Land use modelling connecting spatially explicit data and linear programming: an exploration in Ecuador

Caroline Jongkamp
Maarten van 't Zelfde
Wouter T. de Groot

CML report 163
Department of Environment and Development

Land use modelling connecting spatially explicit data and linear programming: an exploration in Ecuador

Caroline Jongkamp
Maarten van 't Zelfde
Wouter T. de Groot

Instititute of Environmental Sciences (CML)
Leiden University
P.O. Box 9518
2300 RA Leiden
The Netherlands

CML report 163 – Department of Environment and Development

Copies can be ordered as follows:
– by telephone: (+31) 71 527 74 85
– by writing to: CML Library, P.O. Box 9518, 2300 RA Leiden, The Netherlands
– by fax: (+31) 71 527 74 96
– by e mail: eroos@cml.leidenuniv.nl
Please mention report number, name and address to whom the report is to be sent

ISBN: 90-5191-142-4

Printed by: Universitair Grafisch Bedrijf, Leiden

Table of Contents

1.	Introduction	1
1.1	The context of the study	1
1.2.	Case study area: the Ecuadorian Amazon	3
1.3.	Contents and limitations of this report	5
2.	Methodology: review of current practice in land use and deforestation modelling	7
2.1.	Model types	8
2.1.1.	Deforestation models according to Lambin	8
2.1.2.	Economic deforestation models according to Van Soest	11
2.1.3.	Multi-agent and Action-in-Context models	13
2.2.	Towards this report's model on deforestation processes	13
Appendix 2-I	Description of model applications	15
3.	Modelling land use decisions of farm households in tropical rainforests: an application for the Ecuadorian Amazon	17
3.1.	The Action-in-Context framework	17
3.2.	Model development: Core structure of the multi-actor land use model	19
3.2.1.	The farm household submodel	25
3.2.2.	The secondary actor submodel	40
3.2.3.	The tertiary actor submodel	47
3.2.4.	Model linkages	49
3.3.	Choosing a study area	49
3.3.1.	The study area: The Northeast provinces of the Ecuadorian Amazon	49
3.3.2.	Spatial delimitation of the study area	50
3.4	The farm household submodel applied to the Ecuadorian Amazon	51
3.4.1.	Variables and parameters in the Ecuadorian household model	54
3.4.2.	Model equations for the Ecuadorian household model	56
Appendix 3-I	List of symbols in the core model	59
Appendix 3-II	List of symbols in the model application for Ecuador	62
4.	Description of data, data analysis & transformation	65
4.1.	Data & sources	65
4.1.1.	Data on household and farm characteristics	68
4.1.2.	Data on the institutional environment and technology	68
4.1.3.	Data on the natural resource base	70
4.2.	Estimating model parameters from the data	74
4.2.1.	Estimating model parameters in the Ecuadorian farm household submodel	78
4.3.	Results of parameter estimation	94
4.3.1.	Characteristics of three selected sectors in the Ecuadorian Amazon	95
4.3.2.	Values of the parameters	96
4.3.3.	Observations on parameter estimations	98
Appendix 4-I	Description of data sources	99
Appendix 4-II	Flow charts of GIS data transformations	108
Appendix 4-III	Avenue script for calculating travel time between farms and markets	113
5.	Preliminary results from model simulations	115
5.1.	The Ecuadorian farm household submodel applied to three sectors	115
5.2.	Model results: discussion	118
5.3	Recommendations for further model development	119
Appendix 5-I	GAMS program listing	122
Appendix 5-II	GAMS solve results	126
References		127

Acknowledgement

The fruitful discussions and dataset collaboration with the Carolina Population Centre (CPC) of the University of Carolina at Chapel Hill, especially of prof. R. Bilsborrow and dr. L. Murphy is greatfully acknowledged. We further wish to thank the members of the supervisory committee of NRP (prof. B.F. Galjart, prof. E. van Ierland) and the members of the CLUE-team at Wageningen University.

1

Introduction

1.1. The context of the study

Maintaining the tropical rainforests is of vital importance for global biodiversity, national economies and the global climate. But the fate of the rainforest is mainly determined by local people and businesses and only indirectly by global policymakers. So how can one nevertheless create a global policy? One that works? These questions underlie the project “Dynamics of Tropical Forest Depletion and Protection” conducted by the Center of Environmental Science of Leiden University. The research was financed by the National Research Programme “Global Change” (NRP) as a follow-up of the research on “Local actors and global treecover policies” (see De Groot and Kamminga, 1995). This research consists of a quantitative and a qualitative study. The current report is the result of the quantitative study on “multi-actor modelling of land use strategies in tropical forest areas”. The results of the qualitative study are documented in a separate publication (Cleuren, 2001).

The problem and the method

The ever increasing, dramatic destruction of the tropical rainforest (see table 1-1) is due in the first place to the people who make the actual decisions about the use of the chainsaw, the fire torch and the bulldozer. These are local farmers, lumber companies, ministries responsible for infrastructure. Behind these primary actors one finds many other actors exerting an indirect influence, such as regional politicians, NGOs and ministries of agriculture. How can a global policy bridge this long chain of cause and effect between the local and the global level? If we do the decent thing and certify all the tropical hardwood on the world market, wouldn’t the deforestation just carry on, but now for the national markets? If tribal peoples were to gain satisfactory title to their lands, wouldn’t they just go and do what the migrants are doing already? If we were to support capacity building in government, would that increased capacity actually be utilised? It might be that the central powers in the developing countries don’t *want* to maintain the rainforests, which means that capacity won’t make any difference.

Table 1-1. *The rate of loss of tropical rainforest*

	Forest area 1990 (10^6 ha)	*Forest area 1995 (10^6 ha)*	*Forest loss 1990-1995 (10^6 ha)*	*Annual loss (%)*
Tropical Africa	523	505	18	0.7
Tropical Asia	295	280	15	1.1
Tropical South America	936	907	28	0.6

Source: FAO, 1997.

Effective global policy, pursued in terms of bilateral agreements or via global institutions, thus demands a knowledge of the causal chains running between the global level and what actually happens in the rainforest. The research project on "Dynamics of Tropical Forest Depletion and Protection" attempts to establish these connections. Methodologically, we begin with case studies of the primary occurrences and actors in the forest itself. The method then works its way upward, step by step, along the chains of actors up to the global level. Connections are established by examining the choices that the primary actors can make, and the motivations they may have for choosing a given possibility. We then examine which secondary actors exert an influence on these choices, together with their motivations. In exerting their influence, the secondary actors also have a range of options and motivations, etc. This is called an Action-in-Context (AiC) approach (De Groot, 1992).

NRP I results

In the developing countries it is the nation state that determines how much rainforest can be officially felled. This it does by issuing permits. Every extra tree that is felled is cut down illegally. But that is not in fact a serious hindrance to the lumber companies. An illegal tree is only a more expensive tree because officials and politicians have to be paid off to keep their eyes closed when the tree is felled and transported. There are thus opportunities for forestry policy in combating this informal economy. To this end, one has to act on the motivation of the officials, exercising a more effective control, for instance, or increasing the chances of an official promotion, provided his or her behaviour remains uncontroversial, and giving the official confidence that the policy really does work.

If the state or the lumber companies lay down roads – penetration roads – into the forest, then poor peasants might use them to migrate into the forest. If they do so, the central question becomes one of agricultural transition: will they continue with unsustainable logging, or will they adopt a more sustainable system? That is determined by 'blind' markets as well as by government infrastructural policy (feeder roads between the forest margin and the towns), the transmission of agricultural know-how and credit, and the establishment of organisations.

It is often said that governments in developing countries are weak and that capacity building is necessary. But the institutional weakness commonly only holds for the sectoral ministries and the executive services. By contrast, a country's central power, which simultaneously rules over both the economy and the upper echelons of the central ministries, is anything but weak. This central power determines whether the forest is to remain standing or not, via the many lines that radiate to local levels. This brings us to a global policy aimed at increasing the *motivation* of these central powers to protect the forests, rather than their *capacity* to do so. This could be done, for example, by setting up a Global Forest Fund, financed by the Western nations. Such a fund should not pay out for forestry services or promises: it should transfer funds to nations as a whole (in other worlds: to the central powers) per hectare of healthy rainforest. In that way the fund would exert a motivational force at the point where it would have the most effect.

A Global Forest Fund could be financed on the basis of the global value of the tropical rainforest. Funding could be based on ability to pay, on the expected benefits of forest maintenance, on the amount of felling that has already taken place in a given country, and/or on the basis of climate deterioration (CO_2 emission). In all cases it would be the Western nations that carry the lion's share of the financial burden. The fund should not pay out on the basis of *inputs* (projects, promises, etc.), but on the basis of the most common economic principle of payment, viz., concrete *output*. For the fund this means payment per nation, depending on its area of high quality rainforest. Payments would thus be made practically automatically, using satellite images; less area means less money. Whether the country maintains the forest or not, and how this will be done, are left entirely to the country itself. This avoids blockages associated with the principle of sovereignty. Payment to a country means payment to the seat of power, and one cannot then control how the money will be allocated from there. But this is exactly what is important for an effective payment scheme: it is essential to motivate the central powers to maintain the forests. Another criterion for effectiveness is the amount paid per hectare. Provisional estimates indicate that an annual turnover of 15 billion dollars per annum could be effective (De Groot and Kamminga, 1995). That's a lot, but a lot less than the sums mentioned in connection with the stabilisation of CO_2 emissions by means of a global climate fund.

1.2. Case study area: the Ecuadorian Amazon

In the framework of the present project, field research to the Northeast of the Ecuadorian Amazon has made clear how the mechanism of tropical deforestation operates there. The study area was chosen in the provinces of Napo and Sucumbios; in the region around Coca[1]. This area is in the heart of the lower rainforest region. The discovery of oil thirty years ago in this area of Amazonian jungle has wrought a radical change in the region. Western oil companies posses concessions over large areas of rainforest and are the pioneers of forest exploitation. They use heavy equipment to lay down roads and to get the oil wells into production. This is accompanied by the destruction of areas of rainforest and pollution of soil and water by poisonous drilling fluids and oil. In their wake follows an army of impoverished colonists and Quichua indians, in search of land. These are the primary actors in the deforestation and they have established themselves along the oil roads over the last thirty years. Most of them lead a marginalised existence on their plots of low fertility land where they cultivate manioc, plantains and rice for their own consumption, and coffee for the market. The area is "backward" in the sense that the availability of health care, education and other services is far below the standards in the rest of the country. Despite the harsh living conditions, the region is still attractive to colonists because of the employment possibilities in the oil business and the relative abundance of land.

Large cattle ranches are hardly found in this area due to climatic conditions, inadequate knowledge and capital, and poor grazing lands. The peasants do not fell large areas of forest; rather, they work a few hectares to produce their own food, eventually felling the most valuable tree species in the remnant forest for sale to the lumber companies. Forest clearing is slower than in other regions in the Amazon. The year-round rainfall makes it difficult to perform slash-and-burn clearing. A common procedure is the so-called slash-and-mulch method, by which the vegetation is not burned, but cut and left to rot. The rather slow pace of deforestation gives room for interventions geared towards sustainable land use practice, provided that the farmers have the necessary incentives which are vital if they are to develop new and sustainable initiatives.

The commercial extraction of timber does not play a major role. The costs involved in the extraction, transport and processing often do not outweigh the costs. In the face of international pressure, the national forestry service has opted for a hard line, officially strictly regulating the timber trade and granting no logging concessions. In practice, however, the forestry service and the military, both local and central, appear to be more interested in personal gain than in the enforcement of laws and regulations. The result is that official rules are hardly obeyed and that the market sets the conditions for timber extraction. A policy aimed at the very necessary strengthening of the forestry service should be accompanied by arousing an interest in forest maintenance among the ruling power.

[1] The reader is referred to Cleuren (2001) for a detailed report of the field study in Ecuador.

The fate of the Ecuadorian Amazon is to a large extent governed from the capital city, Quito, where the central power views the region merely as an area for the extraction of oil. Though much criticized for its environmental effects, the oil sector can not be denied for its important contribution to national living standards, public budget and foreign exchange. Oil extraction still continues, and so does the construction of new roads into the forest. Only if this process can be halted is there a chance that the process of deforestation can in fact be controlled. However, the lack of political will is a matter of serious concern. The Ecuadorian case study confirms the crucial role played by the central elite in deforestation. The key towards sustainable forest management should therefore be searched in economic stimuli that can change this group's behaviour.

1.3. Contents and limitations of this report

The remainder of this report describes the interaction between the primary, secondary and tertiary actors in deforestation. We have chosen for a modelling approach which makes it possible to quantify the underlying processes and motivational factors. The model was designed with the purpose to evaluate the effect of policy scenarios which slow down the deforestation process. Chapter 2 contains an overview of current research in land use and deforestation modelling. It concludes with observations on modelling techniques that we will employ in our study. Chapter 3 further develops the model for three major actors: the farm household as the primary actor responsible for deforestation; the government as the secondary actor influencing household strategies; and an international organisation which represents the (western) donor community. The model incorporates methods of socio-economic research and GIS techniques. A reduced version of the model is presented for a specific study area in the Ecuadorian Amazon region. Chapter 4 continues with a description of the data analysis required to perform the model calculations. Finally, the Ecuadorian model is tested with data on farm household behaviour in the study area connecting the GIS-data with the linear programming using the GAMS software (chapter 5).

As may be clear from this description, the breadth and multi-agent scope of the first chapters of the report is not maintained up to the last. This is the result of a host of practical difficulties that hit the project especially during its second, empirical phase. The broad and methodological chapters do have their own value for further research, however, which is why they have been kept rather than truncated to exactly fit the empirical scope. Thus, we hope to contribute to the further development of multi-agent and multi-level modelling of land use change, using the traditional and robust tool of lineair programming.

2

Methodology: review of current practice in land use and deforestation modelling

Why modelling?

The choice for a modelling approach was made on the one hand because of the complexity of actors' decisions and their interactions, and on the other hand because of the possibilities that quantitative modelling offers for policy simulations. The causal linkages between actors' behaviour and the motivational factors are the basis for the design of the model equations. After the basic structure of the model has been designed, the model can be validated in a local context. The validation process requires that a thorough data analysis is made, and parameters are estimated to specify the sensitivity of the actors to the local circumstances.

The modelling procedure follows the Problem-in-Context and Action-in-Context frameworks as developed by de Groot (1992). The Problem-in-Context (PiC) approach starts from the environmental problem, in casu destruction and degradation of humid tropical forest. After describing the *problem*, the next step in PiC is to describe its *context*. Problematic actions, which are, in their turn, caused by decisions of actors, and the factors and motivations that facilitate these decisions cause the environmental effect.

This chapter has been written in an early stage of the project as a preparation for the quantitative modelling of land use in tropical forest areas. Many others have used modelling techniques for similar purposes as we do. The integration of both physical and socio-economic factors in quantitative models is – methodologically speaking - the most challenging part of our research because there is little experience in this field. Even though our focus is on the human driving forces of deforestation, we have been determined from the beginning of the research to include geo-bio-physical factors in the modelling, and to do that to the best of our knowledge. In the real world there is a day-to-day interaction between human and physical factors and our aim was to develop a powerful (modelling) tool which is sufficiently robust to improve our insight in the nature of these interactions.

The task that we have set ourselves requires combinations of modelling techniques from various (sub-) disciplines. The following sections provide a review of modelling techniques that are usually applied in land use and deforestation studies, as well as concise discussions on the applicability for the purpose of our research.

2.1. Model types

This chapter is mostly based on three studies which together give an overview on the type of models that have recently been used to analyze the human causes of deforestation. The first is Lambin (1994), the second Van Soest (1994), and the third is Lonergan and Prudham (1994). This section merely reflects the findings of these authors. Lambin focuses on searching a complementarity between monitoring changes in land-use/land cover by remote sensing and the modelling of change processes. As such, the models that he discusses are necessarily spatial. In fact RS and modelling efforts tend to come together only sparsely, yet its combination is one of the most promising "novelties" within the field of GIS.
Lambin is quite brief in his discussion of economic models. This gap is nicely filled in by Van Soest, who focuses particularly on economic models. Lonergan and Prudhan discuss in particular lexicographic goal programming and multi-criteria analysis.

2.1.1. Deforestation models according to Lambin

Lambin distinguishes (Lambin, 1994, page 22) three types of models for processes of land-use and land-cover change: *empirical*, *mechanistic* and *system* models. Empirical models are based on observed relationships between variables. Mechanistic models are based on the modeller's knowledge of the underlying processes by which the system operates. Parameters must be estimated from data, and the suitability of the equations needs to be verified by observations. System models take into account several complex, interacting processes. These models are often difficult to validate and very computer intensive. The models can be made *spatially explicit* to predict changes in spatial structure of the landscape and map the flows that occur between locations. After these and a few more general remarks, Lambin proceeds to the description of more specific model types:

1. **Markov chains:**
   In Markov chain models, *probabilities* are assigned to the transition from state i to state j. For land-use studies, the states of the system are defined as the amount of land covered by various land uses, as percentages of the area of each landscape unit. The transition probabilities are stationary over time. The probabilities can be statistically estimated from a sample of transitions which actually occur and can be observed during a certain time interval. If a_{ij} indicates the transition between state i and state j, then the transition probabilities p_{ij} are estimated as:
   $p_{ij} = a_{ij} / \mathrm{SUM}_j\, p_{ij}$
   In the long run, the predictive power of the Markov chain in its basic form is quite limited because of its assumption of *stationarity*. It does not explain the underlying causes of the transition, nor does it take into account any possible changes in the underlying causes. In order to address the underlying causes, it is possible to incorporate the contribution of exogenous or endogenous variables to the transitions. Hence, p_{ij} could be estimated as a function of the variables $X_1, .., X_n$: $p_{ij} = \mathrm{f}(X_1, .., X_n)$.
2. **Logistic function models** serve to describe trends which progress along an S-shaped (logistic) curve. It assumes that the rate of deforestation is regulated by the density of the deforested areas, increasing in the early stages of deforestation, and decreasing when the forest becomes scarce. A drawback of these models is that the underlying (social and ecological) mechanisms are not

explicitly taken into account, or only very crudely. The focus is on the rate of deforestation. The models are generally not spatially explicit.

3. **(Linear) regression models** postulate a (linear) relationship between the dependent variable (*Y*) and independent variables (X_i) in the form:

   $Y = a_0 + a_1X_1 + a_2X_2 + .. +$ random error

   where the a_i are the regression coefficients, usually estimated by the ordinary least squares (OLS) method. Regression analysis can be conducted by *cross-sectional analysis* or by *panel (time series) analysis*. Lambin sees a weakness in the application of cross-national analysis applied to a wide range of countries in very different situations. He therefore makes a plea for preliminary stratification of the globe into homogeneous zones, on the basis of geographical, ecological or socio-economic criteria. Another weakness of cross-sectional regression models is that they do not allow for cross-boundary effects.

   Many examples exist of studies which used regression analysis to establish relationships between deforestation and explanatory variables such as population density, road density, soil productivity, etc. Lambin mentions a few. Among them is an interesting study on deforestation in the Philippines (Kummer, 1991), which demonstrates the role of roads in opening up the forest area for migrants. Another - more comprehensive - overview can be found in Brown and Pierce (1994).

4. **Spatial, statistical models** combine remote sensing, geographic information systems and statistical methods. They offer cartographic projections and establish correlations between the spatial occurrence of land cover (derived from RS data) with landscape and locational attributes (e.g. from aerial photography). The models deal primarily with the question of which areas are most susceptible to deforestation. A dependent variable, like the "change in land cover" is related to a number of independent variables such as slope, elevation, degree of fragmentation of the forest, proximity to housing, and proximity to roads and streams.

5. **Models of population pressure, agrarian change and deforestation.** Population pressure increases the demand for food, and hence the demand for arable land. This forces people to use forest land for cultivation. On the other hand, people may use techniques which increase the productivity of the land which is currently available.

   This type of models are rather crude and have limited applicability on a local scale. They only explain agrarian changes which are driven by population growth. Non-demographic causes of deforestation are not taken into account. The models are mainly *qualitative*, explanatory, non-predictive and non-spatial.

6. **Models of peri-urban land-use change.**

   This type of models tends to be based on von Thünen's model of the isolated state. The model by von Thünen analyses economic rent and agriculture as a function of the distance to the market. As a result of transportation costs, the landscape is organised around the market in rings of land use of decreasing intensity with increasing distance from the market. In the outmost ring, the natural forest is the optimal type of land use. The von Thünen approach has good explanatory power only where markets play an important role (e.g. in wood exploitation and commercial farming). According to Lambin, the von Thünen model could be improved by using natural and cultural landscape information in a GIS framework and by using land-use/land-cover information extracted from remotely-sensed data.

7. **Econom(etr)ic models**
   A common econometric approach is to specify and estimate supply and demand functions of market goods. Examples of such "market goods" are forest land, labour, forest products, or agricultural cash crops. The standard estimation technique is ordinary least squares regression. Lambin further mentions optimisation theory, partial equilibrium modelling, and general equilibrium modelling.
   The economic models described by Lambin are not explicitly spatial, but he assumes that there is a clear potential for developing spatial economic models of deforestation processes, using concepts from location theory and from regional science.
   A more elaborate sub classification of economic models is given by Van Soest (1994). See the next section.
8. **Ecosystem simulation models:**
   Ecosystem models emphasize the interactions among all components that make part of the system. The dynamic behaviour of the system is then studied through simulation experiments with the model. The simulations can be used for analytical and for predictive purposes. Lambin mentions the IMAGE model, developed by RIVM, as an example of this type of models. Lambin considers it a prerequisite for the development of this type of models that the processes and mechanisms of deforestation in a given situation have been thoroughly investigated, for example through detailed field studies.
   Lambin considers it a disadvantage that most ecosystem models treat the system as spatially homogeneous. Spatial heterogeneities could be introduced more explicitly (see under 9 below), but this is limited by computational constraints on the number of flows to be estimated between spatial units.
9. **Dynamic, spatial simulation models:**
   This model type is an expansion of the ecosystem models mentioned above. They include (i) the spatial heterogeneity of the land surface and (ii) the processes of human decision-making underlying changes in land uses. In this type of models, the flows between adjacent grid cells are explicitly taken into account. Lambin mentions two examples of these models, which are both spatially explicit actor-based models. The first, developed by Wilkie and Finn (1988) applies to the equatorial forest in Zaire. The second, the DELTA model (see e.g. Dale et al, 1993), is developed at the Oak Ridge National Laboratory and has been used to contrast two land use systems in Rondônia, Brazil.

The model types discussed by Lambin each serve a different purpose, and can be used complementary to each other. Lambin evaluates to what extent the model categories are capable of answering the questions as to *why* deforestation takes place, *when* (how fast), and *where* deforestation is likely to take place. His conclusions on which are the main contributions of each model type, as well as the broader class of the models (empirical, mechanistic or system) are presented in table 2.1.

Table 2.1 : Main contribution of each category of models

Types of models *Issues to be addressed*	**Empirical models**	**Mechanistic models**	**System models**
Why?	Regression models	Population pressure & Economic models	Ecosystem models & Dynamic spatial simulation models
When?	Markov chain & Logistic function model	-	
Where?	Spatial, statistical models	Models of peri-urban land-use change	Dynamic spatial simulation models

Source: reproduced from Lambin (1994), page 97, table 1

Lambin advises to start building stochastic models characterized by a simple structure, and to progressively elaborate those models. Though I fully agree with his idea to incorporate stochastic elements, I would rather consider to leave these *out* at the first stage of modelling, and to incorporate them in the elaboration phase. The value of stochastic elements will be evident also in case of deterministic processes, "to represent unpredictable factors associated with deforestation" (quote from Lambin, page VII).

2.1.2. Economic deforestation models according to Van Soest

Van Soest (1994) discusses four model types which are commonly used in applied economic analysis. He distinguishes *statistical analysis*, *input-output models* (and Social Accounting Matrices), *linear programming models* and *general equilibrium models*. Statistical models have been mentioned already in the previous section, therefore I will not repeat the discussion here. I will briefly describe the other three.

10. **Linear programming models**

    The essence of linear programming models is the constrained optimization of an objective function. The objective function specifies the preferences of a decision maker. An example of an objective function is to maximize profits in a production process. The constraints deal with matters such as the production capacity and the availability of labour.

    In a linear programming model, we can make a distinction between *decision variables* (X_1, X_2, ...) of which the values have to be determined by solving the model, and exogenous *parameters* (c_i, a_{ij}, b_j; $i=1,..,n;\ j=1,..,m$) which are used as input to the model. The basic mathematical structure of a linear programming model is:

Maximize: $c_1X_1 + c_2X_2 + ... + c_nX_n$

Subject to the constraints:

$$a_{11}X_1 + a_{12}X_2 + ... + a_{1n}X_n \# b_1$$
$$a_{21}X_1 + a_{22}X_2 + ... + a_{2n}X_n \# b_2$$
$$.$$
$$.$$
$$.$$
$$a_{m1}X_1 + a_{m2}X_2 + ... + a_{mn}X_n \# b_m$$

Van Soest mentions the TROPFORM model as an example of a linear programming model which is built to simulate possible trends that affect deforestation on a global scale (see Jepma, 1995). The model focuses on the production and trade of wood products.

Related model types are the *non-linear programming model*, which has non-linearities in the constraints or in the objective function, and the *integer programming model*, which requires that the optimal values of the decision variables have integer values.

11. **Input-output models and Social Accounting Matrices**

Input-output models are built on so-called *input-output tables*. Generally, these tables are composed for national economies, or for regions. These tables specify the technical relationships between quantities of *inputs* and quantities of *outputs* in the production processes of the economy. The output produced in one sector may be used as an input in another sector, and so on. The quantities of goods required as inputs together constitute the *intermediate demand*. Not all outputs are used in the production process, part of it is directly consumed. This is called the *final demand*. Based on the input-output table, the model can calculate for any value of the final demand how many *intermediate goods* need to be produced. Therefore, it is a useful model to analyze how changes in one sector induce changes in other sectors.

A *Social Accounting Matrix (SAM)* is an extension of the Input-Output table to include social effects of production processes. For example, forest degradation can be included as one of the social effects. With the help of the SAM, it can be seen which of the activities (as specified in the input-output table) inflicts most damage on the forest land. Because all the interlinkages between the sectors are taken into account, it measures not only the direct effects, but the indirect effects as well.

A disadvantage of the I-O and SAM based models is that it assumes fixed relationships between inputs and outputs, even if major changes in demand occur. A second disadvantage is that it assumes that supply adjusts perfectly to demand.

11. **Computable General Equilibrium (CGE) models**

CGE models resemble input-output models in the sense that they explicitly take into account the interlinkages between different sectors in the economy. An important difference is that CGE models can also include non-linear relationships. They usually include non-linear supply and demand functions. Prices can be solved endogenously. Van Soest mentions an example of a CGE model in which forest and non-forest land are included as production factors.

2.1.3. Multi-agent and Action-in-Context models

Further removed from economic science but still largely applying rational choice theory, multi-agent and Action-in-Context approaches stress the needs and opportunity to causally relate various actors with each other, e.g. to explain tropical deforestation.

12. **Multi-Agent Modelling (MAM)**
    Multi-agent models usually work with relatively simple actor equations, focusing rather on the system-level effects or the interplay of may of such actors.

13. **Action-in-Context (AiC) modelling**
    Responding to the need to link the causal strength of actor-based methodology with higher-level cultural and structural variables, De Groot (1992) developed the Action-in-Context framework. One of its characteristics is that actors are related to other actors through influences on these actors' options and/or motivations for action. Thus, 'actor field' structures can be developed in which direct ('proximate', 'primary') actors are causally connected to higher-level ('secondary', 'tertiary' etc.) actors. See also De Groot and Kamminga (1995) for tropical forest examples.

2.2. Towards this report's model on deforestation processes

Many of the models described above could be useful to our own modelling efforts. The most appealing model type is the one which Lambin denotes with the *dynamic, spatial simulation models*. They have everything that could be needed: actor-based hence causally oriented, spatially explicit, linked with remotely-sensed data, stochastic simulation, ... When aiming towards such a comprehensive model, we could use a number of the other model types as well. For example, regression analysis can be used to estimate parameters which are used as inputs in a spatial simulation model. Out of the above lists of model categories I will mention a few which I consider potentially useful for our own modelling efforts:

- dynamic spatial simulation models: the most powerful integrated modelling type
- regression models: useful to estimate parameters and distribution functions
- spatial statistical methods: useful to estimate parameters of an evidently spatial nature
- von Thünen's model: to explain the behaviour of actors vis-à-vis distant markets
- linear programming models: useful to model the decision process of an actor if constraints play an important role

This list is not meant to be exclusive, other model components may be useful as well.

First of all, it is worthwhile to consider which criteria our own modelling efforts should meet. The model should aim to approach the real situation to such an extent that it does support a better understanding of the processes, as well as provide possibilities to evaluate policy measures geared towards less forest destruction. The model should be capable to describe human behaviour in a local context as well as have a certain relevance in a global context. In other words, it should be robust, flexible, and have explanatory power in a local and global sense.

The development of the model is a stepwise process. It is built up similar to the action-in-context approach, starting from the strategies of the primary actors and gradually proceeding to the secondary and tertiary actors. Of course, we need to do a lot of testing and validation with data in between. The modelling efforts and the improvement of our understanding of mechanisms is an interactive process. On the one hand modelling is nothing more than formally describing what we already know, on the other hand, the calculations with models and the outcomes provide a check on whether we actually understood the logic of what is taking place. In this way, it is a steering instrument to guide the collection of data in the field. The knowledge gained in the field should help to improve the models. Data from secondary sources also provide an indispensable input in the improvement and validation of the models.

The model is supposed to be *spatially explicit*. This involves a number of expansions of the model, which deal with (a) the locational aspects of the land and (b) additional variables, parameters and relations which allow us to introduce a certain spatial diversity.

Appendix 2-I

Description of model applications

Allen

The model by Allen (1985) is briefly discussed in Jepma (1995, p.71). It is a forest conservation model of the Dodoma region in Tanzania, used for comparing policy scenarios aimed at the transition from deforestation of public forest areas to the use of plantation wood. The model, a multi-goal programming model, generates an optimal combination of efficient wood production, allocation of labour and conservation of the forest.

CLUE

The CLUE model (Conversion of Land Use and its Effects; Veldkamp and Fresco, 1996, Schoorl et al., 1997) is a spatially explicit, multi-scale land-use change model. It was built up through regression modelling, where the potential driving forces are included by stepwise inclusion according to their performance in the regression model. Applications and further model development of the CLUE model are described in De Koning, 1999; Kok, 2001; and Verburg, 2000.

Hassan and Hertzler

Hassan and Hertzler (1988, discussed in Jepma, 1995) describe a dynamic programming model for Sudan in order to develop an optimal policy to control forest exploitation as a source of energy on the one hand and desertification on the other hand. The model uses real prices of fuel wood, which take into account the costs of logging for future generations.

IDIOM

IDIOM is a global simulation model used by Jepma (1995) for scenario analyses. The IDIOM model integrates TROPFORM (see below), SARUM (a global simulation model developed by the Systems Analysis Research Unit in 1978) and a land use module. SARUM divides the world in 10 geographical, economic and/or political regions (North America, Tropical Latin America, Rest of LA, etc.).

Image 2.0

Image consists of a number of modules, one of which is the carbon cycle module. This module, on its part, contains a *deforestation module* as a submodule (Rotmans, 1990, section 3.7). The submodule takes into account the following processes:

- permanent agriculture
- shifting or pioneer cultivation
- cattle breeding
- logging
- fuelwood gathering
- industrial projects (in Amazonia only)
- reforestation.

Because of the different characteristics of the tropical forest areas, they are subdivided according to the three continents: Latin America, Africa and South-East Asia.

TIMPLAN

A sector simulation model developed by Gane (1986, described by Jepma, 1995) in order to determine the best strategy for forest development on a national level. The model generates projections of the future demand and supply of wood, costs and benefits of the forest sector, earnings in foreign currency and benefits for the national economy.

TROPFORM

This global model describes several factors determining land use and leading (eventually) to deforestation (described extensively in Jepma, 1995, pp. 71-82 and Blom et al., 1990). The emphasis is on the logging industry. The model consists of the following modules:

- consumption module (assumption: all demand for wood products is met)
- spatial allocation module (linear programming)
- standing volume module
- growth reserves module
- deforestation module

The deforestation module considers three types of land use:

- farmland or agricultural land (meets consumption demand for food)
- forest
- the residual areas (urban areas and non-forest nature) are assumed stable.

3

Modelling land use decisions of farm households in tropical rainforests: an application for the Ecuadorian Amazon

This chapter discusses a modelling framework for the analysis of land-use decisions in tropical moist forest areas, with an example from the Northeast Ecuadorian Amazon. The core structure of the model is a multi-actor modelling framework. It takes into account the interaction between primary actors (e.g. the farm households), the secondary actors (e.g. the oil companies and the government) and the tertiary actors (e.g. the international donor community). In section 3.1 it will be explained how the decisions of the actors influence each other through a presentation of the Action-in-Context framework. Section 3.2 discusses the development of the quantitative model. The result is a framework for what we call the “core model”. The core model is built up from a number of submodels which are linked to each other. Each submodel deals with the decision-making process of one type of actor. Section 3.3 contains a geographical delimitation which results in the choice of a particular study area in the Ecuadorian Amazon. Then, in section 3.4. we will present an example of the household submodel as it was applied to the land use decisions of a farm household in the Ecuadorian Amazon.

3.1. The Action-in-Context framework

The core model describes the causal linkages between actors’ behaviour and the motivational factors. The model development follows the Problem-in-Context and Action-in-Context frameworks (see de Groot, 1992). The environmental problem that is at the heart of this study is the destruction and degradation of tropical moist forest. The context of this problem is given by the decisions of the most important actors, as well as the underlying motivational factors. The decision-making processes will be discussed and modelled in the remainder of this chapter. First we will illustrate the influential linkages between the actors in an *Action-in Context* (AiC) framework which comprises, amongst others, the identification of primary, secondary and tertiary actors which together cause the problematic action (see Figure 1).

Figure 1. *Linkages in the Action-in-Context framework*

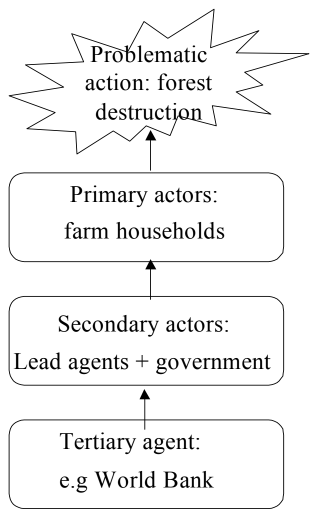


Figure 1 illustrates that the *primary actors* are the ones most proximate to the problematic action. We have chosen to focus on one primary actor in particular: the farm household. We acknowledge that farm households are not the only primary actors, but their role in deforestation is a crucial one in many parts of the world. Other actors, such as governments and so-called "lead agents" (see Rudel, 1993) have the role of *secondary actors*, because their part in the deforestation process is mostly indirect through the encouragement and discouragement of the primary actors. The secondary actors influence the farm households by setting the conditions under which the farmers have to operate. In many cases the primary actors take the actions of the secondary actors as "given", at least in the short run. In the long run the interaction between primary and secondary actors is more complex: through pressure groups, voting and other forms of collective action the farmers do have a certain influence on the actions of secondary actors such as the national governments. In the context of the model we will take this into account only partially and depending on the local context. If important pressure groups exist they will be included as secondary actors. Also, we assume that secondary actors such as governments are to a certain extent willing to pursue farmers' interests in order to strengthen the support from the community. The *tertiary actors* are characterized by the fact that their most important role in the deforestation process is through their influence on the secondary actors. For most tropical moist forest areas we can say that the World Bank, a number of international environmental organizations, international trade organizations and a few key multinational businesses typically have a "tertiary" influence on tropical deforestation.

3.2. Model development: Core structure of the multi-actor land use model

The model development takes the same steps as the Action-in-Context framework: from the primary actors up to higher levels of influence. We can structure the model development along the following steps, which need not necessarily be taken in this exact order:

a) Define the decision-making process for each actor
b) Identify the available data and collect additional data
c) Choose the unit of analysis and link this to a spatial unit
d) Design model structure and equations
e) Estimate model parameters
f) Run the model for a baseline scenario
g) Evaluate the model results for the baseline scenario
h) Design and implement model experiments for alternative scenarios

Steps a)-d) will be discussed in this chapter. Step e) is the subject of chapter 4. Steps f)-h) will be discussed in chapter 5.

Ad a) Define the decision-making process for each actor

Before the behaviour of the actors can be modelled we will first give a more formal description of the decision-making process of each actor. The knowledge on these decision-making processes is obtained from theory and (for specific study areas) from field research, and results in a description of "hypothesized behaviour" of the actor at stake. This hypothesized behaviour is necessarily a generalization of actual behaviour and is the basis for a set of assumptions which together describe how the actor responds to the motivational factors. The assumptions might be verified in the field by means of survey techniques commonly applied in the social sciences, such as sampled surveys or even in-depth interviews. This, however, is quite expensive and often impossible on a limited research budget. In our research we based our assumptions on actor behaviour on "light" surveys and a relatively large component of secondary sources. In this way, a crude understanding was obtained on actor behaviour with respect to the depletion or protection of the rainforest.
First we will formulate a number of assumptions which have been derived from the literature on farm household economics and on deforestation (e.g. Ellis, 1988; Singh, Squire & Strauss, 1986; Rudel and Horowitz, 1993; Brown and Pierce, 1994; Jepma, 1995). [2]

The assumptions regarding the primary, secondary and tertiary actors

The primary, secondary and tertiary actors are defined as follows:

1. The primary actor is the farm household[3]
2. The secondary actors are a lead agent (like a logging company or the oil sector) and the government. The government may have to be split into several divisions, depending on whether these function as relatively autonomous institutions.

[2] In the last section of this chapter a model application for a local context will be discussed. The model assumptions then need to be verified and refined for that particular study area.
[3] For certain study areas this may need to be generalized to include speculators, loggers, … For now, we will concentrate on the farm household as the sole primary actor.

3. The tertiary actors include international agents involved in *policymaking* (e.g. World Bank, international environmental movement), representatives of the *international donor community* (e.g. World Bank), and *trade institutions* (e.g. WTO, ITTO). Typically, international institutions are involved in more than one of these fields.

The assumptions regarding the primary actors: farm households

The following assumptions define the livelihood strategies of farm households in tropical forest areas. They are not meant to describe households in their full diversity and complexity, but they describe the major components of their land use decisions.

4. The objectives of the farm household include at least *(i) food consumption* and *(ii) asset accumulation*.
5. It is assumed that the household prefers family food consumption up to a 'basic needs level' to accumulation of other assets.
6. It is assumed that farm household decisions can be represented by only one decision-maker. Though this assumption may be unrealistic for some households, it is often a necessary approximation of household decision-making if interaction between several decision-makers within a household is complex and impossible to identify properly.
7. The household makes both *operational* decisions (with a short-term effect) and *strategic* decisions (with a long-term effect). For strategic decisions several years may elapse between the decision and the moment that the effect on the household objectives is fully accomplished. Examples are the growing of perennial crops and the clearing of forest land. Expectations on future events (e.g. price development, land tenure perspectives) play an important role in these strategic decisions. Operational decisions are defined as decisions with effects exclusively or predominantly within one year, whereas strategic decisions have effects over the years. We further assume that the household has a planning horizon of 10 years and that strategic decisions take effect within this planning period.
8. Food consumption is satisfied partly by home produce and partly by purchases on the market. The degree to which the farm household is involved in market transactions depends on access to markets as well as on various price and cost factors. The farmer takes these factors into account while making decisions on the production and marketing of food products.
9. Income generating activities are categorized as *farming*, *animal husbandry, extraction of forest products* and *non-farming income generation*.
10. Each of these income-generating activities is subdivided in more specific activities; each has different effects on the household objectives and on the environment (in casu state of the tropical forests). For farming, we need at least a distinction between *food crops* and *cash crops*, as well as *annual* and *perennial crops.* For animal husbandry grazing densities are important, as well as its effect on asset accumulation. Extraction of forest products can be done on-farm (if there is on-farm forest) and off-farm. We assume that the extracted forest products (e.g. wood) are sold in the market. The effect of the extractive activity on the state of the forest, that is degradation from primary to secondary (logged-over) forest, is taken into account. Non-farming activities include at least off-farm wage labour. Non-farming income-generating activities are taken into account for their effect on the household budget and on available family labour.
11. We assume that a well-functioning market exists for all outputs, inputs and food products. Transaction costs for outputs are taken into account, they consist at least of transportation costs between the farm and the nearest market.

The assumptions regarding the secondary actors: lead agents and government

12. The lead agent (see Rudel, 1993) is the actor that enters the forest for commercial motives before households begin to settle. The lead agent builds roads and hires labour. In this way, the lead agents facilitate the settlement procedure. The lead agent is normally active in extraction of natural resources, such as wood or mining products. The objective is profit maximisation. Whether the company has a short-run or a long-run perspective can not be stated in general.
13. The government may also be a lead agent by facilitating the settlement of households, but not necessarily so. It typically has multiple and conflicting objectives, such as protection of nature and generation of foreign exchange or tax revenues. We assume that the major objective of the government is to keep its position in leading the country. All governments have a budget that they use to effectuate policies in the forested areas. Secondly, all governments apply laws, rules and regulations to effectuate policies. The government is regarded as one single decision-maker, who co-ordinates the conflicting objectives of its departments.
14. Other secondary actors may be added depending on their role in a particular study area.

The assumptions regarding the tertiary actors: international agents

The identification of tertiary actors is particularly complex because many international organizations claim to have similar objectives. The World Bank, for example, is primarily an institution that promotes economic development. More and more this objective has bended to make place for a concern with nature and environmental issues. The key objective for the World Bank (as well as for many other international institutions) is now "sustainable development". Actually, this term is just a way of stating that multiple objectives are pursued, where the objectives may even be in conflict with each other. The various organizations each have a particular focus, but it is common that their policies are the result of a complex set of conflicting objectives. For this research we consider a single hypothetical international agent.

15. The tertiary actor is a hypothetical agent, called "International Organization for Sustainable Development", abbreviated as IOSD. Its main objective is "sustainable development".
16. It is assumed that the IOSD has a considerable amount of money available originating from contributions of wealthy nations.
17. The decisions of the IOSD are restricted by external rules such as limits imposed by sovereignty regulations, trade agreements etcetera. These restrictions have to be specified depending on the relevant factors in the particular study area.
18. Other tertiary actors may be added depending on their role for a particular study area. In particular, multinational organizations with an outspoken commercial objective may be influential for certain tropical forest areas.

Ad b) Identify the available data and collect additional data

Both qualitative and quantitative data are required to fill in the details concerning the actors' behaviour. The hypothesized behaviour described by the assumptions above defines only the basic rules along which the actors take decisions. The model and data analysis should point out to what extent the environmental effects are explained by changes in the motivational factors. The search for data then starts with a choice of a study area. Ideally, the study area is chosen such that (a) the dynamics in the study area can be seen as representative for processes in other forest areas, and (b) a large variety of data is available on the natural resource base, the institutional environment, the population and actors' behaviour. In other words, the area must be "interesting" in the sense that conclusions can be projected and used for policy analysis beyond the boundaries of the study area. The availability of data is often a limiting factor. Deforestation is a process which takes place over a long period of multiple decades. Historic information from the early settlements as well as from more recent times is valuable to help understand the changes that take place, and at what speed they are taking place. First an inventory is made of the available secondary data. Also, additional data collection may be required to fill in gaps in the secondary data. Maps of the area may provide valuable information. Data from different sources must be combined and related to each other. Data bases, statistical techniques and GIS tools are required for this procedure. Qualitative studies and field visits are required to understand the actors' behaviour with regard to the forest. For the primary actors the use of forest land is an integrated part of the livelihood strategies, therefore the whole livelihood strategy is subject of study. The purpose of the data collection and analysis is to get a more comprehensive and detailed knowledge on the actors' strategies. This knowledge is the basis for the development of the actor submodel.

The process of data collection and analysis is illustrated in Chapter 4 for a study area in the Ecuadorian Amazon. For now we will proceed with the discussion of developing the "core model structure".

Ad c) Choose the unit of analysis and link this to a spatial unit

Units of analysis

We have chosen the farm household as the unit of analysis for the decision-making of the primary actors. An important aspect is the question "*Who* actually takes the decisions?" Can the household be considered as a homogeneous unit of decision-making? For many decisions the household level can be regarded as appropriate in this respect, for others a higher or lower level is more accurate in practice. A higher level of decision-making might be a village or a group of similar households. Within the household, individuals may take decisions independent of the others. Women are mostly responsible for family health and nutrition, whereas men usually take the primary responsibility for production related decisions. As long as the goals and strategies of different household members are not conflicting, there is no harm done to reality if we treat the household as a collective unit of decision-making. The household is the unit which is distinguished by the "common roof, common pot".

The secondary actor (the government) is considered as one decision-maker; hence one single unit of analysis. However, in local situations we may have to distinguish between local governments and national governments. Local governments may function as autonomous institutions, independent of the national government. Whether or not it is justified to regard them as one decision-maker is a question which needs to be addressed in a study of the local situation. It needs to be identified who takes the "government-decisions"; and – if there are multiple decision-makers – to what extent their motivations and instruments coincide or differ; and to what extent they have a different influence on the options and motivations of the farm households with respect to land use.

The tertiary actor, the hypothetical "International Organization for Sustainable Development" (IOSD) is also considered as one single decision-making unit, and, accordingly, as a single unit of analysis.

Geo-referencing of the unit of analysis

The geo-referencing of the unit of analysis is necessary in order to perform spatial analyses. If the location and size of the household plot is known, and can be identified on a map, the unit of analysis for the primary actor can be linked to a spatial unit.

The unit of analysis for the secondary actor coincides either with the whole study area, or a geographical subdivision depending on whether the influence of the decision-maker(s) can be considered as homogeneous over the whole study area. The tertiary actor has an aggregation level higher than the whole study area, hence the spatial unit can be chosen as the assembly of the tropical forest areas all over the world.

Data with a spatial component are stored and analysed in a GIS. We have chosen to use SPANS-(GIS)-software for most of the spatial data. The SPANS-software is a so-called "quadtree-GIS"; a raster-oriented storage format which can easily be disaggregated or aggregated if a higher or lower level of detail is needed. This may be useful if data need to be exchanged with other models. The choice of a raster storage format (instead of a vector format) for the basic spatial unit has the advantage that each unit (or grid cell) has a standard size and shape. Data from different data-themes can easily be combined and compared. For example, data on land cover and on household characteristics can be stored with the spatial unit to which they belong. However, for some types of data analysis the use of a raster format can be a disadvantage if the data take different shapes. For example, one of the factors influencing household decisions is the distance between the farm and the nearest market. For an accurate calculation of the distance based on road network data it is better to use a vector storage format instead of a raster format. In this particular case, we have used vector-oriented GIS-software (ArcView) for the analysis of the road distances. The results of the analysis were converted to the basic spatial unit for later use in the mathematical model[4].

Besides the storage format, we also have to make a choice for the *size* of the spatial unit of analysis. The actor-oriented approach requires that a spatial unit can be distinguished for each actor/decision-maker in the mathematical model. Therefore, the spatial unit of analysis needs to be as small as the farm size. If the study area contains many farms of different size, then the smallest spatial unit can be chosen as the average of the lowest percentiles of the farm household population. The larger farms can then be represented by 2, 3 or more spatial units depending on their size. The size of the spatial unit, together with the size of the study area, determines the number of separate spatial units that are used in the analysis. If a large area needs to be covered, it may be a better choice to aggregate the smallest spatial units in order to reduce the complexity of the calculations. Of course, this means that some level of detail is lost in the analysis of the smallest farm households. Last but not least, the size of the spatial unit depends on the availability of data. If data are available on farm level, then the farm size is an appropriate choice. If household data and physical data are available only as aggregated data (for example on district level), then the size of the basic spatial unit will necessarily be larger than the farm size.

[4] A full description of the data, as well as of the data analyses is presented in chapter 4 of this report.

The role of the spatial analysis in the modelling procedure is twofold: first the spatial (and non-spatial) data are analysed to generate input for the mathematical model. Secondly, the output generated by the model is used for geographical presentations by making use of the spatial units and their locations in the study area. In this way, we achieved an integration of spatial and non-spatial methods of analysis.

Ad d) Design model equations, estimate model parameters

The knowledge obtained from the data and literature is used for the mathematical formulation of the actor submodels. Each actor submodel describes the decision-making process of an actor. First we will make some general observations on how decisions are being made. A decision maker can be seen as a person who wants to maximize "utility" while being faced with a set of instruments which he can use to obtain his goals. "Utility" is an abstract term which denotes the whole set of material and immaterial values that the decision-maker judges as important for himself and the "unit" that he represents. The most obvious material value is wealth (accumulation of assets). Examples of immaterial values are good health, access to schooling for children etc. The decision-maker weighs all these objectives to the best of his knowledge before he takes actions. He has to take into account all the instruments that he has available. For example, a farmer who has to take planting decisions for the next growing season will take into account the "values" of earning money, as well as the discomfort of spending labour time and the limited land that he has available.

The family of mathematical programming models is particularly useful for describing these kind of actor decisions. For each decision-maker they consist of an objective function and a set of constraints. The objective function is a combination of the goals that the actor wants to maximize, for example the accumulation of assets. The constraints together describe the set of options and limiting factors, formulated as conditions that need to be satisfied. An example of such a condition for a farmer is that the planted area can not be larger than the total amount of land that the farmer has. This type of conditions (constraints) can be formulated in mathematical terms as equalities and as "less-than" or "greater than" inequalities. Together, the objective function and the constraints describe the decision-making process of the actor.

Within the family of mathematical programming models we consider the *linear programming (LP) model* suitable to describe the decision-making of our actors. In the linear programming model[5], the objective function is a linear combination of *decision variables*[6], which can be seen as the instruments that the decision-maker has available.

Steps e) to f) from the estimation of model parameters to the evaluation of the model results for the baseline scenario will be discussed in chapter 4 and chapter 5. Step h), the design and implementation of model experiments for alternative scenarios is mainly left for future research because of the limited scope of this report.

---

[5] Mathematical programming models can be subdivided into *linear* and *non-linear* programming models. In linear programming models both the objective function and the constraints are linear functions of the decision variables. Non-linear models have non-linear function in either the objective and/or the constraints. Examples of non-linear functions are quadratic functions or exponential functions. In general we can say that non-linear models are more complex and less easy to solve than linear models. For our purposes the linear programming model serves well to describe the decision processes.

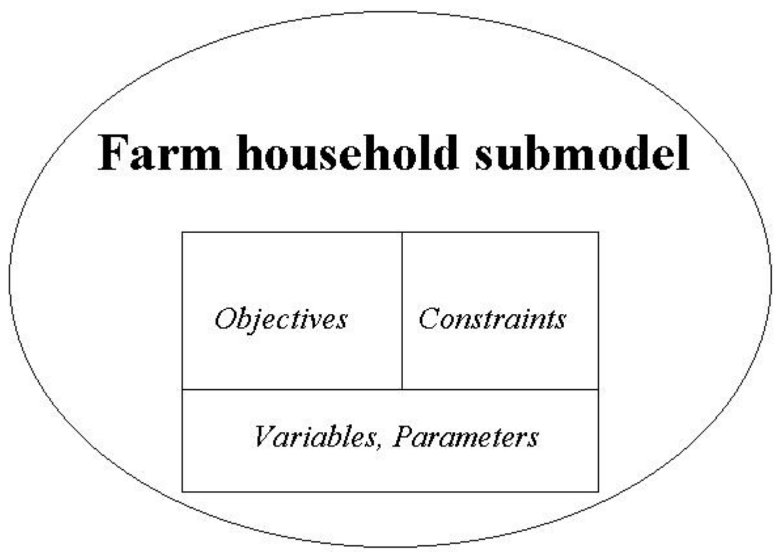
[6] The decision variables, objective functions and constraints are further explained and illustrated in subsection 3.2.1. where the farm household submodel is described.

3.2.1. The farm household submodel

We start out with the submodel for the primary actor: the farm household. The major questions which the model addresses are:

1. For each spatial unit, what is the farmer's expansion strategy towards increasing crop land and pasture at the cost of (primary) forest?
2. On the basis of the major motivational factors, can we identify spatial units currently under forest that are likely to be colonized by farmers in the near future?

Figure 3- 2 Components of the farm household submodel



The decision model is built up from mathematical equations: an *objective function*, and *constraints* (see Figure 3-2). These equations are functions of variables and parameters. The mathematical relations between these elements will be explained later. We will now describe how the objectives, constraints, variables and parameters relate to the activities of the farm household, based on the assumptions described earlier in this section. First we will explain the variables and parameters, then we will proceed to the objectives and the constraints.

Variables and parameters in the household submodel

The *decision variables* in the household submodel represent the choices of the decision maker. The farmer needs to make decisions on the *activities* that he can choose from. In general we can say that the decision describes the choice of the allocation of *inputs* among a number of alternative activities. The values of the decision variables are unknown on beforehand, and need to be calculated by the model.

Activities, decision variables, and inputs

We have defined the categories of farm household activities in Table 3-1. The Table serves as a guideline which represents the most common farming activities; the list of activities may need to be adapted to describe a local context.

Table 3-1. *Activities in the farm household submodel*

ON-FARM LAND USE ACTIVITIES, E.G:
- food production, annual crops
- food production, perennial crops
- cash crop production, annual
- cash crop production, perennial
- land conversion (e.g. forest to arable land)
- animal husbandry
- on-farm logging
- expansion through buying or occupying land
- disposal of land through selling or abandoning

OFF-FARM INCOME GENERATING ACTIVITIES, E.G.:
- off-farm logging
- off-farm labour

NON-INCOME GENERATING ACTIVITIES, E.G.:
- consumption of farm products
- food expenses
- non-food expenses
- saving/investment

The list of activities in Table 3-1 contains only a selection of all activities going on in a farm household. We have chosen this particular set of activities because we consider them relevant for the land use decisions of the household. In the first part of Table 3-1 a distinction is made between food-producing land use activities and income generating land use activities. Food or subsistence crops are those crops that are grown and consumed on-farm. Cash crops are crops that are produced for sale on the market. Even though it may concern edible products (e.g. rice), they are produced primarily for the market. Land conversion includes change from one land use to another, as well as degradation of forest. It was assumed earlier that logging involves degradation from primary forest to secondary forest.

Non-income-generating activities are important too, because they compete with the other activities for the use of inputs. The list of activities is not meant to be exact because local situations may require a somewhat different categorization. In the remainder we will refer to activities with the index *i*. Occasionally we make use of subsets which are denoted as {land use activities}, {off-farm income generating activities}, and {non-income generating activities} according to the categories in Table 3-1. The whole set of activities is denoted by {all activities} (see also Appendix 3-I).

Inputs include the broad categories of *land*, *labour*, and *capital*. These can also be subdivided into smaller categories relevant to the farm household (see Table 3-2). Again, the level of detail and the number of classes for land, labour and capital can be adapted according to the requirements of a local context.

Table 3-2. *Inputs and decision variables in the farm household submodel*

INPUT	*DECISION VARIABLE*
Land	
- Forest land	- Acreage of primary forest used for extraction of wood (and/or other forest products) - Acreage of primary forest converted to secondary forest - Acreage converted from forest to arable land
- Arable land	- Acreage of arable land allocated to annual food crops - Acreage of arable land allocated to perennial food crops - Acreage of arable land allocated to annual cash crops - Acreage of arable land allocated to perennial cash crops - Acreage of arable land for annual crops converted to perennial crops (or vice versa) - Acreage of arable land converted to pasture land
- Pasture	- Size of pasture land - Number of animals on pasture land
- All types of land	- Expand through buying or occupying land - Dispose of land through selling or abandoning
Labour	
- Family labour	- Male labour allocated to activity *i* - Female labour allocated to activity *i* - Other family labour allocated to activity *i*
- Hired labour	- Labour hired for activity *i*

Capital	
- Household capital	- Amount of household capital spent on inputs for activity *i*
- Credits	- Amount of money borrowed for activity *i*

The *decision variables* define the allocation of inputs (land, labour and capital) for the activities summarized in Table 3-1. Examples of decision variables are "the quantity (hours) of family labour used for on farm logging", "the quantity (amount) of household capital spent on non-food consumption", "the quantity (acreage) of arable land allocated to perennial cash crops" or "the quantity (acreage) of forest land converted to arable land". The choices made by the decision maker with respect to the decision variables determine whether activity *i* is carried out and how much of activity *i* is actually being done.
The choices with regard to the decision variables are not unlimited, the availability of inputs is subject to restrictions depending on the local situation. Also, a detailed description of the relationships between the use of inputs and outputs/activities is a necessary part of the model description. These relations will be discussed later in this section. For now we will first proceed to a discussion of the parameters of the farm household model.

The parameters

The parameters represent the exogenous factors that influence the decisions of the farm household. Exogenous factors are the circumstances that are not under the direct influence of the farm household, hence the household considers them as "given" in the context of the decision problem that is described by the model. Examples of exogenous factors are the available size of the land and the composition of the household. Pichón (1993) divided the factors influencing farm household decisions into three groups: household and farm characteristics, institutional environment and technology, and natural resources (see Table 3-3). The same subdivision is useful in our analysis, hence it will be used for categorising the input parameters for the model.

The parameters enter the model through the constraints as we will see below. Appropriate numerical values need to be chosen to represent each of the factors described in Table 3-3. Then, the relationships between the household activities, the decision variables and the parameters need to be formally described and estimated on the basis of available data. This is a crucial step in the whole modelling procedure which has a strong impact on the accuracy of the model. It will be discussed at length later in the description of the model application for Ecuador, but here we will make some general remarks on the role of each of the parameters mentioned above. The basic question for the modeller is: To what extent does each parameter influence activity *i*?

Table 3-3. *Factors influencing the farm household activities*

Household and farm characteristics, for example:

a. Settlers' farm background, ethnicity
b. Household demographic composition
c. Farm size
d. Land tenure

Institutional environment and technology, for example:

e. Road access, distances
f. Availability of technology and agricultural assistance
g. Access to (labour) markets, prices
h. Initial farm size allocated/occupied
i. Land tenure policies
j. Collective action institutions (e.g. cooperatives)

Natural resource base, for example:

k. Soil quality
l. Relief

Source: Adapted from Pichón (1993), page 94

a) settlers background/ethnicity

To what extent do the settler's background and ethnicity influence various land use decisions? In the case of Ecuador it has often been claimed that farmers apply farming practices that they were familiar with in their region of origin. However, the differences between farming practices are small among farmers, even between settlers from outside the Amazon and indigenous farmers. Both qualitative and statistical analysis may point out whether this is a factor of importance that needs to be included or not.

b) Household demographic composition

The size and composition of the household is an important factor in determining land use decisions. It determines the availability of *labour*. Labour is a decision variable that is used for almost all activities mentioned in Table 3-1. The size of the household also determines the food requirements of the household and thus indirectly productive activities.

c) Farm size

The farm size determines the availability of land for production. In the long run the farmer may be able to influence the farm size. In that case the factor *farm size* has to enter the model as a decision variable. In situations where the farmer has no (direct) influence on the farm size it has to be considered as a parameter.

d) Land tenure

The tenure status is an important factor in the degree to which the household is inclined to invest in the land. Titled land can also serve as a collateral for obtaining credits.

The *institutional parameters* are typically those where the government (or other secondary actors) have an important influence. They may enter the government submodel as decision variables but for the farm household they are exogenous factors, and therefore enter the model as parameters.

e) Road access, distances

Road access is an important component in the costs of marketing farm products and access to markets of consumer goods. Households that live far from roads and markets will be more inclined to subsistence agriculture because their net revenues from sales are much lower than those of households living near the market.[7]

f) Availability of technology and agricultural assistance

Farming practices can be influenced by the availability of technology and agricultural assistance. Many governments have subsidized programs for improved seeds, agricultural research and extension. If farmers have access to these facilities they may (gradually) change their farming practices. This parameter is likely to counteract the influence of the settler's background and ethnicity (see under a).

g) Access to (labour) markets, prices

Determines whether labour can be hired and at what costs, as well as the possibilities for household members to be employed in off-farm labour. Also determines whether farming inputs can be purchased, whether farm products can be sold, and consumer goods can be purchased. We have assumed earlier that the household has access to these markets, hence the price is an important determinant for the quantities bought and sold. The market prices, together with the transaction costs (e.g. transportation to/from the market) determine the actual costs/revenues for the farmer.

h) Initial farm size allocated/occupied

Many tropical forest areas have been colonized along well-defined patterns. In the case of Ecuador, for example, the land was divided in plots of approximately equal size where farmers could subscribe for the acquisition of a piece of land. The actual farm size (see also c) depends to a large extent on this initial allocation.

i) Land tenure policies

This parameter is an instrument for the government. For the household it is a parameter that is taken as an exogenous factor. Whether or not land rights can be obtained is important for the household in taking long-term (strategic) decisions.

j) Collective action institutions (e.g. cooperatives)

Collective action institutions are organizations that represent the interests of (groups of) farm households. By being member of a "pressure group", for instance, farmers as a group can exercise a certain influence on government decision making which they would not have as individuals.

k) Soil quality

Soil quality is an exogenous parameter in the short run but may be influenced by the farmer in the long run. The quality is an important to determine to what extent the land is suitable for different activities.

[7] In cases that roads are considered to have a significant *direct* effect as well (hectares of forest felled, fragmentation of the forest etc.), the roads decision-maker is both a primary and a secondary actor.

l) Relief

Hilly land is less suitable for growing crops because it erodes fast, and cultivation requires much more labour than flat land. Steep hills that are currently under forest are therefore more likely to remain forest than flat lands. Not only the total farm size, but also the topography therefore needs to be taken into account when determining the actual production possibilities.

State variables

From the parameters we will now proceed to another category of variables: the *state variables*. State variables depend on the values of the decision variables and the parameters only. Hence, the values of the state variables can be interpreted as results of a decision by the farmer.

Examples of state variables are:

- Total produced output of good *i*, year *t*
- Total gross/net revenues from sales of output, year *t*
- Total purchases of food products, year *t*
- Land value at the end of year *t*
- Amount of money in cash at the end of year *t*
- Total non-farm income, year *t*
- Total expenses on farming inputs, year *t*
- Total expenses on inputs for animal husbandry, year *t*
- Total food expenses, year *t*
- Total amount of investments for future production, year *t*

The relationship between these state variables and the parameters and decision variables is specified by the mathematical functions in the model. For example, the state variable *total food production* is a function of the decision variables *land allocated to food production, labour allocated to food production, capital spent on inputs for food production*, and of the parameters *household demographic composition, farm size, availability of technology* and *agricultural assistance, access to markets and prices of inputs, food and cash products, soil quality,* and *topography.*

The objective function in the household submodel

From farm household theories (see Ellis, 1988; Singh et al., 1986) we know that two types of farm household objectives are commonly distinguished in relation to land use:

(a) For *subsistence households*, the first goal is to have sufficient food (from home produce or bought on the market) to feed the family.

(b) For *market-oriented households*, the family consumption is not a reason for concern; their goal is to earn as much money as possible.

We have chosen to consider both household goals because many households in forest settlements are in fact subsistence households; though they may as well sell surpluses of cash crops on the market. The model can be easily adapted for study areas where households are predominantly market-oriented.[8]

In order to construct a (single-objective) linear programming model it is necessary to define the priority order in the objectives. For poor households it is common that the food consumption objective is the most important. Rich households with ample cash income will not be too concerned with the food consumption objective, and will be more inclined to buy luxury goods. We will assume that farm households in the forest areas are poor and that food consumption indeed has the highest priority, and that *sufficient food for the family* is a requirement that needs to be fulfilled even if this conflicts with the asset accumulation objective. In mathematical terms, the food consumption requirement will be formulated as a constraint rather than as an objective function as we will see later in this section. The constraint is then formulated as: the food consumption needs to be at least the amount required for a healthy and productive life[9].

Because the food consumption objective enters the model as a constraint, the short-term objective function is a function of the second household goal only. The objective function states that the accumulation of household assets should be maximized. Assets can consist of money (savings), consumer goods (both durable and non-durable), or investment goods (e.g. land, cattle). In the short-run, the households (*h)* needs to generate income in order to obtain these assets. Therefore, (introducing the symbol OBJ) the *short term objective* is written as in equation (O-1):

(O-1) $\mathrm{OBJ}(h,t{=}year1) = \mathrm{NETREV}(h,t{=}year1)$
For all $h \in H$.

Where:

OBJ(*h,t=year1)*	the short-run household objective
NETREV*(h,t)*	net revenues from productive activities of household *h*, year *t*

In the long run, another objective has to be taken into consideration. Households may attach some

[8] If food production is dropped from the list of activities then the model reduces to a model for the market-oriented farm household. Consequently, all goods produced are sold and all food products consumed are obtained from the market.

[9] The household goal of sufficient food consumption is here formulated as a "hard constraint" meaning that no deviation is allowed. This is realistic only in situations where food shortages do not occur. In other situations this may be formulated as a "soft constraint" allowing for food shortages, where a penalty function is introduced. The penalty function then reflects the total food shortage. The first objective is then to minimize the penalty function. The second objective is asset accumulation as described above. The model can then be solved by using multi-objective (goal) programming techniques.

value to a good governance of "nature", that is, to keep a certain area under forest. Keeping land under forest is on the one hand in competition with cultivation of the land, on the other hand the forest provides the household with typical forest products such as firewood, construction material, and fodder. Trees also help prevent erosion and may help to keep up the quality of the land by practising agroforestry techniques. Hence, in a long-term strategy forest maintenance and/or agroforestry needs to be included in the objective.

Investments take a particular place in the long-run objective of the farm households. Investments are not an objective as such, but a means to increase future utility. Farm households may invest in material as well as immaterial goods. For example, education can be seen as an investment in human capital, an immaterial good. Farmers who wish to invest in future production possibilities can do so by investments in capital goods, but they might also invest in the quality or quantity of land available to them. Whether or not they will invest in land depends on many factors, notably on the actual scarcity of land and on institutional factors. If land is scarce it will be highly valued by its users which implies that it is a desirable investment. Institutional factors determine whether land can easily be bought and sold, whether property or usefructuary rights can be obtained and traded, etcetera. In the core model structure we will take into account that farmers invest in land in order to increase future assets. The extent to which this applies in a local context will of course depend on the factors mentioned above.

The long-run objective can be expressed as in equation (O-2):

(O-2) $\text{OBJ}(h,t) = \text{NETREV}(h,t) + \text{LANDVAL}(h,t) - \text{LANDVAL}(h,t\text{-}1)$
for all $h \in H,\ t \in \{year2, \ldots year10\}$
Where:
LANDVAL(h,t) the (market) value of farm land *h.*

In the long run the household wants to maximize both the annual net revenues and the annual increase in land value. The short-run and long-run objective functions can be combined into one objective function for the whole planning period. By using weight factors all objectives can be added up. The weight factors reflect the time preference of the household. The values of the weight factors are highest for the first year and lower for each of the successive years expressing that the household values later years as less important than early years. Equation (O-1) defines the household objective function.

$$
\begin{aligned}
\text{(O-3)}\quad \text{MAXIMIZE}\ & weight(h,t{=}year\ 1)*\text{OBJ}(h,t{=}year1) \\
& + weight(h,t{=}year2)*\text{OBJ}(h,t{=}year1) \\
& + \ldots + weight(h,t{=}year10)*\text{OBJ}(h,t{=}year10)
\end{aligned}
$$

Where:
weight(h,t) time preference rate for household *h*, year 1, .. 10.

The constraints in the household submodel

The constraints describe the relations between state variables, decision variables and constraints. In the short run (decisions for year 1) we can distinguish the following categories of constraints:

- Production functions
- Consumption function
- Financial constraints
- Labour constraint
- Land constraint
- Non-negativities

In the following, these functions are described by mathematical functions. The meaning of the symbols are summarized in appendix 3-I at the end of this chapter. Most variables, parameters and equations make use of indices because they need to distinguish between the actors (*h*), the farming activity (*i*), the land quality (*q*) or the "decision year" (*t*). The unit of analysis is the household, where each household *h* occupies one spatial unit.

PRODUCTION FUNCTIONS:

(P-1) $\quad \text{OUTPUT}(h,i,q,t) = \text{A}(h,i,q,t) * yield(h,i,q,t)$

For all $h \in H$, $i \in \{\text{land use activities}\}, q \in Q$, $t \in T$ [10].

Where:

$\text{OUTPUT}(h,i,q,t)$ = output of good *i* on land of quality *q* for household *h*, year *t*

$\text{A}(h,i,q,t)$ = land of quality *q* allocated to activity *i* for household *h*, year *t*.

$yield(h,i,q,t)$ = yield/hectare of product *i* on land of quality *q* for household *h,* year *t*.

Equation (P-1) defines the relation between produced output and the input land (A). Each land use activity has a certain yield per hectare which needs to be estimated. Note that land conversion is also defined as a land use activity. Land that is being converted (for example from annual to perennial crops, from arable land to pasture) usually have lower yields initially, whereas labour requirements are often higher. A typical case of land conversion is land degradation. Land degradation is a consequence of wood extraction (logging). Yields are high in the present year at the cost of future output. The relation between the current and future production possibilities is expressed in the land constraints (see below).

The use of labour and capital in the production process is defined below in equations (P-2) and (P-3). In order to avoid non-linearities in the production functions we have chosen for a formulation where land, labour and capital are used in fixed proportions, although these proportions are allowed to differ

[10] The parameter *yield(h,i,q)* is introduced here as a deterministic parameter. However, yields tend to be highly variable and uncertain in real situations. We will assume that the farm household takes into account the expected values of the yield (estimated as the average from historical data), such that the parameter can be treated as deterministic instead of stochastic. If farmers are not particularly risk-averse the actual strategy can be approximated with deterministic parameters only. In case of strong risk-averse attitudes of farmers this formulation may need to be changed in order to reflect the variability of yields and the risk attitude of the farmer. This would require estimating distribution functions and measuring farmers' risk attitudes.

according to the input requirements for different households, products, land qualities and time (allowing for diversity in household, product, land quality, and technological progress characteristics).[11]

(P-2) $\text{LFAM}(h,i,q,t) + \text{LHIRE}(h,i,q,t) \geq labreq(h,i,q,t) * \text{A}(h,i,q,t)$

For all $h \in H$, $i \in \{\text{land use activities}\}, q \in Q$, $t \in T$.

Where:

LFAM*(h,i,q,t)* family labour in days[12] allocated to activity *i* on land of quality *q,* year *t*

LHIRE*(h,i,q,t)* days of labour hired by household *h* for activity *i* on land of quality *q*, year *t*

labreq(h,i,q,t) labour requirements per ha in working days for activity *i* on land of quality *q*, year *t*

Equation (P-2) formulates the labour requirements per hectare cultivated land.

(P-3) $\text{K}(h,i,q,t) = capreq(h,i,q,t) * \text{A}(h,i,q,t)$

For all $h \in H$, $i \in \{\text{land use activities}\}, q \in Q$, $t \in T$.

Where:

K*(h,i,q,t)* amount of household capital assigned to activity *i* on land of quality *q*, year *t*

capreq(h,i,q,t) capital requirements per ha for activity *i* on land of quality *q*, year *t*

Equation (P-3) formulates the capital requirements per hectare cultivated land.

CONSUMPTION FUNCTION:

(C-1) $\sum_{i \in \{\text{food products}\},q} \text{OUTPUT}(h,i,q,t) + \ni_{i \in \{\text{food products}\}} \text{PURCH}(h,i,t) \geq consreq(h,t)$

For all $h \in H$, $t \in T$.

Where:

PURCH*(h,i,t)* purchases of product *i* for household *h* , year *t*

consreq(h,t) food consumption requirement, depending on nutritional values of food products and household composition, year *t*.

The constraint (C-1) will make sure that enough food is available to the household to meet consumptive needs. The food may either be purchased or produced on the farm. In equation (C-1) all food products are added up without distinguishing between differences in nutritional values. Whether or not this approximates the actual food consumption requirements needs to be verified for each study area. If necessary, the model equation can be adapted to include nutritional values by multiplying the decision variables with appropriate weight factors. The consumption requirements can be estimated from international standards such as from the WHO.

[11] Labour and land are used in fixed proportions though the proportion may differ among households, land of different qualities, products and time. The labour requirements per hectare (parameter *labreq*) can thus be estimated from historical data. Note that the relation between more intensive labour use and output is not specified. If sufficient information would be available in a local context, then the model needs some adaptations to take this into account.

[12] Note that family labour is not subdivided into male, female and other labour. This is done to avoid unnecessary complexity, and is justified if all family members can be involved in all productive tasks. Differences in productivity or engagement in reproductive/household tasks can be accounted for, for example by expressing all labour in "male labour equivalents".

FINANCIAL CONSTRAINTS:

(F-1) NETREV*(h,t)* = REV*(h,t)* + NFINC*(h,t)* - PRODCOST*(h,t)* - FOODCOST*(h,t)* - INV(*h,t*)
for all h∈H, t∈T.

Where:

REV*(h,t)*	total gross revenues from sales of farm output of household *h*, year *t*
NFINC*(h,t)*	non-farm income of household *h*, year *t*
PRODCOST*(h,t)*	total costs of production and harvesting for household *h*, year *t*
FOODCOST*(h,t)*	total costs of buying food for consumption of household *h*, year *t*
INV(*h,t*)	costs of expansion of farmland

Equation (F-1) defines the net revenues as the sum of all family income minus the costs of production, the money spent on food, and investment in farm land.

(F-2) REV*(h,t)* = $\ni_{i\in \{\text{cash crops, forest products, animal products}\},q}$ OUTPUT*(h,i,q,t)* * *price(h,i,t)*
for all h∈H, t∈T.

Where:

price(h,i,t)	the farmgate price that household *h* receives for selling 1 kg of farm product *i* (or pays for buying 1 kg of food product), year *t*.[13]

Equation (F-2) states that the revenues are equal to the value of the sales of the cash crops, forest products, and animal husbandry products.

(F-3) NFINC*(h,t)* = *wage(t)** LFAM*(h,i=off-farm wage labour,t)*
+ *logrev(h,t)** LFAM*(h,i=off-farm logging,t)*
for all h∈H, t∈T.

Where:

wage(t) daily wage, year *t*
logrev(h,t) revenues from off-farm logging for household *h*, per working day, year *t*

Non-farm income is defined in equation (F-3) as the sum of revenues from all off-farm income generating activities. These activities include off-farm logging and off-farm wage labour.

(F-4) PRODCOST*(h,t)* = $\sum_{i,q}$ *(wage(t)**LHIRE *(h,i,q,t)* + K(*h,i,q,t*) *)*
+ *interest(h,t)**CREDIT(*h,t*)
for all h∈H, t∈T.

Where :

interest(h,t) the interest rate for credit

13 It is important to use the farmgate price instead of a market price in this equation, because there may be considerable differences between the two. The calculation of the farm gate price as a function of other parameters (the market price and transportation costs) will be discussed in chapter 4.

CREDIT(h,t)   the credit used to finance expenses in year t

The production costs consist of the costs of hired labour and the costs of the input requirements (K) as defined in equation (P-3), as well as interest on credits.

(F-5) $\text{FOODCOST}(h,t) = \sum_{i\in\{\text{food crops}\}} \text{PURCH}(h,i,t) * price(h,i,t)$
for all $h\in H$, $t\in T$.

Equation (F-5) defines the costs of food purchases.

(F-6) $\text{INV}(h,t) = \sum_i landprice(h,i,q,t) * \text{EXPAND}(h,i,q,t) - \sum_i landprice(h,i,q,t) * \text{DISPOSE}(h,i,q,t)$
for all $h\in H$, $t\in T$.

Where:

EXPAND(h,i,q,t)	expansion of farm land of quality q and land use i, year t
DISPOSE(h,i,q,t)	disposed farm land of quality q and land use i, year t
landprice(h,i,q,t)	the price the farm household pays/receives for buying/selling farmland depending on location, land quality, current land use.

Though the exact relation can not be specified in general, it is known that the landprice depends in particular on the following two parameters:

trancost(h,t)	transportation costs between farm h and the nearest market, depending on distance, means of transport and quality of infrastructure
proright(h,t)	an indicator for the property or usufructuary rights of the household h.

Equation (F-6) specifies the costs of investment in land as well as the revenues when disposing of farm land. Note that (F-6) can only be used if farm land can be bought and sold freely. The price of the land needs to be estimated on the basis of available data. We may expect that the landprice depends not only on the land quality (q) and land use (i), but also on the transportation costs to the nearest market *trancost* and the (eventual) property or usufructuary rights *proright*. Depending on data availability, the relation between landprice and these explanatory factors can be estimated by (linear) regression. However, data availability will often be a limiting factor. The price of land can be estimated if data exist on traded farm lands. This was the case for our study area in Ecuador, but land markets do not always exist in forest areas elsewhere. Forest is often state property, and in that case a legal market value does not exist. However, even if land can not legally be traded the hypothetical price of land may still be a powerful subject for analysis, for example if data can be obtained on farmers' preferences of one piece of land over another. Otherwise we need a somewhat different specification. The parameter *landprice* can be exchanged for a parameter which reflects the labour and costs required to occupy a piece of forest or abandoned land for example.

(F-7) $\text{TOTCAP}(h,t) = \text{TOTCAP}(h,t-1) + (1-consexp(h,t)) * \text{NETREV}(h,t-1)$
for all $h\in H$, $t\in T$.

Where:
TOTCAP(h,t) Total amount of capital available to household h, year t

consexp(h,t) non-food consumer expenses for household *h*, as a fraction of last year's net revenues.

Equation (F-7) defines the total amount of capital available to the household. The net revenues (NETREV) have been specified in equation (F-1). Part of the revenues is used for expenses on consumer goods, the remainder, (1-*consexp(h,t)*), is added to the capital stock.

(F-8) $\text{PRODCOST}(h,t) + \text{FOODCOST}(h,t) + \text{INV}(h,t) \leq \text{CREDIT}(h,t) + \text{TOTCAP}(h,t)$
for all $h \in H$, $t \in T$.

Total expenses for production, food and investment can not exceed the amount of cash. The cash supply consists of the capital stock augmented with credits.

(F-9) $\text{CREDIT}(h,t) \leq credmax(h,t)$
for all $h \in H$, $t \in T$.

Where:
credmax(h,t) the maximum amount of credit that the household is allowed to borrow, depending on collateral (property rights on farm land)

Equation (F-9) defines the upper bound on credits. Moneylenders will allow larger credits to landowners than to people who have only temporary rights on the farm land they are cultivating.

LABOUR CONSTRAINT:

(L-1) $\sum_{i,q} \text{LFAM}(h,i,q,t) \leq famlab(h,t) * workdays$
for all $h \in H$, $t \in T$.

Where:
famlab(h,t) family labour available, in "male equivalent", year *t*
workdays the number of working days per year per person

Equation (L-1) states that the total amount of family labour used for productive activities can not exceed the total amount of family labour available. Only family labour has an upper bound, hired labour is supposed to be available in unlimited supplies[14].

[14] This is a result of the assumption made earlier that a well-functioning labour market exists.

LAND CONSTRAINTS:

(A-1) EXPAND(h,i,q,t) ≤ $maxexpand(h,i,q,t)$
For all $h \in H$, $i \in$ {land use activities}, $q \in Q$, $t \in T$.

Where:
maxexpand(h,i,q,t) the quantity of land available to the farm household for purchase or to be occupied in year *t*.

Restriction (A-1) defines whether the household has possibilities to expand the farm land into the forest or by acquiring land from others. Restriction (A-1) is optional: without the restriction the decision variable EXPAND expresses the *intention* of the farm household to increase the farm size, possibly in other areas.

(A-2) $\mathrm{A}(h,i,q,t) = \mathrm{A}(h,i,q,t\text{-}1) + \sum_{k:k\text{-}i\ conversion} \mathrm{A}(h,k,q,t\text{-}1) + \mathrm{EXPAND}(h,i,q,t)$
$- \sum_{j:i\text{-}j\ conversion} \mathrm{A}(h,j,q,t\text{-}1) - \mathrm{DISPOSE}(h,i,q,t)$
for all $h \in H$, $i \in$ {land use activities}, $q \in Q$, $t \in T$.

Equation (A-2) defines *land conversion* and *expansion*. For example, the area under secondary forest in year *t* equals last year's area under secondary forest *plus* last year's logged over forest (converted from primary to secondary forest) + secondary forest obtained by expansion *minus* this year's converted secondary forest (to arable land, etc) – secondary forest sold or abandoned.

(A-3) $\mathrm{LANDVAL}(h,t) = \sum_i landprice(h,i,q,t) * \mathrm{A}(h,i,q,t)$
for all $h \in H$, $t \in T$.

Equation (A-3) states that the value of farm land *h* can be described as a function of the current land use and the price of land of different quality and land use. If the available data do not allow estimation or approximation of land values and prices, equation (A-3) and the appearance of the land value in the objective function (O-2) have to be dropped from the model.

NON-NEGATIVITIES:

All variables except INV, NETREV and OBJ must be nonnegative by definition.

The core model structure is defined by all equations together: (O-1) to (O-3), (P-1) to (P-3), (C-1), (F1) to (F-9), (L-1), (A-1) to (A-3) and the non-negativities. The solution gives the optimal values for the decision variables. It is not evident on beforehand that a solution exists. The consumption function, (C-1), is crucial in this respect. Depending on the values of the parameters, it is possible that the household does not have sufficient inputs to buy or produce sufficient food. In that case, the farmer will first sell the land and then migrate to another place.

3.2.2. The secondary actor submodel

As explained before, the government has an important role as a secondary actor because they may encourage or discourage the livelihood strategies of farm households. Other actors may be significant secondary actors as well, e.g. when logging companies construct roads or urban elites put pressure on farmers to leave their land. These actors are not considered here. Therefore the relation between government policies and farm household strategies will be the central issue in the secondary actor submodel. The model development requires a good knowledge of the relevant policies in the local context. In this section we will describe the core model, based on common behaviour of governments[15].

The most important objective for the central government is to remain in power. For democratically chosen governments, the only way to remain in power is to earn votes from the population. This means that the government has an incentive to acquire the confidence of the population. They will design policies that visibly serve the benefits of the population. Whether or not the government is concerned with nature depends on whether votes can be earned by protecting nature. The government policies may also be under influence of external pressures from the international community, because international approval of domestic policies often brings in additional money from donor countries and international organizations. An additional budget can be used to increase the popularity among the domestic population. In this way, international pressure may indirectly cause government interest in nature conservation. On the other hand, money spent on nature competes with other expenditures which are more “visible” to the population. Also, law enforcement to protect nature often restricts the possibilities of the population to generate income, which makes it an unpopular policy in the short run. Because of these considerations, we may expect the government to have an ambivalent attitude towards nature.

The major questions that the government submodel aims to address are:

1. How does the government allocate its budget and how does that affect the driving forces for household land use?
2. Which are the laws, rules and regulations that the government uses to influence land use and to regulate the use of national forests?

These two questions address the major instruments of the government: the budget and the power to make laws and enforce people to comply with the laws. As such, they can influence the driving forces of the farm households. First we will look more closely into this matter. As we have seen in the discussion of the household submodel, a number of external or exogenous factors influence the behaviour of the household. In the model, these exogenous factors were represented by the parameters. The household has no direct influence on these parameters. The government, on the other hand, may design policies that change the external factors for the households. In terms of the decision model for the government, this implies that the household submodel and the government submodel are linked to each other through the parameters in the household submodel. The parameters that can be influenced by the government are enumerated in table 3-4:

[15] We have chosen to emphasize the role of the government as a “central power” such as it often found in the presidency and the ministries of defence, domestic and foreign affairs. This may not always be sufficient, because many countries also have strong “line agencies” (e.g. ministries of forestry, agriculture, and education) which may have a strong drive for other objectives such as protection of nature or rural development. Depending on the power of such line agencies in the country under study, it may be relevant to distinguish the government as two separate secondary actors.

Table 3-4. *Household submodel parameters influenced by the government*

credmax(h,t)	the maximum amount of credit that the household is allowed to borrow, depending on collateral (property rights on farm land)
trancost(h,t)	transportation costs between farm *h* and the nearest market, depending on distance, means of transport and quality of infrastructure
intrate(h,t)	the interest rate for credit
landprice(h,i,q,t)	the price the farm household pays/receives for buying/selling farmland depending on location, land quality, current land use.
logrev(t)	revenues from off-farm logging, per working day, year *t*
maxexpand(h,i,q,t)	the quantity of land available to the farm household for purchase or to be occupied in year *t*.
price(h,i)	the price that household *h* receives for selling 1 kg of farm product *i,* depending on the market price and transportation costs to the market.
proright(h,t)	an indicator for the property or usufructuary rights of the household *h*.

In Table 3-4 we selected parameters from the household model which we consider to be important spheres of influence for the government. Particular domains that the government controls are infrastructure, subsidies for the use of farm inputs, and laws with respect to land use. In the description of the government submodel these parameters will come back in a different form. For the government, these factors are no longer exogenous, but part of the choices among different government activities.

Variables and parameters in the government submodel

Similar to the development of the household submodel, we will start with the definition of *decision variables*, then we proceed to the *parameters* and *state variables* before we describe the *objectives* and *constraints*.

Activities, decision variables, and inputs

In table 3-5 we describe the major categories of activities for the government, including some specific examples and their effects on the parameters for the household submodel. We distinguish the broader categories of *investments*, *subsidies* (including price interventions), *taxes*, and *laws*.

Table 3-5. *Activities in the government submodel*

INVESTMENTS IN INFRASTRUCTURE AND REGIONAL DEVELOPMENT, E.G:
- road construction and maintenance (reduces transportation costs for the household *trancost*)
- regional development (increases land value for the household, *landprice*)

SUBSIDIES, PRICES & CREDITS, E.G:
- issue subsidized credits for farmers (reduces the interest rate, *intrate*, increases credit possibilities, *credmax*)
- subsidize farmers inputs (reduces input costs, *capreq)*

TAXES, E.G:
- raise or reduce taxes on income or on activities (e.g. on logging)

LAWS, RULES REGULATIONS, E.G:
- legalizing land for the settlers (increases land value for the household, *landprice*, extends property rights, *proright*)
- protect nature reserves and national parks (reduces the possibilities for households to expand, *maxexpand*)
- restriction of logging activities (reduces logging revenues *logrev* for the household)

From the definition of activities in Table 3-5 we can move straightforward to the definitions of the *decision variables*. Table 3-5 distinguishes four categories of instruments: *subsidies, investments, taxes,* and *laws*. The *decision variables* define the allocation of inputs among these activities:

SUBS(*g,i,t)*	the government subsidy for activity *i* in year *t*
INVEST(*g,i,t*)	the government's direct investment in activity *i* in year *t*
TAX*(g,i,t*)	tax revenues, raised on activity *i,* year *t*
LAW*(g,i,t*)	laws affecting activity *i*, year *t*

The decision variables SUBS, INVEST, TAX concern the allocation of the government budget, hence they are expressed in terms of money. Making laws is a regular task of the government which does not necessarily involve (high) costs. In the decision model we will focus on laws as political choices and we assume that the costs involved do not affect the decision whether or not to adopt a certain law. The variable LAW can be defined as a zero-one variable (value 1 if the law takes effect, value zero if the law does not take effect). If we wish to express the *degree of* law enforcement then the variable can be defined in the range between zero and one. The value then represents the degree to which the law takes effect. For example, a government may adopt a law to protect a certain area from logging activities. The choice of the government is between:

LAW(*i=logging ban,t)* = 0 (meaning that logging is allowed)

Or:

LAW(*i=logging ban,t)* = 1 (meaning that logging is not allowed)

Or:

LAW(*i=logging ban,t)* = α (meaning that logging is partially tolerated, and α is the part of the land which is protected from logging)

The parameters and state variables

Similar to the household submodel, the parameters for the government submodel are those factors that are exogenous to the decision-making of the government. The parameters describe the initial endowments, such as the initial budget, the "exogenous" government revenues, the prevailing laws, taxes and subsidies.

In the household submodel we introduced parameters in the objective function which reflect the time preference. We will do the same for the government submodel. The time preference rate of the government reflects to what extent the government is willing to compromise short-term objectives for the benefit of future objectives. Much depends on the actual political climate. In the extreme situation of a very unstable political climate, for example with poverty, high inflation and an active guerrilla movement, the government will most likely opt for a short-run "survival" strategy. All inputs will be invested to relieve the most severe and urgent problems. On the other hand, a government that is stable, acknowledged and trusted by the population will be inclined to invest in long-term objectives. Nature conservation is such a long-term objective which generates money and votes in the future. Therefore, the "time preference" parameter is very important as an indicator for the government's attitude towards nature.[16]

Other parameters are the current level of the government budget, ongoing commitments with respect to investments, subsidies, taxes and laws, the current state of land use, and current agreements with the international community (represented in the model by the tertiary actor). With respect to the government budget we assume that the government has a regular flow of income from sources that are outside the context of the model.

The state variables are those variables for which the values are determined by the decision variables and the parameters. For a description of the state variables we refer back to Table 3-4, which contains a number of parameters taken from the household submodel. For the government submodel, these household parameters are in fact *variables* because they are a result of the government policies with respect to subsidies, investments, taxes, and laws. In the equations of the government submodel they will appear in the following form:

CREDMAX(*h,t*)	The maximum amount of credit that the household *h* is allowed to borrow
TRANCOST*(h,t)*	transportation costs between farm *h* and the nearest market, depending on distance, means of transport and quality of infrastructure

Etc.

The objective function in the government submodel

The short-term objective is to earn votes from the population, a long-term objective is also to preserve natural resources. Votes can be earned by implementing policies that increase the popularity of the government. Exploitation of natural resources may benefit the population in the short run, but it is harmful in the long run. Therefore, the government will most likely gain votes in the short run at the cost of votes in the long run. We have defined a *popularity index* to measure the effect of the policies. We will assume that the popularity index is a function of the policy measures taken by the government (the decision variables) in the current year and in all previous years (start year τ) that the same

[16] We realize that it is difficult to express time preference in one single quantifiable parameter. The factors that influence the time preference are of a typical qualitative nature. These factors need to be identified and translated into a reasonable estimate to reflect the preference of short-term objectives over long-term objectives.

government was in power[17]:

(O-4) $\text{POPINDEX}(t) = \text{f}(i\in \text{I}, \text{x}\in[\tau,t]\text{: SUBS}(g,i,x), \text{INVEST}(g,i,x), \text{TAX}(g,i,x), \text{LAW}(g,i,x))$

Where:

POPINDEX(*t*) the "popularity index" for the government, a measure which approximates the votes of the population.

τ the first year that this government was in power.

In the long-run the government aims to maximize a weighted function of its popularity indices[18]:

Maximize:

(O-5) $$\text{OBJ(g)} = weight(g,t=year1) * \text{POPINDEX}(t=year1) + weight(g,t=year2) * \text{POPINDEX}(t=year2) + \ldots + weight(g,t=year10) * \text{POPINDEX}(t=year10)$$

Where:

weight(g,t) weight factors describing the time preference of the government.

The long-term objective function is a weighted function of the popularity index over the whole planning period of 10 years.[19] The values of the parameters *weight(g,t)* must be decreasing in time, in order to reflect that later years have a lower priority for the government.

[17] We will not specify the shape of the "popularity function" denoted by f(.) because it requires additional information on the characteristics of the government and study area. The exact nature of the function depends on how the policies affect the livelihoods of the inhabitants of the region. The popularity of the government is also strongly affected by external factors such as economic recessions or natural hazards. Such factors only need to be included in the equation (O-4) as far as they have an effect on the decision variables included in this model. Another issue to consider is that survival of the government may in fact not depend much on its popularity with the farmers in the forest but rather on the urban masses or international sponsors (e.g. of the army).

[18] Instead of maximizing popularity the government objective might as well be formulated as a goal objective, that is, to keep the popularity at a level which is just enough to remain in power or to be re-elected in the next elections.

[19] We have chosen for a planning horizon of 10 years, similar to the household submodel. The planning horizon for governments may as well be shorter. In that case, the weight factors must equal zero for the years beyond the actual planning horizon. For example, the planning horizon reduces to seven years if *weight(g,t=year8)* = *weight(g,t=year9)* = *weight(g,t=year10)* = 0.

The constraints in the government submodel

The constraints for the government submodel need to be defined on the basis of additional information for specific study areas. The constraints together define the boundaries within which the government can choose the decision variables. The limits are determined on the one hand by the available government budget (resulting from tax revenues, debt servicing etc.), and on the other hand by commitments made earlier[20]. Some equations are necessary to define the relations between the state variables and the policies chosen by the government. The values of the state variables are results of the model. These results are important because they may change the institutional or natural environment for the households.

In the following we will define some constraints that we consider relevant independent of the location.

FINANCIAL CONSTRAINT

The financial constraint is the so-called "budget equation", which specifies that the sum of the expenditures and revenues is not allowed to exceed a certain budget.

(F-10) $\sum_i (\mathrm{INVEST}(g,i,t) + \mathrm{SUBS}(g,i,t) - \mathrm{TAX}(g,i,t)) \leq budget(g,t)$

Where:

budget(g,t) the (exogenous) government budget available for policies in year *t*

The government has more money available for investments and subsidies if it raises taxes.

LAND CONSTRAINTS

In the government submodel, we are particularly interested in the size of the forest land. Different from the household submodel, land (allocation) was defined as a state variable in the government submodel because it is an indirect result of a number of government policies (the decision variables) and the initial allocation of the land. The total size of public forest is specified in equation (A-4) as a function of laws (and law enforcement) concerning forest protection and those concerning land rights[21]:

(A-4) $\mathrm{A}(g,i{=}primary\ forest,q,t)) = \mathrm{f}(\mathrm{LAW}(g,i{=}forest\ protection,t), \mathrm{LAW}(g,i{=}landlaws,t))$

Where :

A(*g,i=primary forest,q,t*) the total acreage of high quality forest in the study area[22]

[20] Commitments made by the government consist of, for example, constitutional laws, election programs, or large infrastructure works started by previous governments.

[21] The exact shape of the functions can not be given at this level of abstraction. The equations do specify the decision variables which are relevant in the definition of each of the state variables.

[22] Secondary forest, covering the area of sustainably managed forest, could also be included as a separate category.

The following endogenous factors define the landprice:

(A-5) LANDPRICE*(h,i,q,t)* = f(INVEST(*g,i=infrastructure,t),* INVEST(*g,i=regional development,t),* LAW(*g,i=landlaws,t)*

The value of the farm land of household *h* is a function of government investments in infrastructure (+) and regional development (+), as well as laws defining the land rights and possibilities to sell the land (+/-).

(A-6) MAXEXPAND*(h,i,q,t)* = f(LAW(*g,i=forest protection,t))*

The quantity of land available to the farm household for purchase or to be occupied depends on government laws (and law enforcement) with respect to nature reserves and national parks (-).

(A-7) PRORIGHT*(h,t)* = f(LAW(*g,i=landlaws,t))*

The property right on a piece of land is a function of laws defining property rights on occupied land in the forest areas.

EQUATIONS DEFINING OTHER STATE VARIABLES

The equations (S-1) to (S-7) define the values of the other state variables: CREDMAX,TRANCOST, INTRATE and LOGREV.

(S-1) CREDMAX*(h,t)* = f(SUBS(*g,i=credit,t),*LAW(*g,i=landlaws,t)*

The maximum amount of credit that the household is allowed to borrow is a function of government credit programs and subsidies, and of laws that define property rights for the farm households. A farmer who owns land can use the land as a collateral to obtain credits.

(S-2) TRANCOST*(h,t)* = f(INVEST(*g,i=infrastructure,t))*

Transportation costs between farm *h* and the nearest market is a function of government investments in (road) infrastructure.

(S-3) INTRATE*(h,t)* = f(SUBS(*g,i=credit,t))*

The interest rate for credit is a function of government credit programs and subsidies.

(S-4) LOGREV*(t)* = f(TAX(*g,i=logging,t*),LAW(*g,i=logging*))

Revenues from off-farm logging are a function of taxes on logging (-), and also of laws prohibiting or restricting logging activities (-).

NON-NEGATIVITIES

All decision variables and state variables must be non-negative.

The "core submodel" for the government consists of an objective function (O-5), constraints (O-4), (F-10), (A-4) to (A-7), (S-1) to (S-4), and the non-negativities. This submodel can not yet be solved in its present form because additional information needs to be assembled and integrated in the model. We fully acknowledge that this is a very difficult task, because it requires that government decision-making can be made transparent. The modelling effort does not avoid this problem, but it could be useful indeed in the process of collecting information. Depending on the information that can be obtained, this model version can be used as a starting point from which model equations can be refined and added.

3.2.3. The tertiary actor submodel

The *IOSD submodel* is the key submodel for simulations and scenario analysis for international policies. It can be built up as an extension to the government submodel. By adding conditions to the government submodel (the "tertiary conditions"), the government takes into account the funds and policy conditions from the tertiary actor. The most important instrument that the IOSD uses to implement their policies is the allocation of funds to the policies of national governments. The IOSD extends loans or subsidies to governments in exchange for commitments or performance by the national governments towards sustainable development policies.

The questions that the IOSD submodel addresses are:

1. How can funds be allocated among national governments with the best results in forest protection and sustainable development?
2. Which commitments should be asked from the national governments in exchange for funds?

The first question is another way of saying "where is the most effective place to spend money on forest protection?" It is a relevant question from a climatologic viewpoint: forests are beneficial for storing carbon dioxide, no matter where they are located. The matter then reduces to finding the cheapest place. The cheapest places are normally the places where governments are least under pressure to allow forest destruction. For example countries with an abundance of arable land outside the forest areas, but also stable governments who have developed policies with a long term view. For intercountry comparisons we need much more information on these aspects. The government submodel developed in the previous subsections can be a useful tool to structure the collection and analysis of this information.

TERTIARY CONDITIONS

The key to interventions in the government submodel is the parameter *budget.* If the government gets a certain amount of funds from the IOSD, then the government budget is composed of a national budget and of a contributed funds by the IOSD:

(T-1) $\mathrm{BUDGET}(g,t) = natbudget(g,t) + \mathrm{FUND}(g,t)$

Where:

BUDGET(g,t)	the government budget available for policies in year t (replaces the parameter *budget* in equation (F-10))
FUND(g,t)	the amount of money added to the national budget of government g, year t.
natbudget(g,t)	the national component of the government budget in year t

The second question deals with the conditions under which the IOSD extends funds to the governments. The IOSD can reward "good policies" by granting more funds. An example to reward governments is by means of the "global forest fund" described in chapter 1. Or, in other words, the variable FUND is valued as a linear function of the number of hectares of primary forest:

(T-2) $\mathrm{FUND}(g,t) = \sum_q(\mathrm{A}(g,i{=}primary\ forest,q,t) + \sum_h \mathrm{A}(h,i{=}primary\ forest,q,t))$
$* \ forsubs$

Where:

forsubs the "global forest fund" subsidy for primary forest, in US $ per hectare.

We have seen earlier equation (A4) of the government submodel that the acreage of state forest depend on laws and law enforcement. Indirectly, the acreages of primary forest in private (household) land also depends on variables in the government submodel, such as INVEST(g,i,t), SUBS(g,i,t), TAX(g,i,t), and LAW(g,i,t), because of the linkages between the government submodel and the motivational factors (parameters) in the household submodel.

The value of the parameter *forsubs* depends on the priorities of the IOSD and on the negotiations between the IOSD and the national government. Also, if more conditions need to be added, these can be expressed as additional constraints. The tertiary conditions (T-1) and (T-2) need to be integrated in the government submodel in order to calculate the amount of funds that the government is willing to "exchange" for adopting certain policies that the international community judges necessary for forest protection. Note that this way of modelling does not allow the IOSD to be a real "decision-maker". This is a consequence of the sovereignty principle that foreign actors can not take national decisions, but they can only provide incentives for governments to change policies in a direction that coincides with their objectives.

3.2.4. Model linkages

The core model described in this chapter can be used to evaluate policy scenarios. Basically, this can be done in two different ways. The most "advanced" way is to start with the optimization problem for the government (including the tertiary conditions described in 3.2.3), then to implement the resulting parameters in the household submodel, and finally to evaluate the resulting land use decisions. However, this requires a fully advanced and validated government submodel. This will be difficult especially in case of governments in an unstable political climate. In that case, we can revert to performing "what-if" policy evaluations with the household submodel only. By gradually changing the values of the parameters we can simulate the effects of certain policies. In this way, the household submodel can be used for an ex-ante evaluation of policies. This will improve our insight in the effectiveness of policies which benefit the state of the forests. The question of whether the government is willing to adopt such policies is still unanswered.

The possibilities to use the tertiary actor submodel for policy evaluations is also very dependent on the possibilities to describe government decision-making. In the ideal situation, if we have validated decision models for several regions in the world, including all the government submodels, the model could indeed be used to find an optimal allocation of funds among these different countries and regions. However, such an optimization does not take into account that forest areas have different qualities. It is probably more realistic to use the model for "what-if" scenarios to evaluate the effect of allocating funds under a variety of conditions and policy commitments.

3.3. Choosing a study area

In the previous section we discussed the so-called "core model structure". Before the model parameters can be estimated, we first have to choose a suitable study area. The choice for Ecuador was made because:

(a) this is one of the five areas in which the CML-NOP project carried out field research
(b) There was a substantial amount of secondary data available for estimating parameters.

3.3.1. The study area: The Northeast provinces of the Ecuadorian Amazon.

The field research in Ecuador was carried out in 1997, in the provinces of Napo and Sucumbios in the Amazon region of Ecuador. The major findings of this field research are documented in Cleuren (2001). The knowledge obtained from the field study already provided a great source of information on actor behaviour. We also made ample use of other research carried out in the region, notably studies based on a survey from the Carolina Population Center in 1990 (see Pichón, 1993a, 1993b, 1996; see also Pichón and Bilsborrow, 1991, 1992; Thapa, Bilsborrow and Murphy, 1996). Our own data collection was aimed at making data from secondary sources available and bringing all the data together into our GIS framework.

Within the Ecuadorian Amazon, the provinces of Napo and Sucumbios were chosen because an intensive colonization process is taking place in these provinces. The colonization process accelerated since the 1970s, facilitated by the roles of the oil sector and the government which provided the necessary infrastructure to attract farmers from other parts of the country. Our purpose was to describe the colonization and land use processes in the settlements which arose in the previous decades, in order to understand the mechanisms and conditions under which forests are likely to degrade or disappear.

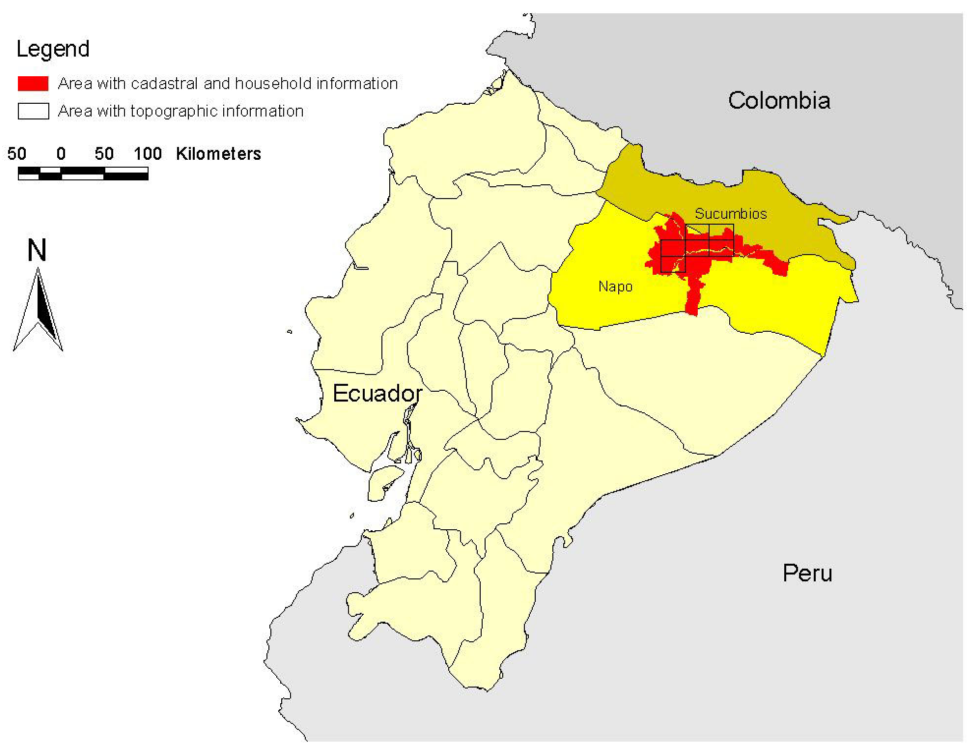
3.3.2. Spatial delimitation of the study area

After we had chosen for the provinces of Napo and Sucumbios as a suitable region for field study and data collection, the study area had to be narrowed down further. The availability of data played an important role in the selection, but we also wanted to choose an area which gives a good representation of the farm household behaviour in the region. We finally chose the district Francisco de Orellana (Coca) in the Napo province (see Figure 3-3). For this area we had three important data sources which we used to validate the model[23]:

a) a cadastral map of the area with the location of the settlements and basic information on the household characteristics
b) topographic maps with detailed data on infrastructure and natural resources
c) a random household survey with detailed information on household demographics and land use.

Figure 3-3 shows the area coverage of the data. The darkest area is the area for which cadastral data and household data (a and c) were available. The topographic information (represented by the six rectangles) (b) was not available for the whole region, therefore we had to delimit the study area somewhat further to the cross-section of the topographic and the cadastral maps.

Figure 3- 3 *Location of the study area in Ecuador*



[23] A more detailed description of the data and sources is presented in chapter 4.

Unit of analysis

The farm household is the unit of analysis. A higher level of decision-making might be the co-operative or 'comuna', which is the first organisation level for migrant households and indigenous farmers respectively. However, it appears that these are very weak and loose organisations, which do not really interfere with the autonomy of the household (see e.g. Pichón, 1993). Therefore, the household is appropriate as the decision-making unit.

Geo-referencing of the unit of analysis

A square kilometre grid was chosen to spatially represent the household plots. The average plot size is 50 ha (250 meters by 2 kilometres, see Pichón, 1993). The square kilometre grid is double the size (100 ha) of the average household plot and has a different shape (square versus rectangle). This difference in size is not particularly important for modelling purposes as long as we carefully scale the relevant parameters in the right proportion. The difference in shape is important as far as we consider the access to the nearest road. In the first line of household plots, the smallest edge is on the roadside. The other side of the plot is situated 2 kilometres from the road. Obviously, it requires more effort to bring goods from the backside of the plot to the roadside than from the front side, but this factor is not taken into account for the modelling purposes.

The choice of a quadtree GIS (SPANS, see also section 3.2) allows the flexibility to use a smaller or larger scale as well, if this would match better with the scale of the input data. However, the square kilometre gridsize matches very well with the scale of both the socio-economic and the geographical data sources. The topographic maps, for example, were scaled 1:50 000, while all other maps had a larger scale.

The study area was projected using the Universal Transverse Mercator (UTM) projection, which contains a square kilometre grid. The UTM projection is *conformal*, meaning that meridians and parallels intersect at square angles. Like every projection, there is a small distortion but this can be considered negligible for our study area[24].

3.4 The farm household submodel applied to the Ecuadorian Amazon

This section discusses an example of the household submodel as applied to the Ecuadorian migrant farm household. The model structure is similar to the core model structure described in section 3.2.1. For the empirical part of the project, the application concentrates on a number of short-term production decisions of the farm households. The aim of this model version was to describe the actual behaviour of farm households with a focus on operational decisions. In future research the strategic decisions can be modelled as well. The present model version well illustrates the process of moving from the core model structure towards an applied model which can be calculated and spatially linked to a specific study area.

Figure 3-3 illustrates how our primary actor, the migrant farm household, can be represented in an Action-in-Context framework. The literature on the Ecuadorian Amazon points out that there exist two important groups of primary actors: the indigenous farmers and the migrant farmers. The indigenous farmers originate from the Amazon region itself, whereas the migrant farmers originate from other regions in Ecuador, notably the Andes provinces. Theoretically, this difference in background would indicate an important difference in land use practices. However, our field research

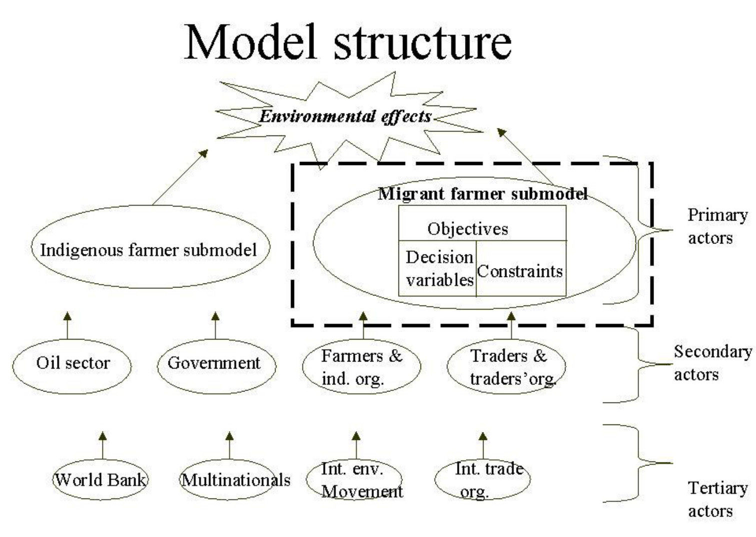
[24] The UTM projection can be used with negligible distortion for study areas which are near the equator and not larger than 6° of longitude (approximately 660 km). The projection used for this study area is UTM 18 South.

as well as other research in the region (e.g. Pichón, 1993) points out that the differences between the two groups of farmers are diminishing in the course of time. For this reason, and also because of data limitations, we have chosen to concentrate this model application on the group of migrant farmers only.

The second layer in Figure 3-3 describes the relevant secondary actors. In the Ecuadorian context the oil sector is a dominant actor, since much of the local economy is built up from oil extraction. The oil sector is an important employer for residents of the region. Both the oil sector and the government play an active role in building and maintaining infrastructure. Law enforcement is a task of the government but the oil sector has an important influence on the government because of the economic interests involved. Much of the government budget is contributed from oil revenues. Farmers and indigenous people in the study area are often member of organizations that represent their interest and which are rather important pressure groups. They may have some influence on policies and may also play an active role in the implementation of policies as far as the interests of their members are served by these policies. Traders and traders' organizations play a role in the market environment for the farm households.

The tertiary actors are the usual conglomerate of international organizations. The international environmental movement is seriously concerned with the environmental effects of the oil extraction activities. Many environmentalists would welcome the complete withdrawal of the oil sector, nevertheless their "votes" are marginal because of the enormous economic interests involved. International policies are best described as geared towards economic development of the region, while reducing the pollution and harm done to primary forests, and protecting the interests of indigenous people.

Figure 3-3 *Actor linkages in the Ecuadorian land use model*



This model application is a reduced version of the core model structure presented in 3.3.

Summarizing, the Ecuadorian household model contains the following characteristics:

A. With respect to the model structure and decision options

- short term land use decisions, decision horizon of 1 year
- yields are exogenous, given (derived from average yields in survey)
- household activities: annual allocation of land to crops and pasture, hiring labour, food purchases versus subsistence production
- "abandoning", or not working certain parts of the land for the next year

B. Farm household characteristics

- household size (including its implications for family labour, family food demand)
- area currently under coffee, manioc, pasture etc.

C. Institutional environment and technology

- farm gate prices
- costs of inputs
- costs of hiring labour, a labour market is assumed to exist

D. Natural resource base

- Soil type

The following factors are not included in the Ecuadorian household model

A. With respect to the model structure and decision options

- long term (strategic) decisions: forest clearing, selling land, …
- endogenous yields as a function of several production factors
- off-farm income generation
- attitudes towards saving / cattle accumulation
- the whole range of functions of livestock: consumption (milk, eggs, carrying loads), profit, capital accumulation, collateral for loans, status.
- sales of on-farm wood

B. Farm household characteristics

- parameter estimates based on a larger number of households
- duration of stay
- initial wealth

C. Institutional environment and technology

- relation between prices and distances from markets
- land values, depending on location, tenure status, soil quality and topography

D. Natural resource base

- relief
- use of firewood, other forest products

Most of the factors which are not included in the Ecuadorian household model could be added with the available data, but this requires more data gathering and analysis.

3.4.1. Variables and parameters in the Ecuadorian household model

Activities, decision variables, and inputs

The activities undertaken by the Ecuadorian farm household are similar to the activities in the core model, as presented in Table 3-1. We have made a few adaptations for a better description of the local context, as presented in table 3-6.

Table 3-6. *Activities of the Ecuadorian farm household*

ON-FARM LAND USE ACTIVITIES:

- annual food crop production: manioc
- annual cash crop production: corn and rice.
- perennial cash crop production: cacao and coffee.
- pasture: cattle and small ruminants
- land conversion: crop land or pasture land to fallow land

OFF-FARM INCOME GENERATING ACTIVITIES:

- off-farm labour

NON-INCOME GENERATION ACTIVITIES:

- consumption of manioc
- purchases of manioc

The activities presented in this table were derived from the survey carried out by the University of

North Carolina (see Pichón, 1993). A few others were left out of the analysis because the data did not allow for a proper estimation of parameters. Examples of crops for which little data was available are plantain and fruits. The lack of data may indicate that these crops are less important than the ones included in the model.

Data on logging activities are virtually non-existent in the study area. It is often said that selling wood is not worth the costs of labour, though other sources (e.g. Cleuren, 2001) claim that logging is done on a substantial scale along the main roads of the colonized area. Without further evidence we will assume that wood is a by-product in the conversion of forest land to arable land and as such it is important in the strategic decisions of the household. This model version describes only operational decisions, therefore logging is not included in the activities of Table 3-6.

The household investments in farm land are represented by investments in pasture. It is often reported that Ecuadorian farmers aim for a high proportion of pasture on their farmland. The literature reports three major reasons for this strategy. First, having a large herd of cattle is supposed to give a certain status to the farmer. Second, cattle can be used as 'savings', especially in the absence of well-functioning financial markets. Third, cattle are supposed to be less demanding on labour input than growing crops. We will take this particular status of pasture into account by attaching a certain value to pasture land, which incites the household to make investments in pasture.

The decision variables are given by:

A*(h,i,q)*	the area of quality *q* where the household *h* grows (harvests) product *i*, $i \in$ {land use activities};
PURCH(*h,i*):	purchased quantities of product *i* for household *h*, *i* =manioc;
LFAM*(h,i,q)*	the amount of family labour allocated to activity *i*, $i \in$ I ;
LHIRE*(h,i,q)*	days of labour hired by household *h* for activity *i* on land of quality *q*.

Where:

i	activities,	$i \in$ {cacao, coffee, corn, rice, manioc, cattle, fallow, primary forest, secondary forest, off-farm labour}
q	land quality,	$q \in$ {black, red, alluvial}

The decision variables are similar to the ones defined in the core model. The variables for credits, expansion and disposal of farm land are not taken into account. In our study area farmers hardly have access to formal credits. Insufficient information was available on the use of informal credits. Expansion and disposal of farm land are strategic decisions which do not play a role in this model version which describes only operational decisions.

The parameters and state variables

The state variables are defined as follows:

FOODCOST*(h)*	total costs of buying food for consumption of household *h*
K*(h,i,q)*	amount of household capital assigned to activity *i* on land of quality *q*
LANDVAL(*h*)	the (market) value of farm land *h.*
NETREV(*h*)	net revenues of household *h*
NFINC*(h)*	non-farm income of household *h*
OBJ*(h)*	objective of farm household h

OUTPUT*(h,i,q)*	output of good *i* on land of quality *q* for household *h*
PRODCOST*(h)*	total costs of production and harvesting for household *h*
REV*(h)*	total gross revenues from sales of farm output of household *h*

As compared to the core model, the index *t* is not necessary as long as the model is reduced to cover one single decision year.

To adapt for the single-year decision-making process we introduced two additional parameter:

astart(*h,i,q*)	the starting value for land allocated to activity *i*
capstart(h)	the starting capital available to the household

The parameter *astart*(*h,i,q*) replaces the decision variable A(*h,i,q,t-1*) from the core model. The parameter *capstart(h)* replaces the variable TOTCAP(*h,t*). For a further description of the parameters we refer to the appendices 3-I and 3-II.

3.4.2. Model equations for the Ecuadorian household model

In designing the model equations we follow the same structure as for the core model, with some adaptations for the specific study area. The model consists of the following equations:

OBJECTIVE FUNCTION

Maximize:

(O-2′) $\mathrm{OBJ}(h) = \mathrm{NETREV}(h) + \mathrm{LANDVAL}(h)$

for all $h \in H$

The Ecuadorian farm household aims to maximize the net revenues from (farming) activities as well as to increase the value of the land.

PRODUCTION FUNCTIONS

(P-1′) $\mathrm{OUTPUT}(h,i,q) = \mathrm{A}(h,i,q) * yield(h,i,q)$

For all $h \in H$, $i \in \{\text{land use activities}\}, q \in Q$.

(P-2′) $\mathrm{LFAM}(h,i,q) + \mathrm{LHIRE}(h,i,q) = labreq(h,i,q) * \mathrm{A}(h,i,q)$

For all $h \in H$, $i \in \{\text{land use activities}\}, q \in Q$.

(P-3′) $\mathrm{K}(h,i,q) = capreq(h,i,q) * \mathrm{A}(h,i,q)$

For all $h \in H$, $i \in \{\text{land use activities}\}, q \in Q$.

The production functions for the Ecuadorian household are similar to those for the core model and need no further explanation.

CONSUMPTION FUNCTION:

(C-1′) $\sum_q$OUTPUT$(h,i=manioc,q)$ + PURCH$(h,i=manioc) \geq consreq(h)$
for all h∈H.

Manioc is the only food product considered in this model. It can be produced on-farm as well as purchased in the market.

FINANCIAL CONSTRAINTS

(F-1′) NETREV(h) = REV(h) + NFINC(h) - PRODCOST(h) - FOODCOST(h)
for all h∈H.

(F-2′) REV$(h) = \sum_{i\in\ \{\text{cash crops, forest products, animal products}\},q}$ OUTPUT(h,i,q) * $price(h,i)$
for all h∈H.

(F-3′) NFINC(h) = $wage$* LFAM$(h,i=off\text{-}farm\ wage\ labour)$
for all h∈H.

(F-4′) PRODCOST$(h) = \sum_{i,q}$ $(wage$*LHIRE (h,i,q) + K(h,i,q) $)$
for all h∈H.

(F-5′) FOODCOST(h) = PURCH$(h,i=manioc)$*$price(h,i=manioc)$
for all h∈H.
(F-8′) PRODCOST(h) + FOODCOST$(h) \leq capstart(h)$
for all h∈H.

The financial constraints (F-1′) to (F-8′) together describe the revenues and expenses of the farm household.

LABOUR CONSTRAINT:

(L-1′) $\sum_{i,q}$ LFAM$(h,i,q) \leq famlab(h)$ * $workdays$
for all h∈H.

The labour constraint is similar as in the core model described in section 3.2.1.

LAND CONSTRAINTS:

(A-2′) $\sum_{i\in\{land\ use\ activities\}}$ A$(h,i,q) \leq \sum_{i\in\{land\ use\ activities\}} astart(h,i,q)$
for all h∈H,q∈Q.

(A-2′′) A$(h,i,q) \leq astart(h,i,q)$
for all h∈H, i∈{perennial crops,pasture},q∈Q.

(A-2′′′) $\sum_{i\in\{annual\ crops\}}$ A$(h,i,q) \leq \sum_{i\in\{annual\ crops\}} astart(h,i,q)$

for all $h \in H, q \in Q$.

(A-3′) $\text{LANDVAL}(h) = \sum_{i \in \{pasture\}, q} \text{A}(h,i,q,t)$
for all $h \in H$.

The land constraints differ from the core model in two respects. To adapt for the single-year decision-making process we introduced parameters *astart*(*h,i,q*) to describe the starting values for land allocation. In the core model these were decision variables A(*h,i,q,t-1*). In the core model, the land allocated to activity *i* has a starting value A(*h,i,q,t-1*) depending on the solution for the previous year. In this model version the starting values have to be derived from the data as exogenous parameters. The second adaptation concerns the land value, LANDVAL(*h*). We have assumed here (for reasons mentioned earlier in this section) that the land value depends on the total area under pasture.

NON-NEGATIVITIES:

All variables except NETREV and OBJ must be nonnegative by definition.

The model application for the Ecuadorian household can be solved as soon as the values for the parameters are known. These values have been derived from secondary data as well as from GIS data. A description of the data and sources as well as a presentation of the estimated parameters is the subject of the next chapter. Then, in chapter 5, the model will be solved and a discussion on the results is presented.

Appendix 3-I
List of symbols in the core model

<u>*HOUSEHOLD SUBMODEL*</u>

DECISION VARIABLES

A*(h,i,q,t)*	land of quality *q* allocated to activity *i* by household *h*, year *t*
CREDIT(*h,t*)	the credit used to finance expenses in year *t*
DISPOSE(*h,i,q,t*)	disposed farm land of quality *q* and land use *i*, year *t*
EXPAND(*h,i,q,t*)	expansion of farm land of quality *q* and land use *i*, year *t*
LFAM*(h,i,q,t)*	family labour in days allocated to activity *i* on land of quality *q*, year *t*
LHIRE*(h,i,q,t)*	days of labour hired by household *h* for activity *i* on land of quality *q*, year *t*
PURCH*(h,i,t)*	purchases of product *i* for household *h*, year *t*

STATE VARIABLES

FOODCOST*(h,t)*	total costs of buying food for consumption of household *h*, year *t*
INV(*h,t*)	costs of expansion of farmland
K*(h,i,q,t)*	amount of household capital assigned to activity *i* on land of quality *q*, year *t*
LANDVAL(*h,t*)	the (market) value of farm land *h*.
NETREV(*h,t*)	net revenues of household *h*, year *t*
NFINC*(h,t)*	non-farm income of household *h*, year *t*
OBJ*(h,t)*	objective of farm household h, year *t*
OUTPUT*(h,i,q,t)*	output of good *i* on land of quality *q* for household *h*, year *t*
PRODCOST*(h,t)*	total costs of production and harvesting for household *h*, year *t*
REV*(h,t)*	total gross revenues from sales of farm output of household *h*, year *t*
TOTCAP(*h,t*)	total amount of capital available to household *h*, year *t*

PARAMETERS

(indices)

h	primary actors, $h \in H$
i	farming activity or product $i \in I$
q	land quality, $q \in Q$
t	time in years, $t \in T$

(index *i*)

I = {annual food crops, perennial food crops, annual cash crops, perennial cash crops, land conversion, animal husbandry, on-farm logging, off-farm logging, off-farm wage labour, consumption of farm products, food expenses, non-food expenses, saving/investment}

(subsets for index *i*)

I ={land use activities, off-farm income generating activities, non-income generating activities} {land use activities} = {annual food crops, perennial food crops, annual cash crops, perennial cash crops, land conversion, animal husbandry, on-farm logging}

{off-farm income generating activities} = { off-farm logging, off-farm wage labour}

{non-income generating activities} = {consumption of farm products, food expenses, non-food expenses, saving/investment}

(index t)
T = {year1,…year10}.

capreq(h,i,q,t)	capital requirements per ha for activity *i* on land of quality *q*, year *t*
consexp(h,t)	non-food consumer expenses for household *h*, as a fraction of last year's net revenues.
consreq(h,t)	food consumption requirement, depending on nutritional values of food products and household composition.
credmax(h,t)	the maximum amount of credit that the household is allowed to borrow, depending on collateral (property rights on farm land)
famlab(h,t)	family labour available, in "male equivalents", year *t* per year
intrate(h,t)	the interest rate for credit
labreq(h,i,q,t)	labour requirements per ha in working days
landprice(h,i,q,t)	the price the farm household pays/receives for buying/selling farmland depending on location, land quality, current land use.
logrev(t)	revenues from off-farm logging, per working day, year *t*
maxexpand(h,i,q,t)	the quantity of land available to the farm household for purchase or to be occupied in year *t*.
price(h,i,t)	the price that household *h* receives for selling 1 kg of farm product *i,* depending on the market price and transportation costs to the market, year *t*.
proright(h,t)	an indicator for the property or usufructuary rights of the household *h*.
trancost(h,t)	transportation costs between farm *h* and the nearest market, depending on distance, means of transport and quality of infrastructure
workdays	the number of working days per year per person
wage(t)	daily wage, year *t*
yield(h,i,q,t)	yield/hectare of product *i* on land of quality *q* for household *h,* year *t*.

DECISION VARIABLES

SUBS(*g,i,t)*	the government subsidy for activity *i* in year *t*
INVEST(*g,i,t*)	the government's direct investment in activity *i* in year *t*
TAX*(g,i,t*)	tax revenues, raised on activity *i,* year *t*
LAW*(g,i,t*)	laws affecting activity *i*, year *t*

STATE VARIABLES

CREDMAX*(h,t)*	the maximum amount of credit that the household is allowed to borrow, depending on collateral (property rights on farm land)
TRANCOST*(h,t)*	transportation costs between farm *h* and the nearest market, depending on distance, means of transport and quality of infrastructure
INTRATE*(h,t)*	the interest rate for credit
LANDPRICE*(h,i,q,t)*	the price the farm household pays/receives for buying/selling farmland depending on location, land quality, current land use.
LOGREV*(t)*	revenues from off-farm logging, per working day, year *t*
MAXEXPAND(*h,i,q,t)*	the quantity of land available to the farm household for purchase or to be occupied in year *t*.
PRICE*(h,i)*	the price that household *h* receives for selling 1 kg of farm product *i,* depending on the market price and transportation costs to the market.
PRORIGHT*(h,t)*	an indicator for the property or usufructuary rights of the household *h*.

PARAMETERS

budget(g,t)	the (exogenous) government budget available for policies in year *t*
natbudget(g,t)	the national component of the government budget in year *t*
forsubs	the "global forest fund" subsidy for primary forest, in US $ per hectare.

Appendix 3-II
List of symbols in the model application for Ecuador

DECISION VARIABLES

A*(h,i,q)*	the area of quality *q* where the household *h* grows (harvests) product *i, i*∈ {land use activities};
PURCH(*h,i*):	purchased quantities of product *i* for household *h, i* =manioc;
LFAM*(h,i,q)*	the amount of family labour allocated to activity *i, i*∈ I ;
LHIRE*(h,i,q)*	days of labour hired by household *h* for activity *i* on land of quality *q*.

STATE VARIABLES

FOODCOST*(h)*	total costs of buying food for consumption of household *h*
K*(h,i,q)*	amount of household capital assigned to activity *i* on land of quality *q*
LANDVAL(*h*)	the (market) value of farm land *h.*
NETREV(*h*)	net revenues of household *h*
NFINC*(h)*	non-farm income of household *h*
OBJ*(h)*	objective of farm household h
OUTPUT*(h,i,q)*	output of good *i* on land of quality *q* for household *h*, year
PRODCOST*(h)*	total costs of production and harvesting for household *h*
REV*(h)*	total gross revenues from sales of farm output of household *h*

PARAMETERS

(indices)

h	actors, h ∈ {fh1,..,fh3}
i	activity
	i ∈ { cacao, coffee, corn, rice, manioc, cattle, fallow, prim_for, sec_for, off-farm labour}
q	land quality, q ∈ { black, red, alluvial}

(subsets)

annuals*(i)*	annual crops : {corn, rice, manioc }
cash*(i)*	subset of *i* : {cacao, coffee, corn, rice }
food*(i)*	subset of *i* : {manioc }
landuse*(i)*	land use activities : {cacao, coffee, corn, rice, manioc, cattle, fallow}
past*(i)*	pasture activities : {cattle}
per*(i)*	perennial crops : {cacao, coffee}

astart(*h,i,q*)	the starting value for land allocated to activity *i*
capreq(h,i,q)	capital requirements per ha for activity *i* on land of quality *q*
capstart(h)	the starting capital available to the household
consreq(h)	food consumption requirement, depending on nutritional values of food products and household composition.
famlab(h)	family labour available, in "male equivalents"
labreq(h,i,q)	labour requirements per ha in working days
price(h,i)	the price that household *h* receives for selling 1 kg of farm product *i,* depending on the market price and transportation costs to the market.

wage	daily wage
workdays	the number of working days per year per person
yield(h,i,q)	yield/hectare of product *i* on land of quality *q* for household *h*.

4

Description of data, data analysis and transformation

The land use model described in section 3.4 was tested with data for the study area in the Napo province in the Ecuadorian Amazon region. Data from a number of secondary sources have been used to estimate the model parameters. This chapter describes the data sources (section 4.1), the estimation methods (section 4.2) and the resulting parameter estimates (section 4.3).

4.1. Data and sources

As we explained in chapter 3, the parameters represent the exogenous factors that influence the decisions of the farm household. We categorized these factors into three groups: the *household and farm characteristics*, the *institutional environment and technology*, and the *natural resource base* (see also table 3-3). In the following we will introduce the data and data sources according to this classification. We recall that the parameters to be estimated from the data can also be classified into these three categories:

Table 4-1. Parameters in the core household submodel

HOUSEHOLD AND FARM CHARACTERISTICS	
capreq(h,i,q,t)	capital requirements per ha for activity *i* on land of quality *q*, year *t*
consexp(h,t)	non-food consumer expenses for household *h*, as a fraction of last year's net revenues.
consreq(h,t)	food consumption requirement, depending on nutritional values of food products and household composition.
famlab(h,t)	family labour available, in "male equivalent" working days, year *t* per year
labreq(h,i,q,t)	labour requirements per ha in working days
INSTITUTIONAL ENVIRONMENT AND TECHNOLOGY	
credmax(h,t)	the maximum amount of credit that the household is allowed to borrow, depending on collateral (property rights on farm land)
intrate(h,t)	the interest rate for credit
landprice(h,i,q,t)	the price the farm household pays/receives for buying/selling farmland depending on location, land quality, current land use.
logrev(t)	revenues from off-farm logging, per working day, year *t*
maxexpand(h,i,q,t)	the quantity of land available to the farm household for purchase or to be

	occupied in year *t*.
price(h,i,t)	the price that household *h* receives for selling 1 kg of farm product *i,* depending on the market price and transportation costs to the market, year *t*.
proright(h,t)	an indicator for the property or usufructuary rights of the household *h*.
trancost(h,t)	transportation costs between farm *h* and the nearest market, depending on distance, means of transport and quality of infrastructure
workdays	the number of working days per year per person
wage(t)	daily wage, year *t*
NATURAL RESOURCE BASE	
yield(h,i,q,t)	yield/hectare of product *i* on land of quality *q* for household *h,* year *t*.

Table 4-1 describes all the parameters in the core household model. The data analysis has also been used to define a proper classification of the indices *h*, *i*, and *q*:

h index for the household as well as for the location of the farm
i index for the products/activities
q index for the quality of land

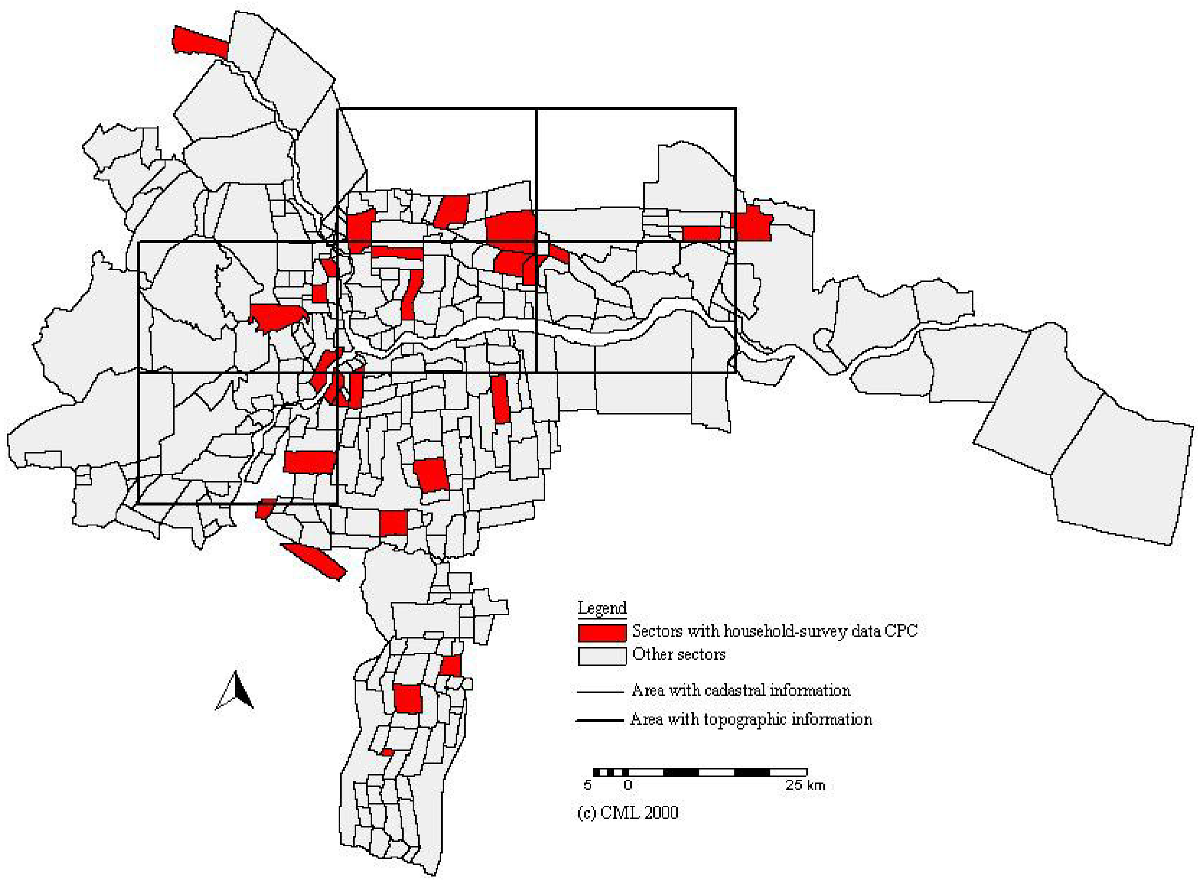
In chapter 3 we discussed both the core model as the general framework for our land use model and the reduced model with the application for the Ecuadorian study area. This chapter will focus on the reduced version, but some attention will also be paid to a description of methodology to estimate model parameters for the core model which do not appear in the reduced version.

The data sources are summarized in Table 4-2. A full description of the data sources is presented in Appendix 4-I.

Table 4-2 Data sources for the study area in the Ecuadorian Amazon

A. Household survey data Carolina Population Center (CPC)
B. Study on production systems in the Napo province by Estrada, Seré, and Luzuriaga
C. Various statistical yearbooks
D. Cadastral map
E. Topographical maps
F. Soil map

Figure 4- 1 Data on household and farm characteristics in the study area



Source: adapted from the cadastral map of the Coca (San Francisco de Orellana) region, IERAC

4.1.1. Data on household and farm characteristics

The data on household and farm characteristics in the study area originate from household survey data collected in 1990 by the Carolina Population Center (source A in Table 4-2) and from the cadastral map (source D in Table 4-2)[25].

Figure 4-1 shows the location of the households for which information is available. The cadastral map contains elementary data of each "sector" such as the number of farms and the surface. A "sector" is either a cooperative or indigenous community ("comuna"). The sectors may contain up to 50 households depending on the size of the area. Each sector has a number which was assigned by the IERAC, the government organisation responsible for registration of the cooperatives and comuna. The number indicates the age of the cooperative or comuna, because the numbers reflect the order in which the sectors have been registered.

In Figure 4-1, the sectors with red colour have been included in the CPC household survey. For each of these sectors, a sample of several households was selected for detailed interviews with both the head of the household and the spouse. For these selected households, the household composition and background is known, as well as many other factors which are relevant for the livelihood strategies. For example, the survey contains information on the current land use and methods of production. Additional information on methods of production and the labour requirements was derived from Estrada et al. (1988, source B in Table 4-2).

4.1.2. Data on the institutional environment and technology

Most of the data on the institutional environment were derived from the topographical maps (source E). Figure 4-2 shows the road network, river network and major towns. The towns have a market function, for farm products, for consumer goods, as well as for wage labour. The distance between the farm and the nearest market is an important parameter both as a determinant of farm gate prices as well as for the attractiveness of the farm location. The location of oil wells can also be considered as an indicator for employment possibilities, but this was not taken into account in the current model application.

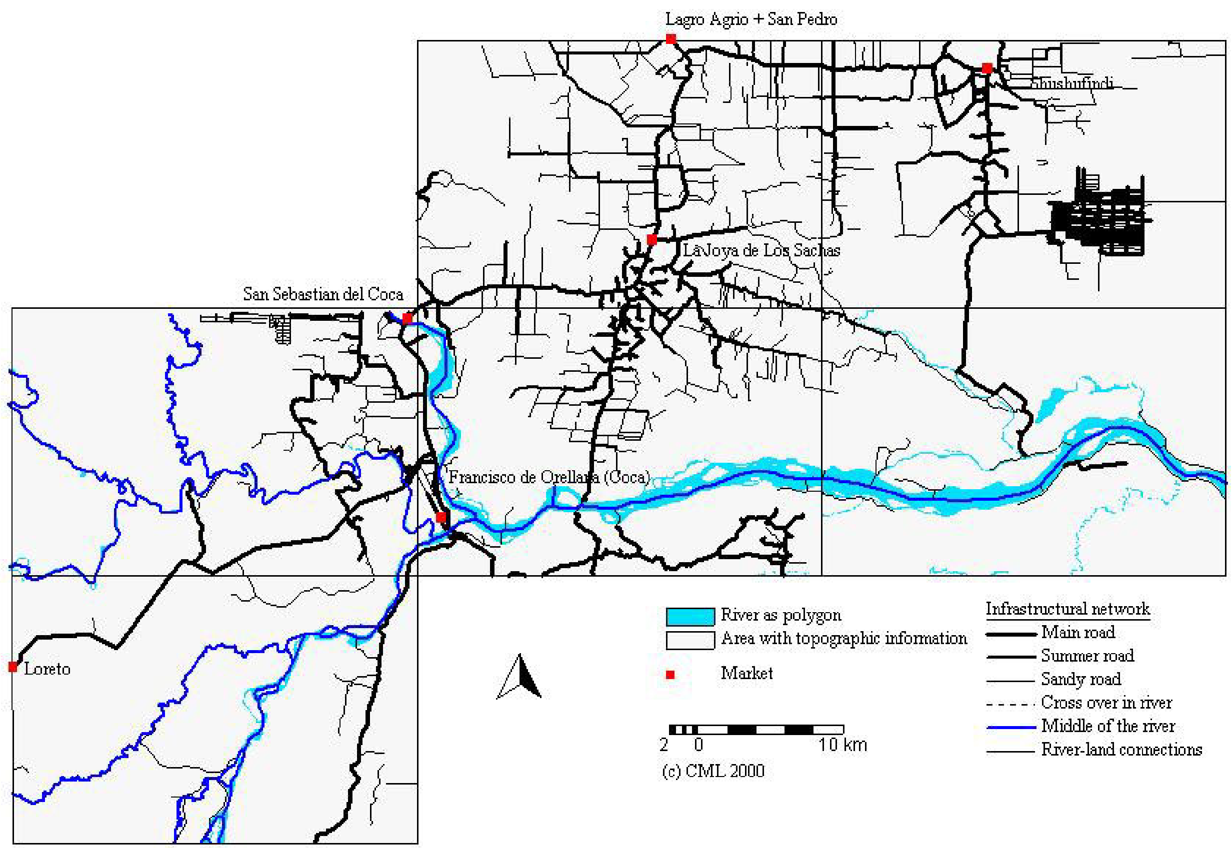
The region has two main markets which are well connected to the outside world. These markets are Coca and Lago Agrio[26]. Lago Agrio is situated in the Sucumbíos province.[27] We assume that the towns of Loreto and Lago Agrio have important markets for agricultural products because they are closest to the capital Quito. On a national scale, Quito is the most important distribution market.

[25] See also Appendix 4-I

[26] Coca is also known as San Francisco de Orellana; Lago Agrio is also known as Nueva Loja.

[27] In Figure 4-2 Lago Agrio is located symbolically on the edge of the study area. In this way the market could be taken into account even if it is outside the study area. San Pedro is a small town on the road between La Joya de los Sachas and Lago Agrio. As such it is an important "transit" market for goods from the region to be transported to the larger market of Lago Agrio. Loreto is also outside the study area, west of Coca.

Figure 4- 2 Data on infrastructure in the study area



Source: adapted from topographic maps, IGM/DMA

Additional information on infrastructure was also taken from the CPC household survey. The survey also collected data on distances and nearest towns. This information can be compared to the map data in order to obtain an indication of measurement errors.[28]

4.1.3. Data on the natural resource base

Physical data which are relevant for the land use decisions are *soil quality* and *slopes*. Soil data can be derived from a 1:500 000 soil map from MAG/PRONAREG (Figure 4-3, also described in Appendix 4-I as source F). According to Pichón (page 135) "the soils of North-eastern Ecuador are primarily ultisols, with small areas of brown-black, flat, volcanic soils and inceptisols (alluvial soils) along flood plains. The scale of the 1: 500 000 soil map of the area is too crude; the apparent homogeneity at the regional level masks the variability at the micro- or farm-level. Amazon soils are for the most part a patchwork, with radical differences in nearly every kilometre."

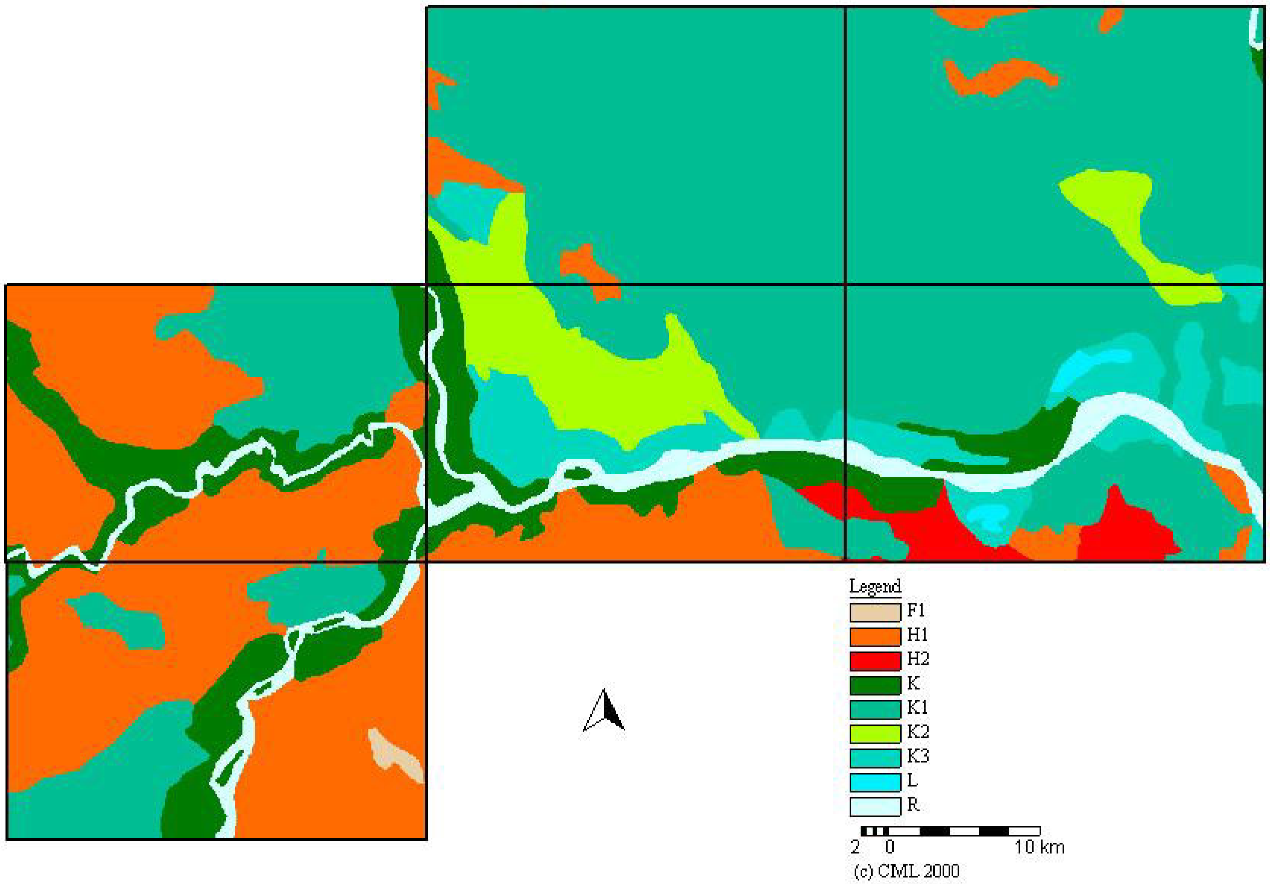
Even though the 1:500 000 soil map is too crude to get a real understanding of farming systems on the household level, we used it as a source of information for the dominant soil types. Figure 4-3 shows a variety of soil classes which occur in the study area. The CPC household survey assembled the quality of the soil as it is perceived by the farmer. The farmers distinguish four main classes: black soils, red soils, alluvial soils, and sandy soils. The black soils coincide roughly with the classes K1 and K2 in Figure 4-3; the red soils coincide with the classes F1, H1, and H2, and the alluvial soils coincide with the classes K3 and K. Sandy soils are not represented in the 1:500 000 soil map. The black soils are most wanted by the farm households, because they are suitable for all kinds of production. The red soils are not very fertile and contain toxic aluminium. The red soils can be used for agricultural production, but not for all products. Alluvial soils are fertile, but suffer from bad drainage (Estrada et al., 1988). Alluvial soils are suitable for annual crops.

The altitude differences can be described in more detail, thanks to the contour lines that we digitised from the topographical maps (source E in table 4-2). Figure 4-4 shows the graphical representation of the elevation classes in the study area. The actual altitude is not an important factor in understanding farm household behaviour, but the steepness and presence of hills is an important limiting factor in production possibilities. The contour lines in the topographic maps made it possible to bring further classifications in the factor "land quality". The CPC survey distinguishes steepness classes *level, rolling hilly,* and *steep.* Of course this is the steepness as perceived by the farmer. A statistical analysis might point out to what extent the "subjective" farm perceptions coincide with the steepness as derived from the topographic maps.

The quality of the "match" between data derived from farmers' perceptions and natural resource data determines whether the socio-economic and physical data can be linked. Farmers' decisions are based on perceptions of the land quality. These perceptions do not necessarily match with the physical data due to measurement and interpretation errors. The physical data cover a larger area than the socio-economic data, therefore we would rather use the physical data as a proxy for farmers' perceptions on land quality in the case of this model application. However, the quality of the "match" should be tested statistically as a part of the validation of the model.

[28] Note that both the map data and survey data are liable to measurement errors. However, we do not expect any systematic biases between distance measurements in both sources. A systematic comparison of distances is advised if a full validation of the model and parameter estimates is carried out.

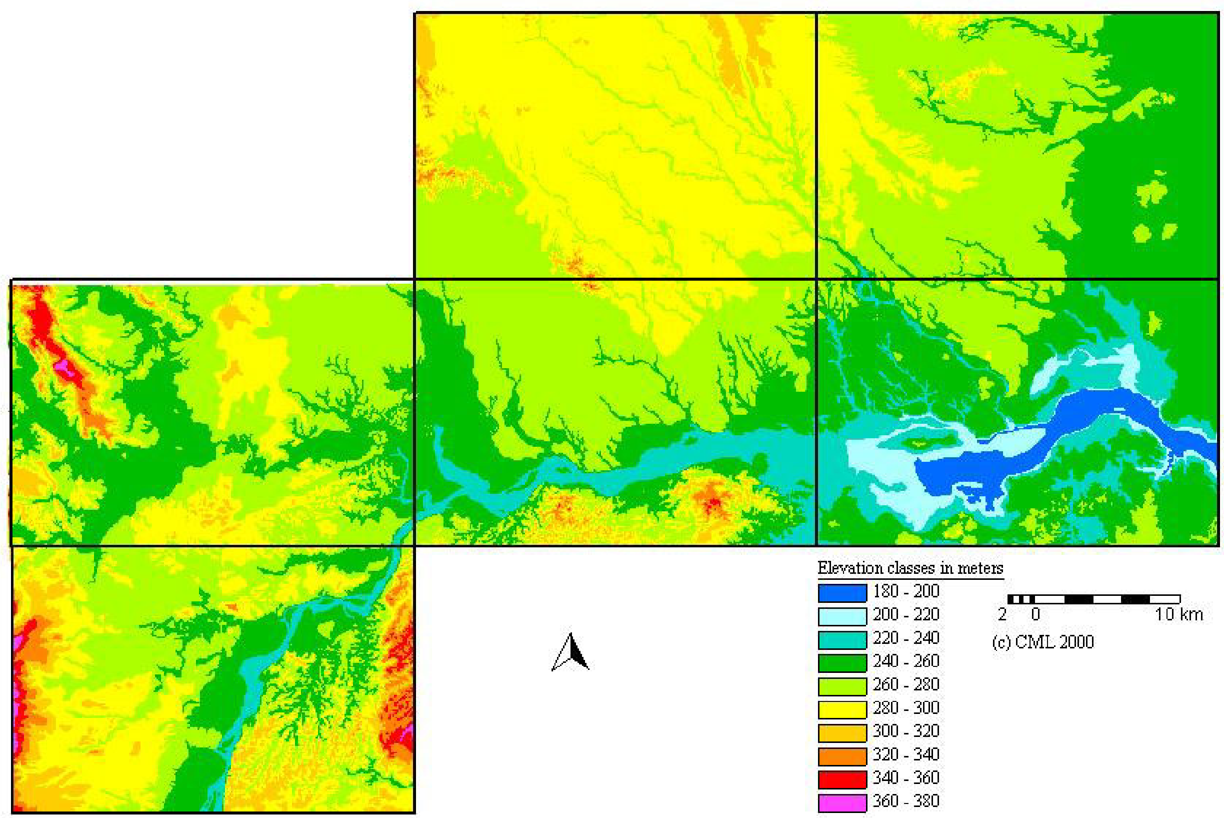
Figure 4- 3 Soil map of the study area



Source: adapted from soil map, MAG/PRONAR

Class	*Larger landscape*	*Landscape*	*Physiography*	*Sedimentary rock*	*Dominant soils*	*Soil characteristics*	*Recommendations for use and management*
F1	Plateaus and derived relief; lower mountains	Deeply dissected plateau	Undulating, hilly or ruptured surface on horizontal foundation	Deep layers of green or grey sandstone with minerals of volcanic origin, predominantly loamy sediments with loose appearance, and grey or yellow colour	Oxic DYSTROPEPTS (brown) tending to HAPLORTOX	Very deep and loamy soils with very low fertility and highly toxic aluminium (brown) Loamy soils with low fertility and toxic aluminium (red)	Integrated agriculture/forestry/ pasture or controlled forestry, protection, ecological and wildlife reserves
H1	Hilly Amazon basin	Relief dissected by hills	Variable slopes below 50%	Layers of loam and eroded sandstone	Oxic or typic DYSTROPEPTS (red)		
H2		Hills with generally swampy intervals	Rounded hills		The same red soils with TROPAQUEPTS in the swamps		
K1	Flat and/or swampy low Amazon basin	Various average and high level plains	Undulating or flat plains, generally with good drainage	Waste material of volcanic origin, shaped with rounded edges, sand, mud	Typic DYSTRANDEPTS or DYSTROPEPTS (brown)	Rather deep soils, loose loamy, with high fertility according to degree of evolution	All cultures with adequate handling
K2		Various low level plains	Undulating or flat plains, with variable drainage		VITRANDEPTS DYSTRANDEPTS Aquic DYSTROPEPTS	Various soils generally very fertile, but with variable drainage	Subsistence crops in viable parts
K3		Very low plains and swampy terraces	Flat plains with generally bad drainage	Deposits	TROPAQUEPTS	Soils generally muddy-loamy, fertile but susceptible to flooding	Limited use because of lack of drainage
K		Complex of terraces of different levels	Flat and saddled area, including unstable river bench	Deposits	Complex of all units "K"		
L		Lake					
R		River					

Figure 4- 4 Elevation map of the study area



Source: based on topographic maps, IGM/DMA

4.2. Estimating model parameters from the data

In this section we will explain the methodology of the parameter estimation from the sources described in section 4.1. The results of the parameter estimation must be interpreted as intermediate results. The purpose was to test the model for its use in explaining current land use for a selection of households in the area. As a first step in the validation process this should result in recommendations for eventual adaptations in the model structure as well as improvements in parameter estimations[29]. As a first step we have chosen to estimate parameters on a selection of 28 households, inhabitants of three sectors in the study area.

This section consists of two sections; in section 4.2.1 we will discuss the methodology of the parameter estimation[30] for the reduced household submodel applied to the Ecuadorian study area, while section 4.2.2. discusses parameter estimation for the more elaborate "core model". In order to calculate the reduced model we need estimates for the following parameters (see also section 3.4 in the previous chapter):

HOUSEHOLD AND FARM CHARACTERISTICS

astart(h,i,q)	the starting value for land allocation
capreq(h,i,q)	capital requirements
capstart(h)	starting capital
consreq(h)	consumption requirement
famlab(h)	family labour available
labreq(h,i,q)	labour requirement

INSTITUTIONAL ENVIRONMENT AND TECHNOLOGY

price(h,i)	product price
wage	wage rate for labour
workdays	number of workdays per year

NATURAL RESOURCE BASE

yield(h,i,q)	yield per hectare

Bu way of preperation, first we will discuss the classification of the indices *h* (for households), *i* (for activities) and *q* (land qualities).

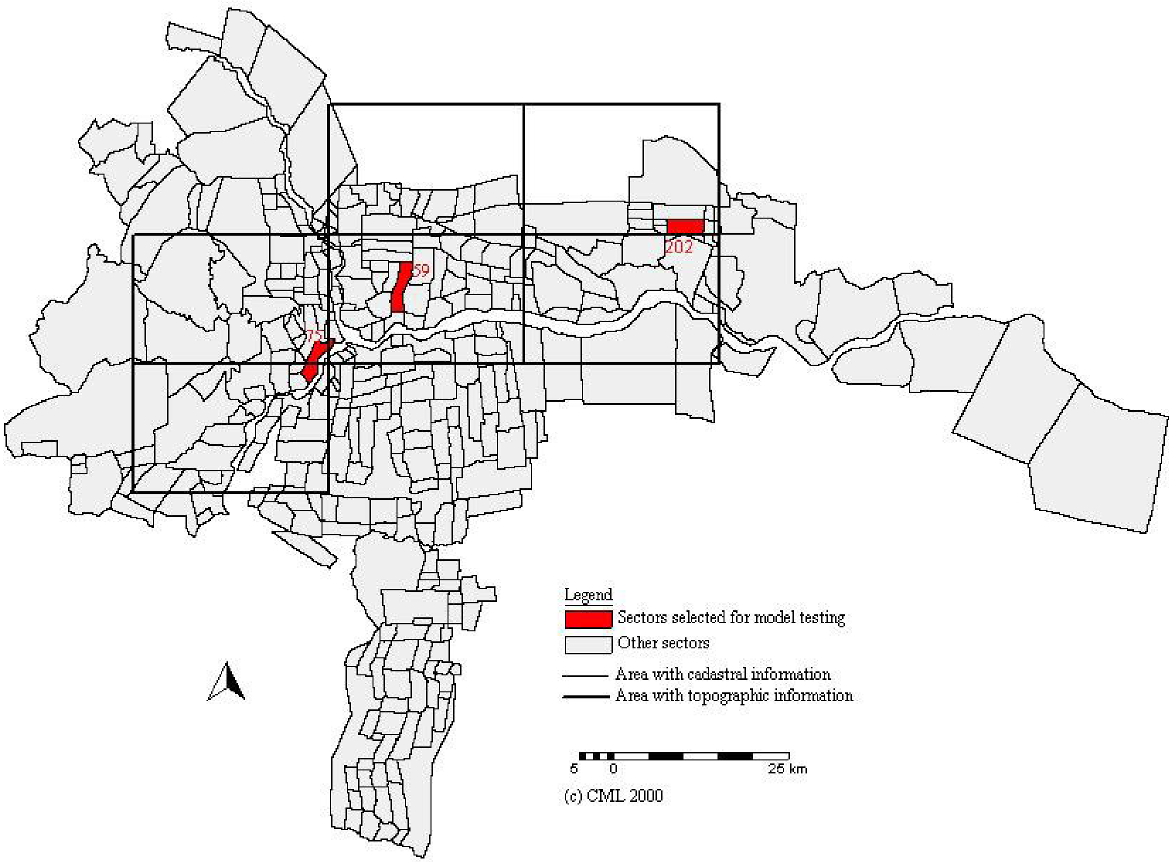
Index h (households)

Initially, three "representative households" have been constructed to test the model. The three "representative households" are called *h59*, *h75*, and *h202*. Their locations are illustrated in Figure 4.5. The numbers represent the cadastral numbers of the sectors. The CPC survey team interviewed 13, 8,

[29] The results presented here are preliminary because they are part of an ongoing process. The validation process should be and will be continued if a follow-up will be given to this research. Considerable attention has been given in this chapter to the directions in which model validation should be proceeded.

[30] The actual results of the calculations are presented in section 4.3.

Figure 4- 5 Three selected cooperatives in the study area



and 7 households respectively for these sectors, and our "representative households" have the mean characteristics of these. For each sector, the parameters are calculated as the average over the survey households in the sector.

The three sectors were chosen from a total of 17 sectors (see Figure 4-5). We have selected these sectors mainly because together they cover many aspects of the diversity which characterizes the population of the whole study area. Sector *h75* consists of 63% natives, *h59* and *h202* consist of 100% migrants. The years of foundation are 1975 (*h59)*, 1970 (*h75)*, and 1983 (*h202*).

Recommendations for validation[31]:
In future model testing the set of households should be extended to the whole set of 17 sectors for which the CPC survey collected data. So far we assumed that for all households the land use strategies can be described by the same model. For example, the factor ethnicity was assumed to play a negligible role in decision-making. This assumption can be tested by separating the indigenous households from the colonist households and testing for differences in crop choice and other land use practices[32].

Index i (activities):

The following ten categories were chosen on the basis of the CPC survey data (see also section 3-3):
Coffee, cacao, plantain, corn, rice, manioc, other crops, animals, fallow, primary forest, secondary forest, off-farm labour.
The selection was based on the available data. The CPC survey also contains information on cultivation of african palm, fruits, and vegetables. These crops plaid a minor role in our selection of households, therefore they were grouped together as "other crops".

Recommendations for validation:

The group of activities can be further extended to include the production of african palm, fruits, vegetables, poultry, small-ruminants, and wood. The CPC survey contains data on these products, but they were left out in the reduced version of the model. African palm is a different category, because it is a commercial crop grown only by 5 households in the CPC survey (Pichón, 1993). These are typically highly specialized farms, unlike the farm household represented by our household submodel. Therefore, we have chosen to leave production of african palm out of consideration. Eventually, the farms growing african palms may be considered as a separate primary actor.

Table 4-3 shows the importance of each crop for the households in the CPC survey. The perennial crops (coffee, cacao, african palm and fruits) are grown as cash crops for the market. Subsistence crops include plantains (grown by 85% of the surveyed households), manioc (59%), corn (49%), and rice (16%). These crops are not "pure" subsistence crops, some households produce them also as a source of income. Vegetables are mostly grown for the market. In our model application we have considered only one subsistence crop: manioc. The three "representative households" sold most of their production of corn, plantain and rice. Therefore we shared these annual crops under the cash crops for this particular case.

[31] Here we will point out the directions in which further parameter estimation should develop in order to improve robustness of the model.
[32] Other explanatory factors are likely to correlate with ethnicity, for example the distance from the market, the year of establishment. Statistical tests should be chosen that are able to account for these correlations.

Data analysis pointed out that timber sales take place on only 12% of the farms and mainly on the farms located next to a road (Pichón, 1993). Often the revenues do not outweigh the costs of hauling the logs to the roadside.

Table 4-3 Farming practices in the study area

Perennial/annual	**Crop**	**Mean proportion of cropped area**	***% of farms***
Perennial crops	African Palm	0.01	1%
	Coffee	0.70	95%
	Cacao	0.03	36%
	Fruits	0.02	52%
Annual food crops	Corn	0.10	49%
	Manioc	0.04	59%
	Others	0.01	4%
	Plantains	0.11	85%
	Rice	0.01	16%
	Vegetables	0.01	1%

Source: Adapted from Table 6.5 in Pichón (1993)

Index q (land quality):

We distinguished the following three land quality classes: *alluvial, black,* and *red*. The categories are derived from the CPC survey. For the three households used for this parameter estimation, the categories match very roughly with the soil map (see Table 4-4). Slight differences are not surprising, because the CPC survey measures farmers' perceptions of soil quality and the soil map has a scale which is actually too rough to distinguish soil differences on farm level.

Table 4-4 Soil types for selected sectors

Sector	***Soil type derived from CPC survey***	**Soil type derived from GIS**
h59	38%black 25% alluvial 25% sandy 13%black/alluvial	20% K1 (black) 80% K2 (black)
h75	100% black	80% K (black/alluvial) 15% K3 (black/alluvial) 5% H1 (red)
h202	71% black 14% sandy 14% black/red	45% K1 (black) 40% K2 (black) 15% K3 (black/alluvial)

Recommendations for validation:

For other sectors, the soil classes in the CPC survey may or may not coincide with the soil categories in the soil map for reasons explained earlier. A correlation analysis should point out to what extent the classifications coincide. Reliable soil data are not available at present, neither from the CPC household

survey nor from the soil map. More reliable information on the soil quality could be obtained through chemical analysis of soil samples.
The land quality index should be extended to include hilliness, because it is an other important determining factor for farming practices. This issue will be further explored in section 4.2.2.

4.2.1. Estimating model parameters in the Ecuadorian farm household submodel

HOUSEHOLD AND FARM CHARACTERISTICS

***astart*(*h*,*i*,*q*), starting value for land**

The starting values for current land use have been derived from the CPC survey. The survey contains data on the number of hectares under coffee, corn, manioc, etc (see variable CCOFFEE etc. in Table A4-1, appendix 4-I). The CPC survey does not make a distinction for land of different quality. For households with only one soil class there is no classification problem. Otherwise, we use the following rules of thumb for the ex-post classification[33]:

a) If the cultivated area is smaller than the acreage of black soil, then the whole cultivated area is allocated as "black soil".
b) If there is no black soil, then food, perennials and pasture are divided according to the following rules:
   b1) food is cultivated on alluvial soils (red only if no alluvial soils exist)
   b2) perennials are cultivated on red soils (alluvial only if no red soils exist)
   b3) livestock (pasture) is on red soils (alluvial only if no red soils exist)
c) If the cultivated area is larger than the acreage of black soil, then the "surplus" of food, perennials and pasture is allocated equally and according to the rules b1, b2, and b3.

For the three sectors used in our model application there was sufficient black soil, so we applied rule a) that all crops and pasture were on black soil.

We were able to cross-check the accuracy of the land use data under crops with the total hectares in crops through the CPC variable CTOTHEC. It appeared that the sum of the data on land use for individual crops was systematically higher than the values for CTOTHEC. This bias is only partly explained by the hectares interplanted (i.e. mixed cropping). We assumed that the farmers estimate for the total hectares in crops is more accurate than the hectares under individual crops, and that the

---

[33] Based on Tobit regression results in Pichón (1993, table 7-3 on page 229) explaining land use as a function of soil type, slope and time as the dependent variables:

Dependent variables Share of farm allocated to:				
	Perennials	Food	Pasture	Forest
Black soil	3.57** (1.71)	2.34** (1.13	2.47 (2.41)	-8.24*** (2.79)
Red soil	2.80 (2.11)	-0.12 (1.41)	1.53 (3.01)	-4.08 (3.45)
Hilly	-0.23 (1.58)	-2.85*** (1.05)	-6.59*** (2.79)	7.24*** (2.60)
Time	0.18 (0.15)	0.13 (0.09)	0.81*** (0.21)	-0.53** (0.24)

farmers apparently have a tendency to overestimate the current land use. We adjusted the averages for each crop proportionally in order to match the data with the totals as given by the variable CTOTHEC.

Recommendations for validation:
Regression analysis should point out whether the initial land use, in particular the size of the area converted from forest to crops, pasture, etc., can be explained by factors such as household composition, year of establishment, and land quality. Regression results from Pichon (1993) point out that the family size and the duration of colonization have a negative effect (5% significance level) on the share of farm land allocated to forest. The distance to the nearest market has a positive effect (1% significance level) on the forested area (see Pichón, Table 7.3). With respect to land quality, the regression analysis done by Pichón points out that hilly land is preferably kept under forest, and black soils are preferably converted for growing perennials or food.

***capreq(h,i,q)*, capital requirements**

Capital requirements for inputs in the production process are derived from the CPC survey. The CPC survey assembled the amounts of money spent on seeds, pesticides, medicines, vaccinations for livestock and mineral salts (see Table A4-1 in Appendix 4-I). The expenses are not subdivided by product. The expenses are highly variable among the households: one of our three "representative households" spent 0.16 $/ha, where the other two spent approximately 12 $/ha on crop production. The use of inputs such as fertilizer and pesticides is not widespread in the North-East of Ecuador. Seeds and organic fertilizer can be produced on-farm as well. One would expect that lower use of capital results in lower yields per hectare, but this is not consistent with the data for our three households. The one with low expenses has similar yields and financial revenues per hectare[34].

We calculated the input costs per ha cultivated and per hectare pasture by dividing total costs of inputs by the total of cultivated land:

For *i* = *coffee, cacao, plantain, corn, rice, manioc, other crops*
$capreq(h,i,q) = \{\text{MPSEED(h)} + \text{MPPEST(h)}\} / \sum_{i,q} astart(h,i,q)$

For *i* = *fallow*
$capreq(h,i,q) = 0$

For *i* = *pasture*
$capreq(h,i,q) = \{\text{MPMED(h)} + \text{MPMINE(h)}\} / \sum_{q} astart(h,i{=}pasture,q)$

Where (see also Appendix 4-I) :

MPSEED	CPC variable for money spent on seeds
MPPEST	CPC variable for money spent on pesticides etc.
MPMED	CPC variable for money spent on medicine and vaccine for livestock
MPMINE	CPC variable for money spent on mineral salt for livestock

[34] Of course, the number of sample households is too small to conclude that "capital does not matter". The importance of money spent on inputs could be evaluated through farming systems studies. More literature search is needed if this factor should be fully taken into account. For now we will use the data from the CPC survey and take any inaccuracies for granted.

Recommendations for validation:
For further validation and model development the capital requirements should be estimated from the whole CPC survey and compared with other secondary sources. For further model development the use of capital inputs should be estimated as a function of influential factors, such as access to credits, the time of settlement of the household, or the state of wealth.

***capstart(h),* starting capital**

The starting capital of the households was not known, but we estimated an indicator for the household wealth from the value of the livestock[35]. The CPC variables MANIMAL1-7, MSOLD1-7 and MPRICE1-7 were used for this purpose:

$$capstart(h) = \sum_i \mathrm{MANIMAL}i * \mathrm{MPRICE}i \,/\, \mathrm{MSOLD}i$$

Where:

MANIMAL	CPC variable for number of animals
MPRICE	CPC variable for revenues from sales of animals
MSOLD	CPC variable for animals sold

***consreq(h)*, consumption requirements**

The consumption requirement depends on the size of the household and on the diet. The only food product considered in our model example is manioc. Further validation of the model should include adding more food products to cover the basic consumptive needs. Also the age composition of the household should be taken into account[36]. Because of the current lack of data on consumption, we have arbitrarily fixed the consumptive need for manioc at 10 units of 50 kg ("quintales") per person per year. The parameter is now calculated as:

$consreq(h) = size(h)*10$

Where size(h) = the household size, estimated from CPC variable for household size HNUMBER (see appendix 4-I).

[35] Livestock includes cattle, horses, mules, pigs, goats, poultry and "other animals".

[36] The CPC variable HAGE can be used to calculate the percentage of youngsters in the household. No data are available on consumption. Theoretically, consumption of home produce could be calculated from the data as production minus sales. However, this would probably result in very inaccurate estimates as the sales data are often lacking. An additional literature search for consumption data will probably have better results.
We will briefly summarize the data on food sales for the three "representative households". The marketed share of total production for each crop has been calculated from the CPC data. The sales of corn are 17%, 99%, and 100% of total production for *h59*, *h75*, and *h202* respectively. The sales percentages for manioc were n.a. (not available/no data), 100% and n.a. For plantain: n.a., 55%, n.a. For rice: n.a., 83%, n.a. It can be seen that not all of the corn, plantain and rice is sold; part of it is probably reserved for family consumption.

Recommendations for validation:
First of all the number of consumption products needs to be extended for a better representation of the diet. Subsistence crops that are not yet included in the model are corn, plantains, and rice. The parameter estimate for consumption requirement should be based on information on the nutritional composition of these products, as well as consumption habits in the study area. Additional scrutiny of secondary sources may provide this information[37].

***famlab(h),* family labour**

The available family labour is estimated as the number of persons in the household of 16 years old and above. We made no distinction between male and female labour. A correction factor could be applied, but this requires additional data on differences in productivity between male and female labour. For this example we assume that differences are negligible and that all tasks can be done by men as well as women. Neither do we distinguish between skilled (or experienced) and unskilled labour.

famlab(h) = number of persons with HAGE(h) ≥ 16.

Where:
HAGE CPC variable for age.

Recommendations for validation:
In the validation process available data on labour productivity should be evaluated and eventually labour of different kinds should be differentiated.

***labreq(h,i,q),* labour requirements**

The labour requirements are derived from a simulation study by Estrada et al. (1988, Table 34, page 72). For each product (*i*) and soil type (*q*) they estimated the labour requirement per hectare on an average parcel. The study includes the products coffee, cacao, plantain, corn, manioc and pasture. We had no data available for rice, therefore we assumed (arbitrarily) that the labour requirements for rice are similar to those for corn.

Recommendations for validation:
If possible, the estimates for labour requirements should be compared with information from other farming systems studies (e.g. Luzuriaga, 1986; Chalá, GTZ/PROFOGAN, 1993).

INSTITUTIONAL ENVIRONMENT AND TECHNOLOGY

***price(h,i),* product price**

The price received by the farmer depends on the price on the nearest market and on the distance between the farm and the market. Many farmers sell their products by the roadside, others bring their produce to the nearest town. The CPC survey contained a question on where the produce was sold: by the roadside, on the market or directly to the consumer. Two out of our three “representative

[37] The World Health Organization (WHO) and the Food and Agricultural Organization (FAO) regularly bring out studies on food and nutrition for specific regions.

households" sold almost exclusively on the market, the third one (*h59*) sold both on the market and on the roadside. The prices collected by the CPC survey are similar for our three households, except for cacao which is sold by *h75* at a price 25% below the price received by *h59*[38]. In our model test we have chosen to use the prices reported by the three households as our price parameters. In a few cases there was no price information available because no sales occurred for these households. In those cases we took the lowest price estimate from the other households. This was done for manioc, plantain, cacao and rice:

price(h,cacao) = *PACACAO(*h*)*

price(h,coffee) = *PACOFF(*h*)*

price(h,corn) = *PACORN(*h*)*

price(h,manioc) = *PAYUCCA(*h*)*

price(h,plantain) = *PABANA(*h*)*

price(h,rice) = *PARICE(*h*)*

Where:
PACACAO, etc. the CPC variable for the price of cacao, etc (see Appendix 4-I, Table A4-1).

For livestock we derived the price per animal from the total value of livestock sold (CPC variables MPRICE) divided by the number of animals sold:

price(h,animals) = $\sum_{1\text{-}7}$MPRICE*i* / $\sum_{1\text{-}7}$MSOLD*i*

Where (see also appendix 4-I):
MPRICE*i* the CPC variable for revenues from sales of cattle, horses, etc.
MSOLD*i* the CPC variable for the number of cattle, horses etc. sold.

Recommendations for validation:
The validation contains two aspects: estimates of the market prices and of transportation costs. The price data obtained from the CPC survey can be compared with agricultural statistics such as those collected by the Ministry of Agriculture. Estimations of transportation costs can be derived partly from survey data, and completed with a GIS based analysis of distances. The methodology of such a distance analysis is described in section 4.2.2. below. The estimates for the price of animals can be refined by distinguishing different types: cattle, horses, mules, pigs, goats and poultry.

***wage*, labour wage rate**

The CPC survey contains data both on the use of day labourers on the farm (see variables

[38] The price differences can be caused, for example, by quality differences, differences in transaction cost or by measurement errors. In the case of our three households we do not have sufficient information to explain the price difference for cacao.

MDAYLAB, MDAYSAL in Table A4-1) and on salaries earned with off-farm work (variable WSALARY and WMONTHS). Also, the CPC survey acquired data on whether the household had problems finding day labourers (variable MHADPROB). These data were combined to make an "educated guess" on the current wage rate. This was not a straightforward issue, because:

- The range of salaries is too large to justify one single wage rate for labour. The range was \$27 - \$50 for the monthly salary of day labourers and \$24-\$1111 for monthly off-farm salaries[39]. The average salaries were \$30 and \$194 respectively. Apparently labour has many different faces in the study area which need to be distinguished.
- Seven out of fourteen households reported that they had problems finding day labourers. This means that the assumption that a well-functioning labour market requires further research.

We adopted a wage rate of 30 \$ per month, based on the average salary of a day labourer.

The wages were calculated from the CPC data as follows:
Monthly wage for day labourer = $\sum_h$MDAYSAL/$\sum_h$MDAYLAB
Monthly wage off-farm labour = $\sum_h$(WSALARY*WMONTHS)/ $\sum_h$WMONTHS

Where:

MDAYSAL	the CPC variable for salaries paid to day labourers, per year
MDAYLAB	the CPC variable for number of day labourers hired (in man months)
WSALARY	the CPC variable for monthly salaries earned off-farm
WMONTHS	the CPC variable for number of months worked off-farm

Recommendations for validation:
Validation of the wage rate estimate should take into account the relation between wages for the most common labour types (e.g. skilled/unskilled/male/female labour). The model may have to be adapted to make a distinction at least between wages for day labourers and for off-farm work. The data for our three selected sectors indicate that wages for day labourers have a much smaller range than off-farm wages. Off-farm wages may need to be related to education, for example, while wages for day labourers can be estimated as the average wage. Also the assumption of a perfect labour market need to be scrutinized, both from the CPC survey as from additional literature search. Another factor not taken into account so far is that many farm households exchange labour with neighbours and relatives. This type of agreements possibly reduces the demand for wage labour. The CPC survey contains data on these types of labour exchange as well, these could be further explored for use in the model.

workdays, **number of workdays per year**

The number of workdays per year for family labour was fixed at 260, based on the assumption that people in the study area work 5 days a week and 52 weeks in a year.

NATURAL RESOURCE BASE

yield(h,i,q)**, yield per hectare**

We derived yield data in units per hectare form the CPC survey as follows:

[39] The amounts in sucres were converted to US dollars at the rate of 900:1 (Compendio Estadistico)

yield(h,cacao,black) = PYCACAO(*h*)
yield(h,coffee,black) = PYCOFF(*h*)
yield(h,corn,black) = PYCORN(*h*)
yield(h,manioc,black) = PYYUCCA(*h*)
yield(h,plantain,black) = PYBANA(*h*)
yield(h,rice,black) = PYRICE(*h*)
yield(h,animals,black) = ($\sum_i$MSOLD*i*) /*astart(h,animals,black)*

Where

PYCACAO etc.	the CPC variable for yield per hectare for cacao, etc (in units: bunches for plantain and quintales of 50 kilograms for the other products)
MSOLD*i*	the CPC variable for the number of animals sold.

Note that we assumed earlier that the soil type currently under cultivation is "black" for all households. Not all yield data were available for all the households. If a certain product was not produced in all sectors, then the lowest value of the other two was taken as a proxy.

Recommendations for validation:
Validation of the yield estimates should be directed towards testing and establishing relationships between the yields per hectare and the use of inputs (capital, labour and natural resources).

4.2.2. Recommendations for additional parameter estimation

As we mentioned before, the Ecuadorian household model is a reduced version of the core model. Some of the parameters in the core model have not been used in this version. This section discusses estimation methods and preliminary results for these other parameters. It concerns the following parameters (see also appendix 3-I):

HOUSEHOLD AND FARM CHARACTERISTICS

consexp(h,t)	non-food consumer expenses for household *h*, as a fraction of last year's net revenues.

INSTITUTIONAL ENVIRONMENT AND TECHNOLOGY

credmax(h,t)	the maximum amount of credit that the household is allowed to borrow, depending on collateral (property rights on farm land)
intrate(h,t)	the interest rate for credit
landprice(h,i,q,t)	the price the farm household pays/receives for buying/selling farmland depending on location, land quality, current land use.
logrev(t)	revenues from off-farm logging, per working day, year *t*
maxexpand(h,i,q,t)	the quantity of land available to the farm household for purchase or to be occupied in year *t*.
proright(h,t)	an indicator for the property or usufructuary rights of the household *h*.
trancost(h,t)	transportation costs between farm *h* and the nearest market, depending on distance, means of transport and quality of infrastructure

NATURAL RESOURCE BASE

(index) q *classification of land quality in soil type and slope-classes*

HOUSEHOLD AND FARM CHARACTERISTICS

***consexp(h,t),* non-food consumer expenses**
The non-food consumer expenses have to be estimated indirectly, because no data are available on non-food expenses. This would require panel data of households who keep administrations of their expenditures. The indirect way to estimate non-food consumer expenses (as a share of net revenues) is to calculate the following[40]:

consexp(h,t) = (net revenues – investments – savings - food expenses)/net revenues
Where:
net revenues = gross revenues – production expenditures

The gross revenues can be calculated as the sum of the following variables from the CPC survey (see Appendix 4-I, Table 4A-1 for a description of the variables):
PVCOFF, PVCORN, PVYUCCA, PVBANA, PVCACAO, PVAFPALM, PVFRUIT, PVVEGET, PVOTHER, PVRICE, NANNINC, WSALARY*WMONTHS, MPRICE, MPRICE1, MPRICE2, MPRICE3, MPRICE4, MPRICE5, MPRICE6, MPRICE5, MAMT1, MAMT2, MAMT3 (if the loan was received in the last year).

Production expenditures can be calculated as the sum of all expenditures for production purposes (see Appendix 4-I, Table 4A-1 for a description of the variables):
MPCHEM, MPLIME, MPSEED, MPPEST, MPMED, MPMINE, MPOTHER, MRELSAL, MNOTSAL, MNEISAL, MDAYSAL, MCONSAL, MOTHSAL, *intrate**TOTDEBT.

Monthly food expenses in the CPC survey are assembled in the variable FFOODCOS. Annual food expenses are calculated as:

Food expenses = 12*FFOODCOS

Where:
FFOODCOS The CPC variable for the money spent on food.

No data are available on amounts of savings and investments. Additional data need to be searched to obtain suitable estimates.

[40] Note that this can at best be an approximation of the actual expenses because the data on income and expenses in the CPC survey was not meant for tracing the complete household budget. If possible, additional data need to be searched for cross-reference.

INSTITUTIONAL ENVIRONMENT AND TECHNOLOGY

credmax(h,t) **, credit maximum; *intrate(h,t),* interest rate**

From the CPC survey it appeared that 18% of the households had ever received a credit (Pichón, 1993). Being able to provide collateral is a prerequisite for receiving credits. Therefore, only households who own land or other fixed assets can apply for credits. Pichón also mentions the high transaction costs involved in receiving credits. The procedures are annoying and time-consuming (more than 5 months according to Pichón). Not only commercial banks give credits to farmers, also a number of NGOs offer subsidized credits to farmers and the governmental "Banco de Fomento" has a subsidized credit programme. However, it appears from our discussions with farmers during our field research that the procedures are similarly tedious for subsidized credits as for commercial credits. In the validation process the role of these subsidized programmes merits a more thorough investigation.
Concluding, we would recommend the following estimate:

credmax(h,t) = 0 if the household is not a land owner
"the average of the variable TOTDEBT in the CPC survey" for those households who do have land of their own.

Where:
TOTDEBT the CPC variable for the total amount of debt.

Furthermore, the costs of obtaining credits must be estimated (including the time involved) and explicitly included in the model as a "transaction cost". Interest rate estimates need to be corrected for inflation.

landprice(h,i,q,t), **the price of land**

The CPC survey contains a variable for the purchase price of the land :
MBOUGHT the amount paid for the farm, in sucres, to another settler[41]

The value of the land depends on many factors such as the percentage of cleared forest, the distance, or the soil quality. We suggest that a regression analysis is carried out, with the following explanatory variables from the CPC survey:
MYEARAQ year of acquisition of the farm
MCERTIF type of title at purchase (certificate of possession/ right of title /nothing)
MLINE number of line on which the farm is located (distance from the road)
MPRIMARY, MSECOND, MWALKED, MBOAT, distance by primary road/by secondary road/ by foot/ by boat.
NSOIL, NTOPOG, soil type and slope.
The distances (or travelling time) between farm and market, as well as soil type and slope, can also be derived from our GIS data. The methods used for the GIS analysis are described later in this section.

[41] In order to compare prices in different years the amounts paid need to be corrected for inflation.

logrev(t), **revenues from off-farm logging**

There is no indication that people are engaged in off-farm logging in the study area. On-farm logging does occur, though it is rarely profitable according to Pichón (1993). The revenues from logging can be estimated from the CPC variable:
MPRICE total revenues received for selling wood in the past year.

maxexpand(h,i,q,t), **land available for expansion**

Expansion by occupying new land is limited in the area. The maps indicate that the study area is not yet fully occupied, but the remaining pieces of land are typically unattractive because the land is of bad quality (hilly or toxic soils) and have very bad road access. Colonization may also take place illegally into areas with a protected (nature) status. Most likely, farmers will expand by buying land from other settlers. More research is needed to study the actual mobility of farmers within the region.

proright(h,t), **property right**
This parameter can easily be estimated from the CPC survey data through the variable:
MTENURE current form of tenure of the farm (certificate of possession/right of title/nothing)

trancost(h,t), **transportation costs between farm and market**

Additional research is needed to estimate costs of transport in the study area[42]. In the Ecuadorian Amazon, the usual means of transport are the boat (usually canoe) for river transport, the pickup or bus for paved and unpaved roads, the donkey and feet for tracks. Pichón considers only two types of transport, mechanised and non-mechanised. Mechanised travel is the type used for primary and secondary roads. Non-mechanised (by foot or canoe) is used for crossing the distance from the household plot to the nearest primary or secondary road. A rough indication of travelling costs can be obtained from our own field observations. For small freights, the bus is the usual means of transport. On top of the bus is ample room for bags and sacs. On a primary/secondary road (e.g. between Coca and La Joya de los Sachas) it takes approximately 4 hours to cross an 80-kilometer distance, including breaks and stops to let passengers out. The fee to be paid was 9000 Sucres in November 1996 (1 US $ ≈ 3300 Sucres). Hence the fare for one person is approximately 2.7 US $ for 80 kms, hence 3.4 US $ per 100 km. The fare includes luggage, though it will probably cost extra to bring larger freights.
Also, we have to account for the time the farmer spends to transport his freight. For distances over 50 kms we can easily assume that it costs one day to travel to the market and back again. The best measure to translate this into costs for the farmer is to take the opportunity costs of labour, or the wage of an unskilled worker. If we use our wage estimate of 30 US dollars per month then one working day (assuming 22 working days/month) costs approximately 1.4 US dollars.

[42] Unfortunately, we did not find empirical studies on transportation costs for agricultural products. A comparative study on transportation costs would be quite valuable also for comparisons in an international context. Transportation costs are known to be crucial factors in decisions on agricultural production, yet the subject has not been given much attention in research so far.

Access to markets, distance to markets

The distance to markets is an important parameter for many farm household decisions. Therefore we have given it considerable attention even though the actual cost analysis is incomplete without having reliable estimates for the transportation costs per kilometre (see above). However, the basis of the cost calculation is a distance analysis based on our GIS framework. The focus of our distance analysis is on *travelling time*. We have chosen travelling time instead of distance in kilometres because it is not the actual distance that matters to the farmer but the time and effort required to travel and bring goods from the farm to the nearest market and vice versa.

The distance analysis finally results in travel time estimates for the study area (see Figure 4-7). This Figure was obtained after a stepwise spatial network analysis. The following steps were undertaken[43]:

A. Digitising the infrastructural network
B. Attach travel time to the network according to road quality
C. Link network data to the farm locations
D. Link network data to the major markets
E. Calculate travelling time between farms and markets
F. Determine nearest market for each farm

A. <u>Digitising the infrastructural network</u>

Spatial data on the infrastructural network were derived from the 1:50 000 topographical maps. No digital version was available, therefore we digitised the data ourselves using Carta Linx digitisation software[44]. The result is presented in Figure 4-2. Transportation takes place on roads and rivers. The following classes were distinguished:

- Main road
- Summer road
- Sandy road
- Rivers

[43] The spatial analysis is illustrated with flow charts in Appendix 4-II.
[44] The infrastructural network was digitised in vector-format. This was necessary in order to analyse distances and perform shortest path algorithms. Despite the difference in format, these vector data are fully integrated in the quadtree framework as well.

Figure 4- 6 Service areas of the markets

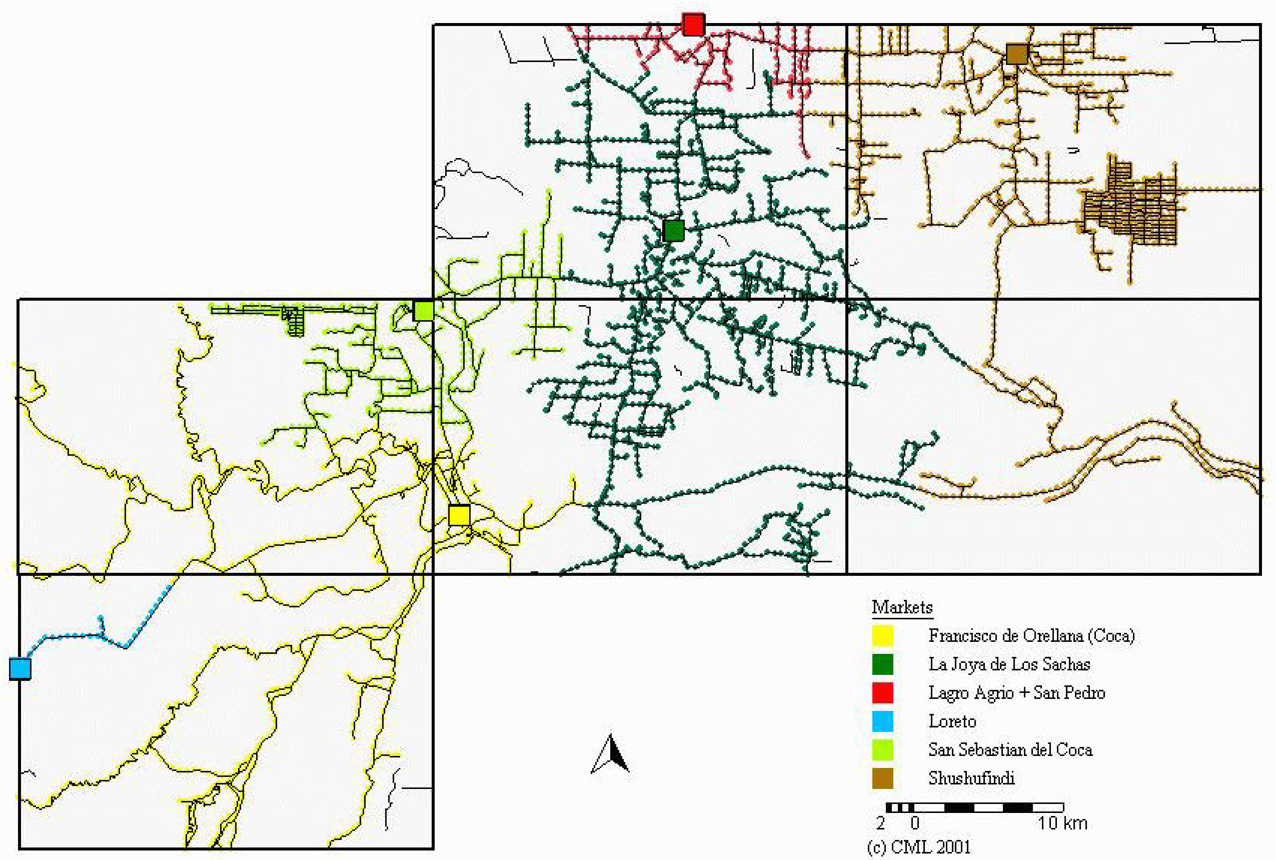
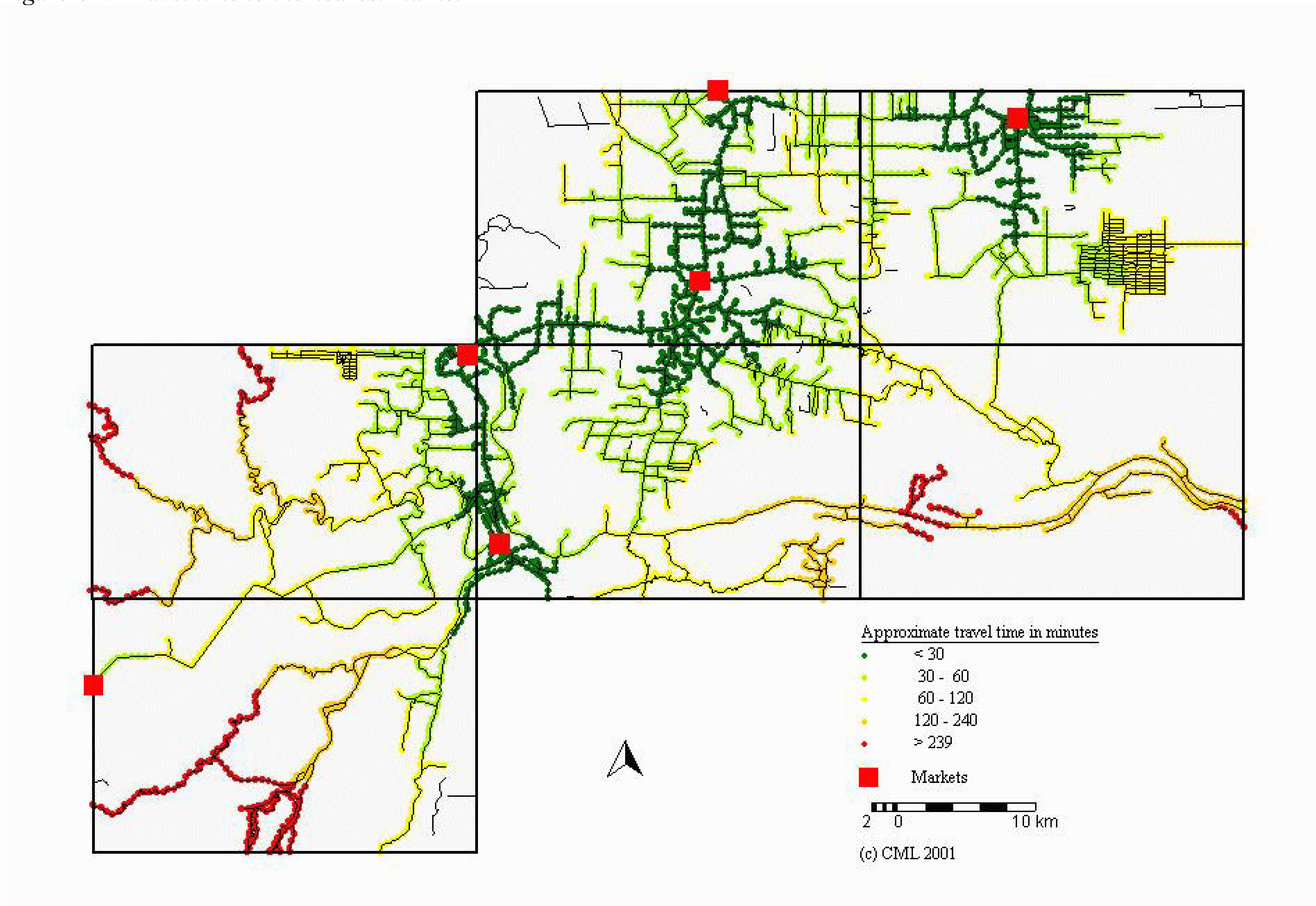


Figure 4- 7 Travel time to the nearest market



The four road classes represent different qualities in the transportation network. It is cheaper and easier to travel a main road than a summer road or sandy road. Summer roads are badly accessible during rainy periods. River transport is very common in the study area, for land that can not easily be reached by road[45].

The locations of larger towns/markets in the study area have also been digitised. The following tows are located within the study area:

- Francisco de Orellana (Coca)
- La Joya de los Sachas
- San Sebastian del Coca
- Shushufindi

Three towns are located outside the study area, but they have been added because they are important market outlets for people living in our study area. Markets in these towns function not only as local market outlets, but also as transit markets for transportation towards the Sierra and Coastal parts of Ecuador:

- Lago Agrio (north of the study area)
- San Pedro (north of the study area on the road to Lago Agrio)
- Loreto (south-west of the study area on the road to Tena)

B. Attach travel time to the network according to road quality

In order to distinguish between the different classes in the transportation network, we added travel time information to the basic network structure according to the following table:

Table 4-4 Travel times for 20 kilometres

Main road	1 hour
Summer road	1½ hour
Sandy road	2 hours
River	3 hours

C. Link network data to the farm locations

In order to integrate the transportation network with the spatial units for the farm locations, we imported the network structure into a SPANS (quadtree) framework. Next, the farm locations have been added as nodes in the network. For each farm the nearest location on the road (or river) was chosen. These point data were exported into Arc-View format in order to perform further network analysis.

[45] Three classes have been added to the topographic maps in order to integrate river connections into the network: the "middle of the river", the "cross-over in rivers" and the "river-land connections". In the topographic maps the rivers were represented by polygons, whereas we need representation by lines in order to connect the rivers with the remainder of the network.

D. Link network data to the major markets

We mentioned before that three important market towns are located outside the study area. Leaving them out of the analysis would be a serious omission, therefore we chose to situate these markets at the edges of our map while accounting for the distance between the edge of the map and the locations of these towns. On the topographical maps, the towns were represented by polygons. However, in order to perform the distance analysis, we needed a point representation for the towns. The conversion was made in SPANS by determining the centroid for the polygons and situating the point data at these centroids. The point data for the markets (see Figure 4-2) were finally converted into Arc View format for further analysis.

E. Calculate travelling time between farms and markets

We calculated the fastest connections between each farm and each market. For this purpose we wrote an AVENUE–script (see Appendix 4-III) which was subsequently solved with Arc View software. The result was a series of point data containing travel times from all roadside farms to all markets. We converted these attribute data back to SPANS format for further analysis.

F. Determine nearest market for each farm

By means of an SQL-query in a relational database (MS-Access) we determined the "nearest" market for each farm location based on travel time. Figure 4-6 presents the "service areas" for each market; the areas which are within closest reach of each of the markets. Figure 4-7 presents the travel time for each location in the infrastructural network. Note that most roadside locations are within 4 hours travel time to the nearest market.[46]

NATURAL RESOURCE BASE

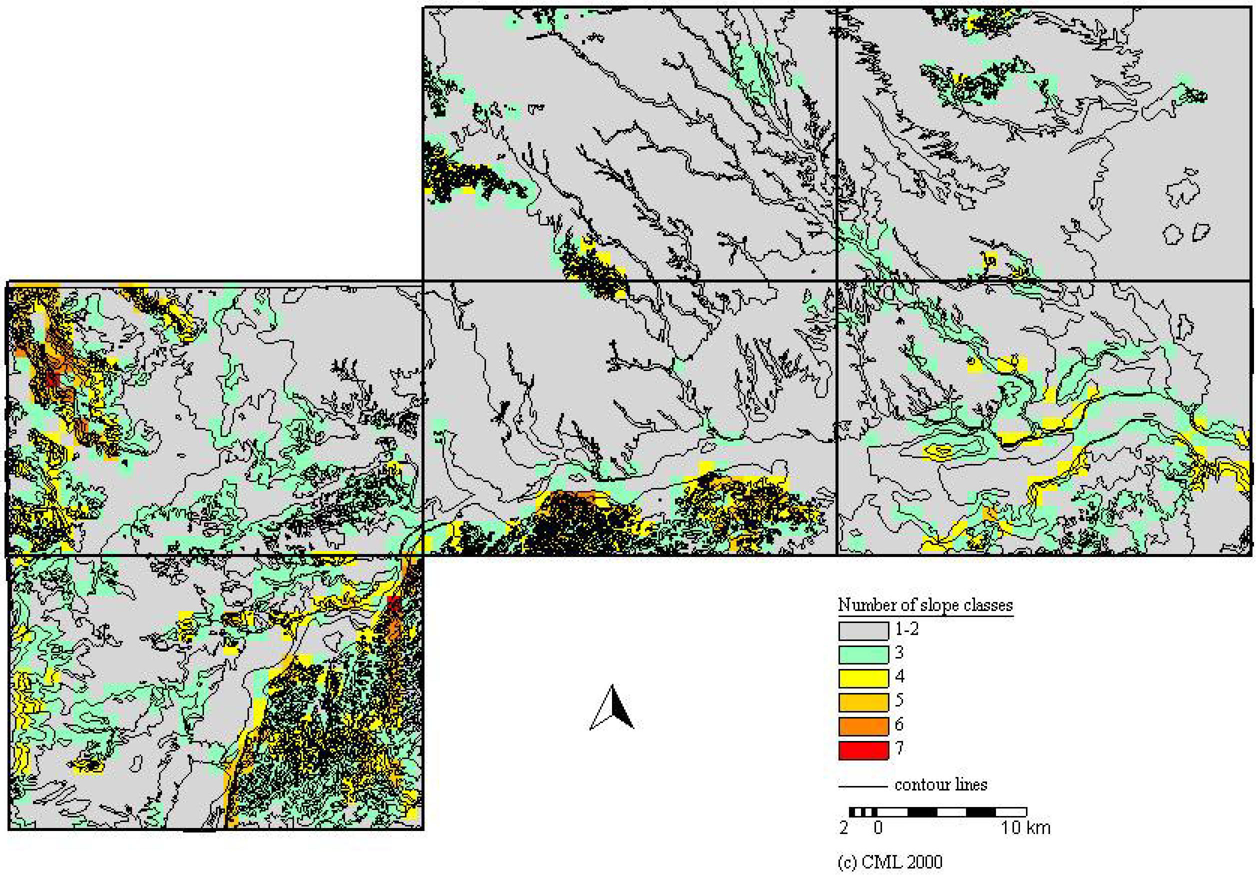
(Index) q, *land quality*

In the previous section we described how the soil classification was derived from the CPC household survey and how these might be linked to the GIS data (see Figure 4-3) for future model development and validation. In absence of detailed physical data, we will assume for the moment that farmers are quite able to distinguish between volcanic and red-hill soils, merely based on the difference in colour (black or red). Alluvial soils can also be distinguished, because they are situated in river beds.

So far, not much attention has been given to the slopes on the farm lands. In the following we will describe our classification method to distinguish slope indicators. The following steps were undertaken to derive the classification illustrated in Figure 4-8 (see also Appendix 4-III):

[46] Note that the walking distance from the farm to the nearest road side is not included in Figure 4-7.

Figure 4- 8 Slope indicators per square kilometre



A. Digitising the contour lines from the topographic maps, overlay with grid cell map

The maps are described in Appendix 4-I. We digitised the contour lines for every 20 meter as polygon data. An overlay was made with a second map containing grid cells of one square kilometre in order to link the data to our base format. This procedure was done in SPANS software using the unique conditions method. The result was a map containing contour classes for each grid cell.

B. Determine the slope-indicator for each grid cell

The attribute data of the contour map were exported to a database file (MS-Access – format). An SQL-query was performed in MS Access to determine the highest contour class for each grid cell. A second SQL query was performed to determine the lowest contour class for each grid cell. A third query calculated the difference between the highest and lowest contour class in each grid cell. The result is a slope indicator: the number of height classes in each grid cell (see Figure 4-3). The slope indicators were brought back to SPANS format for further analysis and map-making.

The slope indicator can be further classified as:
Flat land: changes in altitude between 0 and 60 meters per square kilometre (≤3 height classes)
Hilly: changes in altitude of 60 to 80 meters per square kilometre (4 height classes)
Steep: changes in altitude of more than 80 meters per square kilometre (≥ 5 height classes

As we can see in Figure 4-8, most of the area is "flat", "hilly" and "steep" lands are mostly situated in the area south of the Coca river. Figure 4-8 is very illustrative in understanding why the area south of the Coca has been colonized later than the area north of the river Coca. Hilly and steep lands are not very suitable for agriculture, and therefore these (forested) lands are less endangered by the colonization process. Nevertheless, the colonization process does proceed even in hilly lands for farmers who are less endowed to buy lands of better (farming) quality.

Further model validation should explicitly take these slope indicators into account. Also, statistical analysis should point out whether the GIS based classification described here can be linked with the (subjective) household survey data. The latter reflects the perception of the farmer with respect to the steepness of the land.

4.3. Results of parameter estimation

Before we will present the results of the parameter estimations (section 4.3.2) we will first describe some characteristics of the three sectors from which our "representative households" have been generated.

4.3.1. Characteristics of three selected sectors in the Ecuadorian Amazon

Table 4-5 Characteristics of three selected sectors

Characteristic:	**h59**	***h75***	***h202***
Name	Associacion 12 de febrero	Varios posesionados	Pre-cooperativa Unidos Venceremos
Number of households	50	30	22
Number of households in CPC survey	13	8	7
Nearest town/market	La Joya de los Sachas	Coca (San Francisco del Orellana)	Shushufindi
Distance calculated from: - *CPC survey* - *GIS analysis*	 18 km by road 15 km by road	 9.5 km by boat 5 km by boat	 38 km by road 24 km by road + 4 km by foot
Year of foundation of the cooperative	1975	1970	1983
Ethnicity	Natives 0% Migrants 100%	Natives 63% Migrants 17%	Natives 0% Migrants 100%
Average purchase price of a plot	\$1070	\$278	\$755
Average year of purchase of plots	1983	1990	1982
Average size of plot of households in survey	86.3 ha	54.9 ha	48.1 ha
Primary forest as a share of total farm land	34%	75%	76%
Average value of livestock on the farm	\$1355	\$1471	\$615
Average use of inputs	\$372	\$15	\$72

Sources: GIS analysis and CPC household survey

Sector h59

Sector *h59*, Associacion 12 de febrero, was founded in 1975 by migrant farmers. Many of the early inhabitants sold their land to newcomers. The average year of purchase of the present inhabitants is 1983; the average purchase price was 1070 US dollars. The sector has fairly good access to the nearest town; the travelling distance is approximately 15 kilometres by road. The value of livestock is relatively high: \$1355. It should be noted here that there is one large land- and cattle owner in the sample for sector 59 who dominates the sample in this respect. The share of primary forest on the household plots is low: 34%. This can be explained by the relatively long period of colonization and good access to roads and markets.

Sector h75

Sector *h75,* Varios Posesionados, has a large percentage (63%) of native inhabitants. The cooperative was founded in 1970. The average year of purchase of the present inhabitants is 1990; the average purchase price was 278 US dollars. The price for land is much lower than for the two other sectors. A possible reason for this difference is the access to the nearest town: the distance to the nearest market (Coca) is approximately 5 kilometres by river whereas the other two sectors can be reached by road. Livestock is important for sector *h75*, as it appears from the relatively high value of the livestock ($1471). This may be the result of long time accumulation of capital of the early settlers. The use of inputs is quite low ($15 per household per year), indicating that farming practices are more traditional than in the other sectors. Relatively much land is still under forest: 75%.

Sector h202

Sector h202, Pre-cooperativa Unidos Venceremos, was founded in 1983 and is the youngest of the three. The distance from the nearest town is 38 kilometres by road. respectively. The average plot was bought in 1982 for 755 US dollars. The value of livestock is estimated at $615, the lowest value of the three sectors. 76% of the land is still under forest which can be explained by the "late" colonization and the relatively large distance to the nearest market.

4.3.2. Values of the parameters

In section 4.2.1 we explained the estimation methods for the parameters used in the Ecuadorian farm household model. In this section we will briefly present the parameter estimates for the three selected sectors *h59*, *h75*, and *h202*:

HOUSEHOLD AND FARM CHARACTERISTICS

astart(h, i,black) *"starting value for land"*

	cacao	coffee	corn	manioc	plantain	rice	other
h59	0.7	4.9	0.5	0.1	0.9	0.3	0.6
h75	0.2	2.2	0.5	0.3	1.5	0.0	1.5
h202	0.0	2.0	0.5	0.2	0.7	0.2	0.3

astart(h, i,black) *"starting value for land"*

	fallow[47]	sec_for	animals
h59	0.3	8.1	40.6
h75	3.5	1.3	2.5
h202	0	0.7	6.9

astart(h,prim_for,q) *"starting value for land"*

	black	red	alluvial
h59	29.3	0	0
h75	18.7	0	22.7
h202	32.6	4.0	0

[47] This category was not in the CPC survey, but it was included here because the sum of all categories did not add up to the total hectares on farm.

capreq(h,i,q) *"capital requirements"*

i = coffee, cacao, plantain, corn, rice, manioc, other crops:

$capreq(h59,i,black) = 94/\sum_{i=coffee,cacao,plantain,corn,rice,manioc,other} astart(h59,i,black)$

$= 11.75$ US$/ha

$capreq(h75,i,black) = 1/\sum_{i=coffee,cacao,plantain,corn,rice,manioc,other} astart(h75,i,black)$

$= 0.16$ US$/ha

$capreq(h202,i,black) = 49/\sum_{i=coffee,cacao,plantain,corn,rice,manioc,other} astart(h202,i,black)$

$= 12.56$ US$/ha

capreq(h59,animals,black) = 278/ *astart(h59, animals, black)* = 6.85 US$/ha

capreq(h75,animals,black) = 14/ *astart(h75, animals, black)* = 5.60 US$/ha

capreq(h202, animals,black) = 23/ *astart(h202, animals, black)* = 3.28 US$/ha

capstart(h) *"starting capital"*

h59	7346
h75	235
h202	626

consreq(h) *"consumer requirements"*

*= cons*size(h)*

cons *"consumption requirements per person"*

= 10 units of 50 kgs of manioc

size(h) *"household size"*

h59	8.2
h75	6.8
h202	4.6

famlab(h) *"family labour"*

h59	3.4
h75	2.9
h202	2.6

labreq(h,i,q) *"labour requirements"*

	alluvial	black	red
coffee	71.0	74.9	58.7
cacao	37.5	36.1	34.2
plantain	12.2	13.4	14.6
corn	11.1	26	16.6
rice	11.1	26	16.6
manioc	12.2	13.4	14.6
animals	27.9	23.2	15.6

(for all *h)*

INSTITUTIONAL ENVIRONMENT AND TECHNOLOGY

price(h,i) *"product price"*

	cacao	coffee	corn	rice	manioc	animals
h59	27.8	2.6	3.3	8.9	2.8	91.6
h75	20.9	2.7	3.2	8.9	2.8	10.7
h202	20.9	2.8	3.3	8.9	2.8	70.4

wage *"labour wage rate"*

= *monthlywage*12/workdays*

monthlywage = 30 US $

wage = 1.38 US$/day

workdays *"number of workdays per year"*

260 days/year/male labour equivalent

NATURAL RESOURCE BASE

yield(h,i,black) *"yield per hectare"*

	cacao	coffee	corn	rice	manioc	plantain	animals
h59	4.25	37.7	18.7	10	20	566.7	6.82
h75	5	24.3	24.2	10	20	566.7	38.00
h202	4.25	20.4	15	10	20	566.7	1.28

4.3.3. Observations on parameter estimations

The parameters presented in the previous section will be used as input in the model discussed in chapter 3 (section 3.4.2). In the next chapter we will present the model results for the three "representative households". We will compare the land use decisions as calculated from this model with the average land use observed in the CPC survey.

The parameter estimations presented in the previous section are not necessarily the final results for these three sectors. In fact, it is likely that the first model simulation for the baseline scenario will give reason for further validation of the data estimates. In section 4.2.1 we have indicated how parameter estimates may be improved and which additional data could be useful for validation of the estimates.

Before going at length to search for the best possible parameter estimate, it can be worthwhile to perform a sensitivity analysis with the model first. By adjusting the parameter upwards and downwards by – say – 10 percent and recalculating the model, we can see from the model results how the model households respond to changes in this particular parameter. Obviously, the need for accurate estimates is higher for those parameters which have a higher impact on the decisions.

Appendix 4-I Description of data sources

A. HOUSEHOLD SURVEY CAROLINA POPULATION CENTER (CPC)

For an extensive description of the survey we refer to Pichón (1993). The survey was a part of a research on agricultural settlement, land use, and deforestation in the provinces of Napo and Sucumbíos in Ecuador. The study was carried out by a team of researchers from the Carolina Population Center (CPC) of the University of North Carolina at Chapel Hill. Publications of the research team include a.o. Pichón and Bilsborrow, 1991 and 1992, Pichón, 1993a, 1993b, 1996, 1997; Thapa et al, 1996. The survey contains data on 419 colonist households in 1990; 189 in the Coca district and 230 in the Aguarico district. A number of 159 survey households spread over 19 sectors are located in our study area (see Figure 4-1).
The survey was conducted by a team of 10 interviewers under supervision of Bilsborrow and Pichón of CPC. The result is an extensive and representative source of data on household behaviour. We are very grateful to Prof Bilsborrow who gave us permission to use the data for our modelling purposes. It provides us with a valuable source of information which allowed us to estimate the majority of the parameters for the household submodel. The following table presents a list of items that were useful in our analysis:

Table A4-1 Selected variable from the CPC survey

Name of variable in the CPC survey	Model parameter	Description
HOUSEHOLD CHARACTERISTICS		
FFOODCOS	Consumption expenditures	Amount of money spent on food in the past month
HAGE	*famlab*	Age for each household member
HNUMBER	*size(h)*	Household size
MANIMAL1-7	*capstart*	Number of animals the household has, presently
MAMT1	-	amount of first loan
MAMT2	-	amount of second loan
MAMT3	-	amount of third loan
MCONSAL	Production expenditures	total salaries paid in the year to contract labourers
MDAYLAB	wage	Number of day labourers used on the farm during the past year multiplied by the months worked
MDAYSAL	*wage,* Production expenditures	Total salaries paid during the year to day labourers
MHADPROB	*wage*	Had problems finding day labourers
MMIGRANT	-	Ethnicity: migrant, native or second generation
MNEISAL	Production expenditures	total salaries paid in the year to neighbours and friends
MNOTSAL	Production expenditures	total salaries paid in the year to relatives who don't live in the NE
MOTHSAL	Production expenditures	total salaries paid in the year to other workers

Name of variable in the CPC survey	Model parameter	Description
MPCHEM	Production expenditures	Money spent on inputs: chemical fertilizer
MPLIME	Production expenditures	Money spent on inputs: lime
MPMED	*capreq,* Production expenditures	Money spent on inputs: medicine & vaccine for livestock
MPMINE	*capreq,* Production expenditures	Money spent on inputs: mineral salt for livestock
MPOTHER	Production expenditures	Money spent on inputs: other
MPPEST	*capreq,* Production expenditures	Money spent on inputs: pesticides, insecticides, fungicides, herbicides
MPSEED	*capreq,* Production expenditures	Money spent on inputs: better/new seeds
MRELSAL	Production expenditures	total salaries paid in the year to relatives who live in the NE
MSOLD1-7	*capstart, price*	Number of animals sold last year
MYEARAQ	*landprice*	Year of acquisition of the farm
MYEAREST	-	Year of settlement in the Northeast
MYRLOAN1	-	Year of first loan
MYRLOAN2	-	Year of second loan
MYRLOAN3	-	Year of third loan
NANNINC	Revenues	Total annual income from land leased for agriculture and for cattle
TOTDEBT	Production expenditures	total amount of debt to repay
WMONTHS	*wage,* Revenues	Number of months worked off-farm
WSALARY	*wage,* Revenues	Average salary per month worked off-farm
INSTITUTIONAL DATA		
MACCESS	distance	Name of nearest road/river
MBOAT	*distance, landprice*	River distance
MCERTIF	landprice	Type of title at purchase (certificate of possession/right of title/none)
MFROM	*distance*	Nearest larger town
MLINE	*distance*	Line (1, 2, ..)

Name of variable in the CPC survey	**Model parameter**	**Description**
	landprice	
MPRIMARY	*distance, landprice*	Primary road distance
MSECONDARY	*distance, landprice*	Secondary road distance
MTENURE	-	current form of tenure of the farm (certificate of possession/right of title/nothing)
MWALKED	*distance, landprice*	Distance walked to finca
NATURAL RESOURCE DATA		
NSOIL	index q, landprice	Soil type
NTOPOG	*index q, landprice*	Topography
PYBANA	*yield*	Yield per hectare: plantain (in units/bunches)
PYCACAO	*yield*	Yield per hectare: cacao (in units of 50 kg)
PYCOFF	*yield*	Yield per hectare: coffee (in units of 50 kg)
PYCORN	*yield*	Yield per hectare: corn (in units of 50 kg)
PYYUCCA	*yield*	Yield per hectare: manioc (in units of 50 kg)
PYRICE	*yield*	Yield per hectare: rice (in units of 50 kg)
LAND USE DATA		
CAFPALM	astart	Number of hectares under african palm
CBANANA	astart	Number of hectares under plantain
CCACAO	astart	Number of hectares under cacao
CCOFFEE	astart	Number of hectares under coffee
CCORN	*astart*	Number of hectares under corn
CFRUIT	*astart*	Number of hectares under fruit
COTHER	*astart*	Number of hectares under other crops
CPRIMARY	*astart*	Total hectares in primary forest
CRICE	*astart*	Number of hectares under rice
CSECOND	*astart*	Total hectares in secondary forest
CTOTCROP	*astart*	Total hectares in crops
CTOTHEC	*astart*	Total hectares on farm
CTOTPAS	*astart*	Total hectares in pasture
CVEGET	*astart*	Number of hectares under vegetables
CYUCCA	*astart*	Number of hectares under manioc
MARKETING DATA		
MPRICE	Revenues	total revenues received for selling wood

Name of variable in the CPC survey	Model parameter	Description
MPRICE1	Revenues, *capstart, price*	total revenues received for cattle sold
MPRICE2	Revenues, *capstart, price*	total revenues received for horses sold
MPRICE3	Revenues, *capstart, price*	total revenues received for mules sold
MPRICE4	Revenues, *capstart, price*	total revenues received for pigs sold
MPRICE5	Revenues, *capstart, price*	total revenues received for goats sold
MPRICE6	Revenues, *capstart, price*	total revenues received for poultry sold
MPRICE7	Revenues, *capstart, price*	total revenues received for other animals sold
PABANA	Price	Average price per unit: plantain
PACACAO	*price*	Average price per unit: cacao
PACOFF	*price*	Average price per unit: coffee
PACORN	*price*	Average price per unit: corn
PARICE	*price*	Average price per unit: rice
PVAFPALM	Revenues	total value of sales: african palm
PVBANA	Revenues	total value of sales: plantain
PVCACAO	Revenues	total value of sales: cacao
PVCOFF	Revenues	total value of sales: coffee
PVCORN	Revenues	total value of sales: corn
PVOTHER	Revenues	total value of sales: other crops
PVRICE	Revenues	total value of sales: rice
PVFRUIT	Revenues	total value of sales: fruit
PVVEGET	Revenues	total value of sales: vegetables
PVYUCCA	Revenues	total value of sales: manioc

B. STUDY ON PRODUCTION SYSTEMS IN THE NAPO PROVINCE, BY ESTRADA, SERÉ, AND LUZURIAGA

The study by Estrada, Seré, and Luzuriaga (1988) is an excellent source of information on farming practices in the Amazon basin of the Napo province. In particular, their estimations of labour requirements for production (table 35, page 72) have been useful for our parameter estimations. The data on labour use for the cultivation of crops and pasture include an estimate for the year 1985 and a simulated projection for 1990. The labour requirements are subdivided by product (coffee, cacao, plantain, maize, manioc, trees, and pasture) as well as by soil type (alluvial, volcanic and red-hill).

C. STATISTICAL YEARBOOKS

A number of statistical yearbooks contain data that can be used to complete the micro-level data described above. In particular, we mention the following sources:

Ministerio de Agricultura y Ganaderia (1994) *Ecuador en cifras; Compendio estadistico agropecuario 1965-1993.* MAG, PRSA, AID, Quito, Ecuador.
The data contain a.o.:

- production, surface and yield of major agricultural products
- prices at farm level (national, annual and monthly)
- consumer prices for urban areas (national, annual and monthly)
- profit margin
- production indices
- price indices
- annual imports and exports of agricultural products
- credits

Estimates for national production are given over the period 1965-1993. Most other data cover the period 1980 to 1993.

Almeida, M.G. (1995) *Principales estadisticas forestales del Ecuador.* Dirección general de planificación. Quito, Ecuador.
The data contain a.o.:

- wood extraction (annual)
- exploited surface for wood extraction
- production of wood products: saw wood, pulp, etc
- import and export of wood and wood products
- value of exports

Most data cover the period 1985-1994.

INEC (1996) *Cifrando y descifrando. Sucumbíos* and *Cifrando y descifrando. Napo.*
The data contain a.o.:

- census data on total population by canton 1950-1990
- 1990 census: population by age category, urban/rural
- agricultural production, yields and cultivated area by product 1994
- urban price indices 1994/1995.

Profors (1992) *Diagnostico socio-economico de la provincia de Sucumbios.* INEFAN, GTZ. Quito, Ecuador.
The data contain a.o.:

- infrastructure
- land with and without title
- production of oil
- surface of agricultural area, number of plots
- surface under coffee, cacao, rice, maize, plantain, pasture
- number and production of meat- and milk cows
- changes in land use, 1975 and 1985
- volume of high/medium/low quality wood in 1990
- deforestation in the cantons Lago Agrio, Joya de los Sachas and Yuca

- evolution of forestry products 1982-1989

D. CADASTRAL MAP

We have digitised the cadastral units of a 1:250.000 map from the Ecuadorian agency responsible for land titling (IERAC). From this map we have only used the sectors in zona Francisco de Orellana (Coca). The digitising have been done with Tosca digitising software version 2.0. The accuracy of the digitisation was satisfactory, the Root Mean Square (RMS)-error was lower then 50 meter. The cadastral map is described by the following characteristics:

Table A4-2: Cadastral map of the Coca region

Title	**Source**	**Year production**	**Projection**	**Scale**
Mapa de tenencia de la tierra en la jefatura zonal Coca.	IERAC	1993	Lat/Lon	1:250.000

The digitisation resulted in a file which contains the thematic data (see Table A4-3). The digitised information was converted into SPANS-format via the Arc View Shape-format (SHP/SHX). The projection is UTM 18 south.

Table A4-3: Description of digitized cadastral units

<table>
<tr><th>Name</th><th>Object</th><th>Data type</th><th>Name class variable</th><th>Classes</th></tr>
<tr><td rowspan="6">KADASCOC</td><td rowspan="6">Cadastral Information (some with household survey data LPC)</td><td rowspan="6">Polygons</td><td>CLASS</td><td>Sector number *10</td></tr>
<tr><td>NOMBRES</td><td>Name of sector</td></tr>
<tr><td>LOTES</td><td>Number of Lotes</td></tr>
<tr><td>TYPE_SEC</td><td>Type of sector
0 = No data or non occupied land
10 = Colonizacion
20 = Cooperatives
30 = Indigenous communal lands
40 = Agribusiness
50 = Urban Area
60 = Military property</td></tr>
<tr><td>CLASS1</td><td>Sector number *10 for sectors with household survey information (without = 0)</td></tr>
<tr><td>SELECTED</td><td>Sectors selected for model-testing</td></tr>
</table>

E. TOPOGRAPHICAL MAPS

We have digitised six 1:50.000 topographic maps obtained from the Instituto Geografico Militar/ Defense Mapping Agency (IGM/DMA). These six maps are summarised below in Table A4-4. The digitising has been done with Carta Linx software version 1. The accuracy of the digitisation was satisfactory, the Root Mean Square (RMS)-error was lower then 10 meter.

Table A4-4: Digitised 1:50.000 topographic maps

Title	Source	Year photo	Year production	Projection	Scale
Topographic map 4192 II: Bajo Huino	IGM/DMA	1989	1993	UTM 18 South	1:50.000
Topographic map 4192 I: San Sebastian del Coca (photocopy)	IGM/DMA	1991	1996	UTM 18 South	1:50.000
Topographic map 4292 IV: San Francisco de Orellana (El Coca) (photocopy)	IGM/DMA	1990	1993	UTM 18 South	1.50.000
Topographic map 4292 I: Limoncocha	IGM/DMA	1990	1993	UTM 18 South	1:50.000
Topographic map 4293 III: La Joya de los Sachas (Huamayacu) (photocopy)	IGM/DMA	1990	1993	UTM 18 South	1:50.000
Topographic map 4293 II: Shushufindi	IGM/DMA	1990	1993	UTM 18 South	1:50.000

The digitisation of the topographic maps resulted in seven files which are summarised in Table A4-4. The files have been converted to SPANS via the Map Info export format (MIF/MID). The projection is UTM 18 south.

Table A4-4: Description of seven digitised topographic map files

Name	Object	Data type	Name class variable	Classes
K50BLAD	Boundaries of 6 topographic maps	Polygons	CLASS	1 = Boundary
HEIGHT5T6	Contour classes of height for every 20 meter	Polygons	HOOGKLAS	10 = 180 – 200 m 11 = 200 – 220 m 12 = 220 – 240 m 13 = 240 – 260 m 14 = 260 – 280 m 15 = 280 – 300 m 16 = 300 – 320 m 17 = 320 – 340 m 18 = 340 – 360 m 19 = 360 – 380 m

Name	Object	Data type	Name class variable	Classes
TOWN5T6	Villages and towns	Polygons	CLASS	1 = Town 2 = Comuna 3 = Cooperative 4 = Industry
OILWE5T6	Oil wells	Points	CLASS	1 = Oilwell
RIVLI5T6	Rivers as lines	Lines	CLASS1	1 = River as line 2= River as border polygon
RIVPO5T6	Rivers as polygons	Polygons	CLASS1	1 = Water 9 = Island
ROAD50T6	Roads as network	Lines	CLASS1	1 = Main road 2 = Summer road 3 = Sandy road 4 = Cross over in the river 5 = Middle of the river 6 = River-land connections
			TRANSPORT_FACT	IF class1= 1 then : 1.0 If class1= 2 then : 1.5 IF class1= 3 then : 2.0 IF class1 >3 then : 3.0
			COST	TRANSPORT_FACT* LENGTH

F. SOIL MAP

The soil map was obtained from the Ministerio de Agricultura y Ganaderia (MAG), Programma Nacional de Regionalizacion Agraria (PRONAREG). The map was digitised with Carta Linx digitising software version 1.0. The accuracy of the digitisation was satisfactory, the Root Mean Square (RMS)-error was lower then 80 meter.

Table A4-5: Soil map Napo province

Title	Source	Year production	Projection	Scale
Mapa morfo edafologico, prov del Napo	PRONAREG	1983	Lat/Lon	1:500.000

The digitisation of the soil map resulted in the file described in Table A4-6. The file was converted to SPANS via the Arc View Shape-format (SHP/SHX). The projection is UTM 18 south.

TABLE A4-6 : DESCRIPTION OF DIGITISED SOIL MAP FILE

Name	Object	Data type	Name class variable	Classes
MEDAF	Soils	Polygons	CLASS	Soil-classes as number
			CODE	Soil-classes as code Nr Code 41 = D1 42 = D2 43 = D3 44 = D4 51 = E1 53 = E2 61 = F1 62 = F2 81 = H1 82 = H2 110 = K 111 = K1 112 = K2 113 = K3 114 = K4 115 = K5 116 = K6 998 = L (Lake) 999 = R (River)

Appendix 4-II Flow charts of GIS data transformations

Legend:

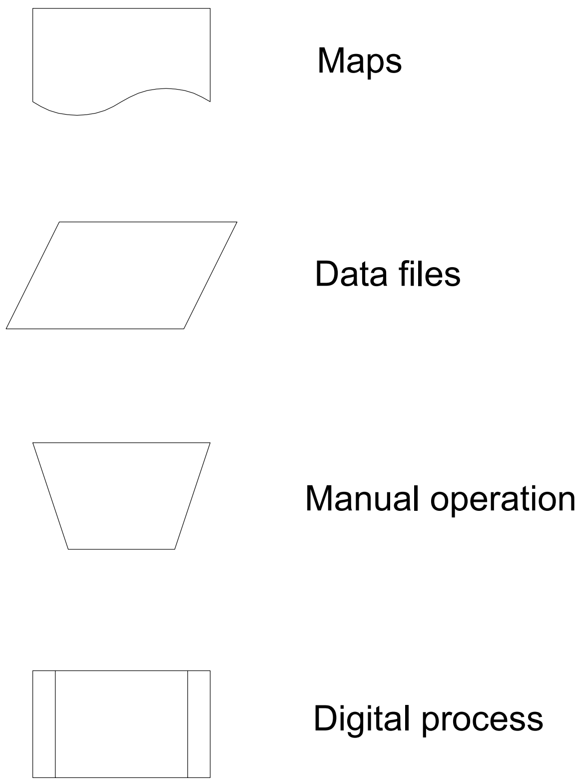


Figure 4- 9 Flow chart: digitization process of maps of the study area

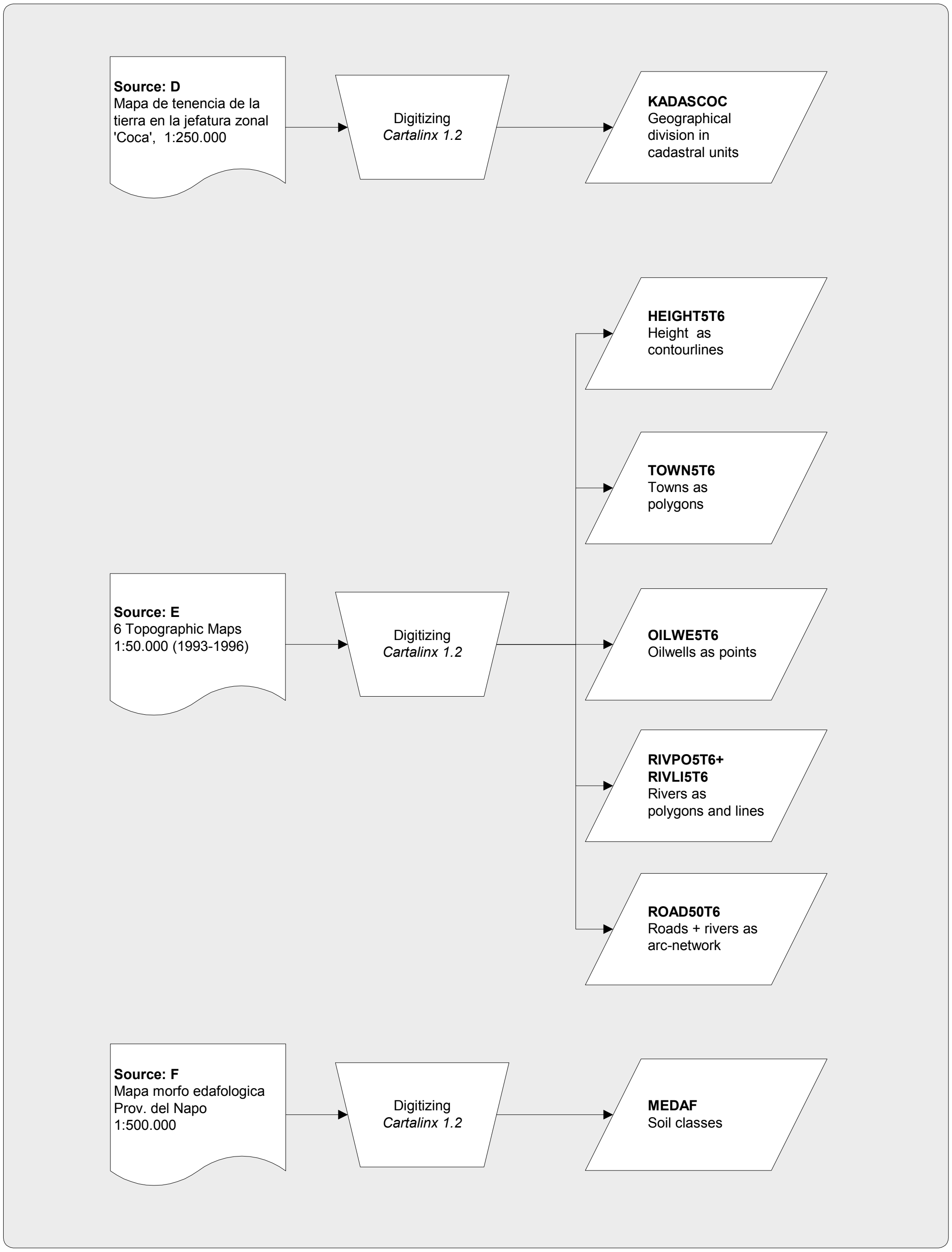


Figure 4- 10 Flow chart: process of data transformations from digitised maps

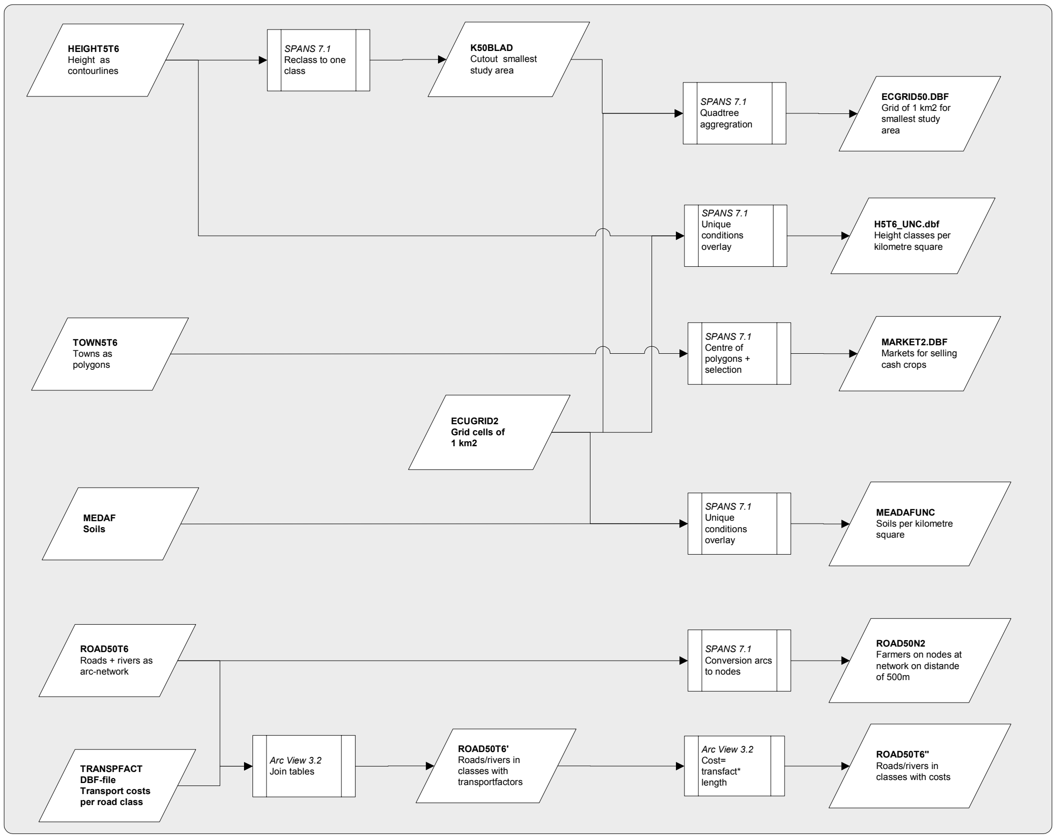
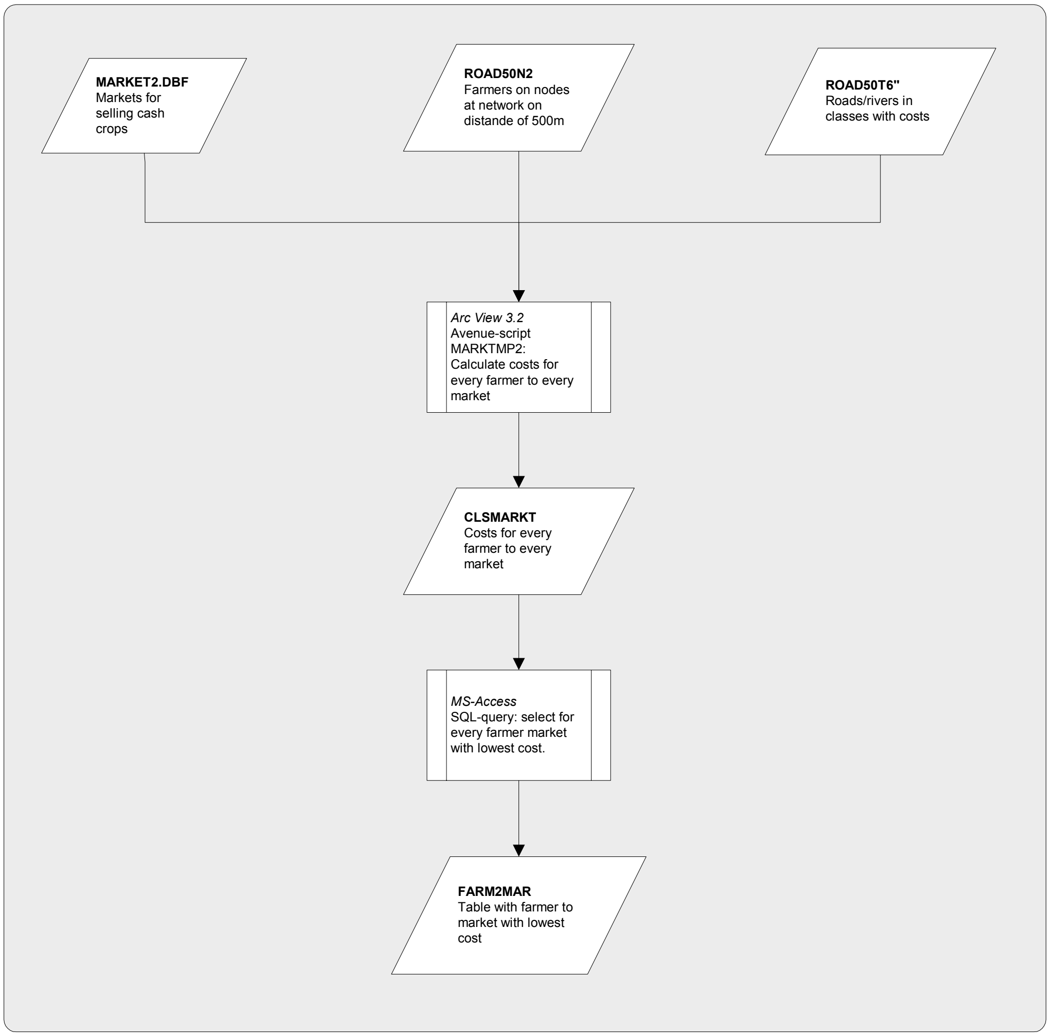


Figure 4- 11 Flow chart: process of calculating travel time between farms and markets



Appendix 4-III
Avenue script for calculating travel time between farms and markets

```
'--------------------------------------------------------------------------------------
' Arcview 3.2 AvenueScript
; Determining the costs/travel time between farms and markets
' CML 2000, Maarten van 't Zelfde
'--------------------------------------------------------------------------------------
'
aProject = av.GetProject
'
'--------------------------------------------------------------------------------------
' Determination of the name of the view
'--------------------------------------------------------------------------------------
'
aView = aProject.Finddoc("Kortste route naar dichtbijzijnde markt")
'
'--------------------------------------------------------------------------------------
'Determination of the name of the network theme
'--------------------------------------------------------------------------------------
'
aStreetFtab = aView.FindTheme("road50t6.shp").GetFtab

'--------------------------------------------------------------------------------------
' Check if the theme is available for networkanalysis
'--------------------------------------------------------------------------------------
'
if (not (NetDef.CanMakeFromFtab(aStreetFTab))) then
     msgBox.error ("Can 't make a network from this Ftab","")
     Exit
end
'
'--------------------------------------------------------------------------------------
' Assignments of the networkobjects
'--------------------------------------------------------------------------------------
'
aNetDef =NetDef.Make(aStreetFtab)
aNetwork=Network.Make(aNetDef)
'
'--------------------------------------------------------------------------------------
' Check if there are network-errors
'--------------------------------------------------------------------------------------
'
if (aNetDEF.HasError) then
     msgBox.Error ("The network has an error", "")
     Exit
end
'
'--------------------------------------------------------------------------------------
' Choice (by user) of attribute field to use as cost-field.
'--------------------------------------------------------------------------------------
'
aSelectedCostField = msgBox.Choice(aNetDef.GetCostFields,
                     "Kies een kostenveld uit:","Netwerk kosten selectie")
'
'--------------------------------------------------------------------------------------
' Assign the costfield for the network
'--------------------------------------------------------------------------------------
'
aNetwork.SetCostField(aSelectedCostField)
'
'--------------------------------------------------------------------------------------
' Choice of the point-table with farmers along the road
'--------------------------------------------------------------------------------------
'
aCustomerFtab=aView.FindTheme("road50n2.shp").GetFTab
aPointField=aCustomerFTab.FindField("Shape")
'
'--------------------------------------------------------------------------------------
' Check if all the farmer-nodes are on the network
'--------------------------------------------------------------------------------------
'
aPointList = {}
aTusList = {}
for each rec in aCustomerFtab
```

```
     p  = aCustomerFtab.ReturnValue(aPointField, rec)
     aTuslist = p.AsList
     for each m in aTusList
        if (aNetwork.IsPointOnNetwork(m)) then
            aPointList.Add(m)
        end
     end
end
'
'-------------------------------------------------------------------------------------
' Choice of the point-table with available markets
'-------------------------------------------------------------------------------------
'
aMarktFtab=aView.FindTheme("Market02.shp").GetFtab
aPointField2=aMarktFtab.FindField("Shape")
'
'-------------------------------------------------------------------------------------
' Check if all the market-nodes are on the network
'-------------------------------------------------------------------------------------
'
aMarktList = {}
aTus2List = {}
for each rec in aMarktFtab
     p  = aMarktFtab.ReturnValue(aPointField2, rec)
     aTus2List = p.AsList
     for each n in aTus2List
         if (aNetwork.IsPointOnNetwork(n)) then
             aMarktList.Add(n)
         end
     end
end
'
'-------------------------------------------------------------------------------------
' Determine the cost of every farmer to every market
'-------------------------------------------------------------------------------------
'
aNumFoundList = aNetwork.FindClosestFac (aPointList, aMarktList, 6, 0, TRUE)
'
'-------------------------------------------------------------------------------------
' Creation of an object "Clsmarkt" with the outcome of above analysis.
'-------------------------------------------------------------------------------------
'
aRoute="Clsmarkt".AsFileName
aThroute=aNetwork.WriteClosestFac(aRoute)

theSrcName = SrcName.Make("Clsmarkt.shp")

if (theSrcName=nil) then
  MsgBox.Error("Invalid Data Source Name", "")
  exit
end
'
'-------------------------------------------------------------------------------------
' Make the new shape-file "Clsmarkt.shp"
'-------------------------------------------------------------------------------------
'
aSearchTheme = aView.FindTheme("Clsmarkt.shp")

if (aSearchTheme=nil) then
     MsgBox.Info ("Even geduld a.u.b.", "Laden van beste route bestand")
     aNewTheme= Theme.Make(theSrcName)
     aView.AddTheme(aNewTheme)
end
'
'-------------------------------------------------------------------------------------
' Make the new shape file visible and close the script
'-------------------------------------------------------------------------------------
'
aTheme=aView.FindTheme("Clsmarkt.shp")
aTheme.SetVisible(true)
aTheme.SetActive(true)
MsgBox.Info("Gefeliciteerd, we zijn klaar", "Eindmelding")
'
'-------------------------------------------------------------------------------------
' End of the Avenue-script.
'-------------------------------------------------------------------------------------
```

5

Preliminary results from model simulations

This chapter presents the model results for three "representative households" in the study area. We will compare the results with the start situation and discuss some implications for future model development.

5.1. The Ecuadorian farm household submodel applied to three sectors

The Ecuadorian household submodel was solved using GAMS modelling software. The model equations (section 3.4.2) were translated into GAMS language[48] and solved with the BDMLP linear programming solver. The outcomes of the model are the short-run decisions for each of the three households. The starting situation for the households were derived from the CPC survey. Table 5-1 presents the initial land use and the optimal land allocation as it was calculated from the model[49].

Table 5-1. *Land use in hectares for the three model households: start situation and model results*

Activity I	**hh59**			**hh75**			**hh202**		
	start	***model***	***dif.***	***start***	***model***	***dif.***	***start***	***model***	***dif.***
Cacao	0.7	0.7	=	0.2	0.2	=	0.0	0.0	=
Coffee (full grown)	4.9	0.0	–	2.2	0.0	–	2.0	0.0	–
Corn	0.5	0.0	–	0.5	0.0	–	0.5	0.0	–
Manioc	0.1	2.4	+	0.3	3.4	+	0.2	1.9	+
Plantain	0.9	0.0	–	1.5	0.0	–	0.7	0.0	–
Rice	0.3	0.0	–	0.0	0.4	+	0.2	0.0	–
Other crops	0.6	0.0	–	1.5	0.0	–	0.3	0.0	–
Total hectares in crops	**8.0**	**3.1**	–	**6.2**	**4.0**	–	**3.9**	**1.9**	–
Total hectares fallow	**0.3**	**5.2**	+	**3.5**	**5.7**	+	**0.0**	**2.0**	+
Total hectares in pasture	**40.6**	**40.6**	=	**2.5**	**2.5**	=	**6.9**	**6.9**	=
Total hectares primary forest	**29.3**	**29.3**	=	**41.4**	**41.4**	=	**36.6**	**36.6**	=
Total hectares secondary forest	**8.1**	**8.1**	=	**1.3**	**1.3**	=	**0.7**	**0.7**	=
Total hectares on farm	***86.3***	***86.3***	=	***54.9***	***54.9***	=	***48.1***	***48.1***	=

Data source: for the start situation, estimated from household data in the CPC survey

[48] A listing of the GAMS programme file is added as Appendix 5-I.

[49] A partial listing of the GAMS output file is added as Appendix 5-II.

In Table 5-2 we present the model results for the hired labour, off-farm labour, food purchases and net revenues.

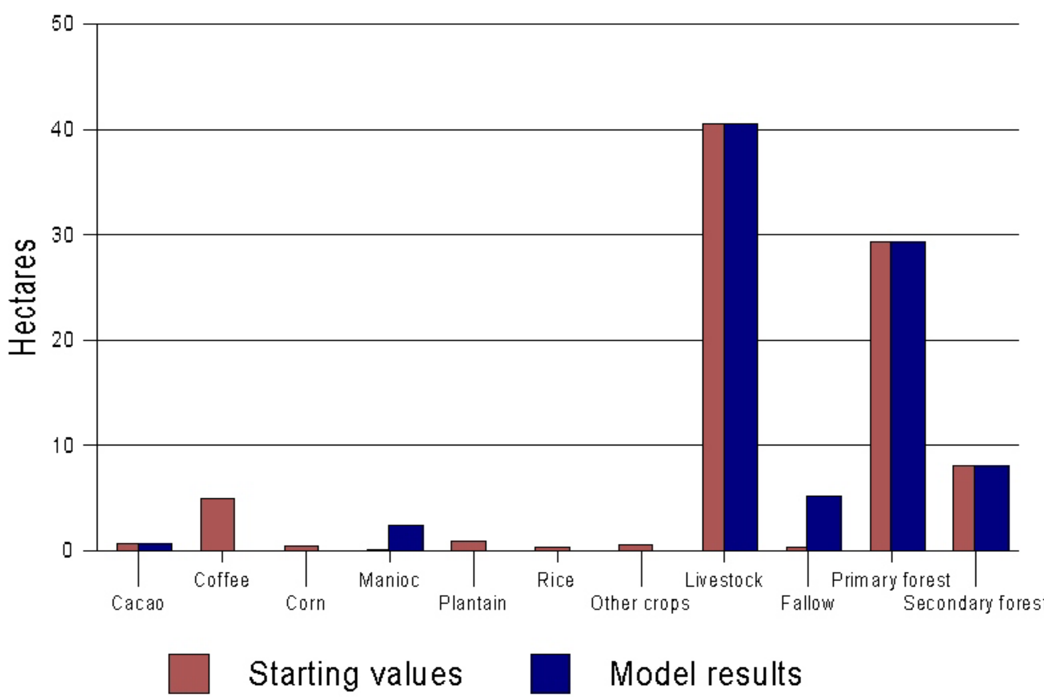
Table 5-2. *Model results: Optimal values for other decision variables*

Variable	hh59	hh75	hh202
Labour hire (LHIRE)	115	0	0.0
Off-farm labour / total available family labour	0 / 884	633 / 754	490 / 676
Food purchases in quintales[50] (PURCH)	34	0	8
Annual revenues in US $ (NETREV)	25012	2116	1331

Model results compared to the starting situation

The figures 5-1 to 5-3 illustrate the cultivation of crops both in the starting situation and as a result of the model calculations.

Figure 5-1. *Land use for hh59*



[50] 1 quintal = 50 kgs.

Figure 5-2. *Land use for hh75*

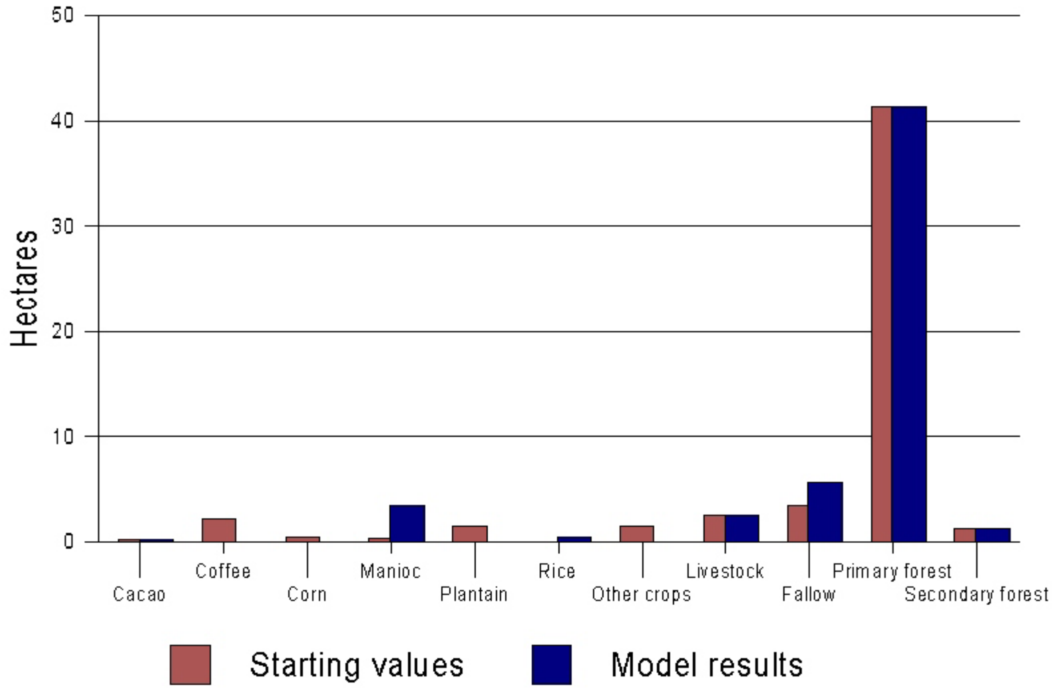
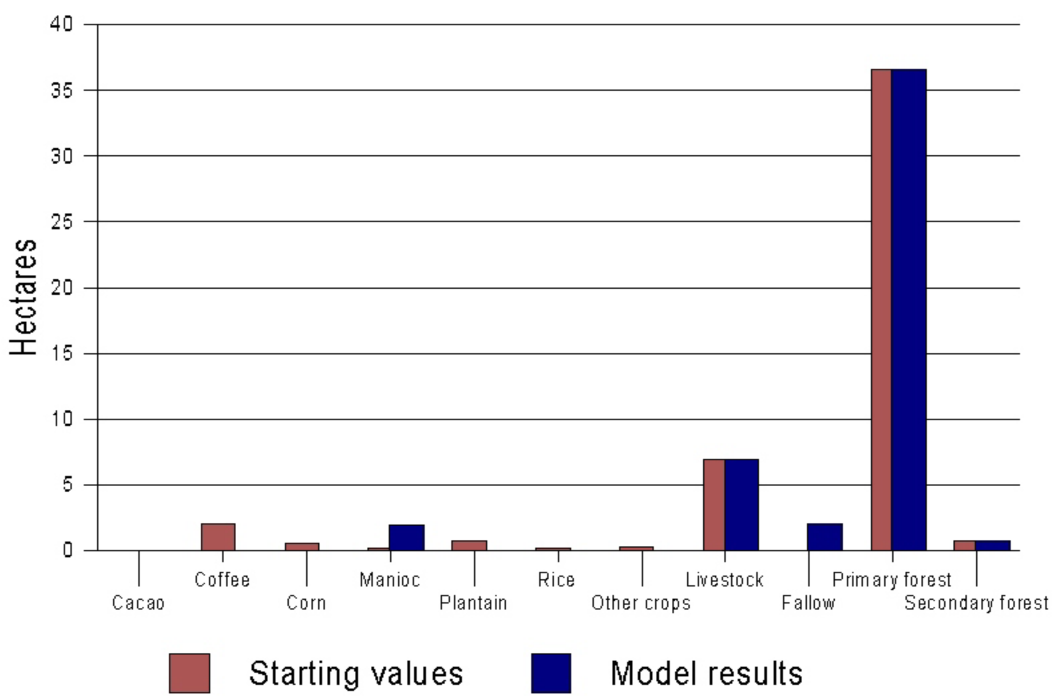


Figure 5-3. *Land use for hh202*



It appears from the Tables 5-1 and 5-2 and the figures 5-1 to 5-3 that the model results differ in a few respects from the starting values calculated directly from the CPC survey. The model results must be interpreted as the optimal strategy for the short run, given the assumptions made on the behaviour of the farmers. The cultivation of annual crops can be changed in the short run. Because of the short-run focus, the hectares under perennial crops and pasture can be changed only into fallow land. Fallow land has no revenues and no costs. The differences between starting values and model results can be summarized as follows:

- All coffee production is abandoned and converted to fallow
- Cacao production is unchanged
- Production of manioc is expanded
- One household expands the rice production, the others stop cultivating rice
- All households stop cultivating corn and plantain
- Pasture is unchanged
- Food is purchased by two households, *hh75* is self-sufficient in food
- Only household *h59* hires labour
- Two households, *h75* and *h202,* have off-farm jobs.

5.2. Model results: discussion

The production of coffee is abandoned by all three households; this means that the short-run revenues do not outweigh the costs of production and labour. This is not surprising if we take into account that coffee growers in the Ecuadorian Amazon do indeed abstain from harvesting during periods when the price is too low (Chalá, personal communication). Prices for coffee are highly variable from one year to another, hence the situation may be different in future years. The production of cacao was profitable in the short run for our three model households. The increased production of manioc is easily explained by the model assumption that manioc is the only component of the household diet. One household (*h75)* produces enough for the family consumption. Household *h75* uses the remainder of the crop land to produce rice; apparently rice is more profitable than the other annual crops. The other two households may wish to expand the production of manioc to meet the consumption requirements, but not enough land is available for annual crops in the short run. They have to buy manioc in the market.

Pasture is obviously an important source of income; all three households keep their land under pasture unchanged. Sector *h59* is the most affluent with respect to pasture land; on average they have 40.6 hectares. This explains the high net revenues of more than 25 thousand US dollars. Because of the substantial area of pasture, household *h59* needs a larger amount of labour input. This explains why this is the only household which hires labour additional to the available family labour. The other two households have a labour surplus, which they capitalize for money by working off-farm.

Expected long run effects

The results so far did not show important land use changes. This is because of the “short run bias” of this reduced model. The hectares under primary and secondary forest are as yet unchanged, because long run motivations need to be introduced before the mechanisms of land conversion can be effectively calculated. The short run results indicate that there is room for expansion as long as there is labour available. All households have fallow land, mostly because they abandoned coffee. This fallow land will most likely be taken into production in future years. Next, the households will gradually expand into areas currently under forest with the support of household capital to cover the costs of clearing the land. Households *h75* and *h202* only had small profits, indicating that the deforestation process may be much slower than *h59*. Household *h59* has substantial revenues which can easily be used for “investments” in forest clearing. On the basis of the short run results we can not yet predict which types of land use will be preferred in the long run. In the long run, we may see a change from coffee production to cacao if the cost-benefit structure for our model year turns out to be representative for the future. Production of food crops may remain important if food crops are expensive in the market or distances are too long; even though the introduction of food crops other than manioc may as well have other effects on the model results. Rice was a promising alternative as a cash crop. This may also be an important product in the long run. Pasture is likely to have an important place in long run household strategies; not only for annual revenues but also because livestock is often kept as a means to accumulate savings.

Observations on the modelling process
Despite the preliminary status of the model, the results so far show that the model is consistent and can be used to generate plausible household decisions. The results demonstrate the complexity of farm household decisions in allocating land, labour and capital, accounting for the initial endowments, the need to feed the family and the desire to accumulate wealth. Even though the real world is much more complicated than what can be described by this reduced model, we may expect that the explanatory power of the model will improve along the way if the stepwise process of “parameter estimation – validation – model improvement” is continued. Calculations with the long run version of the model is likely to increase our understanding of land conversion in particular. Sensitivity analysis will further improve our knowledge of the driving forces of land conversion.

5.3 Recommendations for further model development

The previous two sections described the outcomes of a baseline scenario for only one year. No conclusions can be drawn yet on the consequences for the state of the forest, because long-run motivations of the households still need to be included. In chapter 3 we described the modelling process along the following steps:

i) Define the decision-making process for each actor
j) Identify the available data and collect additional data
k) Choose the unit of analysis and link this to a spatial unit
l) Design model structure and equations
m) Estimate model parameters
n) Run the model for a baseline scenario
o) Evaluate the model results for the baseline scenario
p) Design and implement model experiments for alternative scenario’s

Steps a)-e) were described in chapter 3 and 4. Steps f) and g) have been illustrated in the previous sections. We are not ready yet to proceed to step h), because we have not finished the iterative process of steps d), e), f), and g). Only after the model is found to be a good representative for current household decision-making, we can move to the final stage in the modelling: to design and implement model experiments for alternative scenarios. In the following we will add some comments on the continuation of the modelling process; that is, steps d) to h).

Ad d) (Re-)design model structure and equations

As a part of the model validation, the model structure and equations need to be expanded for a better representation of the "real world". The model was a reduced version of the "core structure" as it was described in Chapter 3. We will briefly recall the actions that are required to expand the model for a more inclusive representation of current household decision-making. At this stage, no other model adaptations need to be taken into account.

The following factors need to be added to complete the household decision model:

- a planning period of 10 years to include strategic land use decisions
- include other household activities such as on-farm logging, forest clearing, selling and buying land, savings and investment (see chapter 3, Table 1).

Ad e) (Re-)estimate model parameters

We mention a few examples of parameter estimations that need to be improved (see also our recommendations in chapter 4):

- parameters need to be added for the long-run decisions such as land clearing, selling land, etc,
- some parameter estimates can be validated or improved by means of additional data sources,
- consumption of food crops need to be estimated from secondary data,
- estimates for land values need to be improved, depending on location, tenure status, soil quality and elevation,
- yields need to be estimated as functions of soil quality and elevation,
- farm gate prices for products need to be estimated as functions of the distance from the nearest market,
- the parameter estimates need to be based on the whole data set of nineteen sectors instead of three[51].

Ad f) Re-run the model for a new baseline scenario

After each round of model adaptations and parameter estimations we need to run the model again and compare the model results with the starting values. A sensitivity analysis on the parameters should point out which parameters are crucial for the land use decisions.

Ad g) Evaluate the model results for the new baseline scenario

If after successive adaptations the model seems fit to describe the actual behaviour (start situations), then we can proceed to a calculation of base line scenarios for the secondary and tertiary actors in a similar way. After these results are satisfactory as well, the model is ready to use for policy evaluations.

---

[51] We recommend that during the fases of model development only part of the whole data set of 19 sectors from the CPC survey is used for parameter estimation. The remainder can be used as a tool for model validation, in the final stage when the model is found to be representative. Because of the iterative process of model development and parameter estimations, it is useful to keep part of the data aside to test the robustness of the model in its final form. For a theoretical discussion on model validation, see Law and Kelton (1991).

Ad h) Design and implement model experiments for alternative scenarios

In the policy scenarios special attention will be given to the Global Forest Fund discussed earlier. Furthermore, the policy scenarios can be used to evaluate the result of specific policy measures and institutional changes, for example:

- changes in prices of cash crops,
- tax changes,
- changes in property rights,
- changes in land prices,
- changes in the quality of infrastructure,
- etc.

The policy scenarios need to be evaluated on the consequences for land use, as well as for household wealth and employment. Depending on the results, recommendations can be made for suitable policies aimed at forest protection while taking into account the vulnerability of the people living in the area.

Appendix 5-I GAMS program listing

```
$TITLE CML household submodel : Ecuador.gms
***********************************************************************************************
*MODEL SPECIFICATIONS
*
* This GAMS program file (ecuador.gms) is structured as follows:
*
* 1. Declarations of Parameters and variables
*
* 2. Assignment of values for the parameters
*
* 3. Definitions of model equations
*
* 4. Subsequent model solves for each household
*
***********************************************************************************************


***********************************************************************************************
* 1. Declarations of Parameters and variables                                                 *
***********************************************************************************************

SETS    h                actors   / hh59, hh75, hh202 /
        i                activity
                         / plantain, cacao, coffee, corn, rice, manioc, other, animal,
                         fallow, off_lab, prim_for, sec_for /
        q                land quality /black, red, alluvial/

        crops(i)         subset of i  / plantain, cacao, coffee, corn, rice, manioc, other/
        food(i)          subset of i /manioc /
        cash(i)          subset of i /cacao, coffee, corn, rice /
        past(i)          subset of i /animal/
        annual(i)        annual crops / plantain, corn, rice, manioc, other/
        per(i)           perennial crops /cacao, coffee/
        luse(i)          land use activities /plantain, cacao, coffee, corn, rice, manioc,
other, animal,
                         fallow, prim_for, sec_for/;

PARAMETER   price(h,i)      price of good i in $ per kg UNC var pacoff etc
            capreq(h,i,q) non-labour costs of producing crops in $ per ha
            capstart(h)   starting capital or value of livestock
                      / hh59     7346
                        hh75      235
                        hh202     626     /

            cons(i)    the minimum required consumption of food crops per person
                        /manioc        10        /

                size(h) household size
**comment:**      UNC survey variable hnumber
                      / hh59     8.2
                        hh75     6.8
                        hh202     4.6     /

        famlab(h) available family labour
**comment:**       defined as the number of persons of 16 years old and above
**comment:**       calculated from UNC variables hnumber and hage
                      / hh59     3.4
                        hh75     2.9
                        hh202     2.6     /
;

TABLE        labreq(i,q) annual labour requirements per ha in working days

              alluvial           black      red
Coffee        71.0       74.9       58.7
Cacao         37.5       36.1       34.2
Plantain       12.2       13.4       14.6
corn          11.1       26          16.6
rice           11.1       26          16.6
manioc        12.2       13.4       14.6
animal         27.9       23.2       15.6


**comment**     source: calculated from Estrada 1988, rice was missing, copied from corn
```

```
;

TABLE           price(h,i)      farm gate price of good i in $ per unit UNC var pacoff etc

        cacao   coffee  corn    rice    manioc  animal
hh59     27.8    2.6     3.3     8.9     2.8      91.6
hh75     20.9    2.7     3.2     8.9     2.8      10.7
hh202    20.9    2.8     3.3     8.9     2.8      70.4

**comment**     source: UNC variable pacoff pacorn payucca pabana pacacao pafruit paveget
parice
**comment**     some price data were not available, because no sales occurred in the survey
**comment**     the lowest price from the others is then taken as a proxy
;

TABLE            cropland(h,q)   quantity of cropland of quality q available to actor h in ha
**comment:**    calculated from UNC survey variable ctotcrop and nsoil
        black
hh59     8.0
hh75     6.2
hh202     3.9


TABLE           totland(h,q)  quantity of land available to actor h in ha
**comment:** UNC survey variable ctothec
       black  red     alluvial
hh59   86.3   0      0
hh75   32.2   0       22.7
hh202  44.1   4.0    0



TABLE            astart(h,i,q)    starting value for land allocated to activity i
**comment:**     UNC survey variable ctotpas


        animal.black  cacao.black  coffee.black  corn.black  manioc.black  plantain.black
rice.black  other.black
hh59       40.6         0.7        4.9           0.5         0.1           0.9            0.3
0.6
hh75       2.5          0.2        2.2           0.5         0.3           1.5            0
1.5
hh202       6.9         0          2.0           0.5         0.2           0.7            0.2
0.3


TABLE yield(h,i,q)     annual yield of good i for household h depending on land quality in kgs
per ha pycoff etc
        cacao.black  coffee.black  corn.black  rice.black  manioc.black plantain.black
animal.black
hh59       4.25          37.7            18.7       10           20      566.7
6.82
hh75       5             24.3            24.2       10           20      566.7             38
hh202       4.25          20.4            15         10           20      566.7
1.28
**comment**     if no data were available on a particular farm, the lowest value of the other
was used


SCALARS
        workdays        number of working days per year available for agriculture
        wage            daily wage
;

VARIABLES       A(h,i,q)        land of quality q allocated to good i by household h
                FOODCOST(h)     costs of buying food
                K(h,i,q)        household capital assigned to activity i on land of quality q
                LFAM(h,i,q)       family labour allocated to activity i
                LHIRE(h,i,q)    labour hired by household h
                LANDVAL(h)      the value of farm land h
                NFINC(h)        non farm income
                NETREV(h)       net revenues
                OBJ             objectives of the farm households
                OUTPUT(h,i,q)   output of good i on land of quality q for household h
                PRODCOST(h)     total costs of production for household h
                PURCH(h,i)      food purchased on the market by household h
                REV(h)          gross revenues for household h
```

```
;

POSITIVE VARIABLES A,  FOODCOST, LANDVAL, LFAM, LHIRE, OUTPUT, PRODCOST, PURCH, REV   ;

*********************************************************************************************
* 2. Assignment of values for the parameters                                               *
*********************************************************************************************
**comment:** capreq(h,i)$crops(i) is the sum of UNC variables mpseed and mppests divided by
ctotcrop
**comment:** capreq(h,'animal') is the sum of UNC variables mpmed and mpmine divided by
ctotpas

capreq('hh59',i,'black')$crops(i) = 94/cropland('hh59','black')     ;
capreq('hh75',i,'black')$crops(i) = 1/cropland('hh75','black')      ;
capreq('hh202',i,'black')$crops(i) = 49/cropland('hh202','black')     ;
capreq('hh59','animal','black') = 278/astart('hh59','animal','black')       ;
capreq('hh75','animal','black') = 14/astart('hh75','animal','black')        ;
capreq('hh202','animal','black') = 23/astart('hh202','animal','black')       ;
astart(h,'prim_for','red') = totland(h,'red');
astart(h,'prim_for','alluvial') = totland(h,'alluvial') ;
astart('hh59','prim_for','black') = 29.3 - astart('hh59','prim_for','red') -
astart('hh59','prim_for','alluvial');
astart('hh75','prim_for','black') = 41.4 - astart('hh75','prim_for','red') -
astart('hh75','prim_for','alluvial');
astart('hh202','prim_for','black') = 36.6 - astart('hh202','prim_for','red') -
astart('hh202','prim_for','alluvial');
astart('hh59','sec_for','black') = 8.1 ;
astart('hh75','sec_for','black') = 1.3 ;
astart('hh202','sec_for','black') = 0.7 ;
workdays = 260 ;
wage = 30*12/workdays    ;


*********************************************************************************************
* 3. Definitions of model equations                                                         *
*********************************************************************************************

EQUATIONS O2i          profit of the farmer and animal accumulation (objective function)
           P1i(h,i,q)   production function output
           P2i(h,i,q)   production function labour constraint
           P3i(h,i,q)   production function capital constraint
           C1i(h,i)     food requirement constraint of the farmer
           F1i(h)       financial constraint net revenues
           F2i(h)       financial constraint gross revenues
           F3i(h)       financial constraint non farm income
           F4i(h)       financial constraint production costs
           F5i(h)       financial constraint food costs
           F8i(h)       financial constraint budget limitation
           L1i(h)       labour constraint family labour
           A2i(h,q)     land constraint
           A2ii(h,q)    land constraint animal
           A2iii(h,i,q) land constraint perennials
           A2iv(h,q)    land constraint annual crops
           A3i(h)       land constraint land value
;

O2i..
OBJ =E= NETREV('hh59') + LANDVAL('hh59')
+ NETREV('hh75') + LANDVAL('hh75')
+ NETREV('hh202') + LANDVAL('hh202')
;

P1i(h,i,q)..
OUTPUT(h,i,q) =E= yield(h,i,q)*A(h,i,q)
;

P2i(h,i,q)..
 LFAM(h,i,q) + LHIRE(h,i,q) =G=labreq(i,q)*A(h,i,q)
;

P3i(h,i,q)..
K(h,i,q) =E=capreq(h,i,q)*A(h,i,q)
;

C1i(h,i)$food(i)..
SUM(q,OUTPUT(h,i,q)) + PURCH(h,i)  =G= cons(i)*size(h)           ;
```

```
F1i(h)..
NETREV(h) =E= REV(h) + NFINC(h) - PRODCOST(h) - FOODCOST(h)
;

F2i(h)..
REV(h) =E= SUM(q,(SUM(i,price(h,i)*OUTPUT(h,i,q))))
;

F3i(h)..
NFINC(h)=E= SUM(q,0.99*wage*LFAM(h,'off_lab',q))
;

F4i(h)..
PRODCOST(h) =E= SUM(q,
SUM(i$crops(i),K(h,i,q)+wage*LHIRE(h,i,q))+SUM(i$past(i),K(h,i,q)+wage*LHIRE(h,i,q)))
;

F5i(h)..
FOODCOST(h) =E= SUM(i$food(i), PURCH(h,i)*price(h,i))
;

F8i(h)..
PRODCOST(h)+FOODCOST(h) =L=capstart(h)
;

L1i(h)..
SUM(i,SUM(q,LFAM(h,i,q))) =L= workdays*famlab(h)
;

A2i(h,q)..
SUM(i$luse(i),A(h,i,q)) =E= totland(h,q)  ;

A2ii(h,q)..
A(h,'animal',q)  =L= astart(h,'animal',q)  ;
;

A2iii(h,i,q)$per(i)..
A(h,i,q)  =L= astart(h,i,q)  ;
;

A2iv(h,q)..
SUM(i$annual(i),A(h,i,q))  =L= SUM(i$annual(i),astart(h,i,q))  ;
;

A3i(h)..
LANDVAL(h) =E= SUM(q,SUM(i$past(i),A(h,i,q)))
;

MODEL farm / ALL/  ;


*******************************************************************************************
* 4. Subsequent model solves for each household                                           *
*******************************************************************************************
A.FX(h,'prim_for',q) = astart(h,'prim_for',q);
A.FX(h,'sec_for',q) = astart(h,'sec_for',q); ;


SOLVE farm USING LP MAXIMIZING OBJ  ;


DISPLAY A.L, LFAM.L, LHIRE.L, NETREV.L, PURCH.L   ;
```

Appendix 5-II GAMS solve results

```
GAMS 2.50.093  DOS Extended/C                   02/09/01 13:18:24  PAGE     47
CML household submodel : Ecuador.gms
E x e c u t i o n

----    268 VARIABLE  A.L           land of quality q allocated to good i by
                                    household h

                     black         red    alluvial

hh59 .cacao          0.700
hh59 .manioc         2.400
hh59 .animal        40.600
hh59 .fallow         5.200
hh59 .prim_for      29.300
hh59 .sec_for        8.100
hh75 .cacao          0.200
hh75 .rice           0.400
hh75 .manioc         3.400
hh75 .animal         2.500
hh75 .fallow         5.700
hh75 .prim_for      18.700                  22.700
hh75 .sec_for        1.300
hh202.manioc         1.900
hh202.animal         6.900
hh202.fallow         2.000
hh202.prim_for      32.600       4.000
hh202.sec_for        0.700

----    268 VARIABLE  LFAM.L        family labour allocated to activity i

                     black

hh59 .cacao         25.270
hh59 .animal       858.730
hh75 .cacao          7.220
hh75 .rice          10.400
hh75 .manioc        45.560
hh75 .animal        58.000
hh75 .off_lab      632.820
hh202.manioc        25.460
hh202.animal       160.080
hh202.off_lab      490.460

----    268 VARIABLE  LHIRE.L       labour hired by household h

                     black

hh59 .manioc        32.160
hh59 .animal        83.190

----    268 VARIABLE  NETREV.L      net revenues

hh59  25011.072,    hh75   2116.205,    hh202  1331.208

----    268 VARIABLE  PURCH.L       food purchased on the market by household h

           manioc

hh59       34.000
hh202       8.000

EXECUTION TIME       =        0.050 SECONDS    0.2 Mb      WAT-50-093
USER: Caroline Jongkamp                             G980603:1118AV-WAT
      University of Leiden, Centre of Environmental Science       DC522

**** FILE SUMMARY

INPUT      D:\CAROLINE FILES\GAMSDIR\ECUADOR3.GMS
OUTPUT     D:\CAROLINE FILES\GAMSDIR\ECUADOR3.LST
```

References

Almeida, M.G.
1995 *Principales estadisticas forestales del Ecuador.* Dirección general de planificación. Quito, Ecuador.

Brown, K. and D. Pierce (eds)
1994 *The causes of tropical deforestation, the economic and statistical analysis of factors giving rise to the loss of the tropical forests.* UCL Press, London.

Caballero, D., and A. Anzules S. (eds)
1992 *Producción agropecuaria en la selva húmeda de la Región Amazónica.* INIAP/IICA/CIID,Quito.

Chalá, V.H.
1987 *Breve diagnostico de la región Amazonica y proyecciones de investigacion para la estacion experimental Napo-Payamino del INIAP.* INIAP.

Chalá, V.
Monitoreo de procesos de deforestacion en la Amazonia Ecuatoriana. Proyecto FAO-Japon.

Cleuren, H.
2001 Paving the road to forest destruction. PhD Dissertation, CNWS New Series nr. 1, CNWS, Leiden University

Dale, V.H., R.V. O'Neill, M. Pedlowski, and F. Southworth
1993 "Causes and effects of land-use change in Central Rondônia, Brazil". In: *Photogrammetric Engineering and Remote Sensing*, 59(6): 997-1005.

De Groot, W.T..
1992 Environmental Science Theory; Concepts and methods in a one-world, problem-oriented paradigm. Elsevier Science Publishers, Amsterdam, 584 pp.

De Groot, W.T., and E.M. Kamminga
1995 *Forest, People, Government. A policy-oriented analysis of the social dynamics of tropical deforestation.* Main report of the project 'Local actors and global treecover policies', CML report no. 120. Centre for Environmental Science, Leiden University, Leiden.

De Groot, W.T, C. Jongkamp and H. Cleuren
1997 "In search of a global policy for the tropical rainforest" In: *Change, research and policy newsletter on global change from the Netherlands.* Nr. 36, May-June 1997.

De Koning, G.H.J.
1999 *Spatially explicit analysis of land use change: a case study for Ecuador.* PhD thesis. Wageningen University and Research Centre, Wageningen.

Ellis, F.
1988 *Peasant Economics, farm households and agrarian development.* Cambridge University Press, Cambridge.

Estrada, R.D., C. Seré and H. Luzuriaga

1988 *Sistemas de produccion agrosilvopastoriles en la selva baja de la provincia del Napo, Ecuador.* CIAT, Cali, Colombia.

FAO
1997 *The state of the World's Forests.* FAO, Rome.

GTZ/PROFOGAN
1993 *Proceso de analysis y mejoramiento de sistemas de produccion agropecuaria y forestal de pequeños y medianos productores*.

Hoogeveen, J.A.M. and D.P. van Soest
1993 *Tropical rainforest Degradation in Cameroon and Ecuador: a socio-economic approach.* Centre for Development Studies/ University of Groningen, series Development and Security.

INEC
1996 *Cifrando y descifrando. Sucumbíos*.

INEC
1995 *Cifrando y descifrando. Napo*

INEFAN/ITTO
1995 *Principales estadisticas forestales del Ecuador*.

Jepma, C.J.
1995 *Tropical deforestation, A Socio-Economic Approach.* Earthscan, London.

Kok, K.
2001 *Scaling the land use system. A modelling approach with case studies for Central America.* PhD thesis. Wageningen University and Research Centre, Wageningen.

Kummer, D.M.
1991 *Deforestation in the postwar Philippines.* The University of Chicago Press, Chicago.

Lambin, E.F.
1994 *Modelling deforestation processes: A Review.* Office for Official Publications for the European Community, Luxembourg.

Law, A.M, and W.D. Kelton
1991 *Simulation modeling and analysis.* Second edition, McGraw-Hill.

Lonergan, S. and S. Prudham
1994 "Modeling Global Change in an Integrated Framework: A View from the Social Sciences." In: W.B. Meyer and B.L. Turner (eds) *Changes in Land Use and Land Cover: A global perspective.*

Luzuriaga, H.
1986 *Diagnóstico Estático de los Sistemas de Producción de la Selva Baja de la Provincia del Napo, Ecuador.* Quito, Ecuador.

Ministerio de Agricultura y Ganaderia
1994 *Ecuador en cifras; Compendio estadistico agropeuario 1965-1993.* MAG, PRSA, AID, Quito, Ecuador.

Pichón, F.J. and R.E. Bilsborrow
1991 "Agricultural Colonization and the Social Dimensions of Deforestation in the Ecuadorian Amazon." Paper presented at the Workshop on the Social Dynamics of Deforestation in Developing Countries by UNRISD, Aberdare, Kenya

Pichón, F.J. and R.E. Bilsborrow
1992 *Land Use Systems, Deforestation and Associated Demographic Factors in the Humid Tropics: Farm-level Evidence from Ecuador.* Paper presented for the IUSSP Seminar on Population and Deforestation, Campinas nov. 92.

Pichón, F.J.
1993 "Colonización y Deforestación en la Frontera Agrícola de la Región Amazónica Ecuatoriana." In Ruiz, L. (ed.) *Amazonía, Escenarios y Conflictos.* Quito, CEDIME: 337-376.

Pichón, F.J.
1993 *Agricultural Settlement, land use, and deforestation in the Ecuadorian Amazon frontier: A micro-level analysis of colonists' land-allocation behavior.* Dissertation University of North Carolina at Chapel Hill.

Pichón, F.J.
1996 "Settler Agriculture and the Dynamics of Resource Allocation in Frontier Environments."Human Ecology, 24 (3), pp. 341- 371.

Pichón, F.J.
1997 "Settler Households and Land-Use Patterns in the Amazon Frontier: Farm-Level Evidence from Ecuador." World Development, 25 (1), pp. 67-91.

Pijnenburg, T.
1991 *Accumulacion de capital en una zona de colonizacion de la provincia del Napo-Ecuador.* MAG/PRONAREG-RUU.

Profors
1992 *Diagnostico socio-economico de la provincia de Sucumbios.* INEFAN, GTZ. Quito, Ecuador.

Rudel, T.K.; with B. Horowitz
1993 *Tropical deforestation, small farmers and land clearing in the Ecuadorian Amazon.* New York, Columbia University Press.

Schoorl, J.M., A.Veldkamp and L.O. Fresco
1997 The Conversion of Land Use and its Effects (CLUE-CR), a regression based model applied to Costa Rica (Pascal version 1.2)."In: *Quantitative Approaches in Systems Analysis* 8: 1-53.

Singh, I., L. Squire and J. Strauss (eds)
1986 *Agricultural household models.* Johns Hopkins University Press, Baltimore.

Southgate, D., R. Sierra and L.Brown
1991 "The cause of tropical deforestation in Ecuador: a statistical analysis". *World development*, Vol.19, No.9, pp.1145-1151.

Southgate, D., M. Whitaker
1992 “promoting resource degradation in Latin America: Tropical deforestation, soil erosion, and coastal ecosystem disturbance in Ecuador”. *Economic development and cultural change.* Vol.40, No.4, pp787-807.

Thapa, K.T., R.E. Bilsborrow and L. Murphy
1996 “Deforestation, Land Use, and Women's Agricultural Activities in the Ecuadorian Amazon.” World Development, Vol.24, No.8, pp.1317-133.

Van Soest, D.
1994 *Tropical rainforest in Cameroon: a socio-economic approach.* University of Groningen.

Verburg, P.H.
2000 *Exploring the spatial and temporal dynamics of land use with special reference to China.* PhD thesis. Wageningen University and Research Centre, Wageningen.

Veldkamp, A. and L.O. Fresco
1996 “CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica.” In: *Ecological modelling* 91:231-248.

Wilkie, D.S. and J.T. Finn
1988 "A spatial model of land use and forest regeneration in the Ituri forest of Northeastern Zaire." In: *Ecological Modelling* 41:307-323.