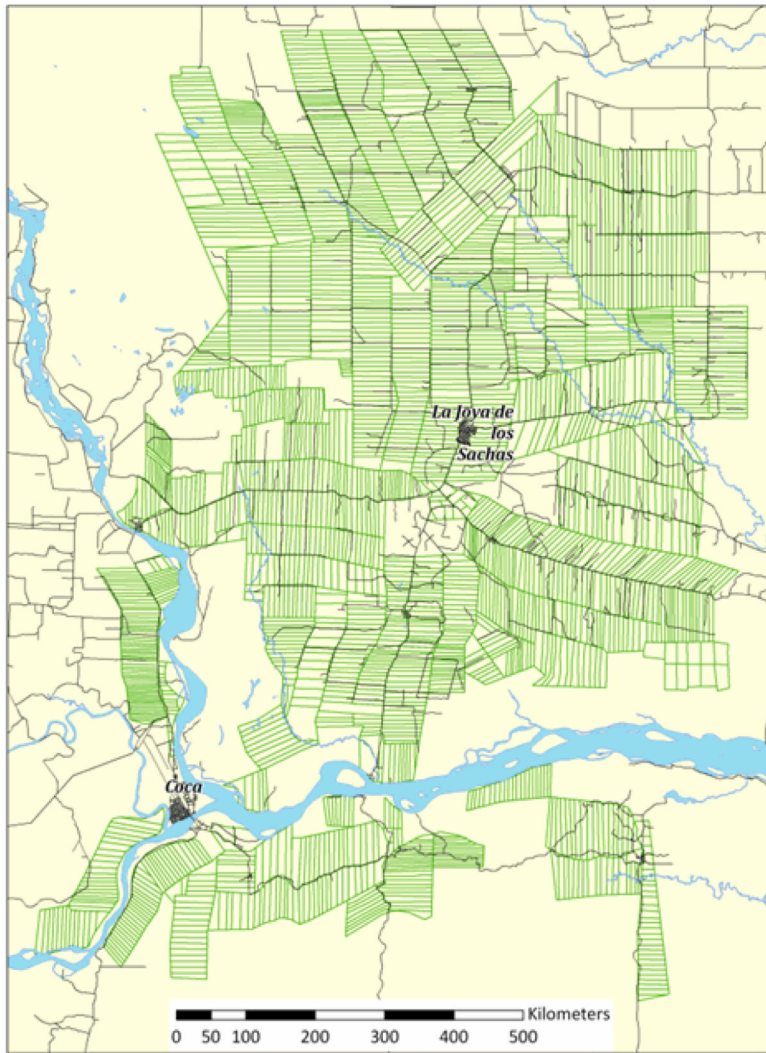
- Bak, P. *How Nature Works*. New York: Copernicus/Springer-Verlag; 1998.
- Barbier, EB. *Natural Resources and Economic Development*. New York: Cambridge University Press; 2005.
- Barbieri AF, Carr DL. Gender-specific out-migration, deforestation and urbanization in the Ecuadorian Amazon. *Global and Planetary Change*. 2005; 47:99–110. [PubMed: 19657469]
- Bilborrow, RE. Cambios Demográficos y Medio Ambiente en la Región Amazónica de los Países Andinos. In: Aramburu, CE.; Bedoya, E., editors. *Amazonia: Procesos Demográficos y Ambientales*. Lima, Peru, Consorcio de Investigación Económica y Social; 2003.
- Bilborrow RE, Barbieri A, Pan WKY. Changes in population and land use over time in the Ecuadorian Amazon. *Acta Amazonica*. 2004; 34(3):635–647.
- Blackman, T. Complexity theory. In: Browning, G.; Halcli, A.; Webster, F., editors. *Understanding Contemporary Society: Theories of the Present*. London: Sage Publications; 2000. p. 139-151.
- Boero R, Castellani M, Squazzoni F. Individual behavior and macro social properties. An agent-based model. *Computational & Mathematical Organization Theory*. 2008; 14(2) 156-17.
- Castella JC, Verburg PH. Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. *Ecological Modelling*. 2007; 202(3–4):410–420.
- Cilliers, P. *Complexity and Postmodernism*. New York: Routledge; 1998.
- Coale, J.; Demeny, P. *Regional Model Life Tables and Stable Populations*. New York: Academic Press; 1983.
- Crawford TW, Messina JP, Manson SM, O’Sullivan D. Complexity science, complexity systems and land use research. *Environment and Planning B*. 2005; 32:857–875.
- Durrett R, Levin S. The importance of being discrete (and spatial). *Theoretical Population Biology*. 1994; 46 363-363.
- ECORAE. *Estadísticas Básicas de la Región Amazónica Ecuatoriana*. Instituto para el EcoDesarrollo de la Región Amazonica Ecuatoriana ECORAE, Quito. 1996



- Pichón, F.; Bilsborrow, RE. Land use systems, deforestation, and demographic factors in the humid tropics: farm-level evidence from Ecuador. In: Bilsborrow, RE.; Hogan, D., editors. *Population and Deforestation in the Humid Tropics*. Belgium: International Union for the Scientific Study of Population, Liège; 1999.
- Pitman NC. A comparison of tree species diversity in two upper Amazonian forests. *Ecology*. 2003; 83(11):3210–3224.
- Schmit C, Rounsevell MDA. Are agricultural land use patterns influenced by farmer imitation. *Agriculture, Ecosystems and Environment*. 2006; 115(1–4):113–127.
- Sirén AH, Brondizio ES. Detecting subtle land use change in tropical forests. *Applied Geography*. 2009; 29(2):201–211.
- Uquillas, J. Colonization and Spontaneous Settlement in the Ecuadorian Amazon. In: Schmink, M.; Wood, C., editors. *Frontier in Expansion in Amazonia*. Gainesville: University of Florida Press; 1984.
- Walsh, SJ.; Messina, JP.; Crews-Meyer, KA.; Bilsborrow, RE.; Pan, WKY. Characterization and modeling patterns of deforestation and agricultural extensification in the Ecuadorian Amazon, Linking People, Place, and Policy. Walsh, SJ.; Crews-Meyer, KA., editors. Boston: Kluwer Academic Publishers; 2003.
- Walsh, SJ. Feedbacks. In: Robbins, P., editor. *Encyclopedia of Environment & Society*. Croton-on-Hudson, NY, Thousand Oaks, CA: Golson Books, Ltd, Sage Publications; 2007. p. 655–656.
- Walsh SJ, McGinnis D. Biocomplexity in coupled human-natural systems: study of population & environment interactions. *GeoForum*. 2008; 39(2):773–775.
- Walsh SJ, Messina JP, Mena CF, Malanson GP, Page PH. Complexity theory, spatial simulation models, and land use dynamics in the Northern Ecuadorian Amazon. *GeoForum*. 2008; 39(2):867–878.
- Wolfram S. Cellular automata as models of complexity. *Nature*. 1984; 311:419–424.



**Figure 1.**  
Study area location: Northern Ecuadorian Amazon.

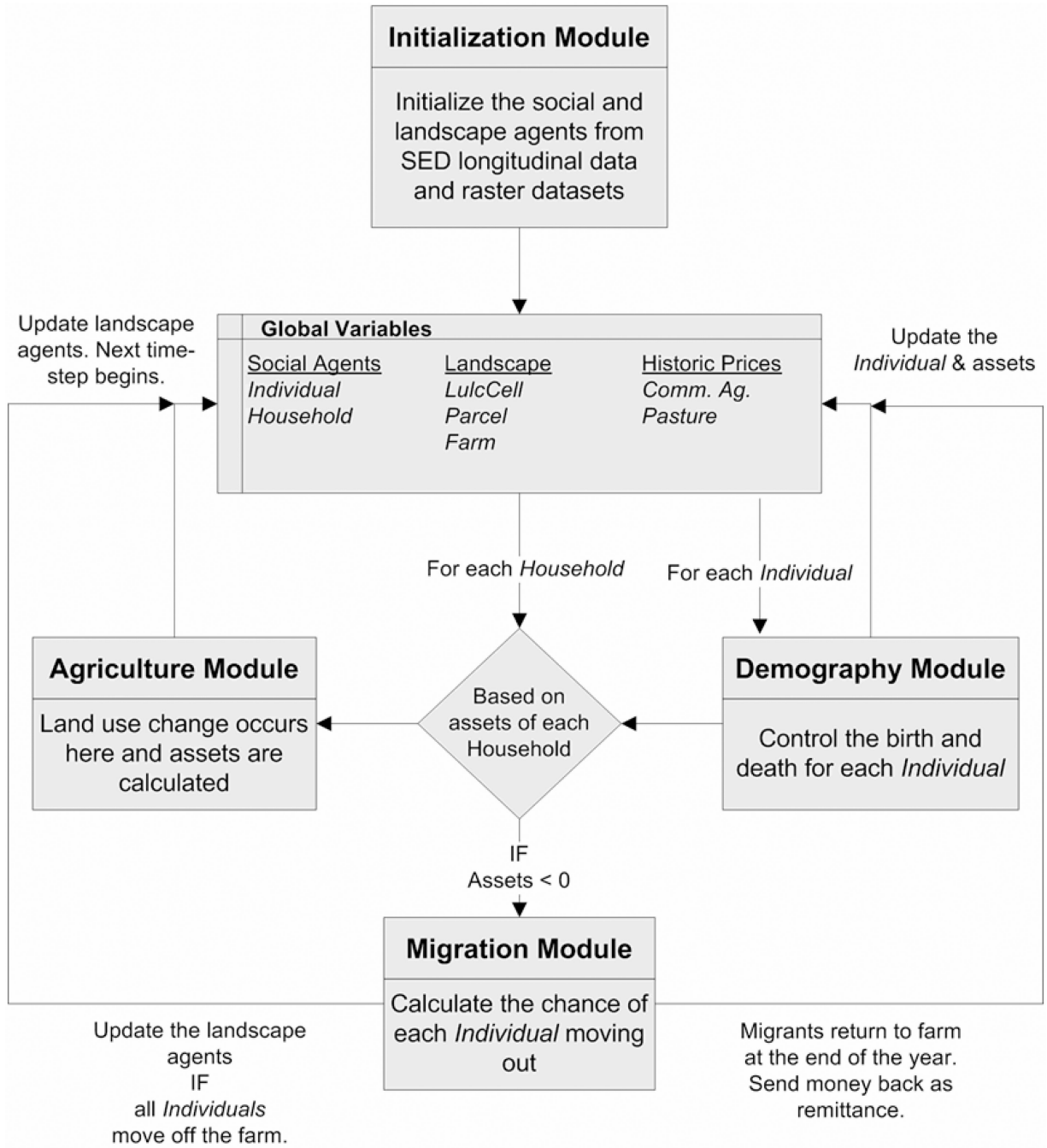


### Reference Map for Ecuadorian Amazon Agent-Based Model

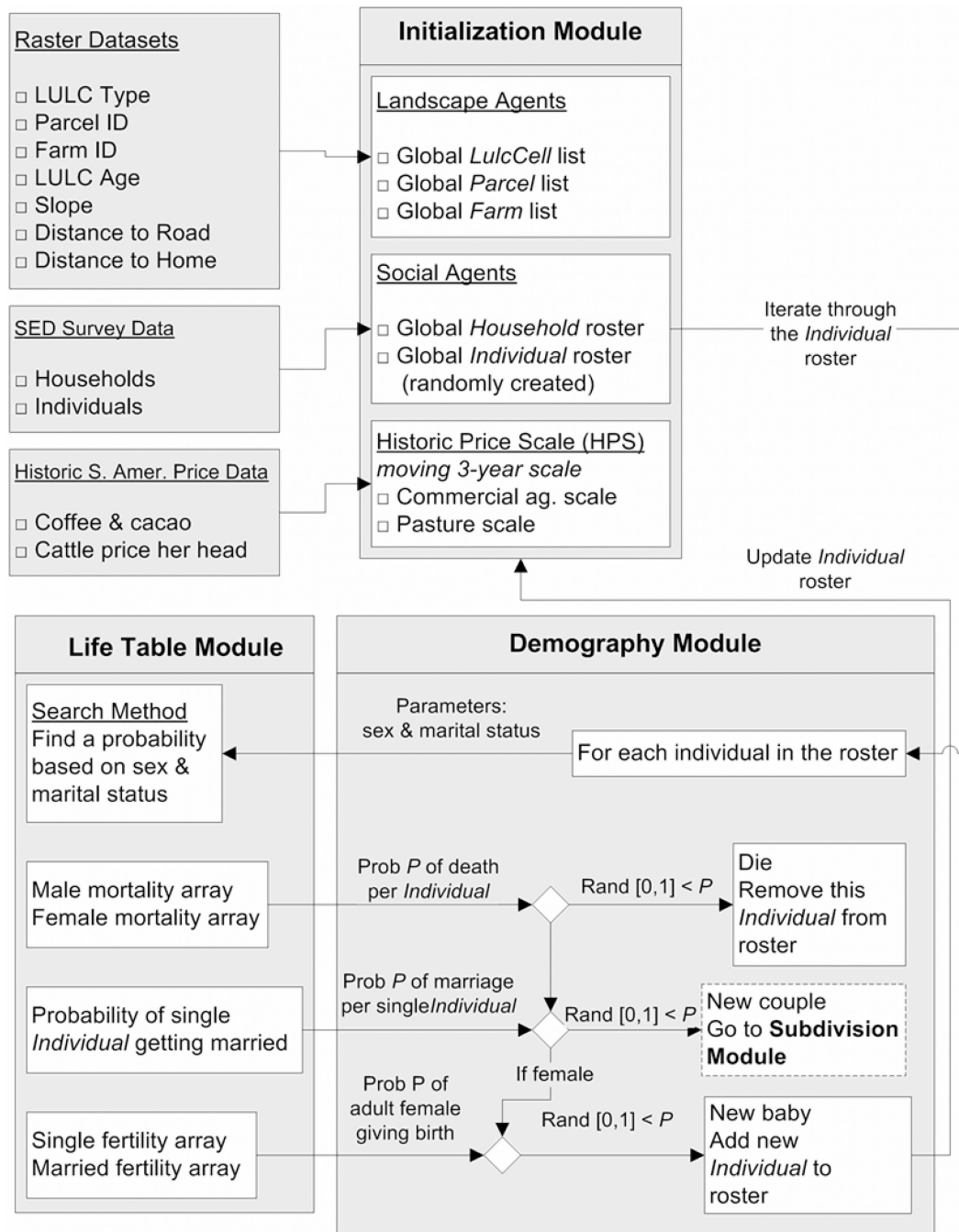


**Figure 2.**  
Delineated survey and non-survey farms in the study area.

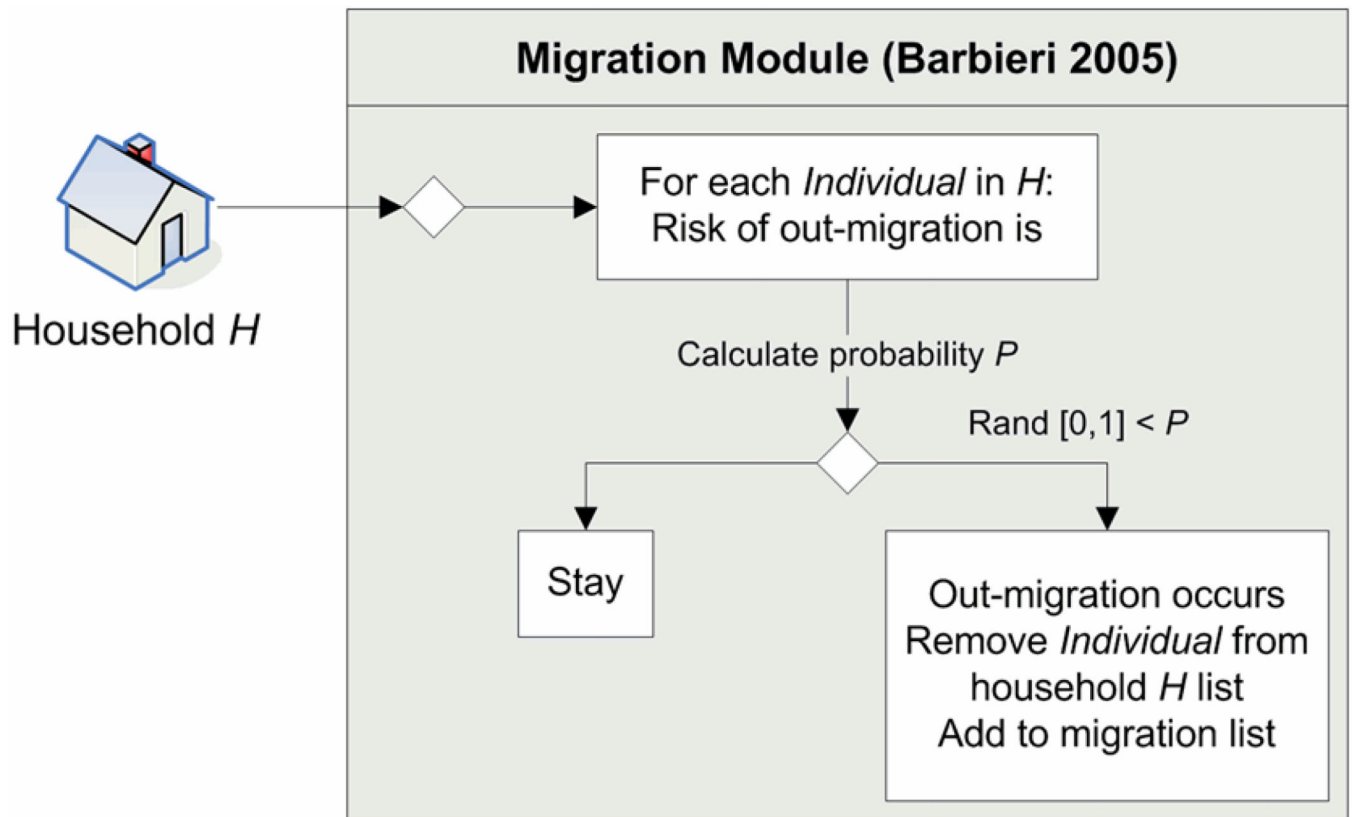




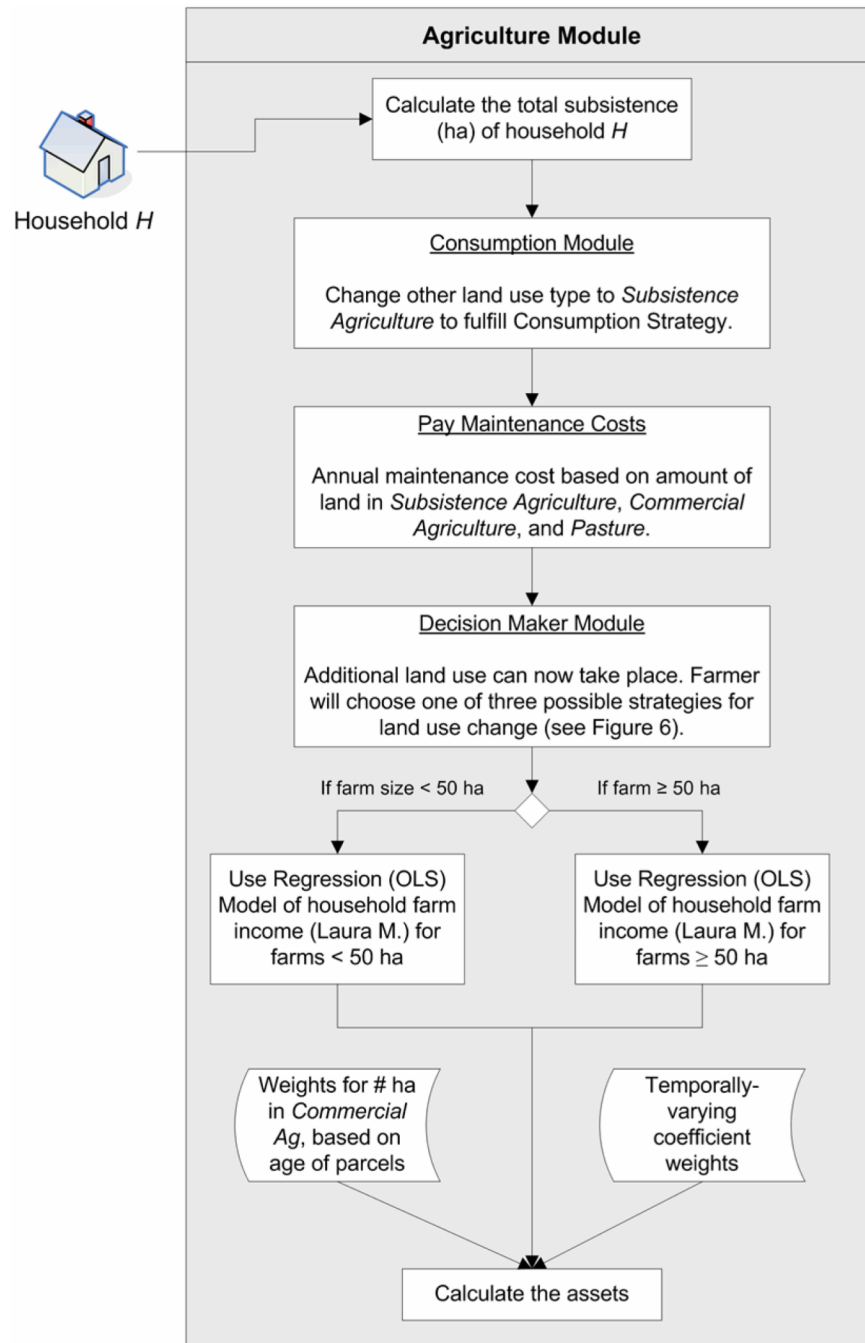
**Figure 3.** Generalized structure of the Ecuadorian Amazon ABM.



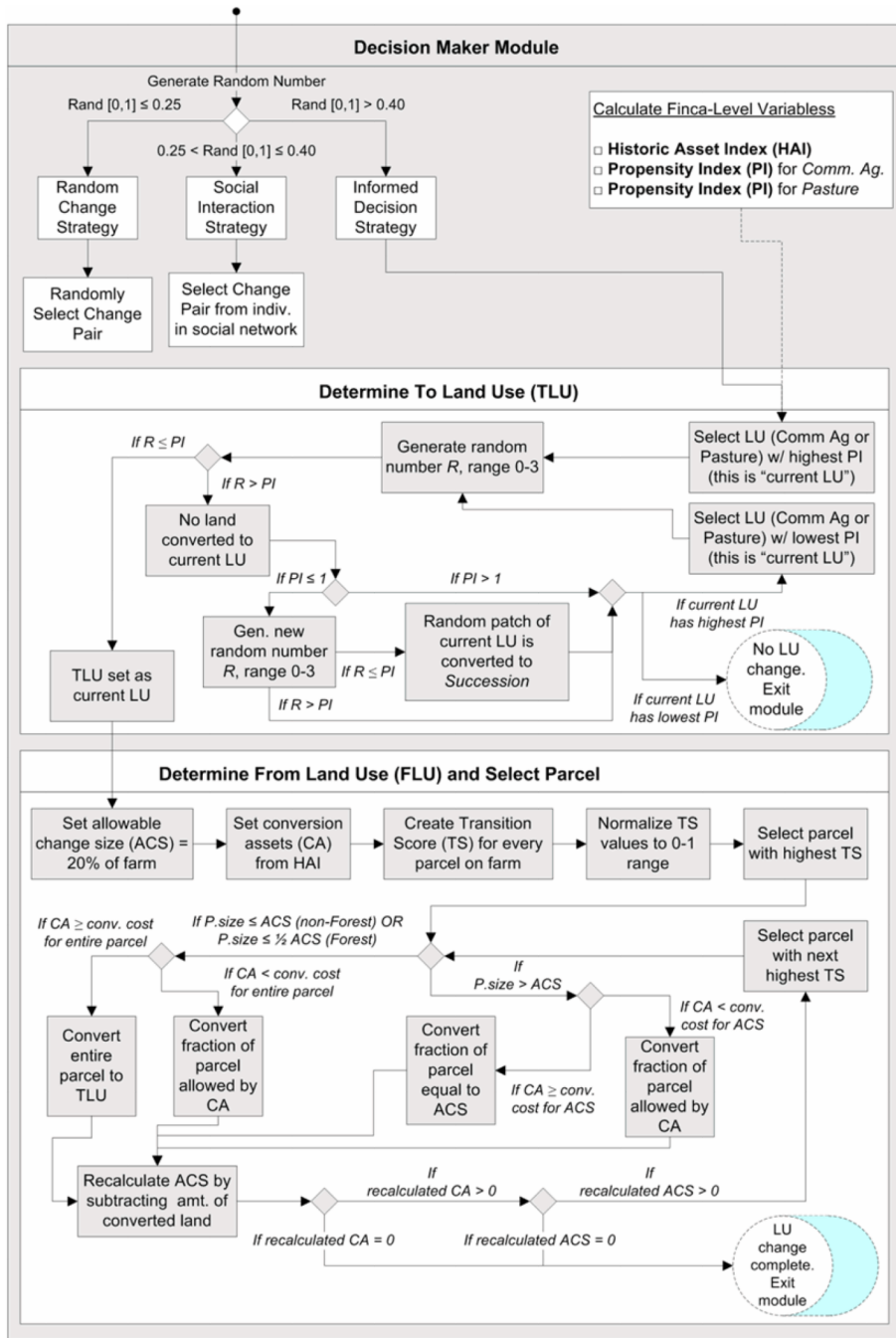
**Figure 4.** Model flow through the Initialization and Demography modules.



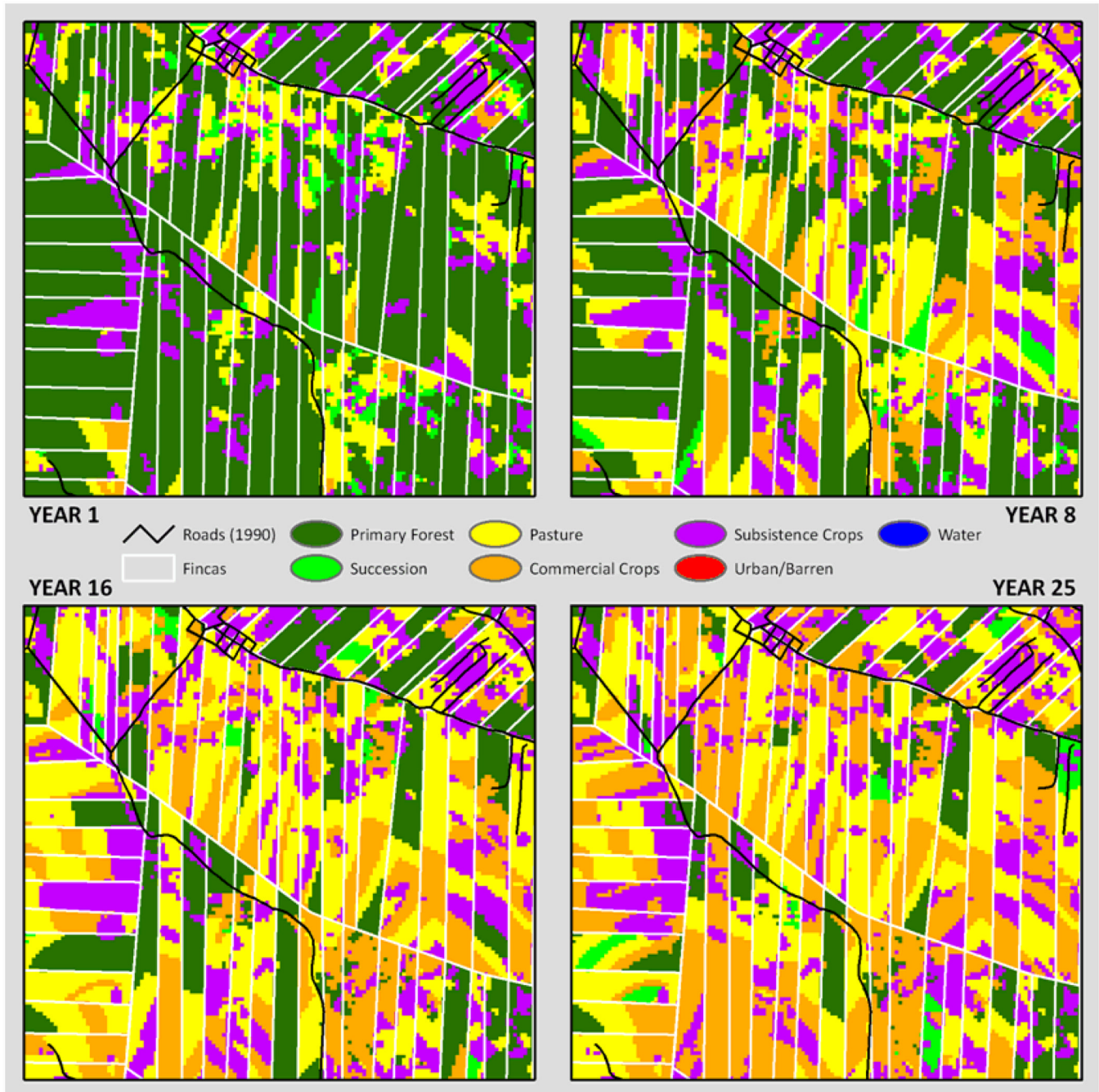
**Figure 5.** Schematic of the Migration Module.



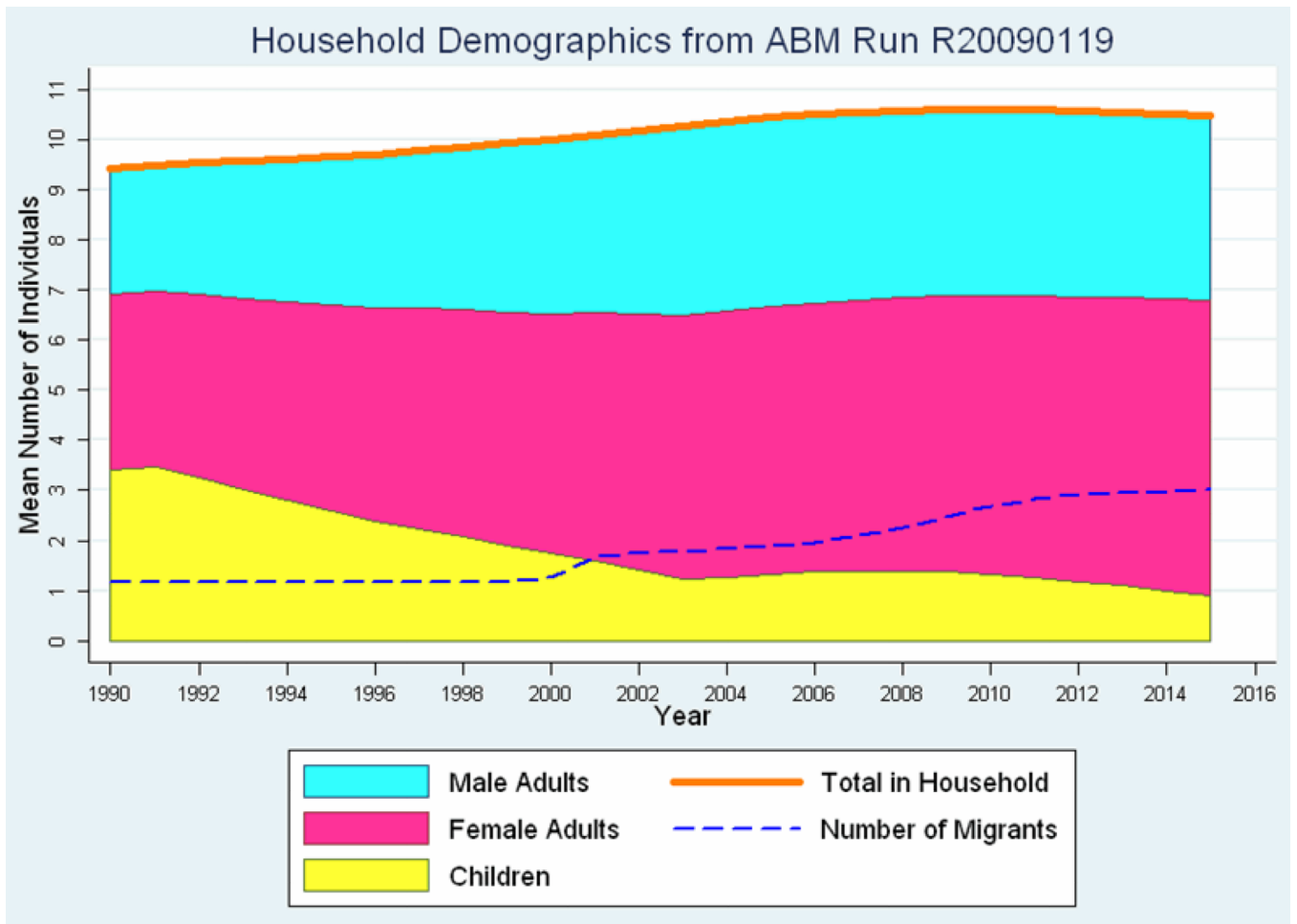
**Figure 6.** Schematic of the Agriculture Module, showing location of Decision Maker Sub-Module.



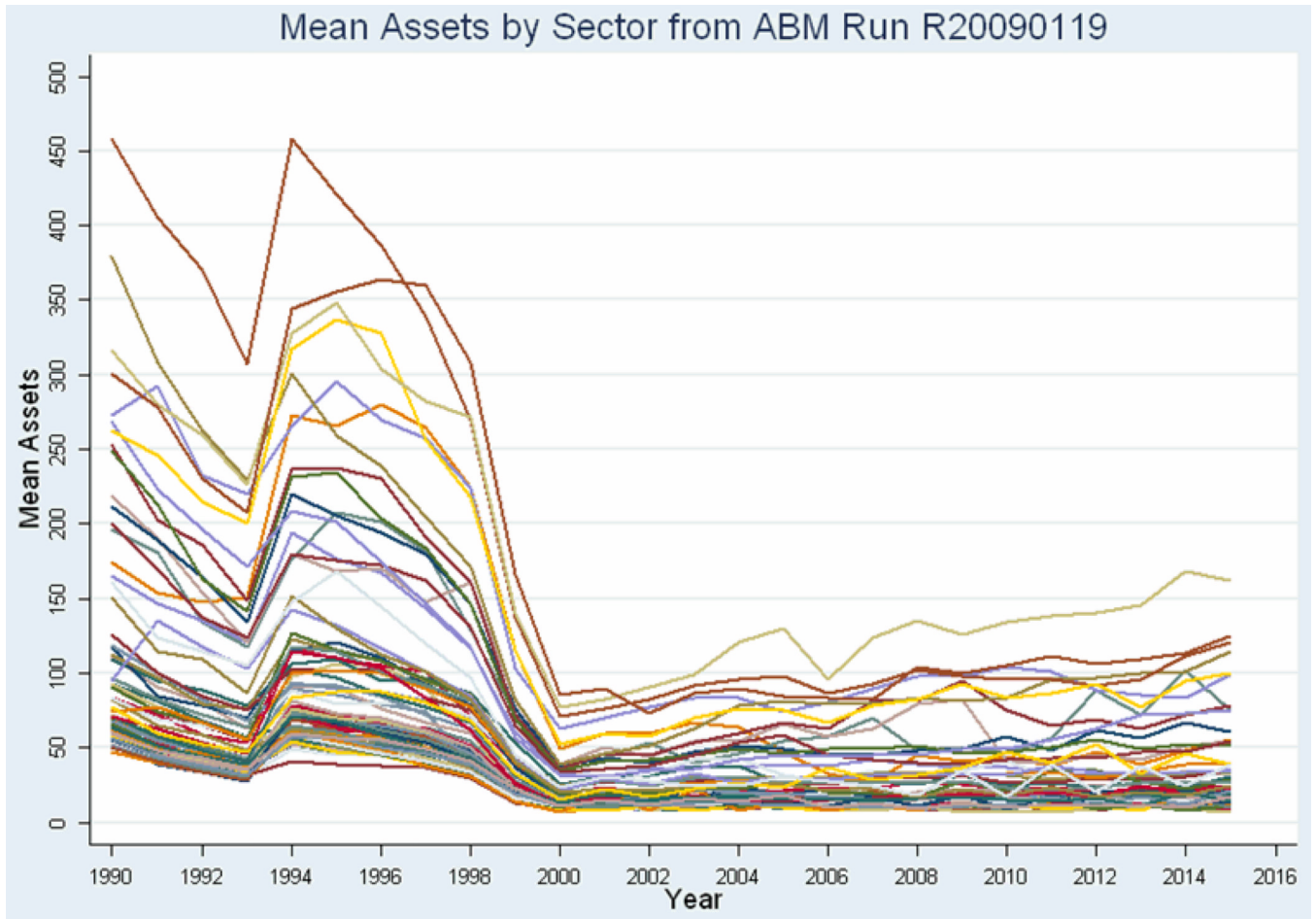
**Figure 7.** Schematic of the Decision Maker Module, including the Find Parcel Sub-Module.



**Figure 8.** An example of LULC change on household farms for a development sector for the period of 1990-2015 – sample images for years 1, 8, 16, and 25 of the simulation period.

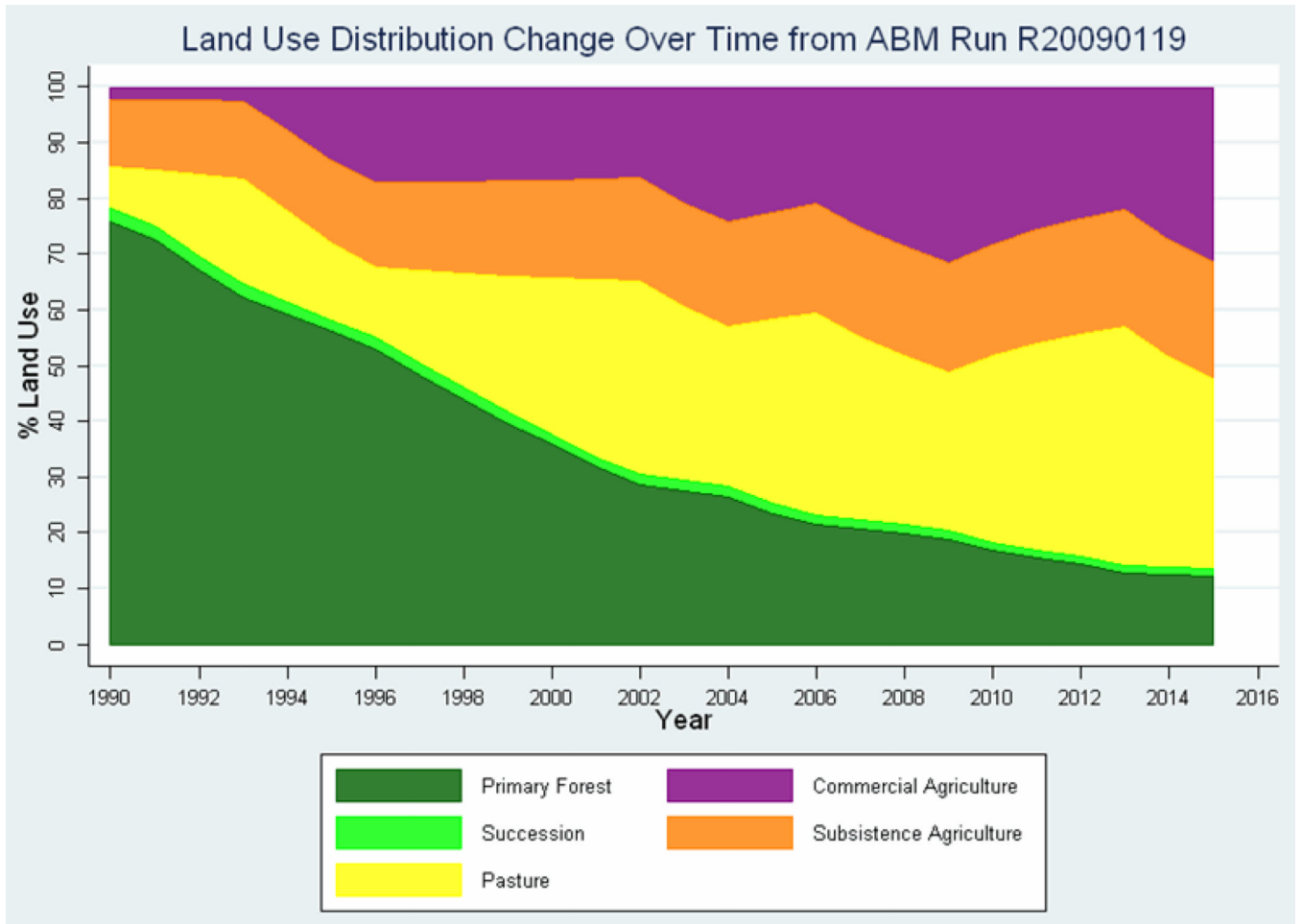


**Figure 9.** Mean number of individuals from 1990-2015 for ABM model run #R20090119.



**Figure 10.**  
Mean assets by development sector for 1990–2015 for ABM run #R20090119.





**Figure 11.** Land use patterns for 1990–2015 for ABM run #R20090119.