Ecology and Development Series No. 68, 2009

Julia Schindler

A multi-agent system for simulating land-use and land-cover change in the Atankwidi catchment of Upper East Ghana

1. Referent: Prof. Dr. P.L.G. Vlek
2. Referent: Prof. Dr. E. Ehlers
Tag der Promotion: 24.09.2009
Erscheinungsjahr: 2009

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Schindler, Julia (2009). A multi-agent system for simulating land-use and land-cover change in the Atankwidi catchment of Upper East Ghana. Ecology and Development Series Bd. 68

ISSN 1864-0443
ISBN 978-3-940124-19-7

Herausgeber:
Paul L.G. Vlek, Manfred Denich, Christopher Martius, Ahmad Manschadi,
Janos Bogardi
Zentrum für Entwicklungsforschung (ZEF)
Abteilung Ökologie und Ressourcennutzung (ZEFc)
Walter-Flex-Straße 3
53113 Bonn
Tel: +49-228-73-1865
Fax: +49-228-73-1889
E-Mail: eds@uni-bonn.de
Druck: Hausdruckerei der Universität Bonn

Für meinen Vater
Hansgeorg Schindler
(1943-2001)

ABSTRACT

Land-use and land-cover change (LUCC), which is a general term for the human modification of the Earth's terrestrial surface, increasingly gains attention in the scientific community, due to its vast global extent and the role it plays in the Earth system functioning. About one third to one half of the global land surface has been modified by humans, and these changes are highly interrelated with many environmental, economic and social processes and problems. However, studies on LUCC processes are often challenged by the complex nature and unexpected behavior of both human drivers and natural constraints. Many studies tend to focus either on the human or the environmental part of LUCC systems, thus neglecting the interrelationships and responses among these two components. Many aspects of complexity can be overcome by a multi-agent based approach, whose design allows an integrated representation of the feedbacks, hierarchies and interdependencies of the coupled human-environment system of LUCC. A multi-agent simulation model (GH-LUDAS - GHana Land Use DynAmic Simulator) was developed to model this coupled human-environment system in a small-scale catchment in Ghana, thereby providing a simulation tool to predict land-use/cover patterns as related to socio-economic indicators. Apart from pure prediction, the aim of the model is to explore alternative future pathways of LUCC under selected policy, demographic and climatic conditions in order to provide stakeholders with support for making better-informed decisions about land resource management.

Multi-agent based modelling is an approach to design computational models for simulating the actions and interactions of autonomous individuals (i.e. agents) in a network, with a view to assessing their effects on the system as a whole. Thus, agent-based modeling can be regarded as a bottom-up modeling approach, as the behavior and interactions of single agents are specified, and complexity is considered to emerge from these specifications. Following this mindset, GH-LUDAS consists of four modules, which represent the main components of the human-evironment system of LUCC. The Human Module consists of collections of human agents (i.e. farm households), which are endowed with a set of attributes and autonomous behavior templates (i.e. the Decision Module), regulating land-use related decisions in response to the human agent's attributes and those of its environment. The Landscape Module consists of collections of individual landscape agents (i.e. land patches of size 30 x 30 m), which are characterized by biophysical attributes and ecological mechanisms, which work in response to human decision-making and natural constraints (e.g. crop yield, land-cover change). The Global-policy Module consists of a range of external parameters, which allow the exploration of alternative future pathways of LUCC, and which relate to attributes of both human and landscape agents. The ability to provide an integrated representation of these components is one of the strengths of this approach, and its flexibility allows the upgrading and modification of processes where these have not yet been considered.

The developed model was applied to a small-scale catchment in Upper East Ghana, the Atankwidi catchment, which covers an area of about 159 km^2. Spatially explicit data were obtained from an ASTER image, digital maps, an extensive land cover inventory and intensive household surveys. Field data were used to specify attributes and calibrate behavioral submodels of households and land patches. Considered external factors were the policies of dam construction and credit access, demographic changes, and rainfall change. Simulation outputs consist of a spatially and temporally explicit land use/cover map, visual graphs, and export

files of selected land-use and livelihood indicators. These convenient output visualization tools, together with the user-friendly interface of GH-LUDAS, allow stakeholders to simulate and analyze selected scenarios, which can serve as a basis for discussion and communication among stakeholders and policy-makers.

Simulation results suggest that, among others, the policy of dam construction had much less effect on average annual income than that of credit provision, although it is the much more costly option in comparison to a credit scheme. Furthermore, a decline in annual rainfall seemed to trigger a shift towards cash cropping and non-farm activities, which could compensate for the losses in harvest caused by decreased precipitation. All simulated spatio-temporal data developed by these simulations can be used for further scientific analyses using GIS and statistical packages, thereby providing a basis for further understanding of local LUCC processes in Northern Ghana.

KURZFASSUNG

Ein agenten-basiertes Modell zur Simulierung von Landnutzungs- und Landbedeckungsänderungen im Einzugsgebiet des Atankwidi in Nordost-Ghana

Landnutzungs- und Landbedeckungsänderungen, die die durch den Menschen verursachte Modifizierung der Landoberfläche der Erde bezeichnen, erfahren zunehmende Aufmerksamkeit in der wissenschaftlichen Welt, aufgrund ihres weltweiten Ausmaßes und der Rolle, die sie für die Funktionsweise der Erde spielen. Zwischen einem Drittel und der Hälfte der Landoberfläche sind bereits durch menschliche Einflüsse verändert worden, wobei diese Änderungen wichtige Wechselbeziehungen mit ökologischen, ökonomischen und sozialen Prozessen und Problematiken aufweisen. Studien, die sich mit Landnutzungs- und Landbedeckungsänderungen befassen, repräsentieren die Komplexität menschlicher Verhaltensweisen und ökologischer Bedingungen oft nur in unzureichender Weise. Viele Studien tendieren dazu, nur eine Komponente des ökologischen Systems, das aus menschlichen wie aus umweltbedingten Prozessen besteht, zu erfassen, und vernachlässigen dabei die Wechselbeziehungen zwischen diesen beiden Komponenten. Der agenten-basierte Modellierungsansatz hat die Fähigkeit, viele Eigenschaften von komplexen Systemen zu integrieren, und ermöglicht die Modellierung von Rückkopplungen, Wechselbeziehungen und skalen-abhängigen Prozessen des ökologischen Systems. In dieser Arbeit wurde ein agenten-basiertes Modell namens GH-LUDAS (Ghana - Land Use DynAmic Simulator) entwickelt, das Landnutzungs- und Landbedeckungsänderungen sowie zugehörige sozio-ökonomische Indikatoren in einem Flusseinzugsgebiet des White Volta in Nord-Ghana simuliert. Das Ziel des Modells ist sowohl die Prognostizierung von Landbedeckungs-/Landnutzungsänderungen als auch die Evaluierung von möglichen Zukunftsverläufen unter gegebenen politischen Maßnahmen, demographischen Veränderungen sowie Klimawandel. Die Simulierung solcher Szenarien kann die Entscheidungsfindungen lokaler Akteure bezüglich Landnutzung unterstützen und als Ausgangspunkt für Diskussionen unter lokalen Entscheidungsträgern dienen.

Der agenten-basierte Modellansatz kennzeichnet sich durch die Modellierung der Aktionen und der Interaktionen einzelner Individuen (i.e. Agenten), deren Spezifikationen in komplexe Phänomene auf Systemebene resultieren. Agenten-basierte Modellierung kann daher als ein ’bottom-up approach’ bezeichnet werden, da die Systembeziehungen nicht auf oberster Ebene spezifiziert werden, sondern von den Prozessen zwischen einzelnen Agenten reguliert werden. Dieser Philosophie folgend, gliedert sich GH-LUDAS in vier Hauptmodule. Das soziale Modul besteht aus einer Kollektion von menschlichen Agenten, die landwirtschaftliche Haushalte repräsentieren, und die mit einer Reihe von Attributen und Entscheidungsalgorithmen ausgestattet sind. Diese Algorithmen, die innerhalb des Entscheidungmoduls spezifiziert sind, regulieren Reaktionen auf persönliche wie auf umweltbedingte Attribute und Prozesse. Das Umweltmodul besteht aus landschaftlichen Agenten, die aus Pixeln von 30 m x 30 m bestehen, und die mit eigenen Attributen sowie ökologischen Mechanismen, die auf menschliche Entscheidungen sowie auf natürliche Prozesse reagieren (z.B. Ernteertrag, Landbedeckungsänderungen), ausgestattet sind. Das globale Modul besteht aus einer Reihe von externen Parametern, die von Modellnutzern reguliert werden können, und die Attribute von menschlichen und landschaftlichen Agenten direkt beeinflussen. Die Fähigkeit, diese Komponenten zu verbinden und miteinander zu integrieren, ist eine der

Stärken des agenten-basierten Ansatzes, und seine Flexibilität erlaubt die Integrierung von Prozessen, wo diese (noch) nicht berücksichtigt worden sind.

Das Modell wurde speziell für das Flusseinzugsgebiet des Atankwidi in Nordost-Ghana entwickelt, das eine Fläche von etwa 159 km^2 aufweist. Räumlich explizite Daten wurden auf der Basis eines ASTER Satellitenbildes, digitalen Karten, einer weiträumigen Bestandsaufnahme von Landbedeckung, und fokussierten Haushaltsbefragungen generiert. Auf diesen Felddaten basierend, wurden die Attribute sowie die reaktiven Mechanismen menschlicher und landschaftlicher Agenten spezifiziert und kalibriert. Die externen Parameter des Modells umfassen Maßnahmen, die Dammbau und Kreditvergabe betreffen, sowie demographische Veränderungen und Reduzierung des jährlichen Niederschlags. Die Ausgabe der Modellsimulationen erfolgt durch eine zeitlich und räumlich explizite Visualisierung von lokaler Landbedeckung/Landnutzung, Graphiken, und exportierbaren Dateien einer Auswahl an Systemindikatoren. Diese Bandbreite von Ausgabemöglichkeiten, in Kombination mit einer benutzerfreundlichen Modelloberfläche ermöglichen beteiligten Akteuren, ausgewählte Szenarien zu simulieren und zu analysieren, und kann zur Diskussion und Kommunikation zwischen Akteuren und Entscheidungsträgern beitragen.

Die Resultate von bereits simulierten Szenarien deuten unter anderem darauf hin, das die Strategie des Dammbaus eine geringere Wirkung auf durchschnittliches Einkommen hat als die Maßnahme der Kreditvergabe, obwohl ersteres die bei weitem kostspieligere Maßnahme darstellt. Desweiteren zeigt sich, dass eine Reduzierung des jährlichen Niederschlags eine Verlagerung auf marktfähigere Agrarprodukte (cash crops) und nichtlandwirtschaftliche Einkommensstrategien auszulösen scheint, die die Reduzierung des Ertrags, verursacht durch die geringere Niederschlagsmenge, kompensieren. Alle simulierten zeitlichen und räumlichen Daten können weiteren wissenschaftlichen Analysen in GIS- und Statistik-Programmen unterzogen werden, und zu einer Erweiterung des Verständnisses von lokalen Landnutzungs- und Landbedeckungsänderungen in Nord-Ghana beitragen.

TABLE OF CONTENTS

1 MULTI-AGENT SYSTEMS FOR SIMULATING LAND-USE/COVER CHANGE

1.1 Introduction

Land-use and land-cover change (LUCC) also known as land change is a general term for the human modification of the Earth's terrestrial surface. Though humans have been modifying land to obtain food and other essentials for thousands of years, current rates, extents and intensities of LUCC are far greater than ever in history, driving unprecedented changes in ecosystems and environmental processes at local, regional and global scales (Ellis, 2007). These changes encompass the greatest environmental concerns of human populations today, including climate change, biodiversity loss and the pollution of water, soils and air. Monitoring and avoiding the negative consequences of LUCC while sustaining the production of essential goods and services has therefore become a major priority of researchers and policy makers around the world (Ellis, 2007).

In order to understand the nature of LUCC, it is important to clarify terminology and definitions used in the field of LUCC research. While land cover is the biophysical state of the Earth's surface and immediate subsurface, the term land use refers to the involvement of both the manner in which the biophysical attributes of the land are manipulated and the intent underlying that manipulation - the purpose for which the land is used (Briassoulis, 2000; Turner et al., 1995). This way, land cover means the physical, chemical, or biological categorization of the terrestrial surface, e.g. grassland, forest, or concrete, whereas land use refers to the human purposes that are associated with that cover, e.g. raising cattle, recreation, or urban living (Meyer and Turner, 1994).

In the analysis of land-use and land-cover change, it is also necessary to depict the term of change in this respect. In land-use/cover research, there are two forms of LUCC: conversion (i.e. the complete replacement of one land-cover/land-use type by another) and modification (i.e. more subtle changes that affect the character of the land cover/land use without changing its overall classification) (Turner et al., 1993). The conversion of forest to crop land is an example of land-cover conversion, whereas the change in the composition or health of a forest can be regarded as a modification within this land-cover class (i.e. forest). Accordingly, changes in land use can be in form of both conversions and modifications. As the replacement of one agricultural type by another (e.g. from rainfed to irrigated agriculture) can be seen as the conversion from one land-use type to the other, modifications of one single land-use type might include, for instance, the intensification of crop production, without changing its land-use classification.

The recognition of the importance of such changes in land use and land cover for the Earth system's functioning already emerged in the mid 1970s, when studies revealed the significance of the relationships between land-cover and climate change. At this time, it was recognized that land-cover change may induce changes in the albedo, and thus modify the surface-atmosphere energy balance, resulting in regional and global climate change (Otterman, 1974; Charney and Stone, 1975). In the subsequent decades, it was discovered that land-cover change does not only modify climate through a changed energy balance, but also through the creation and especially diminishment of carbon sinks, thus contributing to global carbon emissions (Lambin et al., 2003). During the following years, many other consequences of land-use/cover change were identified, showing severe impacts on the ecosystem, including soil degradation, desertification, a loss of biodiversity, declining human health, and the threat to the ability of biological systems to support human needs (Vitousek et al., 1997). As the Earth is a complex system of biogeochemical cycles and energy fluxes, which are largely regulated by the land surface, the understanding and monitoring of processes related to land-use/cover change is crucial to the understanding of global dynamics.

In the following, we will depict the five most well-known forms of LUCC in order to understand the relevance and the magnitude of land-change processes. Deforestation is one of the most commonly recognized forms of land-cover change (Williams, 2003). According to FAO (FAO, 2001), deforestation occurs when forest is converted to another land cover or when the tree canopy falls below a minimum of 10%. On the basis of this definition, it is estimated that the world's natural forests decreased by 16.1 million hectares per year on average during the 1990s (FAO, 2001). Until today, that is a loss of about 5 % of the natural forests that existed in 1990. The reasons for this reduction are manifold and are highly dependent on the region. Whereas in Latin America large-scale forest conversions are mainly due to the expansion of livestock agriculture (Lambin et al., 2003), deforestation in Africa is mainly a result of cropland expansion. In Asia, intensified shifting agriculture, including migration into new areas, and logging explain most of the deforestation (Achard et al., 2002).

The consequences of deforestation for the ecological system are manifold: First, deforestation can lead to soil erosion or impoverishment, especially in tropical areas where soils tend to be thin and nutrient-poor. Second, deforestation is linked to habitat loss, which is a leading cause of species endangerment and biodiversity loss, particularly in humid tropical forests. Third, it affects the hydrological cycle through changes in evapotranspiration and runoff. And last but not least, deforestation, and particularly forest burning, contributes to greenhouse gas emissions that bring about climate change (SEDAC, 2002).

A major trend of global LUCC is the expansion of agricultural land. Currently, agricultural land covers about a third of the global land surface, and has expanded into forests, but also steppes and savannahs, to meet the growing demand for food (Lambin et al., 2003). Such conversions involve a change of the whole local ecosystem, e.g. changing animal habitats and faunas, thus being a direct threat to biological diversity. However, not only the conversion to cropland plays a role in global change, but also the intensity in agricultural management. Historically, humans have increased agricultural output mainly by bringing more land into production. This process of agricultural expansion was gradually replaced in the 1960s by a process of intensification in some regions of the world, i.e. an increase in food production per hectare, being mainly achieved through mechanized tillage, fertilizer use and irrigation. Such agricultural practices contribute to carbon emissions through several mechanisms: the direct use of fossil fuels in farm operations, the indirect use of embodied energy in inputs that are energy intensive to manufacture (e.g. fertilizers), and the cultivation of soils resulting in the loss of soil organic matter (Ball and Pretty, 2002). Furthermore, the use of freshwater for irrigation and the use of fertilizers lead to a modification of the water and nutrient cycles, especially the nitrogen cycle.

Natural vegetation cover has not only given way to cropland, but also to pastures, which are defined as land used permanently for herbaceous forage crops, either cultivated or growing wild (FAO, 2004). The distinction between pasture and natural savannah or steppes is not always clear. However, FAO statistics suggest that most pastures are located in Africa (26 % of the global total of 35 million ha), followed by Asia (25 %) (Lambin et al., 2003). During the last decade, pastures increased considerably in Asia and the former Soviet States, which is mainly due to the tremendous increase in the demand for meat (Mooney and Neville, 2005). To meet the growing demand, total meat production is projected to double by 2020

(Mooney and Neville, 2005). In response to this increase, industrialized animal production systems are proliferating, and consequently result in complex negative externalities with respect to the environmental sustainability of livestock production.

The resulting concentrated waste production from these systems and its effects on terrestrial and aquatic ecosystems is a serious matter, with stored liquid manure producing over 13 million tons of the greenhouse gas methane per year (de Haan et al., 1997). In addition, the massive global trade in grains for animal feed has greatly altered regional water and bio-geochemical balances.

Finally, urbanization can also be ranked among the most well-known frontiers of LUCC. Since urban areas occupy a relatively small fraction of the Earth's surface (i.e. 2 %) (Gruebler, 1994)), this relatively small fraction of urbanized areas may lead to the misconception that urbanization can be ignored in land-change studies (Heilig, 1994). In reality, urbanization affects land change elsewhere at a large scale through strong linkages between urban and rural areas (Lambin et al., 2001). Furthermore, raising living standards of the growing urban population around the world tend to raise the consumption expectations, leading to local and global changes in land-use intensity.

When aggregated globally, such LUCC do not only endanger the biotic diversity world-wide (Lambin et al., 2001) but also contribute to changes in the energy, hydrological and biogeochemical cycles of the Earth's system, thereby leading to climate and ecosystem change, thus affecting the ability of biological systems to support human needs (Vitousek et al., 1997). It is therefore of utmost importance to understand the processes involved, to anticipate future land-use/cover patterns, and to find strategies to mitigate the adverse impacts of such land-use/cover changes. The ability to project future LUCC and its socio-ecological consequences depends on our ability to understand the past, current, and future drivers of land-use and land-cover change (USGCRP, 2003). However, relationships between driving forces and phenomena of LUCC are highly complex and interwoven, thus hampering the establishment of a general theory of these relations. An attempt to derive a theory through the identification of specific typical pathways of land-use/cover change has been made by Lambin and Geist (Lambin and Geist, 2006), based on a review of 132 case studies around the world. However, instead of repeating these pathways and demonstrating typical drivers of land-use change, we will rather focus on the aspects of the complexity that is exhibited by such processes of land-use and land-cover change, as the understanding of this complexity is the first step for a reliable representation of the involved processes.

1.2 The complexity of the coupled human-environment system of land-use/cover change

The complex nature of land-use/cover change is mainly due to the complex way in which humans and the environment interact in response to each other, whereby these interactions are regulated by a wide range of factors influencing land-use decisions at different temporal and spatial scales. Feedback mechanisms among the components of this coupled human-environment system even enhance the level of complexity, possibly resulting in an emergent land-use/cover pattern, which cannot be explained by an analysis of the single constituents of the system (Parker et al., 2003).

As an understanding of the way such a system works is crucial for a reliable analysis or synthesis of land-use/cover change processes, in this chapter this complex nature of land-use systems is characterized. Land-use/cover systems are complex, and the notion of complexity has consequences for the way the system should be described (Kok, 2001). However, complexity science is still in its infancy (Goldstein, 1999), and there is no common definition of complex systems shared by the various involved disciplines (Manson, 2005). With respect to land-use systems, Parker et al. (2003) define complex systems as 'dynamic systems that exhibit recognizable patterns of organization across spatial and temporal scales'. In complexity science as well as in ecological sciences, complexity is often discussed in the two different dimensions: functional and structural complexity (see Bandte, 2007; Lambin and Geist, 2002; Kok et al., 2000). In the following, we will summarize the characteristics of LUCC complexity with regard to both aspects.

1.2.1 Functional complexity

According to Marks (2007), functional complexity of a system is the complexity of the mappings from inputs to outputs, whereby the system itself is regarded as a black box. More precisely, the complexity of the mode of operation of the system is examined by determining the effect of variation of the input on the system output (Bandte, 2007) without considering the internal mechanisms. Within land-use system research, functional complexity thus refers to the complexity in which variations of driving forces (i.e. explanatory factors) of land-use change influence land-use/cover patterns. This complexity is driven by the large variety of explanatory factors, their variation in both time and space, thereby being episodic or progressive, and their high level of interlinkages, thus having a synergetic effect on land-use/cover patterns. In the following, we will outline the complexity of these driving forces for LUCC, and justify the significance of this complexity on land-use/cover patterns through examples.

Multitude of driving forces

Land-use change is always caused by a multitude of interacting factors originating from different levels of organization of the coupled human-environment system (Lambin et al., 2003). At the local level, causes of land-use/cover changes involve a physical action on land cover such as agriculture, forestry and infrastructure construction (Lambin and Geist, 2006). Such proximate causes generally operate at the level of individual farms, households or communities (Lambin et al., 2003; Mather, 2006). At the regional to global level, underlying factors are fundamental forces that underpin such proximate causes, covering a wide range of political, economic, demographic, technological, cultural and biophysical factors. Changes in any of these indirect drivers usually result in changes in one or more of the proximate factors, thus triggering land-use/cover changes (Lambin and Geist, 2006). Due to this wide variety of driving forces operating at different scales and a frequent sensitivity of land-use/cover patterns to any of these forces, the output-input relations of the coupled human-environment system underly a high level of complexity.

Multiple causality in LUCC

Driving forces of land-use/cover change not only include variables from a wide range of factors, but also are highly interrelated with each other. As such, underlying forces do not only influence proximate causes in a mediated fashion, but are often shaped themselves by other factors. For example, population increase in a given area - often considered an underlying cause of land change - may be amplified or modulated by existing or changing social norms or by fertility or resettlement programs, which may in turn be influenced by changes in knowledge and policy at national and international levels (Lambin and Geist, 2006). It is helpful to recognize that some factors concern the motivation to change behavior, while others function in contextual ways, often filtering the effects of other factors (Turner, 1989; Moran, 2005). The interplay and interrelations between such driving forces amplify the complexity of the system functioning, resulting in land-use/cover patterns often difficult to predict.

Temporal and spatial variation of driving forces

Driving forces of land-use and land-cover change are not only highly interrelated, but also can vary both in time and space, whereby the strength of their interrelations is also temporally and spatially variable. An example for the spatial variability of driving forces and their effect on land use is given by Lambin and Geist (2006) who describe a typical pathway of land-use intensification dependent on local market opportunities and population pressure. As such, land scarcity-driven agricultural intensification occurs in economies that are not fully integrated in the market, and is usually linked to population growth and density (Lambin and Geist, 2006). Thus, regional variations in market opportunities and population dynamics may lead to totally detrimental outcomes in agricultural intensification, and ultimately land-use and land-cover patterns.

With respect to the temporal variation of driving forces, climate change and its effect on land-use/cover is a widely cited example. For instance, it has been shown that a reduction in rainfall in West Africa shortens the length of the growing period and has a considerable impact on potential crop yields and their variability (Voortman, 1998), thus having a direct effect on the survival strategy of farming households and ultimately land-use choice. Furthermore, it is important to distinguish between gradual and episodic changes (Lambin et al., 2003). Episodic changes show periods of rapid and abrupt changes and can have a completely different impact on land use than progressive changes. Such short-term changes, often caused by the interaction of climatic and land-use factors, have an important impact on ecosystem processes. For example, droughts in the African Sahel and their effects on vegetation are reinforced through a feedback mechanism that involves land-surface changes caused by the initial decrease in rainfall (Zeng et al., 1999).

1.2.2 Structural complexity

In contrast to functional complexity, structural complexity refers to the level of complexity of the internal functioning of the system (Bandte, 2007). Within ecology and land-use system sciences, structural complexity of LUCC systems is usually described by three characteristics of internal complexity, comprising interdependencies, heterogeneity, and nested hierarchies (Arthur et al., 1997; Epstein, 1999; Holland, 1998; LeBaron, 2001; Manson, 2001). Many examples of these three key sources of complexity can be identified in human-influenced landscapes (Parker et al., 2003). Furthermore, an important feature of LUCC complexity is the evolvement of emergent phenomena at the higher scales of human and biophysical systems. The term 'emergence' refers to system's properties that are not analytically tractable from the attributes of the internal components (Baas and Emmeche, 1997). More intuitively, an emergent property may be defined as a macroscopic outcome resulting from synergies and interdependencies between lower-level system components. In the following, a description of these four key sources of complexity with respect to land-use and land-cover change is given.

Nested hierarchies and scale dependency

It has long been apparent to ecologists that ecological systems are hierarchically structured (e.g. Egler, 1942; Schultz, 1967). Hierarchy, in mathematical terms, is a partially ordered set, which is a collection of parts with ordered asymmetric relationships inside a whole. In less mathematical terms, the system works as an organization of levels at different scales, whereby phenomena at a certain level of scale are explained by processes operating at the immediate lower level, but are, on the other hand, constrained by processes operating at the immediate higher level (Le, 2005). The result is a so-called 'constraint envelope' among the involved hierarchical levels.

An example of such a 'constraint envelope' is the reproduction behavior of a single organism. The internal reproduction process of the organism is determined by the operation and interaction of the single subcomponents of the organism, while the actual reproduction behavior is constrained by characteristics of the whole population made up of all organisms (e.g. population density). LUCC systems are usually described as nested hierarchies among human and natural subsystems, which involve levels consisting of, and containing, lower levels. As such, individual waterways join to define nested watersheds, and assemblies of individual species members aggregate to form communities.

Processes involved in the functioning of the system usually operate along the different levels of this organized hierarchy, whereby processes at the higher levels proceed slower but to a larger extent, and processes at the lower levels proceed faster but to a smaller extent (Easterling and Kok, 2003). In LUCC, such lower-level processes might refer to direct land-use decisions made at the household level, which have an immediate but short-term consequence on the local environment. Higher-level processes, on the other hand, might include the aggregated land-use behavior of the whole population, which influences land-use and land-cover patterns at the landscape level, but at a lower pace. This difference of type and pace of processes induced by the difference of scale is called scale dependency.

Evidence from case studies suggests that these scale-dependent processes are also driven by scale-dependent factors. Variations in explanatory variables of land-use change across scales usually follow a consistent pattern: at farm scale, such explanatory factors comprise mostly social and accessibility variables, at landscape scale such factors might include topography and agro-climatic potential, and at the regional to national scale climatic variables as well as macro-economic and demographic factors can be identified as land-use drivers (Veldkamp and Lambin, 2001). For the establishment of a realistic representation of processes of land-use change, the existence of hierarchies, the scale-dependency of processes, and drivers operating at different scales of this hierarchy need to be considered.

Interdependencies and feedback loops

Interdependencies exist among all components of the coupled human-environment system, both in time and space. These interdependencies exist along the horizontal axis as well as along the vertical axis of the nested hierarchy levels (Lambin et al., 2003). On the human side, land-use decisions might be influenced by both the land-use history of the land manager and those of others (temporal interdependency), and by the attributes of their surrounding environment (spatial interdependency) (Parker et al., 2003). These spatial influences on agent behavior may include flows of information, diffusion of technology, spatial competition, local coordination, social networks, and positive and negative externalities among neighbors (see Case, 1991; Irwin and Bockstael, 2002; Krider and Weinberg, 1997; Lansing and Kremer, 1993; Miyao and Kanemoto, 1987; Parker, 2000; Ray and Williams, 1999). On the biophysical side, spatial interactions may include slope processes, up- and down-stream effects, connectivity of natural habitats and ecological edge effects (Parker et al., 2003).

Webs of interdependencies among system variables and components form a complex network of transforming feedback loops (Eoyang, 1997). These loops carry material, energy and information from one system component to another (Eoyang and Berkas, 1998). Positive feedback loops tend to amplify system behavior, whereas negative feedback loops usually counteract the amplification as stabilizers of the system. An example of a positive feedback loop is the downward spiral of frontier deforestation. Immigrants might clear forest for crop production, which causes the expansion of agricultural activities. This inappropriate use of forest soils often results in land degradation and low soil fertility, which finally amplifies the deforestation process.

Such feedback loops in LUCC systems bring forth that drivers of land-use change can themselves be modified by land-use changes, i.e. they are not purely exogenous but also endogenous to the system (Lambin et al., 2003). For instance, demographic changes can result in changes of land use and land cover, but these changes might influence demographic patterns in turn. In general terms, the changes in ecosystem goods and services that result from land-use change lead to important feedbacks to the drivers of land-use change (Lambin et al., 2003), thus again causing changes in land-use patterns.

Heterogeneity

The consideration of heterogeneity within LUCC systems is often important to ensure a realistic representation of the landscape as well as of the human agents. For example, heterogeneity among land managers can be reflected by differences in values, ability, resources and experience, which might have an influence on land-use decisions. On the environmental side, spatial heterogeneous factors important for land-use decisions might include differences in soil quality, water availability, topography and vegetation (Parker et al., 2003). This heterogeneity of both land managers and the biophysical environment might also change over time, due to interactions among these two components.

When heterogeneity and interdependencies are combined in a model, analytical solutions may be very difficult to obtain. The adoption of a new technology is such an example in which both agent heterogeneity and spatial interdependencies are important (Parker et al., 2003). Here, the spatial heterogeneity is represented by the variability of risk aversion among land managers to adopt the new technology. The information of the success or failure of those land managers who take the risk may spread among the neighboring land managers, the process of which represents spatial interdependency. Thus, the spatial distribution of agent types with different risk aversion over space may influence the spatial extent of adoption. This way, regions of adoption and non-adoption may emerge as a result of local heterogeneity and spatial interdependencies between land managers. In models that feature both heterogeneity and interdependencies, usually many possible stable equilibria exist. These equilibria are usually dependent on the initial state of the model, which is called path dependency. With respect to the example of technology adoption, the presence of a single land manager willing to adopt the new technology is required to initiate a cascade of technology adoption among neighboring land users. This way, two equilibria are possible: one with adoption, and one without, dependent on the initial state of the model in terms of heterogeneity.

Emergent phenomena

If researchers are specifically interested in modeling the complex dynamics of a LUCC system, they also may be specifically interested in understanding the macroscopic, or emergent, phenomena that may result. Emergent phenomena are described as aggregate outcomes that cannot be predicted by examining the elements of the system in isolation. Emergent phenomena exhibit structures that are not explained by lower-level dynamics and typically persist beyond the average lifetimes of entities upon which they are built (Crutchfield, 1994). More intuitively, an emergent property may be defined as a macroscopic outcome resulting from synergies and interdependencies between lower-level system components.

With respect to LUCC, land-use change at the landscape scale can be regarded as the aggregation of the multiple small land-use changes, which reinforce or cancel each other (Lambin et al., 2003). These small changes are the result of the decisions of land managers under certain socio-economic and environmental conditions, which are, in most cases, made independently without a central direction. Thus, land-use change is a complex large-scale spatial behavior that emerges from the aggregate interactions of less complex land managers (Lambin et al., 2003). This way, the behavior of the coupled human-environment system at the landscape scale can be regarded as an emergent phenomenon resulting from low-level actions and interactions, which makes the behavior of the system unpredictable in most cases.

1.2.3 The importance of modeling LUCC

Given the diversity of complexity in which LUCC systems operate, we will argue in this section why a modeling approach can be a useful tool to integrate and consider such complexity, thereby providing a tool to understand and predict land-use/cover changes. The analysis of the multiple interactions of land-use/cover change (see Introduction) with the Earth system suggests that the understanding of the role of LUCC within this system deserves considerable attention. Based on the urgency of monitoring land-use/cover change processes, as they are highly interrelated with bio-geochemical global and regional cycles, soil and forest degradation, and biodiversity, reliable approaches to understand and predict LUCC processes are needed. Based on this background, the two main targets within the LUCC research community can be summarized as follows: i) the projection of alternative pathways in the future, and ii) the development of hypotheses about the functioning of LUCC systems, whereby both require the understanding of involved processes, which underly a high level of complexity.

Although humans build 'mental models' when faced with complex phenomena, the ability to fully capture all aspects of complex systems and ultimately make predictions is limited, as human mental models tend to simplify systems in particular ways (Costanza and Ruth, 1998). Humans base most of their mental modeling on qualitative rather than on quantitative relationships, linearize the relationships among system components, disregard temporal and spatial lags, and treat systems as isolated from their surroundings (Costanza and Ruth, 1998). When problems become more complex, and when quantitative relationships, non-linearities and time and space lags are important, as is the case for LUCC systems, human mental models need to be supplemented. When models are built with consideration of these different aspects of complexity, they can serve as useful tools to understand and predict future land-use/cover patterns.

Reliable projections of alternative pathways into the future are important, as increasing evidence suggests that a proactive land management instead of a reactive one is needed. Proactive management, in contrast to reactive management, which tries to reverse environmental damages that occurred in the past, attempts to find strategies to avoid damage in the future. This current shift to a proactive view is based on the evidence that environmental damage, once done, is very diffcult to undo (Le, 2005), implying that maintaining ecosystems in the face of changes requires active management for a foreseeable future (Vitousek, 1997). Models, in this respect, can serve as useful tools to predict future patterns of land-use/cover, and possibly help to find strategies to mitigate future adverse impacts on the natural resource base, or even enhance the sustainability of the use of these resources.

Apart from the assessment of alternative future pathways of LUCC, the second main target that can be approached by models is to provide a tool to test hypotheses about the LUCC system functioning. Authors within the LUCC research community argue that the understanding of land-use processes still lacks a valid theory (Couclelis, 2001), which also impedes the development of reliable LUCC models. However, although current models might rely on a weak theoretical basis, models in turn are often a useful tool to develop the understanding of LUCC processes, thereby helping to establish a theory for a future generation of models. In contrast to models used to predict future patterns, which try to be as realistic as possible, explanatory models may be hypothetical, thereby focusing on system aspects that are intended to be explored (Parker et al., 2003), thereby ignoring others. Such models may be used to understand the key processes underlying land-use systems (Parker et al., 2003), to test the sensitivity of land-use/cover patterns to variations in driving forces (Veldkamp and Lambin, 2001), and to assess system stability.

1.3 Modeling LUCC

Due to this urgency to project and understand land-use change processes, LUCC modeling has attracted more attention in recent years in research fields related with global environmental issues (Shibasaki, 2003). A range of LUCC models has been developed to meet land management needs, and to better assess and project the future role of LUCC in the functioning of the Earth system (Veldkamp and Lambin, 2001).

As land-use change usually depends on both the physical environment of the involved actors and their socio-economic context, processes of land-use change are often modeled as a function of a selection of variables from both domains, acting as driving forces of land-use change. Such driving forces are important in all land-use change models, but their selection and the quantification of the relations between the driving forces and land-use change is very much dependent on the modeling approach chosen. In this chapter, we will present various types of modeling approaches and their strengths and limitations, and will give a reasoning for using an agent-based approach within this study.

1.3.1 Approaches to modeling of LUCC

There are different approaches to modeling of LUCC. Based on model purposes, underlying theories, types of modeled land uses, and the spatial and temporal levels of analysis, Briassoulis (2000) distinguished five main categories of models: i) equation-based models, ii) system dynamics models, iii) empirical-statistical models, iv) cellular automaton models, and v) agent-based models. In the following, we will give short descriptions of each of these approaches, and analyze their capability to integrate structural complexity.

Equation-based models

Equation-based models are models that capture system characteristics by identifying system variables and describing the system with sets of equations relating these variables (Sun and Cheng, 2002). The evaluation of these equations produces the evolution of the system characteristics over time (Huigen, 2003). As equation-based models tend to make extensive use of system-level characteristics (Huigen, 2002), the integration of heterogeneous and interacting low-level entities is generally not considered (Sun and Cheng, 2002). Interaction usually takes place among the system-level variables, although literature review indicates that hierarchies or different levels of organization can possibly be integrated to some extent (e.g., Enge-Rosenblatt et al., 2007). Another major drawback of such models is that a numerical or analytical solution to the system of equations must be obtained, also limiting the level of complexity (e.g. feedback loops) that may practically be built into such models (Parker et al., 2003).

System dynamics models

System models represent stocks and flows of information, material and energy as sets of differential equations linked through intermediary functions and data structures (Gilbert and Troitzsch, 1999). Such models, which are usually broken into discrete time steps, can represent human and ecological interactions, thus allowing feedbacks to operate within the system. Although these kinds of models can address the shortcomings of equation-based models in terms of representing feedbacks and dynamic processes, they also operate at an aggregated level (Parker et al., 2003). As such, heterogeneity and interactions are only considered at a very coarse temporal and spatial resolution. However, similar to equation-based models, such models offer the possibility to integrate hierarchical structures.

Empirical-statistical models

The application of statistical techniques to derive the mathematical relationships between dependent variables and sets of independent variables is widespread in modeling socio-economic and other systems of interest (see Colenut, 1968; Lee 1973). Empirical-statistical models find a set of best-fit model coeffcients that express a statistical relationship between a dependent variable (e.g. LUCC) and a series of independent variables (representing drivers of LUCC). Multiple linear regression techniques are generally used to extract transition probabilities among the states of the landscape (Briassoulis, 2000), which are dependent on the selected drivers. The strengths of such an approach are the ability to provide information on the key drivers of LUCC and the ability to enter and analyze data at various scales.

The disadvantage of such statistical models is that they cannot be transferred spatially in the sense that a regression model that fits well in the region of the variable space usually performs poorly outside that region. Furthermore, these models require a data set on the rates and quantities of change. Thus, these models are only suited to predict changes in land-use intensity where such changes have been measured over the recent past (Briassoulis, 2000).

With respect to the representation of structural complexity, such models can take into account spatial heterogeneity and interaction (Parker et al., 2003) at a single hierarchical level of organization (e.g. Furrer et al., 2007). However, feedbacks across scales and system components cannot be effectively modeled (Parker et al., 2003).

Cellular automaton models

Cellular automaton models consist of a regular grid of cells, each in one of a finite number of states, where cell transitions are based on the state of the current cell and the states of neighboring cells. Such 'neighbors' can be very broadly defined, and may include multi-scale influences. These models are very strong at representing local spatial processes of LUCC, but on the other hand they may place too much emphasis on the local interactions, and not sufficiently represent the human behavior regarding land use. Although cellular modeling techniques offer greater flexibility for representing spatial and temporal dynamics, they have limited ability to reflect feedback mechanisms, as these dynamics are built on stationary transition probabilities (Parker et al., 2003). Apart from this drawback, some extension forms of cellular automata can take into account heterogeneity of the modeled landscape, integrate levels of hierarchy (see Adamides et al.,1992), and consider interaction processes spatially and across hierarchy levels.

Agent-based models (ABM)

Most significant, none of the above modeling techniques can represent the impacts of heterogeneous, autonomous and decentralized human decision-making on the landscape (Parker et al., 2003). Many of the limitations faced by other modeling techniques with respect to a realistic representation of complexity can be overcome by ABM models.

Agent-based models of land-use/land-cover change (ABM/LUCC) usually consist of two key components. The first is a cellular model that represents the landscape under study. This cellular model may draw on a number of specific spatial modeling techniques, such as cellular automata, spatial diffusion models, and Markov models. The second component is an agent-based model (ABM) that represents human decision-making and interactions (Parker et al., 2003). As such, an agent-based model consists of autonomous decision-making entities (agents), an environment through which agents interact, rules that define the relationship between agents and their environment, and rules that determine sequencing of actions in the model. Agent-based models are usually implemented as multi-agent systems (MAS), a concept originated in the computer sciences, which allows for a very efficient design of large and interconnected computer programs.

The potential of MAS/LUCC models is their capacity to represent the co-evolution of human/environmental systems regarding land-use/cover change, by integrating human-related processes with those of nature. By modeling such underlying processes, the emerging dynamics and complexity of this coupled human-environment system can be represented within the model. Furthermore, all aspects of structural complexity can be easily integrated by using MAS, including the heterogeneity on both the landscape and human side, environmental and human hierarchical levels, and spatial and temporal interactions among all components and across hierarchical levels. Furthermore, feedbacks within and between the environmental and Human Module can be effectively integrated.

1.3.2 Opportunities for MAS/LUCC

In general, the benefit of MAS over other modeling techniques is due to its ability to capture emergent phenomena, its ability to provide a natural description of a system, and its flexibility, i.e. the easiness with which processes and components can be integrated (Bonabeau, 2002). The latter quality is perhaps the greatest advantage of MAS/LUCC models. Because these types of models do not need to be solved for closed-form analytical solutions, details critical to the system under study can be easily built in. Furthermore, such flexibility allows researchers to design and execute experiments to explore alternative causal mechanisms, by modifying system processes and components (Parker et al., 2003).

In contrast to many other LUCC modeling approaches, the human and the environmental part as well as their interrelations can be effectively modeled with MAS. Other modeling approaches tend to focus on either part of the LUCC system, thus neglecting the interactive nature of the coupled human-environment system of LUCC. Within MAS, land-use change rather emerges from the interactions among various components of the LUCC system, which then feeds back to the subsequent development of those interactions. Thus, agent-based modeling has the ability to represent the dynamic and non-linear pathway of land-use/cover change.

Furthermore, agent-based models do not impose the relationships among system components, but rather represent individual behavior, which results into emergent patterns at system level through interactions (Huigen, 2003). This way, complexity is modeled from the bottom-up, which makes MAS models being increasingly recognized as useful tools for building a sound theoretical framework to deal with the complexity of LUCC (van der Veen and Otter, 2001; Bousquet and Le Page, 2004). Apart from this ability to capture complex system behavior, MAS can provide a natural description of the human-environment system. Its architecture makes it possible to map the concepts and structures of the real world into the model in ways that preserve natural objects and connections (Bonabeau, 2002). Especially the rapid development in spatial information technology (e.g. GIS, remote sensing) facilitates a realistic specification of the environmental component. New MAS computer platforms (e.g. NetLogo) allow the integration of such a GIS database for landscape specification. User-friendly programming platforms facilitate the programming of agent action and interaction, and allow model users who are not familiar with the model code to easily specifiy model parameters and run simulations.

Due to these strengths, MAS/LUCC models have been recognized as a promising tool to address the complexity of the coupled human-environment system in LUCC modeling (Parker et al., 2003). Within the LUCC research community, recent progress has been made from abstract MAS/LUCC models to more comprehensive presentations of real-world land-use systems. The flexibility in the specifications of agents allows the incorporation of social and ecological processes, and models and approaches of many disciplines can be integrated within MAS. This interdisciplinarity aims at improving a realistic representation of the LUCC system, as land-use/cover change involves the interplay of social, economic and environmental processes.

However, although this approach fulfills many of the requirements for reflecting real-world processes, this approach also has some drawbacks, which will be analyzed in the next section. Furthermore, all of the above models have their strengths, and the choice of the modeling approach is highly dependent on the nature of the object of investigation. Finally, based on the analysis of the shortcomings and strengths of ABM and its suitability for our purposes, we will argue why we decided to use a multi-agent-based approach to study land-use/cover change phenomena in our study area, a small-scale catchment in Upper East Ghana.

1.3.3 Challenges of multi-agent systems for studying LUCC

Although it has been argued that MAS is highly suitable for modeling complex LUCC, there have still been many challenges in its application for real-world land-use systems. Due to the high level of flexibility in the specification and design of MAS, a researcher may easily be trapped in modeling causal and non-causal factors, drivers and processes, important and irrelevant (Huigen, 2003). In addition, model outcomes have to be treated with caution, as 'in every case of simulating complex adaptive systems, the emergent properties are strictly dependent on the rules preprogrammed by the investigator' (Fogel et al., 1999). Thus, an in-depth investigation and understanding of the circumstances and their relevance to land-use processes in the study area needs to be obtained beforehand to avoid a biased selection and design of drivers and processes.

The second challenge of MAS models - if they are meant to be realistic - is the great effort involved in programming and data acquisition, as the behavior of single individuals needs to be modeled explicitly, being mostly dependent on a wide range of factors. Relevant and sufficient data are usually not available and have to be collected. Furthermore, as agent-based models aim to explicitly represent human decision-making, the problem of modeling a highly complex, dynamic spatial environment has shifted to the problem of modeling highly complex, dynamic decision-making units interacting with that environment and among themselves in highly complex, dynamic ways (Couclelis, 2001). This way, the computational and modeling effort of MAS might exceed that of other approaches.

Third, the validation and verification of agent-based LUCC models is a difficult endeavor. Due to the huge parameter space, the model outcomes cannot be captured easily and thus cannot be easily analyzed and validated by formal methods (Huigen, 2003). Furthermore, alongside the increase in computational power and the increased ease of programming, the complexity of models has increased manifold. This increased complexity and the lack of available data for validation hamper the assessment of the degree of realism of MAS models. Therefore, assumptions underlying the functioning of the model have to be clearly stated and justified.

1.4 Problem statement and research objectives

As we have discussed the urgency of predicting and understanding future land-use and land-cover change and the subsequent needs for reliable simulation models, the target of this study is to develop an operational LUCC model, which, in order to serve as a tool for testing the impact of policy interventions, should represent land-use processes and their relation to policies in a realistic way. Since farmers in Africa directly depend on the natural resource base for their living, the prediction of future land-use/cover patterns and related income patterns in Africa is an issue of major importance. In order to investigate the nature of LUCC and related ecological services, we selected a study area in Northern Ghana, the Atankwidi catchment in the Upper East Region, as a case study for land-use related problems and prospects in West Africa. Due to the reliance of local farmers on ecosystem services, both future LUCC and income structures need to be assessed. Furthermore, in order to be able to mitigate negative externalities of the local use of natural resources and to enhance their sustainable use, the impact of policy interventions on future land-use and income structures also needs to be estimated. Therefore, the goal of this study is to develop a realistic simulation model for land-use and land-cover change and income structures for the Atankwidi catchment of Upper East Ghana, which can be used to explore alternative pathways into the future caused by policy interventions.

The choice of the modeling approach for this endeavor not only depends on the limitations and strengths of the various techniques, but also on the scale of analysis, comprising spatial resolution and extent. As in agricultural areas the decisions made by man are the main influences on land-use/cover patterns (Mander and Jongman, 1998), it is advantageous to directly simulate the decisions of land managers, resulting in a model resolution at farm level. However, with such a fine-scale resolution, the spatial extent of the area under observation is usually limited to small areas. The study area fulfills this requirement, as with an area of 159 km^2 and a population size of 6400 households it is relatively small, thus allowing such an individual-based approach. Due to these reasons and the potential strengths of MAS models, we decided to use an agent-based approach for modeling LUCC in the study area.

As we have discussed, an agent-based approach is the most appropriate method if the explicit representation of human decision-making and a realistic representation of the structural complexity of the land-use system is desired. However, the major challenges of the agent-based approach lie in the realistic representation and calibration of the coupled human-environment system as found in the real world. The main research objective of this study is, therefore, as follows:

To develop a realistic agent-based model for simulating the complex LUCC pathway in a semi-arid catchment in the Upper East Region of Ghana, thereby generating an operational tool to explore the impact of policy interventions on future land-use/cover patterns and income indicators.

The achievement of this goal indeed involves a model development process that includes sequential steps. First, a parameterized framework representing the structure and functions of the coupled human-environment system underlying LUCC has to be formulated. Next, relevant local socio-economic and ecological processes need to be identified and empirically parameterized using observed data. Finally, these processes need to be integrated into the parameterized framework in order to obtain an operational MAS/LUCC model, which can be used to explore the potential impact of local land-use related policies on land-use/cover and livelihood. The interrelated sub-objectives are therefore:

1. To build a parameterized MAS/LUCC framework for modeling the evolution of the coupled human-environment system in the study area, whereby land-use/cover and socio-economic dynamics are self-organized from interactions among farming households and land patches, under the influence of certain policies and other external circumstances,
2. To calibrate and verify land-use decision-making sub-models of the farming households (i.e. human agents) in the study area,
3. To calibrate and verify sub-models representing relevant biophysical dynamics of land patches,
4. To develop an operational MAS/LUCC model based on the parameterized framework, by integrating the calibrated decision-making and ecological dynamics sub-models, in order to explore the likely outcomes (in terms of land-use/cover and socio-economic features) of selected policy alternatives and other external factors.

1.5 Outline of thesis

This thesis consists of seven chapters. This chapter gives an introduction in global phenomena and problems related to land-use and land-cover change, identifies the complex nature of such changes, and discusses the strengths and limitations of current approaches. A justification is given for the application of the agent-based approach for modeling land-use/cover change in the study area, and the related research objectives are outlined.

Chapter 2 clarifies technological concepts and methods of MAS and establishes a conceptual framework for detailed technical work. First, basic concepts of the agent-based approach are elucidated using land-use-specific examples. These concepts comprise the concept of agents, agent environment, and agent architectures. Following the multi-agent mindset, a conceptual framework for the coupled human-environment system underlying LUCC is presented, serving as a basis for detailed descriptions in later chapters. Third, a brief description of the study area is given. The chapter ends with the discussion regarding the selection of NetLogo, a MAS computer platform (Wilenski, 1999).

Chapter 3 deals with the first specific objective. It formulates the first principles and architecture of the MAS/LUCC framework, named GHana - Land- Use DynAmics Simulator (GH-LUDAS). The chapter consists of two parts. In the first part, the four main modules of the model as derived from the conceptual framework are described in detail, including the Human, Landscape, Decision-making and Global-policy Modules. The range of land-use-relevant variables on both the landscape and the human side is described in detail, and the structure and sub-routines of the Decision Module are presented.

Furthermore, the range of variables of the Global-policy Module, whose values are set externally by the model user, and their integration into the coupled human-environment system is described. The initialization of the model is presented, i.e. the setup procedures at the start of the simulation runs, and the simulation protocol describing the sequence of routines during model run. The architecture of GH-LUDAS and the simulation protocol are presented using textual, graphical and algebraic languages prior to any empirical calibration. These calibrations will be conducted and justified in the subsequent chapters.

Chapter 4 deals with the second specific objective, the calibration and verification of the decision-making processes of human agents. The study area is described with respect to land use and socio-economic conditions in order to make the subsequent specifications of the decision-making sub-models more comprehensive. Based on the findings from the area description, the human agents (households) are categorized into typical groups according to livelihood structure and strategy, using data condensation (Principle Component Analysis) and classification (k-mean Cluster Analysis) techniques. Finally, land-use decision-making sub-models are developed, being partly dependent on the previously derived agent groups, using spatial regression analysis (m-logit regression). The coefficients obtained through the application of these statistical techniques are directly fed into the model in order to calibrate the final operational MAS/LUCC model GH-LUDAS.

Chapter 5 presents the specific objective 3, i.e. the determination of land-use-relevant landscape-specific attributes and the calibration and verification of relevant dynamic ecological models. The detailed description of the biophysical setting of the study area serves as a basis for the further model specifications. The land-use-relevant landscape attributes are then described and visualized, including local land-cover patterns, biophysical attributes and spatial accessibility. Furthermore, the sources and data processing techniques for the determination of these attributes are given. Finally, the biophysical sub-models are developed, being confined to land-use-type specific productivity functions, a livestock dynamics model, and a land-cover transformation sub-model. Both the spatial patterns of the landscape attributes and the biophysical sub-models are fed into GH-LUDAS in order to obtain an operational MAS/LUCC model.

In Chapter 6, GH-LUDAS as a decision support tool, and the identification, simulation and analysis of selected scenarios are presented. Based on an analysis of the environmental, demographic and policy setting of the study area, the external parameters of GH-LUDAS are specified. The setting of these parameters allows stakeholders and researchers to test their assumptions through simulation-based analysis. For these purposes, the use of GH-LUDAS as an operational tool for decision support and research is presented, including a summary of its internal structure, and model input and output. Selected scenarios are specified and analyzed. The range of external parameters allows specifications in policies of dam construction and credit access, as well as in demography and climate change. For each of these families of parameters, scenarios have been selected and compared to a baseline scenario, which reflects the policy settings as they were in 2006. Finally, the sensitivity of these factors to the LUCC system is presented and analyzed.

2 MULTI-AGENT SYSTEM ARCHITECTURE

2.1 Introduction

Multi-agent systems (MAS) are a relatively new sub-field of computer science - they have only been studied since about 1980, and the field has only gained widespread recognition since about the mid 1990s. However, since then, international interest in the field has grown rapidly. This is partly due to the belief that agents are an appropriate software paradigm to understand and build a wide range of artificial social systems (Wooldridge, 2002). Multi-agent-based simulation is nowadays used in a growing number of areas, where it is progressively replacing other techniques (e.g. micro-simulation, object-oriented or individual-based simulation techniques) (Drogoul et al., 2003).

This is due, for the most part, to the fact that MAS can cope with very different models of 'individuals', ranging from simple entities to more complex ones. The easiness with which modelers can handle different organizational levels of representation (e.g., individuals and groups) within a unified conceptual framework is also particularly appreciated, with respect, for instance, to cellular automata (Parker et al., 2003). During the last decade, the approach has been applied to more and more scientific domains: sociology (Pietrula et al., 1998; Goldspink, 2003), biology (Resnick, 1995; Drogoul et al., 1995), physics (Schweitzer and Zimmermann, 2001), chemistry (Resnick, 1995), ecology (Huberman and Glance, 1993), and economy (Ben Said et al., 2002).

In the field of ecosystem management, access and use of natural and renewable resources are key issues. Scientists working in this area need to examine the interactions between ecological and social dynamics. For many years, this question has been indeed examined either exclusively from the angle of 'an ecological system subject to anthropogenic disturbance', or from the angle of 'a social system subject to natural constraints' (Bousquet and Le Page, 2004). With the shift to the agent-based paradigm, the interactions between the social and the ecological components, as well as their heterogeneity, are taken into account (Bousquet and Le Page, 2004). These human-nature interactions as well as their heterogeneity play a major role in the coupled human-environment system underlying LUCC, which can be appropriately addressed by the agent-based methodology.

In this chapter, we will clarify the concepts underlying the agent-based approach in order to understand the further steps of model conceptualization, specification and implementation. Furthermore, we will review recent advances in computer platforms for MAS in order to provide a basis for the selection of a suitable package for our work. Finally, we will present a conceptual MAS framework of the coupled human-environment system underlying LUCC.

2.2 Multi-agent system concepts

There are many different definitions of an agent and multi-agent systems. Here, we present the definition given by Ferber (Ferber, 1995 and 1999) because it seems to be the more meaningful one for researchers in ecology and environmental sciences. A multi-agent system consists of the following components:

- An environment (E), that is usually a space.
- A set of objects (O), which are situated in E.
- An assembly of agents (A), which are specific objects (a subset of O) representing the active entities in the system.
- An assembly of relations (R) that link objects (including agents) to one another.
- An assembly of operations (Op) making it possible for the agents of A to perceive, produce, transform, and manipulate objects in O.
- Operators with the task of representing the application of these operations and the reaction of the world to this attempt at modification, which we shall call 'the laws of the universe' (e.g. productivity as a result of land management decisions and land cover change).

To make this definition more comprehensive, we give examples for each of the concepts from the perspective of the coupled human-environment system underlying LUCC. The environment E is simply the landscape under study where agents and other objects are located. While agents refer to decision-making entities - here represented by farmers, or more specifically by farming households - the non-agent objects include features such as houses, markets, rivers or farm plots, which all possess a certain location within the environment E. The relations among objects - including agents - can be manifold: Relations among among agents might refer to social interaction with respect to land use, whereby relations among agents and non-agent objects might, for instance, refer to tenure rights of an agent to a certain piece of land. The operations of an agent including perception, production and transformation of objects can be interpreted in the way an agent perceives his environment and takes certain actions according to these perceptions and own conditions. These actions might include the choice of land use type, the decision to do irrigation farming, or the choice of land management. Operators - or 'the laws of the universe' - then might include the model of crop productivity, being partly dependent on previous actions of the household agent, or it might include the natural as well as the human-induced transformation of land cover (e.g. natural vegetation growth, tree logging).

2.2.1 Concept of environment

In any MAS, agents are situated in an environment, therein searching for information, interacting with each other, and possibly modifying it. The representation of such an environment is highly dependent on the objectives of the study. Russell and Norvig (1995) gave an overview of the range of possible environment classes as follows:

- **Accessible vs. inaccessible**
  An accessible environment is one in which the agent can obtain complete and accurate information about the state of the environment. Modeled real-world environments are usually accessible to some degree only. The more accessible an environment is, the simpler it is to build agents to operate in it.

- **Deterministic vs. non-deterministic**
  A deterministic environment is one in which the outcome of any action is defined, i.e. there is no uncertainty about the state that will result from performing an action. The physical world can be regarded as non-deterministic with respect to particular properties.

- **Static vs. dynamic**
  A static environment can be assumed to remain unchanged except by the performance of actions by the agent. A dynamic environment is one that has other processes operating in it, and which hence changes in ways beyond the agent's control. The physical world is a highly dynamic environment.

- **Discrete vs continuous**
  An environment is discrete if its states are represented in a countable way (i.e. a discrete scale). For example, the landscape environment is discrete in land-use/cover types and continuous in many biophysical properties, such as surface slope, moisture and biomass.

As the real environment is a highly inaccessible, non-deterministic, dynamic and continuous environment, such properties should be incorporated in a model that tries to simulate real-world processes, such as land-use and land-cover change. Thus, in GH-LUDAS, we consider these real-world properties. For instance, the inaccessibility and non-determinism of our environment is represented by a limited sphere of influence for each agent, in which the agents have limited control over the results of their actions. Furthermore, GH-LUDAS can be regarded as partially dynamic, as land-cover transformation processes take place even without agent interference. Finally, the model environment is continuous to some extent, as objects and agents do exhibit dynamic state variables and actions at a continuous scale, which results into an uncountable number of environment states.

2.2.2 Concept of agent

In MAS literature, there is no universally accepted agreement about the definition of the term agent. However, there is a general consensus that autonomy is central to the notion of agency, being confined by the following definition given by Weiss (1999: page 32): *An agent is a computer system situated in some environment, that is capable of autonomous action in this environment in order to meet its design objectives.* The term autonomy here refers to the ability of agents to act without the intervention of other agents or other systems. Such actions of an agent are a result of the agent's perceptions of the environment, and, if designed as such, also of the agent's own state (see Figure 2.1).

Furthermore, it is important to mention that this definition of agency refers only to 'agents' in general, and not to 'intelligent agents'. According to Weiss (1999: page 32), an intelligent agent is one that is *capable of flexible autonomous action in order to meet its design objectives*, where flexibility means:

- **reactivity:** intelligent agents are able to perceive their environment, and respond in a timely fashion to changes that occur in order to meet their design objectives;
- **pro-activeness:** intelligent agents are able to exhibit goal-directed behavior by taking the initiative in order to satisfy their design objectives;
- **social ability:** intelligent agents are capable of interacting with other agents in order to satisfy their design objectives.

With respect to the coupled human-environment system of land-use/cover change, these abilities can be interpreted in the following way: Humans can be regarded as reactive agents, as they adapt to changes within their environment, such as climate or ecosystem change, in order to meet and maintain their design objectives, which might include economic and social welfare. Second, the human seeking to maintain or improve the personal condition clearly behaves in a goal-directed manner, in that decisions to be made are deliberately chosen to meet such personal objectives. With respect to land use, land-management decisions are closely related to the personal objectives of the farming household, e.g. ability to survive, improvement of living conditions. Finally, interactions among farmers play a role in land-use systems, with knowledge transfer and competition being two major characteristics of such agent interaction. Knowledge transfer refers to the diffusion of agricultural land-management practices or new agricultural technologies through the population by communication and observation, which has a direct impact on land-use patterns. Competition, on the other hand, can be interpreted as the way in which farmers compete for natural resources, e.g. agricultural land, pastures, forests for timber logging, etc.

In GH-LUDAS, all these attributes were considered for farming agents, which are endowed with both reactive and goal-directed behavior. Regarding social interaction, tech-

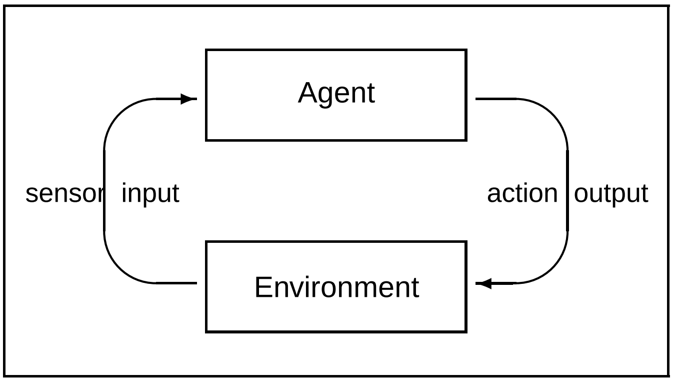


Figure 2.1: Agent-Environment Interaction

nology diffusion has been considered through neighbor effects, i.e. the transfer of knowledge by neighboring farming households. Competition is not represented directly by agent-agent interactions, but is mediated through the use of land, thus resulting in competition for land among households.

2.2.3 Agent architecture

Following these definitions of agent and environment and the concept of agent perceptions resulting in actions, a function that implements such agent mapping from perceptions to actions is required. Such a function is called agent architecture. The literature usually cites the following five different types of architecture (Russell and Norvig, 1995):

- Simple reflex agents
- Model-based reflex agents
- Goal-based agents
- Utility-based agents
- Learning agents

In the following, we will give short descriptions of each of these architectures, and justify the selection of architecture to be implemented in GH-LUDAS.

Simple reflex agents

The agent architecture of simple reflex agents consists of simple 'if-then' rules (or condition-action rules) reacting to environment conditions perceived by the agent, and resulting in certain actions. Figure 2.2 gives the structure of a simple reflex agent in schematic form, showing how the condition-action rules allow the agent to make the connection from percept to action. Such reflex decision-making mechanisms are suitable for representing reactive behavior of both human and biophysical agents. For human agents, the application of reflex decision-making assumes that people do not (or cannot) calculate any anticipated values of alternatives, but rather react in a timely fashion according to their daily routines to select directly options based on current conditions (Cioffi-Revilla and Gotts, 2003; Haggith, 2002).

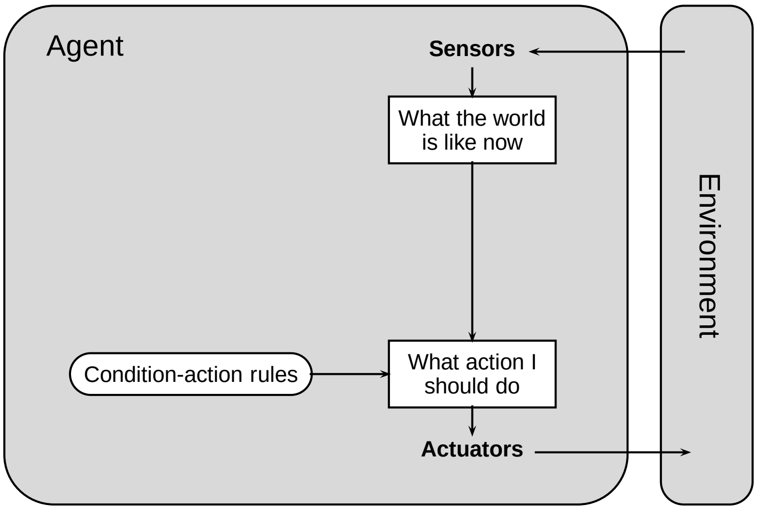


Figure 2.2: Reflex-based agent architecture

Model-based reflex agents

The simple reflex agent described above will work only if the correct decision can be made on the basis of the current perception. Such an architecture can be problematic, because the sensors do not always provide access to the complete state of the world. In such cases, the agent may need to maintain some internal state information in order to distinguish between world states that generate the same perceptual input but nonetheless are significantly different. Updating this internal state information as time goes by requires two kinds of knowledge to be encoded in the agent architecture. First, we need some information about how the world evolves independently of the agent. Second, we need some information about how the agent's own actions affect the world. Figure 2.3 gives the structure of the model-based reflex agent, showing how the current perception is combined with the old internal state to generate the updated description of the current state.

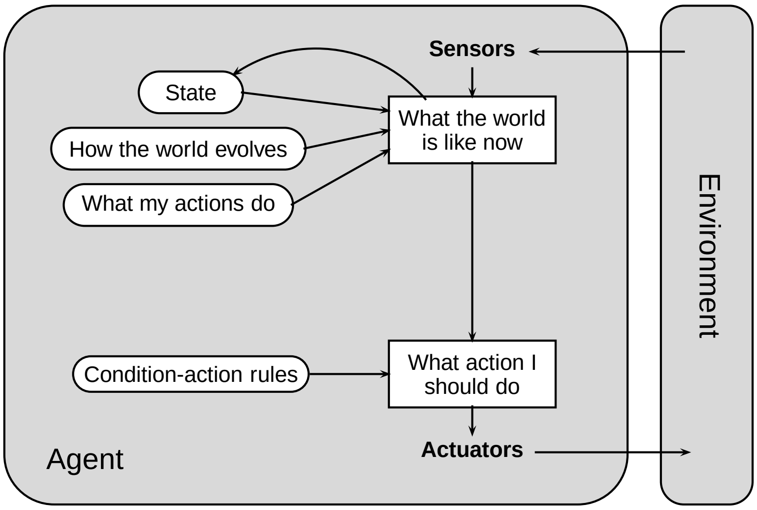


Figure 2.3: Model-based reflex agent architecture

Goal-based agents

Knowing about the current state of the environment is not always enough to decide what to do. The right decision is dependent on the goals of the agent. In other words, to arrive at the desired decision, the agent needs some sort of goal information which describes situations that are desirable. The agent program can combine this with information about the results of possible actions (the same information as was used to update internal state in the reflex agent) in order to choose actions that achieve the goal. Sometimes this will be simple, when goal satisfaction results immediately from a single action; sometimes, it will be more tricky, when the agent has to consider long sequences of actions to achieve the goal. Searching and planning are the subfields of Artificial Intelligence devoted to finding action sequences that do achieve the agent's goals. In Figure 2.4, the internal mechanism of such goal-directed behavior is depicted.

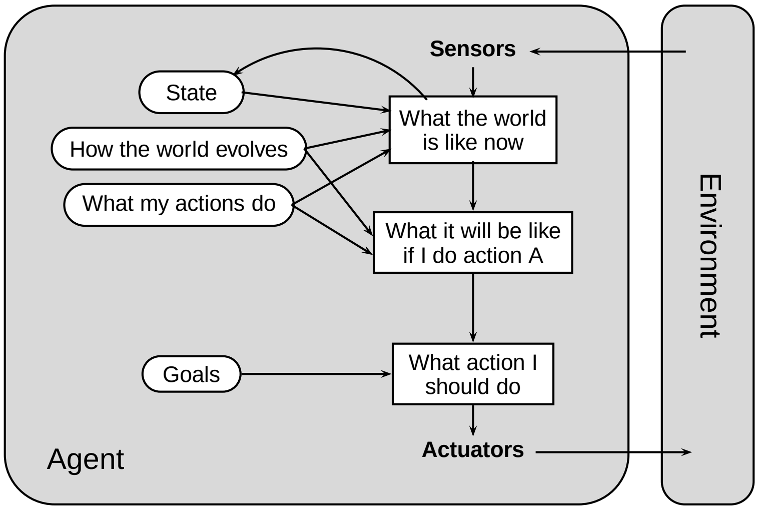


Figure 2.4: Goal-based agent architecture

Utility-based agents

Goals alone are not really enough to generate high-quality behavior. For example, there are many action sequences that will make the agent achieve its goal, but some are quicker, safer, more reliable, or cheaper than others. Goals just provide a crude distinction between 'happy' and 'unhappy' states, whereas a more general performance measure should allow a comparison of different world states (or sequences of states) according to exactly how happy they would make the agent if they could be achieved (see Figure 2.5). The customary terminology is to say that if one world state is preferred to another, then it has higher utility for the agent. Utility is therefore a function that maps a state onto a real number, which describes the associated degree of happiness.

A complete specification of the utility function allows rational decisions in two kinds of cases where goals have trouble. First, when there are conflicting goals (e.g. benefit maximization and risk minimization) only some of which can be achieved, the utility function specifies the appropriate trade-off. Second, when there are several goals that the agent can aim for, none of which can be achieved with certainty, utility provides a way in which the likelihood of success can be weighed up against the importance of the goals.

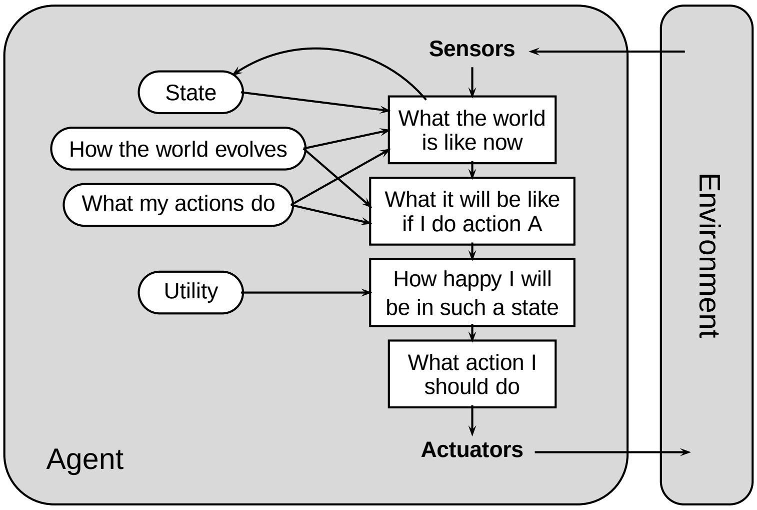


Figure 2.5: Utility-based agent architecture

Learning agents

Turing (1950) noted the huge amount of work it takes to program an intelligent machine, and concluded that it would be easier to build learning machines and then to teach them. Another advantage of learning agents is their adaptability to unknown environments, and the improvement of their behavior with time. The learning agents use a feedback, called critic, to learn which perceptions of the environment are desirable, and in consequence, how to behave (Figure 2.6). This means that agents' learning consists of improving their future performance based on their past critic, by optimizing their behavior such as to maximize their utility when the world continues evolving as it has. This kind of learning makes agents discover that some kind of (but not exactly) condition-action rules always do the same thing, based on their current knowledge.

A problem arises here: after some learning time, agents are always going to do the same things because of these discovered rules, though the agents are not sure that these actions are optimal, while they might have a better performance if they had a wider knowledge of their environment. In fact, they should try to do very different actions than those prescribed by their learning process. This exploration of new actions is insured by the problem generator.

These architectures are presented in ascending order of complexity and ability to represent real-world intelligent agents: Learning agents are surely more realistic than utility-based agents, and utility-based agents are more realistic than goal-directed agents, etc. Al-

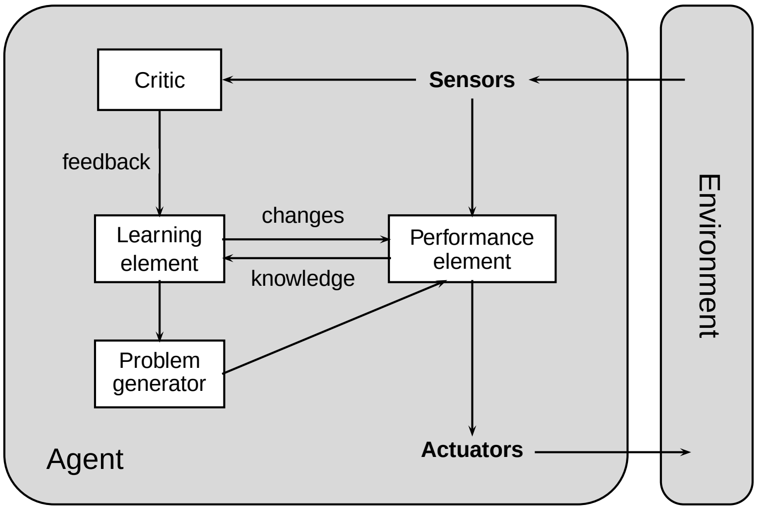


Figure 2.6: Learning agent architecture

though learning agents might be the most realistic architecture for human agents, the implementation of learning mechanisms can cause a dramatic decrease in computing speed (Russell and Norvig, 1995). To compare the computing speed of learning and utility-based agents, which is usually regarded as the second most realistic agent architecture, both approaches were implemented in a simple separate agent-based model. The comparison of both showed that even one of the simplest machine learning algorithms for agents, the k-nearest neighbor algorithm, had a 10-fold lower computing speed than the utility-based approach. Thus, for implementation in GH-LUDAS, the utility-based approach was chosen in order to keep the computing speed within a reasonable range. However, as there is a debate about modeling agents that behave in a way to achieve highest possible utility (i.e. purely rational behavior), random errors within these decisions have been included to ensure bounded rational behavior. Bounded rational behavior allows agents to choose actions with lower utilities than the optimal one (see section 2.4.2). The reflex-based architecture is also highly suitable for modeling state transitions of biophysical agents (Le, 2005). As such, the model of land-cover transformation for biophysical agents within GH-LUDAS was designed as a rule-based mechanism, determining the conversions among land-cover types during time (see section 5.3.5).

2.2.4 Relations, Operations and Operators

Relations among objects (including agents) in multi-agent systems can be manifold. In general, a relation consists of a database, which can be described by a matrix of the length of the total number of objects. Each matrix value is an item of a tuple of possible relational values. For instance, the tuple might comprise two options ('friend', 'no friend'), and each pair of objects is then assigned the respective value. Relations may not only exist among agents (e.g. in the form of social networks), but also among agents and non-agent objects, and among non-agent objects themselves. For instance, a relation among an agent A and an object O might include the right of A to modify O, and a relation among non-agent objects could be their distance to each other, which might have an influnce on their internal mechanisms. Furthermore, relations do not necessarily remain static and can be modified during time through system performance.

A second important characteristic of multi-agent systems concerns the use of operations, which enable agents to perceive, produce, transform, and manipulate objects. In multi-agent systems, perceptions represent the knowledge base an agent has about objects. The knowledge base consists of a collection of data about objects (including agents) accessible to the agent, which can be objective or subjective. While objective knowledge comprises data about the real state of objects, subjective knowledge can result from a mechansim which distorts the perception of the real state of objects. Furthermore, the set of perceived objects, both agent and non-agent objects, does not necessarily comprise the whole set of objects, but can be confined to subsets individually for each agent. In addition, the range of data about these objects accessible to the agent can be limited. For instance, the knowledge about relations among the agent and other objects can be fully, partly or not accessible to the agent.

If agents are not endowed with a memory mechanism, which enables them to record past data, the knowledge base of an agent is confined to the perceptions of only the current state of objects (including himself). If an agent is endowed with such a memory mechanism, he can record past states, actions and reactions of himself and other objects. Even the knowledge base of an agent can be accessible to other agents, which might result in situations of 'full knowledge' (agent A knows that agent B knows that agent A knows, etc.), which are often studied in game theory.

Based on this individual knowledge, agents make decisions according to their agent architecture as described in the previous section. The range of possible actions resulting from these decisions comprises the deletion, the creation or modification of objects. The creation of agents might be caused by a mechanism of reproduction, while the deletion of agents might be due to dispatch. The modification of objects can include a spatial displacement or an alteration of the objects' internal state.

The combined ability of object perception and the (possible) subsequent manipulation, regulated by the decision-making architecture, represents the set of operations of agents. Non-agent objects react to these operations via operators, which Ferber (1995, 1999) calls 'the laws of the universe'. Such operators update changes in states of objects, which can be due to agent intervention or agent-independent processes, or both. However, such operators are not only confined to non-agent objects. Agents can also be subjected to 'laws of the universe', for instance to processes such as ageing or death. Changes resulting from operators can further be perceived by other objects.

In summary, not only the internal architecture of agents is of concern in multi-agent systems, but also the defined webs of interrelations among objects, including relations, perceptions, actions and reactions. The high flexibility of multi-agent systems in designing these interrelationships is one of the great benefits of this approach, and ensures its applicability to many research domains and areas. Multi-agent systems have been used to study cell communities, ant colonies, animal flocking, strategic military problems, etc As the modeling of multi-agent systems relies on the specifications of agents' behaviors and interactions, which result in emergent properties at the level of the system, agent-based modeling can be considered as one of the few bottom-up modeling approaches.

2.3 Computer platforms for MAS

The use of agent-based models models (ABMs) or individual-based models (IBMs) for research and management is growing rapidly in a number of fields. For example, DeAngelis and Mooij (2005) documented a steady, sharp increase in the number of ecology publications using IBMs starting in about 1990. This growth is partly due to the ability of these models to address problems that conventional models cannot, and partly to the growing number and quality of software platforms for agent-based modeling (Railsback et al., 2006). In this chapter, we review the most widely used computer platforms for agent-based modeling, based on a study of Railsback et al. (2006), and give a justification of the platform employed in our study, which is NetLogo.

The most commonly used software platforms for agent-based modeling comprise the Swarm (based on Objective-C or Java language), Repast, MASON, and NetLogo (based on Java language). The first three plaforms belong to the 'framework and library' platforms, which were designed to make the design, implementation, and use of ABMs more accessible and efficient. Swarm in particular was designed as a general language and toolbox intended for widespread use across scientific domains. Swarm's developers started by laying out a general conceptual approach to agent-based simulation software. Therefore, Swarm was implemented as a framework - a set of standard concepts for designing ABMs - along with a library of Objective-C software implementing this framework. Repast was started as a Java implementation of Swarm but has diverged significantly from Swarm. One objective of the Repast project was to make it easier for inexperienced users to build models, including a built-in simple model and interfaces, which support the process of model construction for beginners. MASON is being developed as a new Java platform, designed as a smaller and faster alternative to Repast, with a clear focus on computationally demanding models with many agents executed over many iterations. Design appears to have been driven largely by the objectives of maximizing execution speed and assuring complete reproducibility across hardware.

These framework and library platforms have succeeded to a large extent because they provide standardized software designs and tools without limiting the kind or complexity of models they can implement, but they also have well-known limitations. Tobias and Hofman (2004) recently reviewed Java Swarm and Repast (along with two less-used platforms), ranking them numerically according to well-defined criteria. In their study, they indicate that important weaknesses include difficulty of use; insufficient tools for building models, especially tools for representing space; insufficient tools for executing and observing simulation experiments; and a lack of tools for documenting and communicating software.

The most recent development of MAS plaforms is the appearance of MAS packages. Differing from the framework and library plaforms, the MAS package is a collection of primitives assembled with a standardized common user interface and provides a new environment for MAS modeling. NetLogo (Wilenski, 1999) is one among few new MAS plaforms.

Its primary purpose was to provide a high-level platform that allowed students down to the elementary level to build and learn from simple ABMs. However, recent versions of NetLogo now contain many high-end capabilities (behaviors, agent lists, graphical interfaces, etc.) and it is quite likely the most widely used platform for ABMs. Of all main currently used platforms, NetLogo is the highest-level platform, providing a simple yet powerful programming language, built-in graphical interfaces, and comprehensive documentation. It is designed primarily for ABMs that contain mobile individuals in a grid space with local interactions. According to a recent evaluation of Railsback et al. (2006), NetLogo is highly recommended, even for prototyping highly complex models.

In contrast to the other platforms, NetLogo almost completely separates the processes of implementing and displaying a model. The modeler writes a program (in NetLogo language) for behavior of agents and the gridded space on a 'Procedures' page. On a separate 'Interface' page, the modeler can design an automatic animation of agent locations on the space. Graphs and parameter controllers can be added to the interface via graphical and menu-driven tools, along with simple statements in the software telling the interface when to update. In the other programming platforms, the processes of implementation and displaying of the model are not separated, with the instantiation of the display or 'animation window' requiring several programming steps. Furthermore, the procedures of the motions of agents on the display have to be implemented using lower-level operations in these platforms, whereas in NetLogo agent motion can be simply implemented using a built-in method that moves agents to a new location.

As users are highly interested in monitoring outcomes of the model runs, it is also useful to compare the strengths and weaknesses of the various platforms in producing graphics of output indicators, output files and statistics. Histograms are particularly useful for ABMs, because they can output the full distribution of some characteristic over all the agents. Repast and Swarm have built-in histogram classes that are relatively easy to use, while MASON does not yet provide such a class. In NetLogo, histograms are created using drag-and-drop and a menu on the interface page. Then, a simple code statement specifies when the histogram is updated. Regarding the provision of output files, Objective-C Swarm and Repast provide built-in classes to facilitate output of data to a file, and data recording actions can be scheduled just like any other action, so that they take place at known times. Java

Swarm and MASON do not provide file writing tools, so a Java class for file output must be used. NetLogo provides simple primitives for opening and writing to files, although their ability to format and control output is limited; for example, there is no way to overwrite a file instead of appending to it. As far as statistical calculations are concerned, Swarm has a powerful tool for collecting summary statistics, and NetLogo also includes primitives that provide all common statistics. Repast's 'DataRecorder' provides only an average, whereas MASON even lacks tools for any summary statistics.

The most significant weakness of NetLogo is the slow speed in model execution, whereas in most of the other aspects this platform exceeded the capabilities of the other platforms, offering a convenient programming environment at the same time. Although execution speed is relevant for the choice of an appropriate software platform, it has to be considered that the most time-consuming part is, nevertheless, the modeling process, and not the execution of model runs. As such, the implementation of the model in Java and C programming languages is much more time-consuming than the use of NetLogo primitives. Therefore, the time spent by the model runs using NetLogo is leveled out by the comparably short time spent for model development. Moreover, the rapid development of high speed CPU mitigates the low speed of NetLogo excution. By virtue of this argument and the other advantages as outlined before, we decided to use NetLogo as a software platform to implement our MAS/LUCC model in this thesis.

2.4 GH-LUDAS: A proposed conceptual framework for modeling LUCC

In this section, we will present a conceptual framework for the MAS/LUCC model developed in this thesis, called GH-LUDAS, in order to provide an understanding of the further specifications in the subsequent chapters. This framework follows the synthesis of the coupled human-environment as proposed by Haggith et al. (2003) and Freudenberger (1995). This framework has already been used in the FLORES model (see Haggith et al., 2003), which aims to capture the interactions between rural communities living at the forest margin, thereby serving as a tool to explore the consequences of alternative policy options. It aims to model the dynamics of the interactions between the biophysical and socio-economic components of rural communities at the forest margin. The 'glue' that binds the biophysical and socio-economic components together is human decision-making at the local level, which influence the performance of the biophysical components. Due to its high level of generality, this conceptual framework can also be applied to the study of land-use/cover change. As such, this type of framework has been applied to the study of land-use/cover change in the uplands of Vietnam (see Le, 2005), and now finds applications in several land-use studies at the Center for Development Research (ZEF) in Bonn.

Like the framework as proposed by Le et al. (2008), the conceptual framework of GH-LUDAS comprises four modules, namely the human, the landscape, the decision, and the Global-policy Module (see Figure 2.7). The design and interrelations of these four components are briefly described in the following.

2.4.1 Landscape module

The landscape environment (E) is usually implemented as a grid consisting of congruent cells, whereby each of the human agents is located on a specific grid cell within E. The non-agent objects are usually implemented as grid values for each cell within E, i.e. each type of object is represented by an own variable with values for each specific cell of E. For instance, the object of houses might be represented by a variable of its own, being 1 for cells covered by houses, and 0 for other cells. Relations thus exist among human agents and cells, e.g. ownership of a cell, whereby human agents operate on these same cells through a set of operations Op (see section 2.2.). Operators (section 2.2) then define the internal mechanisms and responses to these human actions on the landscape cells, e.g. internal ecological processes. The landscape environment is represented by a collection of landscape agents, i.e. intelligent congruent land patches (30 m x 30 m) with their own attributes and internal sub-models of relevant ecological processes (i.e. Operators). The attributes are represented by state variables of each patch, including the specific land-use and land-cover type, biophysical attributes (e.g. topography), accessibility variables (e.g. distance to river), tenure variables (e.g. owner), and yield variables indicating the total yield produced on the respective patch. Whereas topographical and accessibility variables are static in time, the variables of land use/cover, tenure and yield are dynamic over time and space.

Relevant ecological processes encoded within the architecture of landscape agents comprise agricultural production, land-cover transformation, and livestock dynamics. The agricultural productivity models consist of functions calculating the yield for a single patch

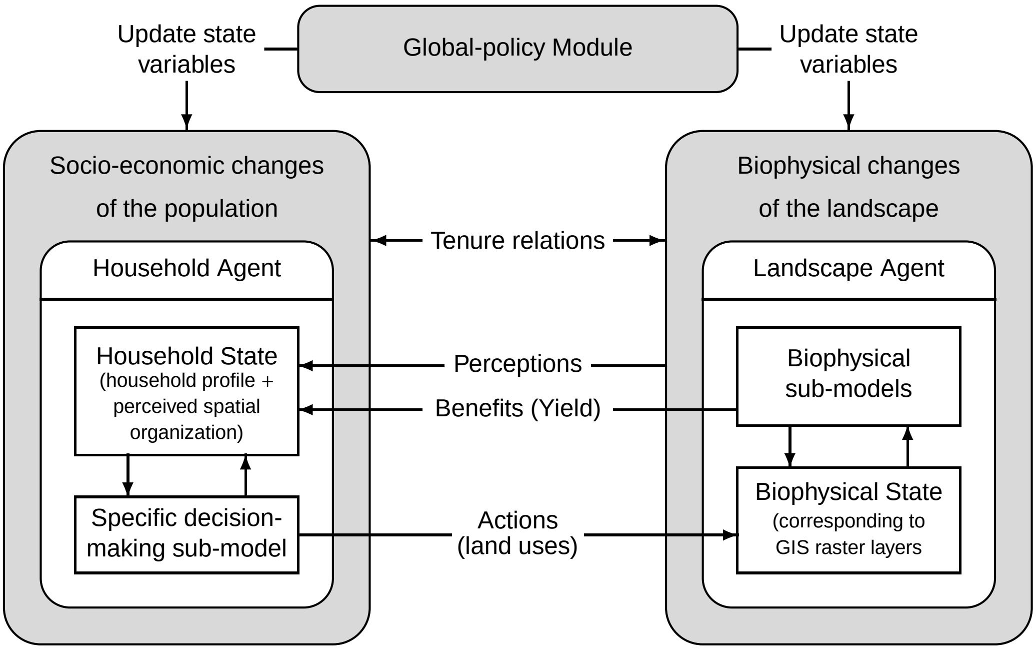


Figure 2.7: Conceptual framework of GH-LUDAS

seasonally, in response to its current state and the input decisions of the land manager (i.e. the household agent that is cultivating the patch), thereby updating the patch variable of yield response. The land-cover transformation model built into every landscape agent enables it to change its categorical variable of land cover, due to natural growth and changes in land use. Within the livestock dynamics model, the total number of livestock is determined in response to forage productivity, which, in turn, is dependent on annual rainfall and land-cover patterns.

2.4.2 Human and Decision Module

The Human Module is considered in terms of *household agents*, i.e. heterogeneous farming households with their own state and decision-making mechanisms about land uses (i.e. the Operations Op). The state variables of the household include a Household Profile and a spatial perception radius within the landscape, called Landscape Vision. The Landscape Vision consists of a collection of landscape agents located around the compound house of the household agent, on which the agent has full information and can set actions. The Household Profile comprises a list of household variables, such as age, household size, income, land resources, and the household's access to certain policies. Generally, the variables of the Household Profile as well as the policy-related variables change over time, but in response to different factors. Whereas some variables undergo a natural change (e.g. age), others are updated in response to agricultural activities (e.g. income). Policy-related variables change according to the values of the policy parameters, which are set by the model user.

The decision-making mechanisms are represented by a separate module, integrated in the architecture of the human agent. The mechanisms, which are based on the concept of the utility-based agent architecture, works by taking inputs from the household profile, policy-related variables, and the state variables of the perceived landscape patches. The decisions modeled by the decision-making mechanisms mostly represent choices among a discrete set of options (e.g. the choice among several land-use types for a given patch), using a utility function to assess the benefit of each option. Utilities for each choice are calculated using multinomial logistic (m-logit) regression, which can be formally expressed as:

$$Utility_p = \frac{e^{\alpha_p + \sum_i \beta_{ip} V_i}}{\sum_q e^{\alpha_q + \sum_j \beta_{jq} V_j}} \tag{2.1}$$

where $Utility_p$ is the utility of option p, having a value between 0 and 1, α_p a constant, and β_p the so-called preference coefficient of option p. When designing purely rational agents, the option with highest utility would be chosen by the agents. However, as purely rational behavior is rightly regarded as unrealistic, the choice models are designed to also consider options with a lower utility, thus allowing bounded rationality of household agents. This way, within GH-LUDAS, the utilities are interpreted as probabilities between 0 and 1, such that option p is only selected with a probability of $Utility_p$.

The Decision Module is universal for all household agents, in terms of its logical sequence. However, as the agent's state and the preference coefficients of the utility functions are individual-specific, decision outcomes result in a highly diverse pattern, thus representing heterogeneity among land users with respect to land-use decisions.

2.4.3 Human-environment linkages and interactions

Human-environment linkages are mainly characterized by tenure relations and a percep-tion-response loop (Figure 2.7). Tenure relations between household agents and landscape agents consist of rules determining the household access to land resources (e.g. ownership and use rights over land). Ownership is a tenure relation applied specifically to an individual household, i.e. the holder of the land. Village territory is a tenure relation applied specifically to a group of household agents, i.e. those households that share the same village.

The perception-response loop involves the flows of information and matter among the human and the environmental modules. Perception corresponds to the perceived spatial status of the Landscape Vision of a specific household, which is fed into the decision model, together with household-specific data, to calculate the anticipated benefits of certain land-use actions. Based on these calculations, the household agent responds by setting actions on his perceived environment, represented by decisions of land-use type and agricultural inputs. Subsequently, the state variables of the considered patches are updated, either directly (e.g. land-use type), or indirectly through the application of biophysical sub-models (e.g. yield response, land-cover transition). Finally, these updated state variables are fed again into the household's perception, thus forming an annual loop of perceptions and actions.

2.4.4 Global-policy Module

The Global-policy Module represents relevant factors that are set externally by the model users, and are thus not a result of the internal mechanisms of the model. These external parameters consist of parameters describing the rainfall regime (e.g. annual precipitation), the population dynamics of the household agents (e.g., carrying capacity, growth rate), and parameters of some relevant policies (i.e. household access to credit and construction of new dams). These factors directly modify either landscape-related variables and household-related variables, or alter the interaction modes between household and the environment (see Figure 2.7). For example, parameters regulating the access to credit directly updates the policy-related variables of the household, whereas dam construction affects state variables of the landscape through changing the biophysical variable of land cover and irrigability. Through the perception-response loop, such changes of state variables on either the human or the environmental side are carried through the model, thus significantly modifying the functioning of the whole system.

This proposed agent-based architecture allows integration of diverse human-, environment- and policy-related factors into farmers' decision making with respect to land use and presentation of subsequent accumulated outcomes in terms of spatial and temporal patterns of the natural landscape and population. Furthermore, aspects of the dynamics and structural complexity exhibited by land-use systems are reflected by this framework, including the representation of heterogeneous landscape and household agents, spatial and temporal interactions among these agents, and the consideration of feedback loops such as the perception-response loop. The representation of nested hierarchical levels and scale-dependent processes was also considered on both the landscape and the human side. Due to the complexity of the integration of hierarchies within the model, this aspect was not presented in this section, but will be outlined in the main chapter of model description (Chapter 3).

2.5 Materials and methods

The framework described above is a general framework for modeling LUCC, independent of the specific conditions of the study area. However, further specifications of the model will highly depend on the local conditions and processes in the study area. Thus, within the following sections, we will give a short description of the study area, justify its selection, and present the sources and generation methods for the data required for model implementation.

2.5.1 Selection of the study area

The study area comprises the Ghanaian part of the Atankwidi catchment in the Upper East Region of Ghana; the Atankwidi is a tributary of the White Volta located between Navrongo and Bolgatanga, with its upper reach in Burkina Faso (Figure 2.8). The catchment lies at $10°31'30''N$ latitude, and $0°56'0''E$ longitude, covering an area of 275 km^2, whereby the Ghanaian part covers an area of 159 km^2.

The catchment comprises the villages of Kandiga, Sirigu, Yuwa, Zoko and parts of Sumbrungu and Mirigu. This area was inhabited by 41.091 people in 2000 (Ghanaian Population Census, 2000). Out of these, 47 % were males, leaving about 53 % females.

This difference in male and female numbers is mainly due to a higher migration rate among the male population, confirming the hypothesis that migration is part of the survival strategy among males in terms of income generation. The major activities of the local people are confined to agriculture, livestock rearing, and non-farm activities such as trading or handicrafts. As most of these livelihood activities in the area are highly dependent on the services of land and water resources, any changes in the land productivity and pattern of land use and land cover are thus highly interrelated with the living conditions and well-being of the local population.

Apart from human influence, local land use and land cover is considerably dependent on climatic conditions. The study area falls within the Sudan-Savannah climate zone, which is characterized by a distinct rainy season lasting approximately from May to September, and a dry season from October to April (Martin, 2005). Land-use and land-cover patterns differ widely between the two seasons, with most of the agricultural activities confined to the rainy season. Within this season, the major part of the land surface is covered by small-farm agriculture, with patches of grassland that are mainly used as grazing plots for local livestock. Only 8.3 % of the land surface can be categorized as bare land in this season, being inappropriate for agricultural use. Due to the extensive use of land for agricultural purposes, the forest area has shrinked to only 3.1 % of the land surface, and mostly consists of 'sacred groves' along the river, i.e. holy forest patches traditionally protected, and forested hills with steep slopes. In the dry season, cultivation is only possible with irrigation, mostly being confined to small areas along the riverside, where groundwater tables are relatively high. The main irrigation technologies comprise bucket irrigation and pump irrigation either using hand-dug wells or large dugouts to reach the groundwater table. Due to the harsh climatic conditions in this season, bare soils are prevalent in the remaining area.

This study area was chosen for the following reasons. First, the area is located in one of the poorest regions of Ghana, which implies that a reliable evaluation of the impact of policy interventions on local socio-economic and ecological conditions can be of importance to ensure a sustainable improvement of local living conditions. Second, the local land-use patterns and socio-economic conditions are representative for other similar areas in the Upper East Region, which makes the results transferable to other areas to some extent. Third, other studies covering the hydrological settings and dynamics have been conducted in the study area, which could provide interesting results when used in combination with the results in this study. The findings of hydrology and groundwater availability can be compared to the actual agricultural water consumption. Finally, as data had to be collected during several field surveys, a good research infrastructure gave the final turn for selecting this area.

We defined the extent of the study area using both natural and institutional boundaries. In the north, the study area is restricted by the border to Burkina Faso, while the south part is confined by the drainage area, major roads, and village borders, which coincide almost completely. It was important to delineate the study area along village borders to ensure that local farmers do not, or very rarely, use land outside the study area. But since an exact map of such village borders was not available, finally the drainage area for the catchment was chosen to represent the spatial extent of the study area.

2.5.2 Biophysical characteristics and data generation

Biophysical characteristics (e.g. climatic, soil, and water-related factors) of the environment are usually important drivers of land-use/cover change. In order to integrate biophysical drivers in GH-LUDAS, relevant biophysical drivers needed to be identified, described, and mapped for further use in GH-LUDAS. In the following, a description of biophysical conditions in the study area is given, followed by a presentation of data sources and data processing methodologies.

Climate

The study area falls within the Sudan-Savannah climate zone, which is characterized by high temperatures and a mono-modal rainfall distribution with a distinct rainy season lasting approximately from May to September, and a dry season lasting from October to April (Martin, 2005). In the rainy season, south-west monsoon winds are prevalent, coming from the Atlantic Ocean, thus being responsible for humid and wet conditions during the rainy season period. These winds reach their maximum northern extent in August (Yaro, 2000). In the dry period, north-east trade winds blowing from the Sahara desert - called the 'Harmattan' - result in warm, dusty and dry conditions, and reach their maximum southwards extent in January. The long-term mean annual rainfall in Navrongo is 990 mm as calculated from monthly rainfall data for the years 1961-2001 (Martin, 2005). Regarding agriculture, the single rainfall

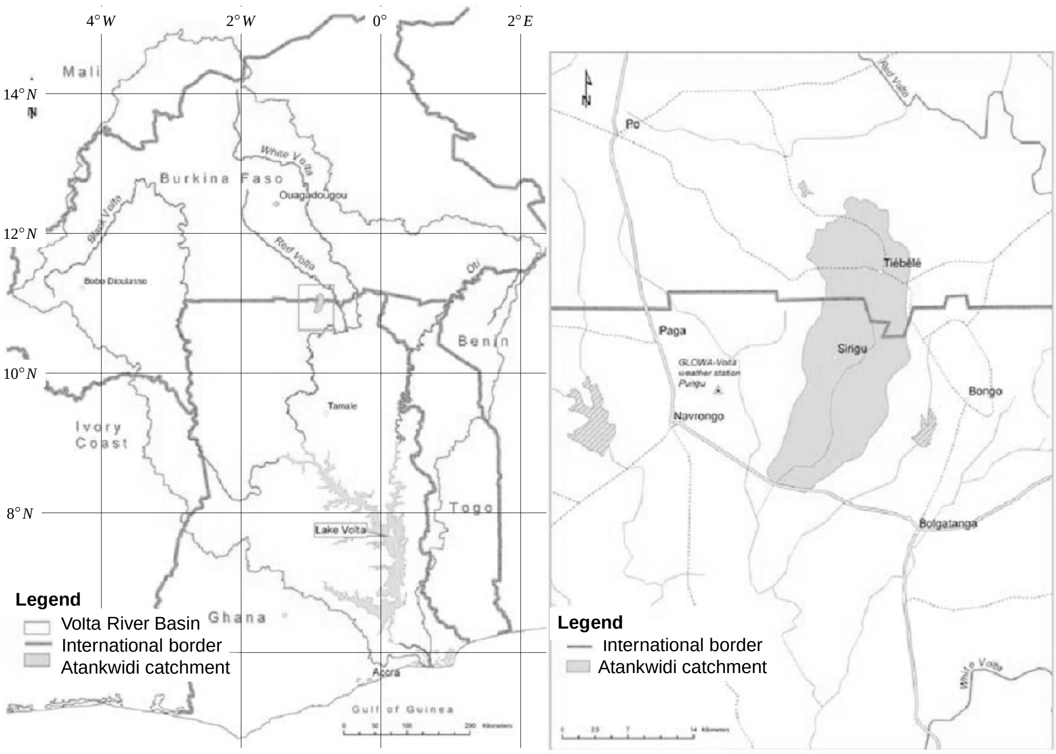


Figure 2.8: Location of the study area

regime received in this area limits full utilization of the physical capability of the people, as most of them are employed only during the short wet season and unemployed for the rest of the year (Yaro, 2000).

Temperatures are considerably higher than in the rest of the country, with mean monthly temperatures ranging between 18° C and 38° C. Temperatures are high throughout the year, with the lowest daytime temperatures coinciding with the peak of the rainy season, while the lowest night-time temperatures occur in December and January, caused by the Harmattan wind. The Harmattan period records the highest diurnal range of temperature, as nights are cool while days are very hot as a result of the absence of clouds. Vapor pressure during this period falls considerably to less than 13 000 hPa, and relative humidity rarely exceeds 20 % during the day but may rise to 60 % at nights (Report by Department of Geography and Resource Development, 1992).

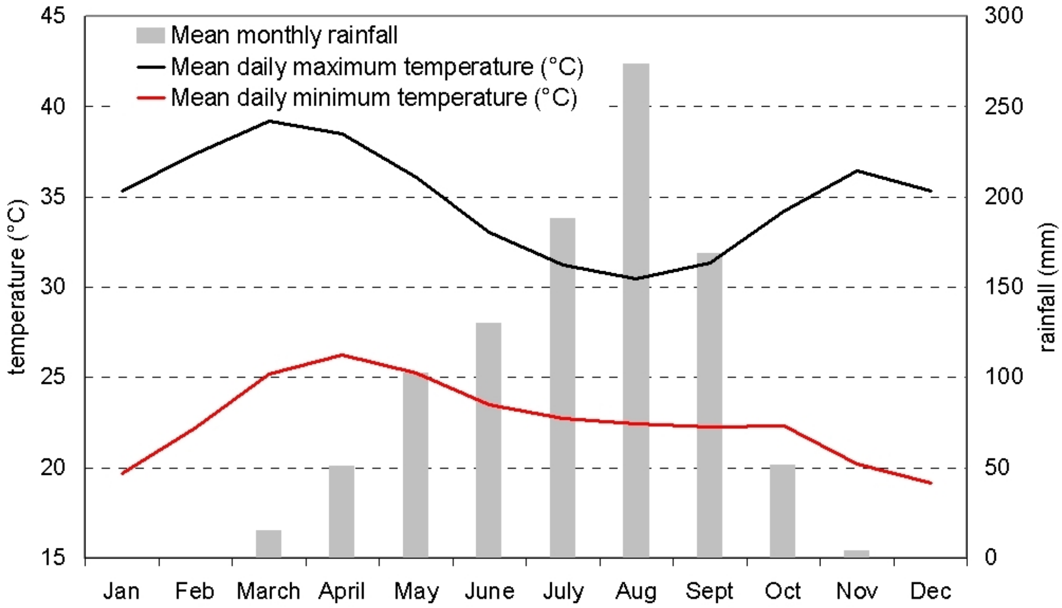


Figure 2.9: Annual temperature and rainfall pattern in the study area

Soils

According to soil maps of the Ghanaian Soil Research Institute in Kumasi, there are six soil associations prevalent in the study area: The associations of Tanchera, Kolingu, Nangodi, Kupela-Berenyasi, Bianya, and Tongo, and the Siare-Dagare Complex along the river banks (Figure 2.10). Following the FAO soil classification system, these associations can be grouped into three soil types, namely Lixisols (Tanchera, Kolingu, Nangodi and Bianya), Leptosols (Tongo and Kupela-Berenyasi), and Luvisols (Siare-Dagare Complex), which developed over granites, sandstones and Precambrian basement rocks, respectively (Martin, 2005).

The soils over granites and sandstones have mainly light topsoils varying in texture from coarse sands to loams, and heavier subsoils varying from coarse sandy loams to clays with a variable amount of gravel. Soils developed over basic rocks and most of those in the valley bottoms have heavier topsoils and subsoils (Adu, 1969). For about five months of the year, the soils receive a total rainfall of about 1000 mm, whilst for the remaining seven months they dry out almost completely. This alternation of wet and dry conditions causes intense leaching of nutrients out of the topsoils and promotes the irreversible hardening of

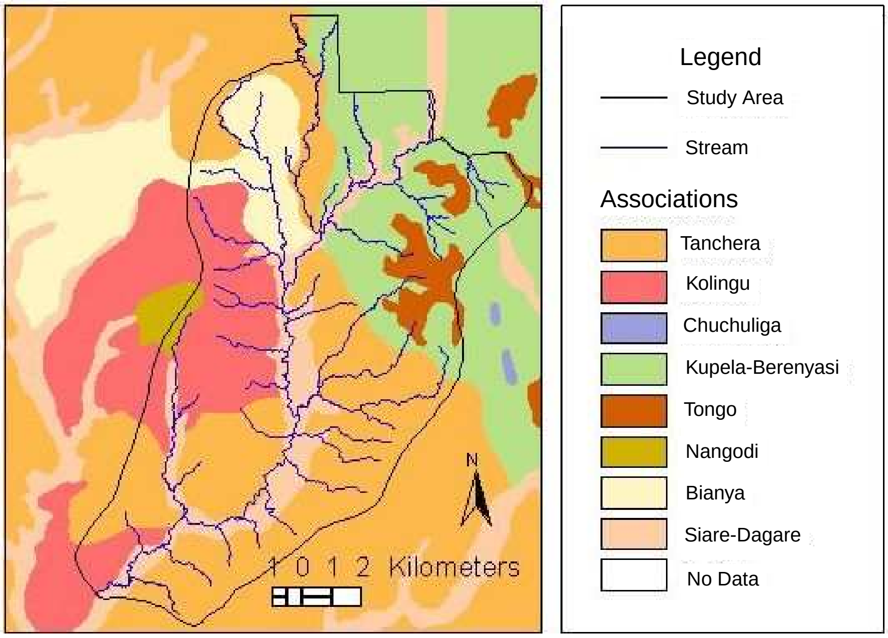


Figure 2.10: Soil associations in the study area

the subsoils, which leads to the development of iron pans.

Vegetation

The study area is a typical savannah parkland, which is a savannah landscape highly modified by agricultural use and settlements, thus being an extreme anthropogenic landscape. The natural tree flora has been severely depleted, apart from small forest patches, mostly consisting of 'sacred groves' along the river banks. Almost every natural tree species, except those with economic or social value, has been systematically eliminated from the farming areas. Such economic tree species include *Vitellaria paradoxa* (55.5%), *Diospyros mespiliformis* (15.5%), *Acacia albida* (9.5%), *Bombax costatum* (2.5%), *Parkia biglobosa* (2.0%), and *Mangifera indica* (2.0%). According to field interviews, these tree species are usually not cut down during land preparation, which is why they became more common over time, giving the impression of planted trees.

Groundwater

Groundwater levels in the study area vary between 1 to 29 m below ground (Martin, 2005),

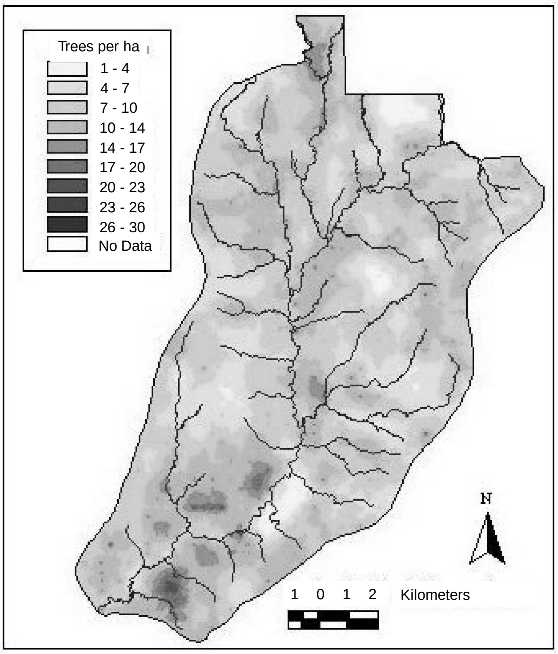


Figure 2.11: Spatial pattern of tree density in the study area

whereby high water tables during the dry season allow irrigated cultivation, mainly by using hand dug wells and dugouts. Except for irrigation, groundwater is withdrawn by boreholes for domestic purposes, such as drinking, cooking, washing, for watering livestock and for building and repair of loam compounds. Use of groundwater for irrigation is currently minimal (Martin 2005). Based on estimations by Martin (2005), total groundwater abstraction in the study area amounts to 167,000 m^3/y (28 %) through hand dug wells/dugouts, and 427,000 m^3/y (72 %) through boreholes. This equals a total groundwater abstraction of 3.6 mm/y. A long term average groundwater recharge of 60 mm/y compares to the total current groundwater abstraction of 3.6 mm/y in the study area (Martin 2005). Groundwater recharge is therefore currently not a limiting factor for groundwater resources development. However, spatial variations of groundwater table and recharge play a decisive role for irrigation-related land-use choices, e.g. the search for suitable land for irrigation.

Data sources

In GH-LUDAS, climate was considered in terms of its temporal but not its spatial variability, as the study area can be assumed to be uniform in terms of climatic conditions, due to the area's relatively small size. Instead, we considered long-term changes in annual precipitation, as this climatic factor plays a major role for local agriculture. These long-term rainfall data, i.e. the annual decrease in precipitation in mm/y, averaged for the next 30 years, were derived from the IPCC data distribution center (www.ipcc-data.org). The values were calculated based on monthly means of daily precipitation (mm/d) within the period 1960 - 2100 as computed by the CSIRO-Mk2 model for each of the four IPCC SRES scenarios. CSIRO-Mk2 is a global grid-based model, with a spatial resolution of 625 km by 350 km. Based on the computed annual rainfall reduction for the pixel the study area is part of, annual rainfall for the next 30 years was calculated and included in the calculations of biomass and crop productivity in GH-LUDAS.

A soil map of the six soil associations in the study area was derived from Adu (1969), which was scanned and digitized. Using this map, a soil fertility and a soil texture map were generated by assigning a specific fertility and texture value to each of the soil associations, respectively. The fertility class, ranging from 'Very Good' to 'Very Poor', and the topsoil textural class of each soil association was extracted from Adu (1969), the latter of which was ranked based on the USDA textural classification, which identifies 12 major soil classes, and 9 further classes for loam and clay (see Brown, 2003). According to this rank, each textural class was assigned a value between 1 (i.e. coarse sand) and 21 (i.e. clay). Accordingly, each fertility class was assigned a value between 1 and 5, representing the five fertility classes ranging from 'Very Good' (5), over 'Good' (4), 'Moderately Good' (3), 'Poor' (2) to 'Very Poor' (1).

A land-cover map was generated based on a ground-truth data set and two satellite images of the study area, including a Quickbird image (DigitalGlobe 2007), and an ASTER image (USGS and Japan ASTER Program, 2007), which can both be acquired from the GLOWA-Volta Project Geo-database at the Center for Development Research (ZEF) in Bonn (www.glowa-volta.de/results_geoportal.html). To interpret these scenes in terms of land cover, a ground-truth survey was conducted in the study area in August 2006. Within the course of this survey, over 1100 GPS points were taken and assigned one of the main land-cover classes 'grassland', 'cropland', 'forest', 'bare land' , and 'water'. The range of these classes had been identified within a 3-days preliminary land-cover survey. The ground-truth survey itself was carried out in daily field visits, whereby the starting point of the GPS measurement was selected on the map prior to each visit to ensure a uniform coverage of the study area by GPS points. From each starting point, measurements were taken every 100 m along all four bearings up to a distance of 3 km to the starting point. Based on this ground-truth data set and the satellite images, supervised classification was applied to generate a local land-cover map (for details see section 5.3.1).

Spatial data on groundwater recharge and groundwater level were derived from time series simulations of a version of the WaSiM-ETH water balance model for the Atankwidi catchment by Martin (2005). For the year 2004, simulated groundwater recharge (in mm/month) for each month and groundwater table (in m below ground) for each day were used. These data, which were produced by WaSim in binary code with a resolution of 100 m x 100 m, were converted to GIS raster layers with the same resolution. Using the map calculator in ArcView GIS 3.2, average monthly groundwater recharge (mm/month) and average groundwater table (m below ground) were calculated and mapped for both seasons.

Topographic features of the study area were derived from a digital elevation model by Le (2006) for the Atankwidi catchment, which had been downscaled from USGS SRTM Elevation data (at the resolution of 92.53 m) to resolutions of 15 m and 30 m. The DEM is available at the GLOWA-Volta Project Geo-database at the Center for Development Research (www.glowa-volta.de/results_geoportal.html. Maps of topographic features for the study area were calculated from the DEM using the surface procedure in ArcView, comprising elevation, upslope contributing area, slope degree, and wetness index. The definition and relevance for land-use/cover change of each of these factors is given in section 5.2.1.

2.5.3 Population characteristics and data generation

The small river basin of the Atankwidi is inhabited by a mainly rural population that in their majority belongs to the Kassena and Nankana ethnic groups. The three main religious groups in the study area comprise the Christian, the Islamic and the traditional religions. Traditional religion is the most common form of worship in the region (46.4 %), followed by Christianity (28.3 %) and Islam (22.6 %). To date, the chieftaincy institution has matured throughout the region, and each village is headed by a chief normally nominated from among a royal family. The chieftaincy system is characterized by a strong hierarchical structure, i.e. political power is exercised through hierarchical levels of authority, from the chief over section and clan heads down to sub-clan heads.

At the lowest level, authority is exercised by the compound head, who is the person in charge of a sub-unit of the clan living together in a compound. In the study area, these compounds are not clustered together, but are rather evenly scattered all over the catchment. The compounds usually give shelter to several households (in average 3.1 households) of the same family. With a total population of about 41000 people (in 2000) and an average of 7.2 persons per household, about 5700 households lived in the study area in 2000. The age structure in the study area is characterized by a large portion (about 48 %) of children (i.e. persons under 18). This large fraction results in a very low mean age of the population (about 24 years), while the mean age of the household head lies higher at 46 years. Education levels seem to have increased strongly during the last decades, as 95 % of persons under 18 have attended at least primary school, while 75 % of the household heads, who mostly belong to the next higher generation, have never been formally educated.

In average, each household owns 2.4 ha of land, which amounts to 0.34 ha/person. Of this area, 68 % is cultivated during the rainy season in average, while the remaining area is left bare as grazing land. The average total gross income from rainfed cultivation amounts to 930 US $, while further 260 US $ are generated by non-farm activities during this season. In the dry season, average total gross income amounts to 510 US $, while in this season, the variation in income is much higher than in the rainy season. This is due to the fact that a part of the households (38 %) is engaged in irrigation, which is a highly profitable activity. In average, about 756 m^2 are cultivated by these households through irrigation.

Most of this information was derived from the data set generated during this study. To obtain these data needed for the implementation and design of GH-LUDAS, two socio-economic surveys were conducted in the study area. In the following, the identification of the relevant survey unit, the sampling strategy and the survey design and realization are presented.

Identification of survey unit

As family relations are highly intervowen in the study area, the family unit for the survey had to be appropriately defined. For our study, the relevant family unit should represent the decision-making entity regarding land use and other activities. Although the compound head, who is the head of the whole compound family, is in charge of the entire land, it made more sense to consider the single household as the relevant unit for the socio-economic surveys, since field investigations had shown that the compound land was usually divided among these households and the decisions about land use were mostly independent, apart from social influence. However, due to interwoven family relations among the inhabitants of a compound and a complex land-tenure system, it was difficult to define the term of household appropriately. The Ghanaian Survey Department usually defines a household as 'the number of people eating from the same pot within a compound'. But this definition was problematic for our study, since family members, each in in charge of own land, were found to still 'eat from the same pot'. Therefore, we defined a household as all people who are dependent on the person who decides about and manages a piece of land, whom we will call the household head. Dependent people are then those who are fed by the yield from the household head's land, and who do not manage own land. Thus, a household is defined by all persons who 'eat from the same *plot*'. This definition of household was then used for identifying household members, their activities and their contribution to household income during the interviews, which were conducted with the respective household head in all cases. Given the homogeneity in livelihood conditions (i.e. housing, food availability, etc.) of the population, the sample size was set to 200 households, which had to be chosen from different compounds in order to meet about 5 % of the compounds in the study area.

Sampling strategy

Since the later data analysis would be based on statistical methods, it was necessary to choose a random sampling strategy. However, not the full set of these 200 households was chosen in a random way, as a part of this sample was specifically dedicated to the assessment of policy impacts. Within GH-LUDAS, these policies include the strategies of dam construction and credit access, which were identified to be the most relevant policy interventions with respect to land use (see Chapter 6). But since functioning dams were absent in the study area, and access to credit minimal, parts of the sample were not chosen randomly, but related to the access to these policies. This way, the sample was split into 140 households to be selected in a random way, 30 households that had once obtained credit, and 30 households from the neighboring Anayere Catchment, where there were operational dams.

To identify the 140 households randomly, a spatial sampling method was chosen, as lists of names of household heads in the study area were not available. In order to make sure that the composition of the sample would reflect the overall composition of the population in the study area, a stratification method for the sample had to be applied. For this, the study area was divided into 8 units, demarcated by major roads and the main river to serve as landmarks for the sampling, to be directly carried out in the field. For each of these units, the single compounds were digitized using a high-resolution Quickbird image, which made the compounds easily identifiable. According to the number of compounds in each unit, the percentage of households selected was calculated for each unit. By using this strategy, an equalized representation of the population was ensured, and, based on this stratification method, random households could be identified in the field (Figure 2.12).

Survey design and realization

As the climatic conditions cause differences in land-use behavior and livelihood strategies between the dry and the rainy season, two socio-economic surveys were conducted, one for each season. The dry-season-related survey was conducted in July 2006, while the rainy-season-related survey was conducted after the final harvest in late November 2006. In both surveys, the same set of selected households (200 households) was interviewed, and the same questionnaires were used for all households. The main targets of the two surveys were the generation of a household-based data set and a plot-based data set. The purpose of the generation of a household-based data set was to characterize household agents in terms of their household state (e.g. household assets, livestock, etc.) and their decision-making sub-models, while the plot-based data set was used to characterize the biophysical state (e.g. land use) and the biophysical sub-models (e.g. agricultural productivity) of landscape agents.

Dry-season survey

The main goal of this survey was to develop i) the basis of the plot-based data set for each household, i.e. to record land-use type, location and size of each cultivated plot (in either season 2006), ii) to collect data on management, agricultural input (i.e. labor, chemicals) and yield for each plot cultivated in the dry season 2005/2006, and iii) to record engagement in

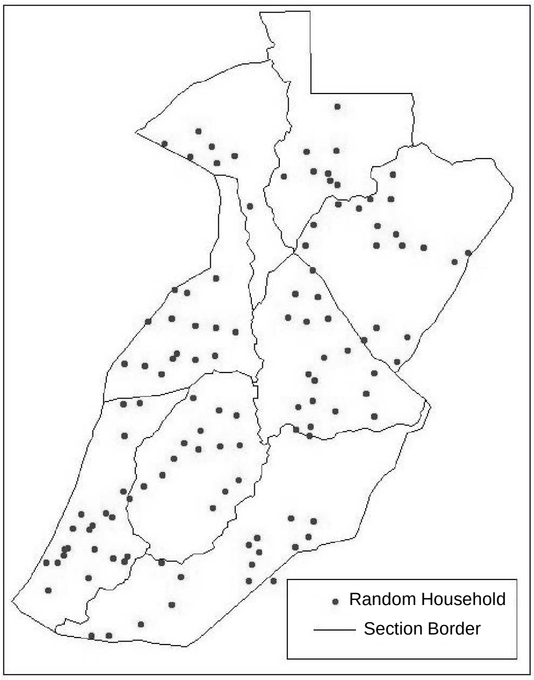


Figure 2.12: Locations of 140 randomly selected households in the Atankwidi catchment

and labor allocation to each of the income-generating non-farm activities. Since information from farmers about their plot sizes turned out to be unreliable, the single plots were measured by GPS, i.e. waypoints were taken by walking around the plot, and its size finally calculated using the XTools extension of ArcView GIS 3.2. In total, 814 plots were measured, accounting to about 4 plots per farmer in average.

Both the relevance to land-use change and the applicability of the questionnaire were examined before finalizing the questionnaire. The relevance of information was assessed with help of LUCC modeling expert Dr. Quang Bao Le (Center for Development Research), followed by the necessary modfications of the contents of the questionnaire. The way of data acquisition and the form of questions was improved under assistance of social scientists, local experts, and field experiments. The survey itself was carried out by four enumerators, who had been educated and trained in field exercises in the proper application of the questionnaire and the use of GPS units for plot measurement.

To ensure a stratified distribution of the random 140 interviewees within the catchment, the households were contacted and selected according to a specific random sampling procedure one day before the interview. According to the number of households to be interviewed in each unit (see above), interviewees were selected randomly within each unit in the field. Using the demarcations of the unit for field orientation, households were selected systematically every 1.2 km along certain bearings by using GPS on a motorbike. Compromises had to be made due to unpassable rivers and rocky areas and also due to the necessity to avoid large distances between the households, as the enumerators were only equipped with bicycles. The remaining 60 households, which comprised households with access to credit and reservoir cultivation, were organized by contact persons. Here, a random approach was impossible, due to the low number of eligible candidates and the fact that information on credit and those who obtained some was strictly confidential; thus these persons had to be organized by a confidant.

Rainy-season survey

While the focus of the dry-season survey was mainly on dry-season activities and plot measurements, the contents of this second survey had a broader scope and were more extensive than that of the previous one. The decision to shift the main interview part to the second questionnaire was based on the fact that farmers in the study area were usually less occupied after the end of the rainy season, which ensured a more relaxed interview atmosphere and thus a higher reliability of information. The range of questions in this questionnaire covered i) plot-based data for the last rainy season (e.g. management, labor input, crop yield), ii) income data (e.g. from non-farm activities), iii) livelihood data (e.g. demographic structure of household, household assets), and iv) policy access (e.g. extension service, credit access). The kind and range of questions within these blocks were selected according to the experiences during an informal interview campaign conducted before the survey.

In order to ensure an accurate recording of the plot-based information, a reliable method for the identification of the single plots during the interviews had to be developed. Detailed digital maps of the plots of each household had been prepared prior to the survey, which facilitated the communication between the interviewer and the interviewee regarding the plot identification. These maps were developed on the basis of the GPS measurements of the first survey, using ArcView GIS 3.2. In order to increase the identifiability of the plots other objects like streams and roads were also mapped, serving as additional reference features. Finally, the mapped plots were labeled with different colors, each color representing a specific land-use type, which further eased the description of certain plots during the interviews.

The six enumerators who conducted the survey were trained to use these maps properly, i.e. to identify the bearings of the various plot locations and to indicate them to their interview partners. The training also comprised exercises in plot description in terms of size, land use or distance to river or roads to enhance plot-based communication, including training in the use of the questionnaire in the field as well as in supervised ’dry runs’. The interviewees were localized with help of a GPS unit and contacted one day in advance to make appointments for the next day’s interviews.

3 SPECIFICATION OF GH-LUDAS

3.1 Introduction

One weakness of MAS is that it is not possible to establish a mathematical proof of the obtained results (Bousquet and Le Page, 2004; Axtell, 2000). However, the model's credibility can be enhanced through several strategies. The first strategy is to assess the relevance of the hypotheses of the model. As such, assumptions underlying the model should be clearly stated and justified. We will follow this strategy throughout the model specifications (Chapters 3 to 5). In addition, we present descriptions of the conditions and practices as observed in the study area, thereby enhancing the credibility of the model assumptions.

The second strategy is to provide a rigorous presentation of the structure of the model (Le et al., 2008; Bousquet and Le Page, 2004) to provide a transparent model description, such that the internal mechanisms can be easily retraced. This way, the specifications of the model focus on two aspects: i) system architecture and ii) system implementation (Cioffi-Revilla and Gotts, 2003). Accordingly, we will present a fully parameterized architecture of GH-LUDAS based on the conceptual model described in Chapter 2, and will outline the simulation protocol for this architecture, including the initialization of the model and the time-loop procedure run during simulation. We will elaborate the system architecture and model implementation as follows:

- The **Human Module** represents the system of human population in which farming households are treated as human agents, endowed with agent-specific variables, parameters and connected to a model of land-use decision-making (Decision Module).

- The **Landscape Module** represents the system of the landscape environment in which congruent land patches are considered as environmental agents, endowed with own parameters and biophysical sub-models.

- The **Decision Module** is a decision-making routine integrated into the human agent simulating household-specific land-use behavior.

- The **Global-policy Module** is an external module in which model users can set the values for selected policy and demographic factors under IPCC climate scenarios, thereby exploring alternative pathways for land-use/cover and related socio-economic conditions.

- The **simulation protocol of GH-LUDAS** which delineates the sequence of sub-procedures during simulation runs.

3.2 System of human population: the Human Module

The Human Module represents the human part of the coupled human-environment system underlying land-use and land-cover change. The dynamics of this module emerge from the local interactions between household agents and their immediate environment. Since these dynamics are scale dependent, with different processes acting at different levels, the human system is designed as a hierarchy of three interrelated levels of organization: household agent, groups of household agents, and the whole population (see Figure 3.1). The process of land-use and land-cover change at the highest level of the whole population is then the result of the interactions at lower levels, which represent real-life individual (and group) land-use behavior.

The household agent represents individual farming households within the study area (section 2.5.3). The structure of an individual household agent comprise four components: i) a data set of household variables (called Household Profile), which play a role in the land-use decision-making processes and other model routines, ii) a rule set defining the changes within this set of these household variables (called Internal Rules) iii) the agent's Landscape Vision, a subset of the whole landscape in which the agent can act on and interact with other agents, and iv) the Decision Module, a complex of procedures mimicking decisions a farming household has to make, e.g. land-use choice or the decision to get involved in irrigation farming. These decisions are dependent on both the Household Profile, the policy parameters as well as on the state of the agent's Landscape Vision.

Groups of household agents are collections of household agents with a similar livelihood typology, thus being assumed to have a similar land-use behavior. This group-wise land-use behavior is represented by group behavior parameters, which have been derived by empirical group data. According to the group an agent belongs to, the group behavior pa-

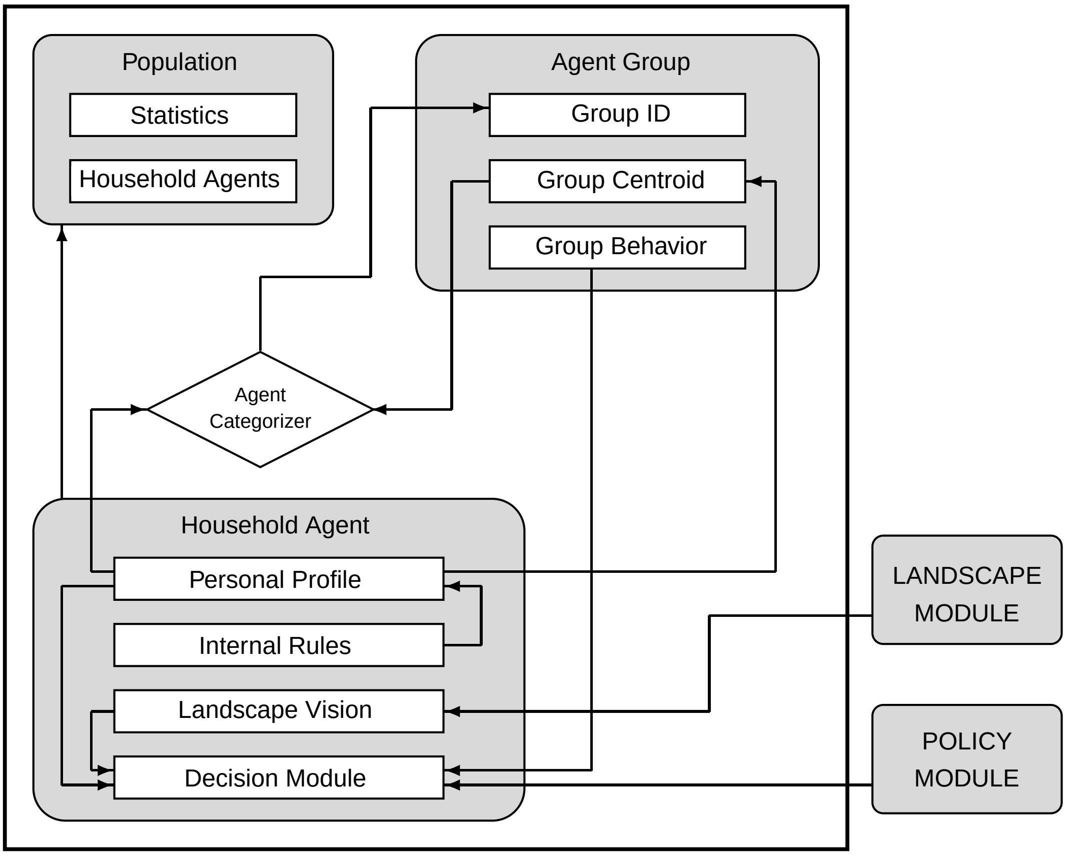


Figure 3.1: Integration of the Human Module in GH-LUDAS

rameters are fed into the agent's Decision Module. Moreover, household agents can change their agent group, and thus their land-use behavior. At the end of each model step, which is represented by one year, a household agent is allocated to the group that has the highest similarity with the agent. If an agent changes his agent group, he will also adopt the new behavior parameters, which will in turn affect his decision structures. Thus, the agent groups play a crucial role in this model of land-use/cover change, as they represent the change in land-use preferences among household agents during time.

The population is the collection of all agents, and its pattern is the emerging result of the processes at the lower levels of the hierarchical system. Statistic procedures are calculated to analyze the characteristics of the population during time, such as mean total gross income and the Gini Index of income distribution.

3.2.1 Structure of the household agent

As already outlined above, the structure of the household agent is as follows:

Household Agent = (Household Profile, Internal Rules, Landscape Vision, Decision Module)

In the following, we will describe all of these four components in detail, and introduce the range of variables used within the model of the household agent.

Household Profile

The Household Profile ($H_{profile}$) includes seven sub-types of variables: social identity and livestock ($H_{soclive}$), human resources (H_{human}), land resources (H_{land}), financial resources (H_{income}), environmental variables (H_{env}), irrigation variables (H_{irr}), and policy-related attributes (H_{policy}):

$$H_{profile} = \{H_{soclive}, H_{human}, H_{land}, H_{income}, H_{env}, H_{irr}, H_{policy}\}$$

The social identity and livestock factor ($H_{soclive}$) includes age of the household head (H_{age}), village code ($H_{village}$), the number of wives of the household head (H_{wives}), the number of cattle belonging to the household (H_{cattle}), the livestock index ($H_{livestock}$), and the group membership (H_{group}):

$$H_{soclive} = \{H_{age}, H_{village}, H_{wives}, H_{cattle}, H_{livestock}, H_{group}\}$$

The agent's human resources (H_{human}) consist of household size (H_{size}), labor availability (H_{labor}), the dependency ratio (H_{depend}), and $H_{pool\ dry}$ and $H_{pool\ rainy}$, which are the labor pool in the dry respectively in the rainy season (in labor days). The dependency ratio is the ratio of labor availability and household size, representing the composition of workers and non-workers in the household:

$$H_{human} = \{H_{size}, H_{labor}, H_{depend}, H_{pool\ dry}, H_{pool\ rainy}\}$$

Household land resources H_{land} comprise six variables including total area owned by the household ($H_{holdings}$), total area owned per capita ($H_{holdings\ per\ cap}$), cultivated area in the dry season ($H_{cult\ dry}$), cultivated area in the rainy season ($H_{cult\ rainy}$), and land-use composition vectors for each of the two seasons ($[H_{\%\ i\ dry}]$, i = (1 . . . N)) and ($[H_{\%\ i\ rainy}]$, i = (1 . . . M)):

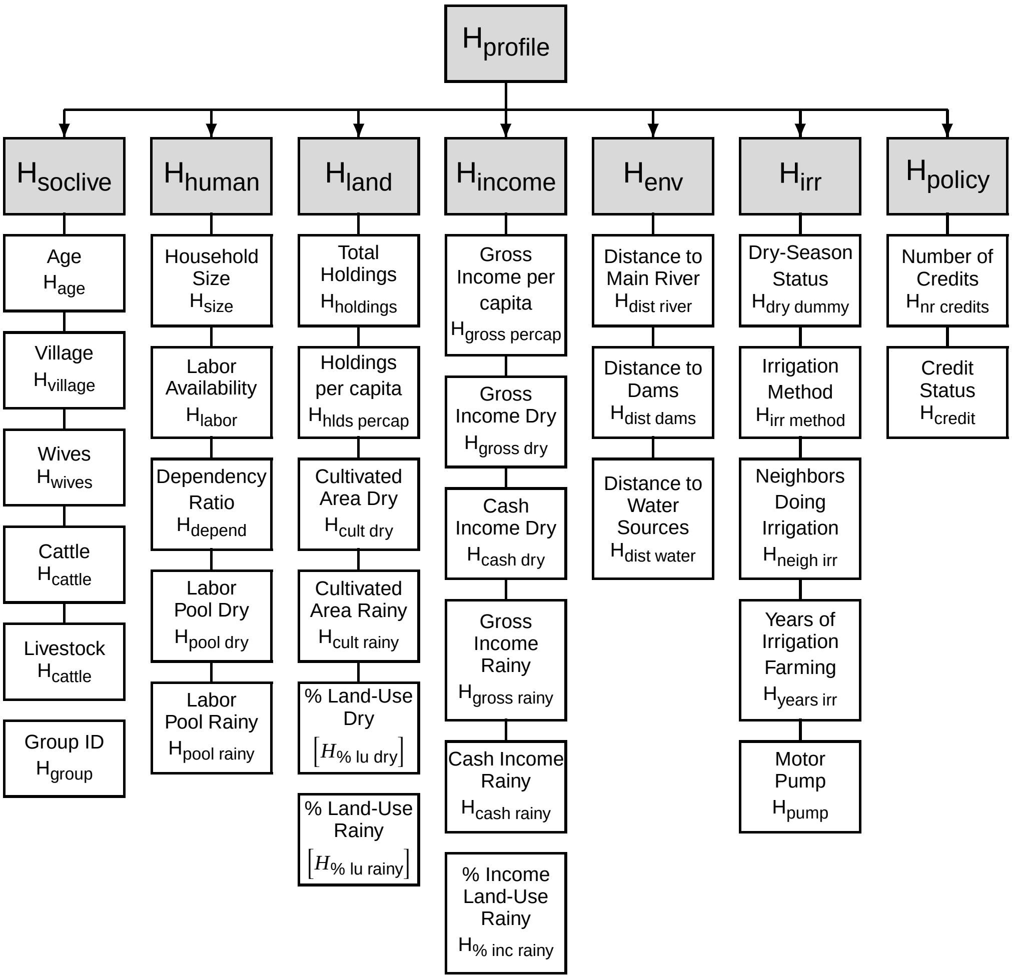


Figure 3.2: Household Profile

$$H_{land} = \{H_{holdings}, H_{holdings\ per\ cap}, H_{cult\ dry}, H_{cult\ rainy}, [H_{\%\ i\ dry}], [H_{\%\ i\ rainy}]\}$$

where i indexes the dry-season respectively the rainy-season land-use types.

The factor of financial resources of the household (H_{income}) comprises total gross income per capita ($H_{gross\ per\ cap}$), gross and cash income in the dry season ($H_{gross\ dry}$) and ($H_{cash\ dry}$) respectively, gross and cash income in the rainy season ($H_{gross\ rainy}$) and ($H_{cash\ rainy}$) respectively, as well as an income composition vector of income from rainy-season cultivation ($[H_{\%j}]$, $j = (1 \ldots M)$), with j indexing the rainy-season land-use types:

$$H_{income} = \{H_{gross\ per\ cap}, H_{gross\ dry}, H_{cash\ dry}, H_{gross\ rainy}, H_{cash\ rainy}, [H_{\%j}]\}$$

The environmental variables (H_{env}) include distances from the compound of the household to main river ($H_{dist\ river}$), to dams ($H_{dist\ dams}$), and to water sources in general ($H_{dist\ water}$), which represents the distance to irrigable areas and is calculated as the minimum of $H_{dist\ river}$ and $H_{dist\ dams}$:

$$H_{env} = \{H_{dist\ river}, H_{dist\ dams}, H_{dist\ water}\}$$

The state of a household agent regarding irrigation (H_{irr}) includes five variables: i) a dummy variable variable ($H_{dry\ dummy}$) indicating if the farmer is inherently capable of doing irrigation, ii) a second variable reporting the kind of irrigation technology ($H_{irr\ method}$) for those households with ($H_{dry\ dummy} = 1$), ranging from bucket irrigation, pump irrigation to dam irrigation, iii) a variable indicating the percentage of household heads practicing irrigation among the five nearest households, iv) a variable representing the number of years the farmer has practiced irrigation ($H_{years\ irr}$), and v) a dummy variable (H_{pump}) indicating whether the household owns a motor pump.

$$H_{irr} = \{H_{dry\ dummy}, H_{irr\ method}, H_{neigh\ irr}, H_{years\ irr}, H_{pump}\}$$

The variables of H_{policy} include two variables: i) the credit status H_{credit}, a dummy variable indicating whether the household has obtained credit in the current year, and ii) $H_{nr\ credits}$, the number of credits the household has obtained so far.

Internal Rules

During model run, most of the model variables are subjected to changes over time. Changes in the performance of the household module involve i) modifications of variables of the Household Profile of agents, and ii) the creation and deletion of agents. The Internal Rules only comprise simple rules defining the changes of household variables, while the deletion and creation, which involve more complicated mechanisms, are described in the subsequent section.

It is important to understand the kinds of changes the variables of Household Profile undergo over time. We can categorize these variables into four categories: i) variables that undergo no change, ii) variables whose changes are due to the effects of household agent activities during simulation (e.g. changes in gross annual income and/or land resources), iii) variables whose changes are defined by natural events, independent of the agent's actions (e.g. the increase of the age of the agent), and iv) changes that are defined by settings outside the system, e.g. policies. The only Household Profile variables that undergo no changes are village code and distance to main river. All variables that are among the sets of H_{land}, H_{income} and H_{irr} belong to the second category and are thus subjected to the internal changes within the system. However, changes of variables within the third and fourth category have to be modeled explicitly, since they are not a result of human-environmental interactions. This task will be accomplished by the procedures of the Internal Rules:

Variables of the third category that undergo natural changes include H_{age}, H_{wives}, H_{cattle}, $H_{livestock}$, H_{labor}, and H_{depend}.

The rule for the changes in age is simple. The age of the household head H_{age} will increase by 1 after each time step, until the upper bound max_{age} is reached. The rule is as follows:

$$^{t+1}H_{age} = \begin{cases} {}^{t}H_{age} + 1 & \text{if} \, {}^{t}H_{age} < max_{age} \\ \text{die} & \text{if} \, {}^{t}H_{age} = max_{age} \end{cases} \tag{3.1}$$

All other variables of this third category are also event-driven phenomena, but they are affected by many causes that are beyond the scope of our study. It is, therefore, reasonable to proximate stochastically the values of these household attributes within uncertainty ranges of the values of the previous time step. For all these variables, the kind of rule follows the same pattern. We will exemplify this pattern by the example of H_{cattle}:

$$^{t+1}H_{cattle} = round({}^{t}H_{cattle} - \sigma_{cattle} + random(2 \cdot \sigma_{cattle})) \tag{3.2}$$

where $^{t+1}H_{cattle}$is the number of cattle at time step $t + 1$, $^{t}H_{cattle}$ the number of cattle at time step t, and σ_{cattle} the standard deviation for H_{cattle} calculated from empirical household data sets. The random command determines a random number within $[0, 2 \cdot \sigma_{cattle}]$. Thus, $^{t+1}H_{cattle}$ lies within an uncertainty range of $[-\sigma_{cattle}, \sigma_{cattle}]$ around the value of $^{t}H_{cattle}$.

Below we will give the rules for all other variables of the third category, following the same kind of rule as in the example for cattle. As some of the variables are regarded as integers, they need a round command to ensure integer outcome values.

$$^{t+1}H_{\text{wives}} = round(^{t}H_{\text{wives}} - \sigma_{\text{wives}} + random(2 \cdot \sigma_{\text{wives}})) \tag{3.3}$$

$$^{t+1}H_{\text{size}} = round(^{t}H_{\text{size}} - \sigma_{\text{size}} + random(2 \cdot \sigma_{\text{size}})) \tag{3.4}$$

$$^{t+1}H_{\text{labor}} = round(^{t}H_{\text{labor}} - \sigma_{\text{labor}} + random(2 \cdot \sigma_{\text{labor}})) \tag{3.5}$$

$$^{t+1}H_{\text{depend}} = round(^{t}H_{\text{depend}} - \sigma_{\text{depend}} + random(2 \cdot \sigma_{\text{depend}})) \tag{3.6}$$

where σ is the standard deviation of the single variable derived from the empirical data set. (The annual variation of the livestock index $H_{\text{livestock}}$ will be determined by the specific biophysical sub-model of livestock dynamics.)

Variables of the fourth category comprise exclusively variables that are set externally, i.e. policy access variables. Variables that are counted among this set comprise distance to dams $H_{\text{dist dams}}$, distance to water sources $H_{\text{dist water}}$, current credit access H_{credit}, and number of credits received so far $H_{\text{nr credits}}$.

Since new dams can be added to the initial settings of the landscape as a policy, the distance to dams for households also has to be changed automatically. For this, a routine checks the distances to the various dams, and finally chooses the minimum. The procedure can be described as follows:

for all dams : set current-dist-dam (distance from house to dam) if ($H_{\text{dist dams}}$ > current-dist-dam) [set $H_{\text{dist dams}}$ current-dist-dam]

The distance to water sources distance to water sources is then defined as the minimum of the distance to dams and the distance to the main river:

$$H_{\text{dist water}} = \min(H_{\text{dist dams}}, H_{\text{dist river}})$$

The percentage of households obtaining credit is given outside the model as a policy parameter, whereas the amount of credit is fixed, and the period of credit provision is set to 2 years. This was the observed pattern within the study area, and cannot be changed within the model, since possible effects of a different credit pattern cannot be derived from the empirical data set. Within the model, credits are given randomly within the population of household agents, whereas those with a lesser number of credits obtained so far are favored. The variable of H_{credit} can therefore be determined as follows:

$$ {}^{t}H_{\text{credit}} = \begin{cases} 1 & \text{if } {}^{t}credit = \text{true} \\ 0 & \text{if otherwise} \end{cases} \tag{3.7} $$

where ^{t}credit denotes whether a household was chosen to access credit in time step t. Changes in the number of credits that households obtained $H_{\text{nr credits}}$ are calculated accordingly:

$$ {}^{t+1}H_{\text{nr credits}} = \begin{cases} {}^{t}H_{\text{nr credits}} + 1 & \text{if} \quad {}^{t+1}credit = \text{true} \\ {}^{t}H_{\text{nr credits}} & \text{if} \quad \text{otherwise} \end{cases} \tag{3.8} $$

The other two components of the household agent structure, Landscape Vision and Decision Module, will be described in later chapters. The Landscape Vision, as an integral part of the multi-level organization of the landscape, will be handled within the description of the patch-landscape module. The Decision Module will be outlined in a separate section of this chapter (section 3.4).

Creation and deletion of agents

Agents who reach their maximum age (see equation 3.1), are deleted. If agents within the same compound id, i.e. living in the same compound, exist, all land belonging to the dead agent is equally distributed among these. If no such agents exist, a new agent is created within this compound who inherits the land;

for all patches with (P_{owner} = dead agent), set P_{owner} = new agent

Apart from land, the new agent inherits the values for all variables, that are household related (e.g. cattle amount, household size, ownership of motor pump), while personal variables (e.g. number of wives, age, years of irrigation experience) are assigned values from a random agent with age under 30. Variables concerning the agent's livelihood strategy (e.g.

group id) are classed among personal variables and thus obtain their values from the random agent.

But agents are not only created as successors for deleted agents, but are also created in the course of population growth. In each time step, the population of households is recalculated, based on the logistic growth function:

$$P(t) = \frac{CP_0e^{rt}}{C + P_0(e^{rt} - 1)} \tag{3.9}$$

where $P(t)$ is the population size at time step t, P_0 the initial population size at time 0 (i.e. the year 2006), and C and r parameters. In each time step, $P(t) - P(t-1) + D(t)$, new agents are created, where $D(t)$ is the number of agents deleted in time step t without successor. These agents are allocated randomly to the compounds of the study area, i.e. to patches with $P_{compound} = 1$. The locations of these patches had been determined prior to the development of GH-LUDAS (for details see section 3.6). These new agents adopt all their variable values from another random agent under age of 30. To ensure that all new agents obtain land, these agents are given priority within the moving phase of land acquisition (see section 3.4.2), where agents search for new patches. That is, new agents are allowed to search for unused patches before any other agent. If any of these unused patches are not owned by anybody, the ownership of these patches is transferred to the new agent. This is the first mechanism that ensures the ownership of patches. The second mechanism consists of the inheritance system as defined above. In case an agent (without successor) dies, the land is equally distributed among the other compound members, including the formerly new agent.

All these mechanisms are geared to observations in the study area. The inheritance system as described here ensures both inheritance with a successor and without, which both happens. Although in cases of a dissolved household, i.e. cases without a successor, the available land is not equally distributed among the remaining households, but is usually distributed according to internal family hierarchies, the approach of equal portions was the most straightforward method to describe the complicated inheritance structure.

3.2.2 Structure of the household agent group

The household agent group is a collection of household agents with similar socio-economic features and is thus assumed to exhibit similar decision-making behavior. The separation of these groups is based on so-called grouping criteria, which form a subset of the set of Household Profile variables. After each time step, every agent is assigned to the group with the most similar values among the grouping criteria. According to the group the agent belongs to, he is endowed with the group-specific set of behavioral parameters. The identification process of the range and values of these parameters as well as of the range of grouping criteria will be outlined in detail in Chapter 4. Following this mindset, the structure of the household agent group can be formally expressed as follows:

$$\text{Household Agent Group} = \{\mathrm{G}_{\mathrm{id}}, \mathrm{G}_{\mathrm{cat\ coeff}}, \mathrm{G}_{\mathrm{behavior}}\}$$

where G_{id} is the group identification code, $\mathrm{G}_{\mathrm{cat\ coeff}}$ the categorizer coefficients of grouping criteria, and $\mathrm{G}_{\mathrm{behavior}}$ the set of group-specific behavior parameters.

Categorizer coefficients and Agent Categorizer

The set of grouping criteria is designed to represent the differences among the agent groups, whereby each group has its own set of categorizer coefficients that serve as weights for these criteria. These coefficients play a role in the routine that assigns an agent to a certain agent group, called the Agent Categorizer. The Agent Categorizer is an automatic classification procedure that categorizes all agents into their nearest groups after each time step. It consists of an m-logit model, which calculates the distance of each agent to each group, and an assignment procedure, which finally assigns the agent to his 'nearest' group. The distance of an agent A to group g $Dist_g$ is calculated as:

$$Dist_g = \frac{e^{\alpha_g + \sum_i \beta_{ig} V_i}}{\sum_h e^{\alpha_h + \sum_j \beta_{jh} V_j}} \tag{3.10}$$

where V_i are the values of the grouping criteria of agent A, α_g a constant, and the β_{ig} the categorizer coefficients for group g (The range of grouping criteria as well as the values of the categorizer coefficients as a result of the m-logit model are presented in Chapter 4).

According to the calculated distances, agent A is then assigned to the group with minimum $Dist_g$:

$$H_{\text{group}} = g \text{ with}(Dist_g = min_h\{Dist_h\}) \tag{3.11}$$

After the agent has been assigned to his nearest group, the agent will adopt the new behavior template of the group, as oulined in the following section.

Group behavior

The agent group behavior G_{behavior} consists of a vector of behavior parameters that are identical for all group members:

$$G_{\text{behavior}} = \{[Labd_g], \sigma_{dg}, [Labr_g], \sigma_{rg}, [\beta_{ig}], \sigma_{ig}, [\%dry_{jg}], \sigma_{jg}\} \tag{3.12}$$

where $[Labd_g]$ and $[Labr_g]$ are the vectors of labor allocation percentages in the dry and rainy season respectively, $[\beta_{ig}]$ a vector of preference coefficients used for the m-logit model of land-use choice for the rainy season, and $[\%dry_{jg}]$ a vector of percentages of dry-season land-use types of the cultivated area in the dry season; σ_i is the respective standard error for each vector.

The labor allocation vectors consist of the labor allocation percentages, which represent the percentage of the total labor pool allocated to a single activity by a household. The range of activities is the same for both seasons, and comprises cultivation, trading, food processing, handicrafts, migration and other income-generating activities (e.g. white collar jobs). During focused interviews with local farmers and field observations, these six activities have been identified to be the main income-generating activities among the local population.

Whereas the choice among rainy-season land-use types is modeled by an m-logit regression with β_{ig} being the respective preference coefficients (see section 2.4.2), a m-logit model was not used for predicting the choices among dry-season land-use types. Instead, simple group-specific percentages of land-use types were used, since the available data set about dry-season farming was not large enough to set up an m-logit model for dry-season land-use choice. Moreover, differences in cropping patterns among the two land-use types within this season were so small that explanatory variables could not adequately reflect these differences. Therefore, group-wise percentages of the land-use types of the cultivated area were used, which turned out to be a more robust approach. The identification of the decision variables as well as the calculation of the preference coefficients will be outlined in Chapter 4 for both seasons.

All group behavior parameters were determined by statistical analysis of group-wise empirical data sets, being the same for all group members. However, within the model, the behavior parameters for a single agent are generated by random values around the fixed parameters of the group, bounded by the related standard error σ_g. For instance, the preference coefficient for land-use type i for the rainy season will be a random value within the range $(\beta_{ig} - \sigma_{ig}, \beta_{ig} + \sigma_{ig})$. Such slight deviations of the average group behavior together with individual Household Profile variables ensure a heterogeneous decision behavior even among agent group members.

3.2.3 Population

The population class is the collection of all household agents, together with a database of statistical parameters about the population. In land-use and land-cover change research, not only the changes in land use or cover, but also the related changes in the socio-economic structure of the population need to be monitored. This will be covered by various statistical parameters about income patterns during the simulation runs. The class of population can therefore be formally expressed as:

$$\text{Population} = \{\{\text{Agents}\}, \text{Stat}\}$$

where Stat consists of the following population performance indicators: i) overall average income per household, ii) overall average annual income per capita, and iii) the Gini Index of household income distribution. The Gini Index is a statistical measure to describe the degree of disparity within a pre-defined population, and is most often applied to measure the equity of income distribution (Gakidou et al., 2000). The values of this coefficient lie within the range of 0 and 1, and the higher the value, the higher the inequality. Mathematically, the Gini Index is the standardized area between the Lorenz Curves of a uniformly distributed population and the observed population (Dorfman, 1979). The Lorenz Curve of income is a graph that for the bottom x % of households shows what percentage y % of the total income they have. The percentage of households is plotted on the x-axis, the percentage of income on the y-axis. If the curve is a diagonal line, the population is in a state of total equity (see Figure 3.3). An unequal distribution will result in a curve below the diagonal. The Gini Index is then calculated as the ratio of the area between the two curves and the area below the diagonal.

3.3 System of the environment: the Landscape Module

The Landscape Module represents the state and processes of the environmental part of the coupled human-enviroment system of land-use/cover change. Just as the Human Module is represented in the form of a three-fold hierarchy, this module is also conceptualized as an organization of three levels: the landscape agent or patch, the Landscape Vision, and the entire landscape (see Figure 3.4). The landscape agents are represented by congruent land patches of size 30 m x 30 m, consisting of two main components: the patch's state variables and the internal ecological sub-models. The state variables comprise both biophysical/environmental attributes (e.g. soil texture, distances), which are independent of human actions, and variables which are related to the human part such as land tenure and use. The internal ecological sub-models consist of i) productivity functions for all land-use types of both seasons, ii) a land-cover transformation model, which regulates the conversion of one land-cover type to the other, and iii) a livestock dynamics sub-model.

As already outlined in the previous section, the Landscape Vision is the environment of a household agent in which he sets actions. Each household agent has his own Landscape Vision, which, in multi-agent-based terms, consists of a set of landscape agents located around the compound patch of the household agent. Within this environment, the household agent has (limited) insight into its features and attributes, makes land-use decisions and creates impacts on this environment. These impacts are accumulated over time and aggregately result in spatio-temporal dynamics of the overall landscape (Le, 2005).

The entire landscape is the collection of all landscape agents or patches, being the emergent result of both the changes and interactions of the single landscape agents. Due to these interactions, which can be either direct or indirect, i.e. mediated through household agents, the change of the entire landscape is not only the sum of the single changes of the patches, but must be rather regarded as an emergent phenomenon created by the interactive

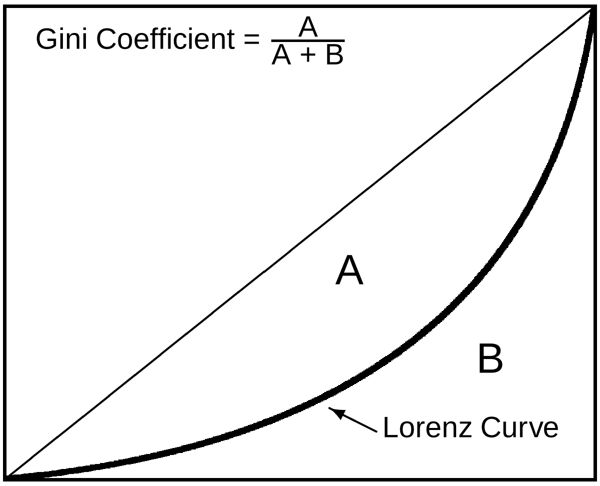


Figure 3.3: Lorenz curve and Gini index

collective of landscape and household agents.

3.3.1 Structure of the landscape agent

The structure of the landscape agent can be formally expressed as:

$$\text{Landscape Agent} = \{\text{Patch Profile, Eco-Sub-models}\}$$

where the Patch Profile is the state of the landscape agent, including both human-related and biophysical variables, and Eco-Sub-models is the collection of all ecological sub-models including the productivity functions and the land-cover transformation model. A detailed specification of these components is given below.

Patch Profile

The set of state variables of a patch consists of six components: biophysical variables ($P_{biophys}$), environmental variables (P_{env}), tenure properties (P_{tenure}), the land-use/cover status (P_{status}), yield (P_{yield}), and irrigation-related parameters (P_{irr}):

$$\text{Patch Profile} = \{P_{biophys}, P_{env}, P_{tenure}, P_{status}, P_{yield}, P_{irr}\}$$

Biophysical conditions comprise the following variables:

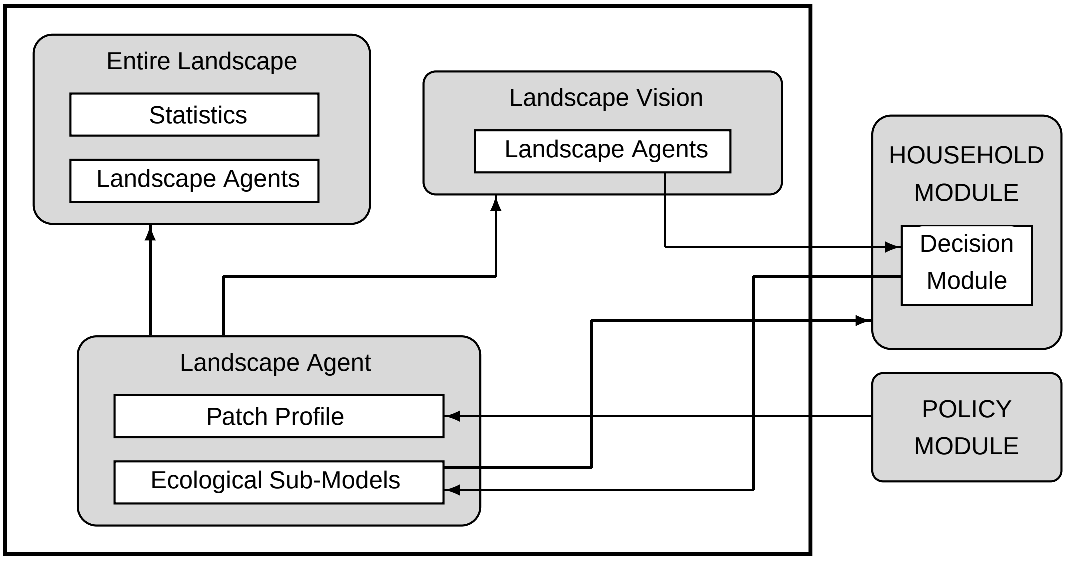


Figure 3.4: Integration of the Landscape Module in GH-LUDAS

$$P_{\text{biophys}} = \{P_{\text{soil fertility}}, P_{\text{soil texture}}, P_{\text{gwl}}, P_{\text{gwr}}, P_{\text{wetness}}, P_{\text{upslope}}\}$$

where $P_{\text{soil fertility}}$ and $P_{\text{soil texture}}$ are soil type and soil texture respectively, P_{gwl} and P_{gwr} are average groundwater depth and groundwater recharge, respectively, during the dry season. P_{wetness} is the topographic wetness index, and P_{upslope} is the upslope contributing area.

The environmental variables exclusively comprise distances to environmental features:

$$P_{\text{env}} = \{P_{\text{dist river}}, P_{\text{dist dams}}, P_{\text{dist water}}, P_{\text{dist border}}\}$$

with $P_{\text{dist river}}$ being distance of the patch to the main river, $P_{\text{dist dams}}$ distance to dams, and $P_{\text{dist water}}$ distance to water sources, i.e. main river and dams. Thus, this variable is just the minimum of the two previous ones, similar to the calculation of the equivalent variable for household agents. $P_{\text{dist border}}$ is the distance to the national border to Burkina Faso in the north.
The tenure properties of the patch can be summarized as follows:

$$P_{\text{tenure}} = \{P_{\text{owner}}, P_{\text{dry-user}}, P_{\text{rainy-user}}, P_{\text{dist user}}\}$$

where P_{owner} indicates the household agent who owns the patch. But since the user of the patch does not necessarily need to be the owner, we also included the variables of $P_{\text{dry-user}}$,

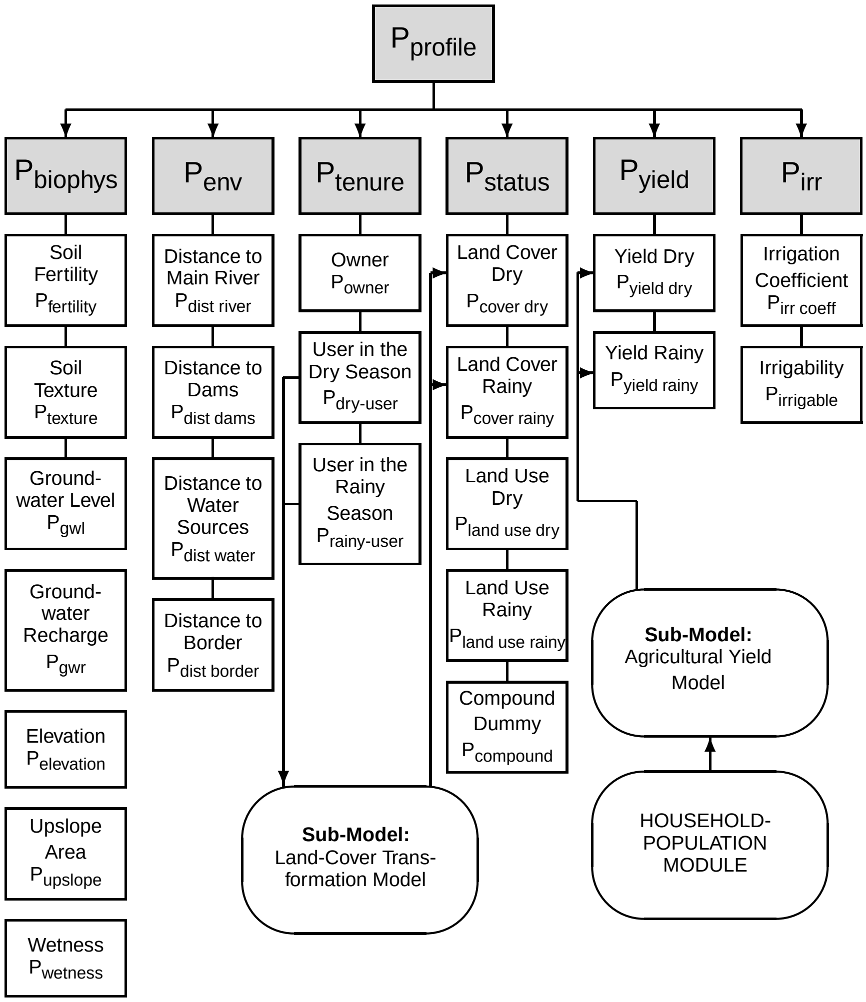


Figure 3.5: Patch Profile

indicating the agent who uses the patch in the dry season, and $P_{rainy\text{-}user}$, the agent who uses it in the rainy season. If the patch is not used or owned by anybody, the variables will get the value 'nobody'. $P_{dist\ user}$denotes the distance of the patch to its rainy-season user (this variable is only needed for the rainy season).

The land-use/cover status of a patch P_{status} comprises the following variables:

$$P_{status} = \{P_{cover\ dry}, P_{cover\ rainy}, P_{land\ use\ dry}, P_{land\ use\ rainy}, P_{compound}\}$$

with $P_{cover\ dry}$ and $P_{cover\ rainy}$ indicating the land-cover type in the dry and rainy season, respectively, $P_{land\ use\ dry}$the land-use type in the dry season, and $P_{land\ use\ rainy}$ the land-use type in the rainy season. If a patch is not used during a specific season, the value of the land-use type is set to 0 for that season. $P_{compound}$ is a dummy variable, indicating whether a compound house is present on the patch.

The yield status of the patch simply reports the amount of yield in the local currency (Ghanaian Cedis) from the dry and the rainy season:

$$P_{yield} = \{P_{yield\ dry}, P_{yield\ rainy}\}$$

The category irrigation includes the following two variables: $P_{irrigable}$, being a dummy variable indicating if a patch is irrigable, and $P_{irr\ coeff}$, which is called the irrigation coefficient, with values between 0 and 1 indicating the irrigation potential of a patch. The calculation of this coefficient, the irrigability, as well as a detailed explanation of the other biophysical variables will be given in Chapter 5.

Ecological sub-models

As mentioned in the introduction, there are three kinds of ecological sub-models to be built into the model of the landscape agent: productivity functions for each land-use type, a livestock dynamics model, and a land-cover transformation model. For further details, see Chapter 5.

i) Agricultural productivity functions

The agricultural productivity functions are patch sub-models calculating the variables $P_{yield\ dry}$ and $P_{yield\ rainy}$ in response to variables of the Patch Profile and the user's land-use decisions. Since the importance of biophysical attributes and the kind of land management differ between the two seasons with respect to crop productivity, a yield model for each season was developed. Although the range of variables differ between the two seasons, the general form of the function is the same (see section 5.3.3).

ii) Livestock dynamics model

The livestock dynamics model is a sub-model to calculate the variable of $H_{livestock}$ in response to random annual variations and forage availability, the latter being dependent on both rainfall data and land-use behavior. The livestock index of a household is basically modeled as being dependent on the livestock index of the previous year (with a random error), reflecting changes in the stock due to sale, death, diseases, etc. The forage availability on the other hand restricts the total number of livestock within the study area, thus reducing the total number of livestock equally for all households, if the carrying capacity with respect to forage availability is reached (see section 5.3.4).

iii) Land-cover transformation model

The land-cover transformation model is a model to simulate the conversion of one land-cover type to another, whereby two variables describe land-cover distributions, one for the rainy season, $P_{\text{cover rainy}}$, and one for the dry season, $P_{\text{cover dry}}$. For the establishment of the model, changes of both variables should thus be analyzed and modeled if necessary. The range of land-cover types for both seasons comprises 'rock', 'water', 'bare land', 'grassland' and 'cropland'. Changes among these land-cover types are driven by both anthropogenic influence (land-use change) and natural processes independent of human interference (e.g. grass growth), which both need to be considered in the analysis. In section 5.3.5, the full land-cover change analysis and the parameterization of the subsequent land-cover transformation model will be presented.

3.3.2 Entire landscape

The entire landscape is the collection of all landscape agents, together with a database of statistical spatial parameters:

$$\text{Entire landscape} = \{\{\text{Landscape Agents}\}, \text{Spatial-Stat}\}$$

The spatial statistical database Spatial-Stat comprises descriptive statistics about land-cover and land-use evolving over time. Percentages of the different land-use types of the total cultivated area are computed for both seasons, as well as the simulated land-cover fractions of the total area under study. The temporal dynamics of these parameters can be observed via graphs on the simulation interface of the GH-LUDAS model.

3.4 Structure of the Decision Module

The Decision Module is an ordered collection of procedures regulating the agent's behavior regarding his livelihood strategy and land-use decisions. Although this module is introduced here as an autonomous part of the model, it is in fact an integral part of the household agent, governing the behavior of that agent. It works as a scheduled programme of procedures reacting to parameters from the household agent and his Landscape Vision, resulting in agent-specific reactions and actions on the environment.

As the kinds of decisions to be made differ among the two seasons in the study area, the Decision Module was designed to consider these differences. Thus, it was divided into two subsequent collections of routines, one for each season, starting with those for the dry season. The general scheme of the two main routines is similar, starting with the labor allocation among the various income-generating activities, followed by the cultivation of its own patches. If labor and cash are still available after the full utilization of its own patches, the agent will search for new patches, and finally, the income from cultivation (through productivity functions) and other income-generating activities will be calculated. As cultivation in the dry season is only possible via irrigation, two additional decision sub-models precede the dry-season procedures, including the decision to irrigate, and the choice of irrigation technology.

3.4.1 Dry-season procedures

The dry-season procedures (D_{proc}) can be structurally expressed in the form of the following consecutive routines:

$$D_{proc} = [D_{irr}, D_{method}, D_{labor}, D_{static}, D_{moving}, D_{income}]$$

where D_{irr} is the decision for irrigation farming, D_{method} the choice of irrigation technology, D_{labor} the labor allocation procedure, D_{static} the static phase of cultivating own patches, and D_{moving} the moving phase for opening new patches, and D_{income} the income-generating procedure for the dry season.

Irrigation choices (D_{irr}) and (D_{method})

The procedures of irrigation decision and method choice are examined here in combination, since they form a nested hierarchy of decisions with respect to irrigation-related decisions. For modeling this decision procedure, we decided to use a two-fold nested m-logit model. The first sequences of the m-logit model will simulate the general decision of a household agent to engage in irrigation farming, and the second will then simulate the choice of irrigation technology, if the decision on irrigation in the first step is positive. This two-fold nested decision is taken by each household agent in each time step at the beginning of the model run, and is independent of the group of agents.

For the first sequence of the nested m-logit model, we employed household-specific data reflecting the economic capability of a household to afford irrigation farming, including financial capital, human resources, land and knowledge. The second sequence then regulates the choice of irrigation technology, which is a choice among three alternatives: dam irrigation (in case a dam is located within the Landscape Vision) and two riverine irrigation methods, i.e. the use of hand dug wells via buckets, or dugouts via motor pumps. The choice of these three options is based on the following indicators: i) the financial capacity of a household, since the three options require varying monetary investments, ii) the availability of a dam within an acceptable distance, and iii) the personal history of the considered household agent regarding irrigation method and practice. The range of variables used for both levels of this nested decision m-logit model will be outlined and justified in detail in Chapter 4, together with a presentation of the calculated m-logit coefficients.

Labor allocation procedure (D_{labor})

Within this labor allocation procedure, the total dry-season labor pool of the household is allocated to the various production lines, including cultivation, trading, food processing, handicrafts, migration and other income-generating activities (e.g white collar jobs). The percentages of labor allocated to these various production lines are defined by the household agent group, reflecting the production strategy of the livelihood or agent group the household belongs to.

But since the amount of labor allocated to cultivation rather depends on the decisions of the household regarding irrigation than on the agent group, the amount of labor is defined by the irrigation choices: If no irrigation is practiced during the dry season, no labor is allocated to the cultivation production line. If, on the other hand, the decision to do irrigation farming is positive, a certain amount of labor is allocated to this production line, depending on the choice of irrigation method. The reason for this differentiation is that in the study area the labor input requirements vary highly among the irrigation methods, with bucket irrigation being twice as labor-intensive as pump irrigation. After the amount of cultivation labor has been set, the spare labor pool is allocated to the five remaining production lines, as pre-defined by the agent group. The total labor pool for the dry season is the number of labor days per household spent on income-generating activities (i.e. the six production lines). For the base year 2006, which is the starting point of the model, this labor pool was calculated from field data for each household, i.e. the dry-season-based survey (see Appendices B for questionnaire). For each subsequent year, the labor pool in the model is recalculated dependent on the value of the preceding year:

$$^{\mathrm{t+1}}H_{\mathrm{pool\ dry}} = {}^{\mathrm{t}}H_{\mathrm{pool\ dry}} - \sigma_{\mathrm{pool\ dry}} + random(2 \cdot \sigma_{\mathrm{pool\ dry}}) \tag{3.13}$$

where $\sigma_{\mathrm{pool\ dry}}$ is the standard deviation of $\mathrm{H}_{\mathrm{pool\ dry}}$. The labor pool represents the labor allocated to income-generating activities and is allowed to lie beneath the labor capacity of the household. Using this approach, underemployment in terms of an incomplete use of labor capacity of the household, is considered, as is the case for many households in the study area.

Static phase ($\mathbf{D_{static}}$)

Since it is natural to first cultivate own patches and then look for other patches, this procedure precedes the routine of borrowing new patches. Since patches considered for cultivation during the dry season need to be irrigable, these patches have to be located either within an irrigable dam area, or within the irrigable area along the main river. The determination of this riverine irrigable area will be given in Chapter 5. This way, a household agent either owns no irrigable patches, patches along the river, patches along a dam, or both of the latter two. Interviews with local farmers suggest that if households own both dam and riverine patches, the dam patches will be preferred, as they are less labor intensive and more cost effective. Therefore, the virtual household agent is programmed to put dam patches under cultivation, before shifting to riverine land holdings, regardless of the irrigation choice made within the irrigation decision procedures. These decisions are considered to only play a role in the moving phase of the agent when searching for new land patches. In the following, we will denote the set of owned patches along a dam as $H_{\text{area-dam}}$, and the set of owned irrigable and riverine patches as $H_{\text{area-river}}$.

Before we present the algorithm of the D_{static} procedure, we have to introduce the concept of how to determine the size of the area a single household is able to cultivate. The size is dependent on two factors: the financial resources of the household, as irrigated cultivation is associated with relatively high costs for fertilizer purchase and maintenance of the irrigation system, and labor resources. Since the requirements of labor and input capital vary highly among the three irrigation types, the possible number of patches to be cultivated is calculated for each irrigation type separately, depending on the available financial and labor resources of the household. This calculation is based on a linear regression for each type, with explanatory variables of cultivation labor pool and income:

$$^{\mathrm{i}}C_{\text{max}} = {}^{\mathrm{i}}a + {}^{\mathrm{i}}b_1 \cdot H_{\text{labor-pool dry}} + {}^{\mathrm{i}}b_2 \cdot H_{\text{income}} \tag{3.14}$$

where $^{\mathrm{i}}C_{\text{max}}$ is the maximum number of patches to be cultivated by a household, with i indexing the type of irrigation method. This parameter $^{\mathrm{i}}C_{\text{max}}$ calculated by this linear regression model then serves as the upper limit for the model of cultivation.

Thus, the number of owned irrigable patches and the number of maximum possible cultivated patches $^{\mathrm{i}}C_{\text{max}}$ serve as upper bounds for the number of cultivated patches within the procedure of D_{static}. However, regarding the cultivation along dams, another limiting factor plays a role, which is represented by the policy of area limitation:

In GH-LUDAS, the maximum dam area a single household agent is allowed to cultivate (called Lim_{Dam}) can be specified outside the model as a policy parameter. Thus, this parameter serves as another limiting factor for the number of cultivated patches if these are located along a dam. Following this mindset, the cultivation of own land holdings or the static phase can be structured as follows:

1. Set Used-Patches 0

2. Set the irrigation method i to dam.
3. Calculate the number n of owned patches actually cultivated by the household:
   $n = \min(\text{count}(H_{area\text{-}dam}), {}^{i}C_{max}, Lim_{Dam})$
4. Select n random patches from the set $H_{area\text{-}dam}$
5. For each of these n patches choose its land-use type
6. Set the input parameters of labor and fertilizer, dependent on the type of land-use
7. Set the irrigation method i to the riverine method with the highest probability
8. Set Used-Patches Used-Patches + n
9. Calculate the number n of owned patches actually cultivated by the household:
   $n = \min(\text{count}(H_{area\text{-}river}), {}^{i}C_{max} - \text{Used-Patches})$
10. Select n random patches from the set $H_{area\text{-}river}$
11. For each of these n patches choose its land-use type
12. Set the input parameters of labor and fertilizer, dependent on the type of land-use
13. Set Used-Patches Used-Patches + n

Moving phase (D_{moving})

The moving phase is similar to the static phase as depicted above, apart from the fact that the choice of irrigation method and the Landscape Vision play a role in this procedure. As the Landscape Vision is the environment a household agent can act upon, the agent will only search for irrigable patches within his individual Landscape Vision. In the following, we will denote the set of irrigable dam patches within the Landscape Vision not used by another agent as LV_{dam}, and the set of irrigable riverine patches within the Landscape Vision not used by another agent as LV_{river}. Regarding the choice of irrigation method, it is a natural assumption that a household agent can change his choice of irrigation method during his search for irrigable patches. For instance, if the agent chooses a riverine irrigation method, but patches along the river are no longer available, he will shift to a dam if one is located within his Landscape Vision. It can also be the other way round, i.e. an agent first chooses dam irrigation, but then has to shift to riverine irrigation if no dam patches are located within his Landscape Vision. Thus, two different procedures are presented here, dependent on the first choice of irrigation technology. These two procedures of D_{moving} are similar to the mindset of the procedure D_{static} and can be summarized as follows:

If ($H_{irr\ method}$ = dam) run the following procedure:

1. Set the irrigation method i to dam.
2. Calculate the maximum number n of patches actually cultivated by the household:
   n = min(count(LV_{dam}), $^{i}C_{max}$ - Used-Patches, Lim_{Dam})
3. Select n random patches from the set LV_{dam}
4. For each of these n patches choose its land-use type
5. Set the input parameters of labor and fertilizer, dependent on the type of land use
6. Set Used-Patches Used-Patches + n
7. Set the irrigation method i to the riverine method with the highest probability
8. Calculate the maximum number n of patches actually cultivated by the household:
   n = min(count(LV_{river}, $^{i}C_{max}$ - Used-Patches)
9. Select n random patches from the set LV_{river}
10. For each of these n patches choose its land-use type
11. Set the input parameters of labor and fertilizer, dependent on the type of land use
12. Set Used-Patches Used-Patches + n

And If ($H_{irr\ method}$ = well or motor pump) run the following procedure:

1. Set the irrigation method i to the riverine method with the highest utility
2. Calculate the maximum number n of patches actually cultivated by the household:
   n = min(count(LV_{river}), $^{i}C_{max}$ - Used-Patches)
3. Select n random patches from the set LV_{river}
4. For each of these n patches choose its land-use type
5. Set the input parameters of labor and fertilizer, dependent on the type of land use
6. Set Used-Patches Used-Patches + n
7. Set the irrigation method i to dam.
8. Calculate the maximum number n of patches actually cultivated by the household:
   n = min(count(LV_{dam}), $^{i}C_{max}$ - Used-Patches, Lim_{Dam})
9. Select n random patches from the set LV_{dam}
10. For each of these n patches choose its land-use type

11. Set the input parameters of labor and fertilizer, dependent on the type of land use
12. Set Used-Patches Used-Patches + n

Income generation procedure (D_{income})

Since cash income plays an important role within the coupled human-environment system, as it serves as the financial basis for land-use related investments, the income generation procedure is designed as a routine to calculate both cash and gross income. However, it is assumed that cash and gross income for the non-farm activities are identical, including the activities of trading, food processing, handicrafts, migration and other activities, since a differentiation among cash and gross income for these activities is a difficult issue and reliable information was not available during the surveys.

The same is valid for the generation of gross income for livestock, as it was not possible to measure the net annual gross income of an animal stock. But as the sale of livestock was captured during the household surveys, at least the annual cash income of this production line could be measured. The seasonal cash income from livestock $H_{\text{inc live dry}}$ was calculated using linear regression based on the amount of livestock, i.e. livestock index:

$$H_{\text{inc live dry}} = a_{livedry} + b_{livedry} \cdot H_{\text{livestock}} \tag{3.15}$$

where $a_{livedry}$ and $b_{livedry}$ are parameters calculated using the statistical analysis programme SPSS.

Using a similar approach, the income of the non-farm activities is generated based on the amount of labor allocated to the various production lines:

$$H_{\text{inc trad dry}} = a_{traddry} + b_{traddry} \cdot H_{\text{lab trad dry}} \tag{3.16}$$

$$H_{\text{inc food dry}} = a_{fooddry} + b_{fooddry} \cdot H_{\text{lab food dry}} \tag{3.17}$$

$$H_{\text{inc arts dry}} = a_{artsdry} + b_{artsdry} \cdot H_{\text{lab arts dry}} \tag{3.18}$$

$$H_{\text{inc migr dry}} = a_{migrdry} + b_{migrdry} \cdot H_{\text{lab migr dry}} \tag{3.19}$$

$$H_{\text{inc others dry}} = a_{othersdry} + b_{othersdry} \cdot H_{\text{lab others dry}} \tag{3.20}$$

The gross income generated by the production line of cultivation ($H_{\text{gross inc cult dry}}$) is simply calculated as the sum of yield of all cultivated patches:

$$H_{\text{gross inc cult dry}} = \sum_{\text{all cultivated patches}} P_{\text{yield dry}} \tag{3.21}$$

whereas the calculation of cash income from this production line follows a different approach: Since crops cultivated in the dry season mainly serve as cash crops and are sold out to traders or at markets, the cash income of dry-season cultivation is modeled as a linear regression function based on the gross income of cultivation as presented above:

$$H_{\text{cash inc cult dry}} = a_{cultdry} + b_{cultdry} \cdot H_{\text{gross inc cult dry}} \tag{3.22}$$

where $a_{cultdry}$ and$b_{cultdry}$ are the parameters of this regression.

As the policy of credit access plays a role in the study area in the generation of additional income, it must be a factor for this income model: Additional financial resources allow a household to generate more income per labor unit, which has to be considered in this routine. This additional income per labor unit for each production line was derived from the empirical household data set, including both households that had had access to credit and households that had not. This additional income was then added to the incomes for each production line of a household once the household had access to credit: This procedure can be depicted as follows:

$${}^{\text{credit(1)}}H_{\text{inc i dry}} = {}^{\text{no credit(1)}}H_{\text{inc i dry}} + a_i \cdot H_{\text{lab i dry}} \tag{3.23}$$

where i indexes the production line, ${}^{\text{credit(1)}}H_{\text{inc i dry}}$ is the income generated by the access to (the first) credit, ${}^{\text{no credit(1)}}H_{\text{inc i dry}}$ the income generated without credit, and a_i is the line-specific factor of additional income per labor unit.

The empirical data set did not provide any information about the income structures of households that had access to credit more than once. However, it is a natural assumption that the additional income generated by an additional credit declines with the number of credits already obtained, i.e. the effect of each additional credit wears off. This decline in the effect of additional credits is regulated by the global-policy parameter called the credit deflating factor, which has values between 0 and 1. In the case of the value 0.5 for this factor, the effect of credit on income is only 50 % as strong as the effect of the previous credit on income. Thus, the income converges against a certain limit, with the number of credits obtained $H_{\text{nr credits}}$ increasing. Mathematically, this relationship can be expressed as:

$$\left(\frac{{}^{\text{credit(n)}}H_{\text{inc i dry}}}{{}^{\text{no credit(n)}}H_{\text{inc i dry}}} - 1\right) \cdot Credit_{\text{def}} = \left(\frac{{}^{\text{credit(n+1)}}H_{\text{inc i dry}}}{{}^{\text{no credit(n+1)}}H_{\text{inc i dry}}} - 1\right) \tag{3.24}$$

where *i* indexes the production line, *n* denotes the number of credits already obtained, and $\text{Credit}_{\text{def}}$ the credit deflating factor. This equation can be transformed such that the income for the $n + 1^{\text{th}}$ credit can be calculated:

$$\begin{aligned} &{}^{\text{credit(n+1)}}H_{\text{inc i dry}} = \\ &= \left(\left(\frac{{}^{\text{credit(n)}}H_{\text{inc i dry}}}{{}^{\text{no credit(n)}}H_{\text{inc i dry}}} - 1\right) \cdot Credit_{\text{def}} + 1\right) \cdot {}^{\text{no credit(n+1)}}H_{\text{inc i dry}} \end{aligned} \tag{3.25}$$

3.4.2 Rainy-season procedures

Similar to the dry-season procedures, the rainy-season procedures (called R_{proc}) can be structurally expressed in the form of the following consecutive routines:

$$R_{\text{proc}} = [R_{\text{labor}}, R_{\text{static}}, R_{\text{moving}}, R_{\text{income}}]$$

where R_{labor} is the labor allocation procedure, R_{static} and R_{moving} the static and moving phase of cultivation, and R_{income} the income generating procedure for the rainy season.

Labor allocation procedure (R_{labor})

In this procedure, the labor pool for the rainy season is allocated to the different production lines, which comprise the same range of activities as in the dry season. As for the dry season, the percentages of labor allocated to the various activities are defined by the agent group, which reflects the livelihood strategy of the household in question, whereby the total annual labor pool $H_{\text{pool rainy}}$ is calculated accordingly. Analogous to the dry season, the provision of credit leads to a small shift of labor allocation by a factor that has been derived statistically from the empirical data set. This procedure is equivalent to the dry-season equation, but with rainy-season specific parameters, which were identified using SPSS.

Static phase (R_{static})

Compared to the dry season, financial resources play a lesser role for cultivation during the rainy season. Therefore, the maximum area C_{max} a household is capable of cultivating is modeled as only being dependent on the available labor pool for cultivation. This way, C_{max} can be formulated as follows:

$$C_{max} = H_{labor\ cult\ rainy} / I_{lab\ mean}$$

where $H_{labor\ cult\ rainy}$ is the available labor for cultivation, and $I_{lab\ mean}$ is the empirical mean of labor input for a single patch. Since only patches with the land cover 'cropland' or 'grassland' are suitable for cultivation, patches that are covered by bare land or forest have to be ignored during the routine of R_{static}. Thus, we will denote the set of patches owned by a household covered by either grassland or cropland as H_{area}. Furthermore, it was observed that a farmer usually prefers to continue cultivating the patches that have been used the year before. The reason for this is that he usually reserves grassland holdings for the feeding of his livestock. Therefore, within this routine, first all patches with the land cover 'cropland' will be selected until all patches have been cultivated. Then the procedure will start selecting grass patches. The procedure R_{static} can be summarized as follows:

1. Set Used-Patches 0
2. Calculate the number n of owned patches actually cultivated by the household:
   $n = \min(\text{count}(H_{area}), C_{max})$
3. Select n random patches from the set H_{area}
4. For each of these n patches (with preference of patches covered by cropland) choose its land-use type
5. Set the parameters of labor input and management, dependent on the type of land-use
6. Set Used-Patches Used-Patches + n

Moving phase (R_{moving})

In the moving phase, the household agent searches for new patches within his Landscape Vision, if labor is still available. According to field observations, a farmer usually tries to continue to cultivate plots he already asked for during the last season. Therefore, the moving phase can be separated into two sub-routines: In the first, the household agent will try to continue cultivating the patches he has already acquired; in the second, he will scan his Landscape Vision for new patches, and if he is successful, mark them as being borrowed. These two procedures can be summarized in the following; LV_{area} denotes the set of still unused patches within the Landscape Vision suitable for cultivation (i.e. either grassland or cropland) and H_{borr} the set of patches borrowed by the household and not yet used by any other household:

1. Calculate the number n of patches actually cultivated by the household:
   n = min(count(H_{borr}), C_{max} - Used-Patches)
2. Select n random patches from the set H_{borr}
3. For each of these n patches choose its land-use type
4. Set the parameters of labor input and management, dependent on the type of land use
5. Set Used-Patches Used-Patches + n
6. Calculate the number n of patches actually cultivated by the household:
   n = min(count(LV_{area}), C_{max} - Used-Patches)
7. Select n random patches from the set H_{area}
8. For each of these n patches (with preference of patches covered by cropland) choose its land-use type
9. Set the parameters of labor input and management, dependent on the type of land-use
10. Set Used-Patches Used-Patches + n

Income generation procedure (R_{income})

Analogous to the dry-season procedure, both cash and gross incomes are calculated. The equivalent equations are as follows:

$$H_{\text{inc live rainy}} = a_{liverainy} + b_{liverainy} \cdot H_{\text{livestock}} \tag{3.26}$$

for cash income of livestock, and

$$H_{\text{inc i rainy}} = a_{irainy} + b_{irainy} \cdot H_{\text{lab i rainy}} \tag{3.27}$$

with i indexing the production lines as in the dry-season procedure, and a_{irainy} and b_{irainy} being the respective parameters. Equivalently, the gross income from cultivation $H_{\text{gross inc cult rainy}}$ is calculated as:

$$H_{\text{gross inc cult rainy}} = \sum_{all\ \ cultivated\ \ patches} P_{\text{yield-rainy}} \tag{3.28}$$

with $P_{\text{yield rainy}}$ being the yield of a single patch, as calculated by the land-use-specific productivity functions.

Regarding the calculation of cash income from cultivation, a different approach is needed, because the pattern of crop sale is distinct from the dry season. Most of the harvest is not sold, but stored and mainly used for consumption during the months after harvest. Nevertheless, some of the crops such as rice and groundnuts can be considered as cash crops to a limited extent. This way, the amount of sold harvest is not dependent on the total gross income of cultivation as in the dry season, but merely on the type and amount of cultivated crops. Thus, the function of cash income for this season was designed as follows:

$$H_{\text{cash inc cult rainy}} = a + \sum_{i} b_i \cdot \text{Cult-Area}_i \tag{3.29}$$

where Cult-Area_i is the total cultivated area of the land-use type i of the household. This way, the amount of cash income reflects the pattern of the choice of cash land-use types and non-cash land-use types. For the rainy season, the impact of the first credit on the different income-generating activities is modeled in the same way as for the dry season, but with the corresponding parameters:

$$^{\text{credit(1)}}H_{\text{inc i rainy}} = {}^{\text{no credit(1)}}H_{\text{inc i rainy}} + \cdot a_i \cdot H_{\text{lab i rainy}} \tag{3.30}$$

where *i* indexes the type of production line, and ai is the additional income per labor unit generated by the first credit. The income generated by further credits is then calculated using the same algorithm as in the dry season.

3.5 Global-policy Module

This module represents policy parameters in the form of tunable parameters that the model user can set according to scenarios he wants to explore. Within the model, these parameters are accessible by both landscape and household agents, and are therefore also called global parameters. The policy and other external factors included in GH-LUDAS to be tested for their impacts comprise

1. Dam construction to improve possibilities for dry-season irrigation
2. Credit access regulations to test the effects of credit schemes on the combined livelihood and land-use/cover pattern
3. Population dynamics and IPCC rainfall scenarios.

A justification of the choice of these policies will be given in Chapter 6, together with a detailed description of the policy situation in the study area. In this section, we will only provide an overview of the parameterization of these policies and their relations with the other model components.

3.5.1 Dam construction policy

In case an institution is interested in providing an area with one or more dams, several considerations have to be made. First, the biophysical conditions of the area have to be examined to decide where and whether these conditions allow the construction of a dam. In addition, the location of the dam should be selected according to the socio-economic conditions of its potential users. Otherwise, the dam will possibly not be utilized fully to its capacity if its potential users do not have the ability or resources to do irrigation farming. Therefore, the selection of the location of the dam, which should be directed towards a maximum benefit for all its users, is a critical issue.

Second, the size of irrigation capacity and number of dams to be constructed have to be carefully determined. As in some situations the construction of a single large dam could match the socio-economic needs of the population, in other cases a collection of several scattered small-scale dams is required. Thus, it is necessary to evaluate scenarios of different combinations of size and number of dams.

Third, to provide a maximum of potential users with the possibility to engage in irrigation farming, a regulation of area limitation could be taken into consideration, i.e. the prescribed maximum area one household is allowed to cultivate along a dam. The selection of this parameter is also a critical issue, as it should ensure a maximal number of dam users on the one hand, but also a full utilization of the irrigation capacity on the other. According to these considerations, in GH-LUDAS, the following parameters of the policy of dam construction $Policy_{dam}$ have been included:

$$Policy_{dam} = \{Dam_{lim}, Dam_{number}, [Dam_{loc}, Dam_{size}]\}$$

where Dam_{lim} denotes the size of maximum cultivated area, Dam_{number} the number of dams, and Dam_{loc} and Dam_{size} the location and size for each of the single dams.

In GH-LUDAS, the single dams can be inserted into the landscape on the user interface via a mouse click, and a slider allows the user to define the size of the dam. Another slider defining the maximum cultivated area can be set according to the scenarios to be explored.

In the model, these parameters are linked to the landscape as well as to the Human Module (Figure 3.6). On the household side, the locations of the dams (Dam_{loc}) regulates the distance to dams and water sources, while Dam_{lim} defines the upper limit for dam cultivation for the household (section 3.4). On the landscape side, the size and location of the dam modify the parameter $P_{irrigable}$ of some of the landscape patches: The parameter $P_{irrigable}$ of those patches that are located within the irrigable perimeter around the dam will be set to 1.

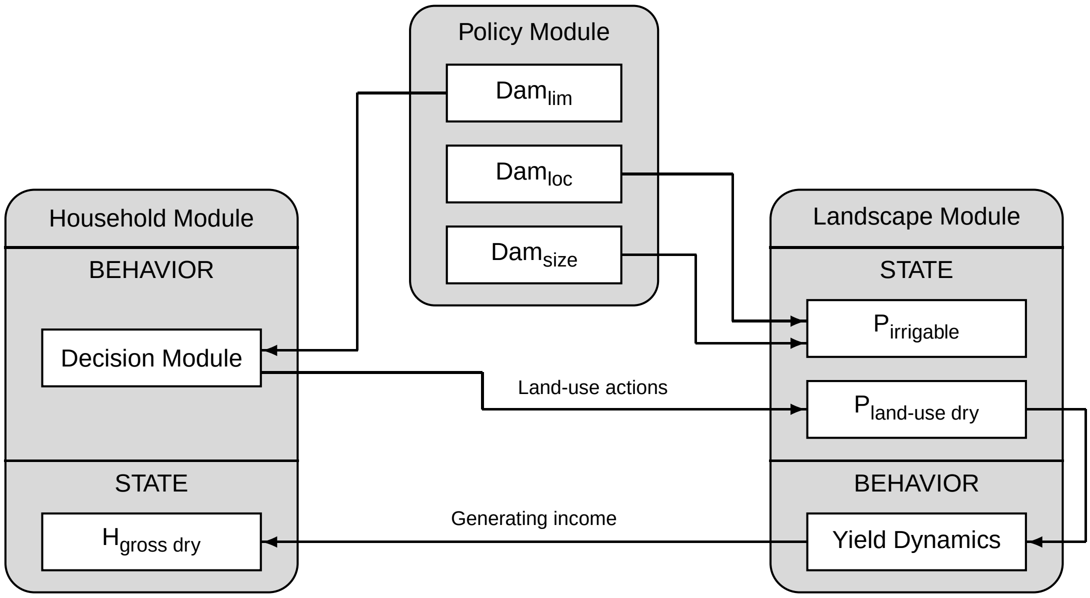


Figure 3.6: Integration of the dam construction policy in GH-LUDAS

3.5.2 Credit access policy

Access to credit directly affects land-use-related household decisions, thus possibly exerting an influence on the local land-use and land-cover patterns. It was observed during field interviews as well as by statistical analysis of the empirical data set, that farmers with access to credit schemes change their focus regarding their activities. They may intensify some of their income-generating activities with higher income generation possibilities (e.g. trading, irrigation), while some of the less productive activities (e.g. food processing) might be reduced. The additional income generated by these investments of labor and cash stimulated by the credit may be reinvested in land-use-related and other activities, thus gradually changing the livelihood strategy and decision-making processes.

In the study area, the credit scheme managed by the Ministry of Food and Agriculture (MOFA) allows a credit of 200 000 Cedis (about 20 US $) per household. Since this credit amount obtained by local farmers is constant, the possible effects of a different credit rate cannot be assessed from the empirical data set. Thus, in GH-LUDAS, the credit rate must presently be regarded as constant at 20 Cedis. The same is valid for the period of credit access, i.e. the number of successive years a household obtains this amount from the credit scheme, which was observed to be constant at 2 years.

Nevertheless, the annual rate of households supplied with credit can be modified as a parameter within the model. Apart from that, the credit scheme can be manually switched to a different kind of scheme than the one observed in the study area, called the 'revolving credit' scheme. The idea of this kind of scheme is that, once the credit has been distributed among the population, it will be handed round until a certain period of time has elapsed. In other words, the credit a household obtains from the scheme at the beginning of this period will not be paid back to the scheme, but to another household. This household will then pay back the credit to a third household, and so on, until a certain period has elapsed. Then, the last household will pay its debts back to the donor. We will call the period of time the credit remains within the population as the 'revolving credit period'. The parameters defining the credit scheme policy $Policy_{credit}$ in GH-LUDAS can therefore be summarized as follows:

$$Policy_{credit} = \{Credit_{perc}, Credit_{scheme}, Credit_{rev\ period}, Credit_{def}\}$$

where $Credit_{perc}$ is the annual percentage of households supplied with credit, $Credit_{scheme}$ a dummy variable defining which kind of scheme is activated, $Credit_{rev\ period}$ the parameter of revolving credit period, which is only called by the model if the scheme is of the revolving type, and $Credit_{def}$ the credit deflating factor (see section 3.4).

As the effects of credit access on the environment are only of an indirect nature, the direct linkages of this policy to the other system components are among these policy parameters and parameters of the household agents (Figure 3.7), and the parameters of this policy directly change the household variables H_{credit}, $H_{nr\ credits}$ and $H_{gross\ income}$. Changes in any of the policy parameters result in a change of income, and ultimately show indirect effects on land-use choice and land productivity.

3.5.3 Population dynamics and climate change

Other external variables of the Global-policy Module, which are not related to policies, include parameters describing population dynamics and the choice among possible future rainfall scenarios. As no reliable population data for the study area were available, due to unreliable and insufficient population surveys (only 4 surveys in 1965, 1975, 1984 and 2000), no reliable model could be established for projecting future population numbers. Instead, the

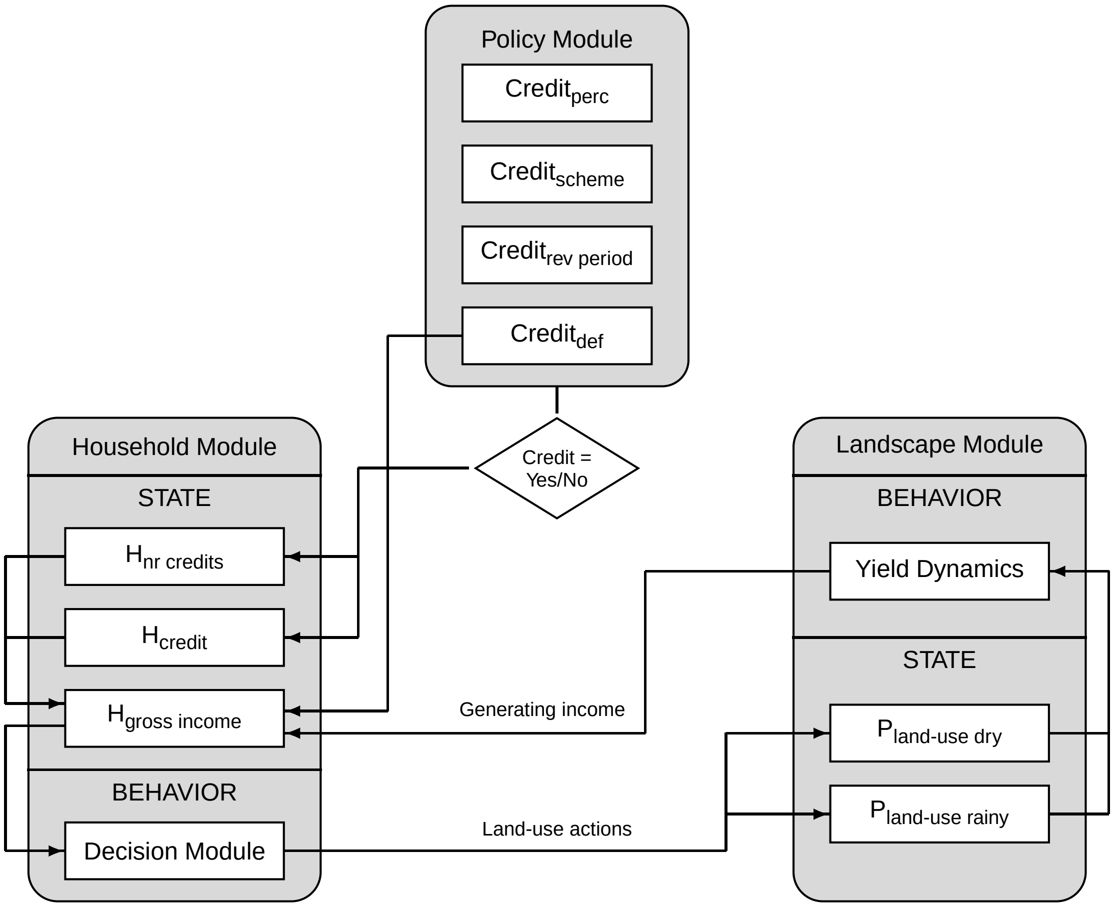


Figure 3.7: Integration of the credit access policy in GH-LUDAS

parameters describing local population dynamics were chosen to be set externally. To represent these dynamics we chose one of the most widely used models for population growth, the logistic growth model, which can be expressed as:

$$P(t) = \frac{CP_0e^{rt}}{C + P_0(e^{rt} - 1)} \tag{3.31}$$

where $P(t)$ is the population size at time step t, P_0 the initial population size at time 0, and the carrying capacity C and the growth rate r parameters describing the convergence behavior of the population. For $t \to \infty$, the population size converges against the carrying capacity C with growth rate or 'speed' r. These two parameters are set externally by the model user, according to the scenarios population growth to be explored (Figure 3.8). New agents are created in each time step, dependent on the logistic growth model and the number of agents that were deleted due to the ageing process incorporated in the model.

Finally, scenarios of future annual rainfall can be selected, based on local climate data as simulated by the IPCC (International Panel on Climate Change), which is the leading research group with respect to global climate assessment. The annual data of the rainfall scenario selected by the model user are fed into the productivity functions for rainy-season land-use types. Furthermore, a model is developed (see Chapter 5) to calculate the forage availability for local livestock based on rainfall data in order to determine the annual carrying capacity for local livestock. This way, in GH-LUDAS, a decrease or increase in crop and forage productivity due to changing rainfall patterns indirectly influence land-use choice and livestock dynamics and thus livelihood strategies (Figure 3.9). The details of the integration of rainfall data into crop and forage productivity are given in Chapter 5.

3.6 Simulation protocol of GH-LUDAS

Within this section, the two main parts of the model will be outlined: The setup procedure of the model, and the main time-loop of sequential procedures during simulation. The setup procedure is a routine that simulates the whole landscape with all its household agents and their attributes before any model run. The goal of this procedure is to simulate as closely as possible the state of the coupled human-environment system as it was in 2006, which was the year of data collection. The time-loop procedures, on the other hand, represent the dynamic part of the model, consisting of a collection of sequential procedures, which will be run in each time step representing one year.

3.6.1 Setup procedure

The setup procedure is a routine to implement the state variables of landscape and household agents, and to visualize the current land-cover patterns in the view of the model. In this section, we will first describe the routine of landscape implementation, and subsequently the setup of household agents within this landscape.

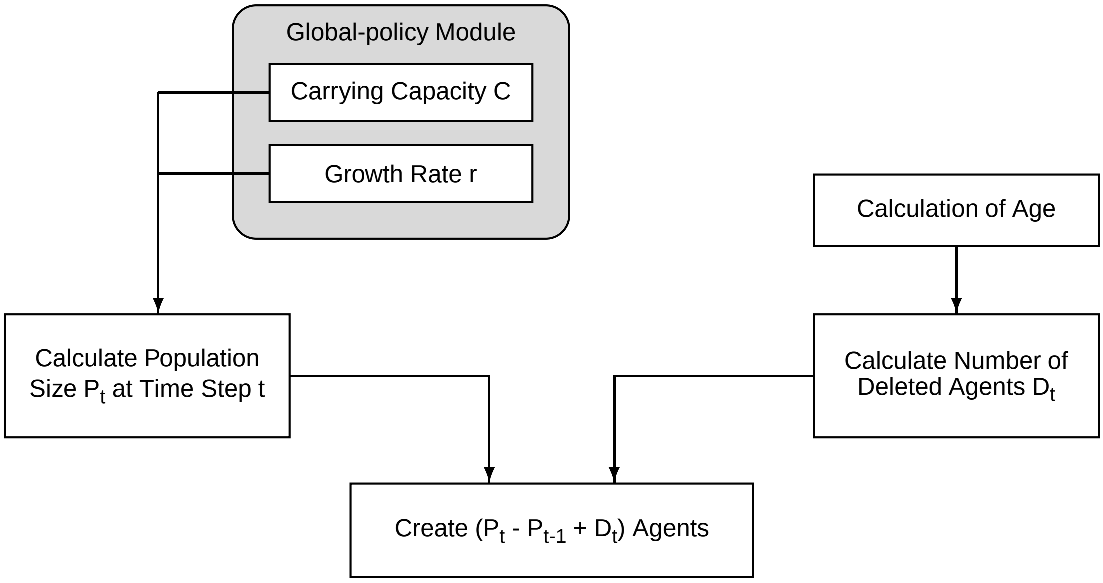


Figure 3.8: Regulation of agent population in GH-LUDAS

Landscape setup

The setup procedure for the landscape can be structurally described by the following successive steps:

1. The implementation and visualization of current land-cover patterns in the study area, based on the analysis of satellite images
2. The assignment of patch-specific variables to all patches located in the study area
3. The allocation of dams to this landscape via mouse click, if the examination of this policy is desired by the model user.

As this section mainly deals with the implementation of the model, we will only give a short explanation of how these patch-specific attributes have been derived, and focus on the way of implementation. The sources and derivation of these attributes will be described in detail later in Chapter 5.

The land-cover pattern of the year 2006 was derived from two satellite images using the ERDAS package. The first image with a higher resolution, served as the basis for the digitization of the main river and its tributaries, while the second provided the basis for the

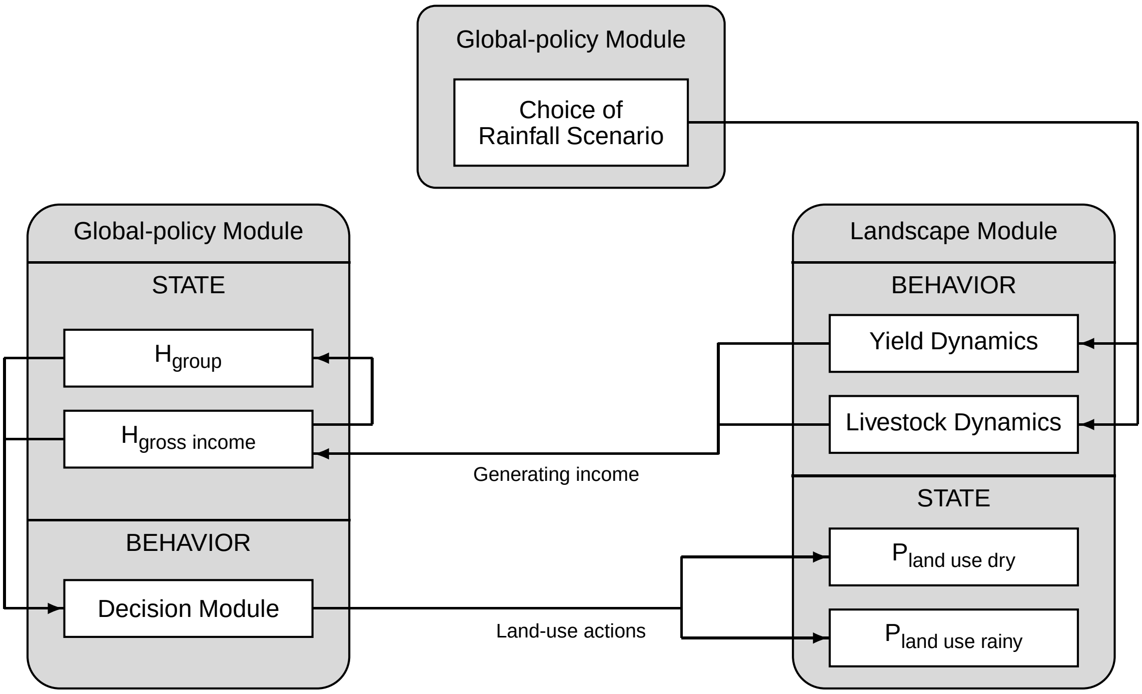


Figure 3.9: Integration of rainfall change in GH-LUDAS

classification of all remaining land-cover types. These two images were then converted to ascii files, which store a single value per pixel, representing one patch of the landscape of 30 m x 30 m. These ascii files can then be easily read by NetLogo, whereas each patch of the view is assigned its specific value of land-cover. Within the view, these different land-cover types were then visualized by different colors.

While the land-cover patterns are visible within the view, the other patch attributes are only stored but not visualized. These variables include institutional attributes ($P_{village}$, $P_{compound}$), distances ($P_{dist\ river}$, $P_{dist\ dams}$, $P_{dist\ water\ sources}$), and all biophysical variabes ($P_{wetness}$, $P_{upslope}$, $P_{elevation}$, $P_{soil\ fertility}$, $P_{soil\ texture}$, P_{gwl}, P_{gwr}). The irrigation coefficient $P_{irr\ oeff}$ and the dummy variable $P_{irrigable}$ are then calculated from this data set (see Chapter 5). All other variables were derived from different sources such as maps, GIS layers created by previous studies of the study area, and satellite images. In the same manner as the land-cover data, the data were also converted to ascii files to be read by NetLogo.

The last procedure is only called if the user wishes to implement dams within the model. Within the view, the dam can be inserted by the user via mouse click, and its irrigation capacity can be set specifically for each dam. This way, each dam has its own specific irrigation capacity. Each inserted dam consists of the dam itself and its respective irrigable area. First, the procedure creates a dam as a circle around the selected patch, while the size of the circle is defined by the irrigation capacity, and converts the land cover of these patches to 'water'. Second, the irrigable area is created along the direction of minimal elevation ($P_{elevation}$), with the number of patches pre-defined by the value of irrigation capacity. Finally, the dummy variable ($P_{irrigable}$) is set to 1 for all patches within this irrigable area.

Household agents setup

The setup procedure for household agents can be structurally described by the following successive steps:

1. The import of the set of 200 interviewed farmers, together with their specific household variables
2. The multiplication of these 200 households to populate the landscape to its actual population size
3. The calculation of distance variables for all household and landscape agents
4. The allocation of land holdings for each household agent

In the first step, to ensure a reliable reproduction of the real population, copies of those households that had been interviewed during the field surveys will be created. These household agents are endowed with the same set of variables as the interviewed farmers, and are located within the respective village of the catchment. Within each village, they are distributed on the compounds as digitized by a high-resolution satellite image, i.e. on patches with the dummy variable being $P_{compound} = 1$. The range of imported variables comprises all attributes that are of relevance for the next time step of simulation, including institutional and social attributes (e.g. $H_{village}$, H_{age}, etc.), labor resources (e.g. H_{labor}), financial resources (e.g. $H_{gross\ rainy}$, $H_{gross\ dry}$, etc.), and land resources (e.g. $H_{holdings}$).

These variables are imported as text files into NetLogo, each storing 200 values, one for each household. Just like the ascii files, these files can be easily called by NetLogo, assigning each value to its respective household agent. After the creation of the set of these 200 agents, the population will be augmented by creating copies of these basic agents until the actual population size is reached. These new agents are allocated to the same village as their original, and distributed within the different compounds in the respective village. The actual population sizes for each of the villages were calculated from statistical data sets provided by the Ghanaian Survey Department.

In the third step, when all agents have been created, the distances of these agents to landscape features such as main river and dams are calculated. Furthermore, if dams have been inserted into the landscape, the distance to dams are updated for all landscape agents.

Finally, since virtual household agents should also own patches as in reality, this procedure allocates land holdings to each of the agents. The sizes of these land holdings are given by the holding variables of the agents, as called by the first procedure. The location of these patches is given by the land-allocation procedure, which works as a loop. In each loop, each agent is allowed to select one single patch, and the procedure will be run until all agents are assigned their specific amount of land.

The loop itself runs as follows: As long as patches within the Landscape Vision are still available (i.e. P_{owner}= 'nobody'), the called agent will mark a random patch within this vision as his. If no patch within the Landscape Vision is available, the agent will select a random patch within the same village, and if none of these are available, the agent will select a random patch from the whole catchment. The design of this procedure avoids a biased pattern of distances of owned patches to their respective owners.

3.6.2 Time-loop procedure

The time-loop procedure consists of a collection of sequential routines, which will be run in each time step (Figure 3.10). The policy parameters apart from the location and size of the dams, as well as the parameters of population growth, are usually set before simulation, but can also be modified during the simulation, if this is of interest to the model user. The time-loop starts with the update of the population, i.e. deletion and creation of household agents, allocates credit to this updated population, and then starts with the annual production cycle, beginning with the dry-season procedures and ending with those for the rainy season. Finally, agent and landscape variables are updated according to the results of these procedures. The main steps of this time-loop procedure are outlined in the following:

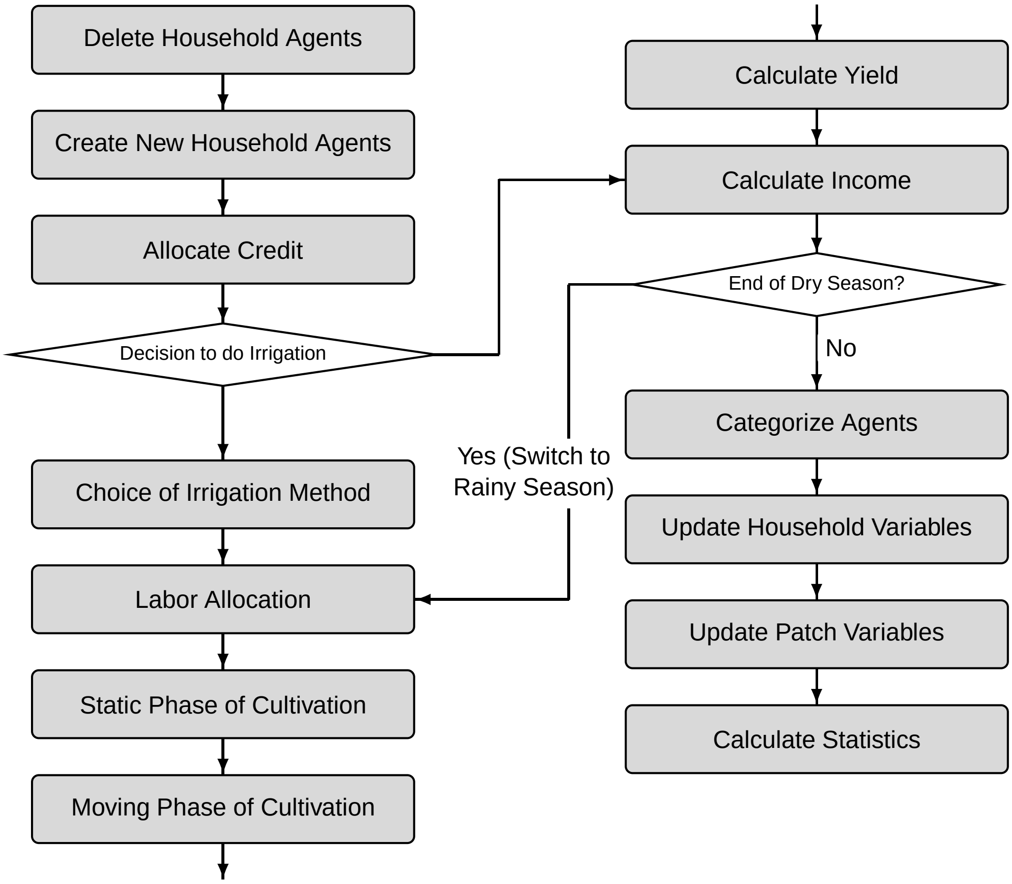


Figure 3.10: Time-loop procedure

1. Update of age and deletion of household agents. In this step, the age of the household agent is updated, and if the maximum age is arrived, the agent is deleted.
2. Creation of new household agents. This procedure creates new household agents according to the new population size, as calculated by the parameters of population growth, and the number of deleted agents without successor..
3. Allocation of credit. According to the annual credit access rate, agents are selected randomly to obtain credit, whils those agents are preferred that had obtained a lesser number of credits do far.
4. Decision to do irrigation. In this step, each household agent generally decides between doing irrigation and not doing it. This procedure is dependent on both the household's

state and the biophysical attributes of the landscape (see section 3.4.1; procedure D_{irr}).

5. Choice of irrigation method. If the decision to do irrigation is positive, the agent will decide here about the irrigation method he is going to use (see section 3.4.1; procedure D_{method})
6. Labor allocation for the dry season. In this step, the dry-season labor pool will be allocated to the various production lines, dependent on the group the agent belongs to. Furthermore, for each credit the agent had obtained, a shift in the labor allocation is executed.
7. Static phase of dry-season cultivation. Here, the agent starts cultivating his own irrigable patches, by deciding about land-use type and input of fertilizer and labor. The procedure runs as long as the required labor and cash resources are available.
8. Moving phase of dry-season cultivation. In this step, the agent will start searching for new patches, but with the same land-use related decisions as in the static phase. The procedure runs until the combined labor and cash resources are exhausted, or until all irrigable patches within the Landscape Vision of the agent are under use.
9. Calculation of dry-season yield. This procedure calculates the yield of each irrigated plot in the local currency, using productivity functions (see section 3.3.1).
10. Calculation of dry-season income. In this step, the cash and gross incomes for each production line are calculated. Furthermore, the gross income is augmented according to the credit access of the household and the credit deflating factor.
11. Labor allocation for the rainy season. Similar to the dry season, the rainy-season labor pool will be allocated to the various production lines, being dependent on both the agent group and the credit access patterns of the household.
12. Static phase of rainy-season cultivation. Here, the agent starts cultivating his own patches, by deciding about land-use type, management, and input of labor. The procedure runs as long as the required labor resources are available.
13. Moving phase of rainy-season cultivation. This procedure is similar to the static phase, apart from the fact that the agent now shifts to new patches, if labor is still available. Once the agent has borrowed a patch from another agent, he will try to continue using it in the next time step.

14. The calculation of rainy-season yield. This procedure calculates the yield of each cultivated plot in the local currency, using productivity functions (see Chapter 5).
15. The calculation of rainy-season income. Equivalent to the rainy season, in this step the cash and gross incomes for each production line are calculated, also being dependent on the credit access pattern of the household.
16. Agent Categorizer. After the season-specific procedures have terminated, the agent categorizer will allocate each agent to its nearest group, while the values of the grouping criteria for each group are updated according to the mean criteria values of the group members.
17. Update of household variables. According to the group the agent has been assigned to, the group-specific household variables will be updated. Furthermore, all other household variables that are the result of the previous procedures will be updated for the next time step.
18. Update of landscape variables. This routine, called the land-cover transformation procedure, will update the land-cover type for those patches that had undergone a land-cover change during the simulation of the previous procedures.
19. Statistical calculations. Finally, statistical parameters will be generated for both the landscape and the population. On the population side, mean annual income as well as the corresponding Gini Index are calculated, and on the landscape side, land-cover and land-use fractions are calculated for both seasons.

4 LAND-USE DECISIONS BY HETEROGENEOUS HOUSEHOLD AGENTS

4.1 Introduction

Land-use dynamics, which involve decisions of land users, are major determinants of land-cover changes. Thus, the critical element in land use is the human agent, who takes specific actions to his own calculus or decision rules that drive land-cover change (Lambin et al., 1999). However, in order to give a meaningful representation of such human agents, heterogeneity regarding land-use decisions among these agents needs to be considered (Rand et al., 2002). The importance of diversity in agent behavior in complex systems (see Chapter 1) suggests that it is worth an effort to characterize the observed heterogeneity in an agent population (Fernandez et al., 2003). Some recent studies have shown that differences in the livelihood background of the human agents usually result in different patterns of land-use behavior (e.g. Le, 2005; Caviglia-Harris and Sills, 2005; Soini, 2006). Therefore, any classification approach to derive typical agent groups for land-use choice should be based on a meaningful representation of agent livelihoods.

In general, the livelihood of humans comprises resources or capital, ranging from human, natural, social, physical to financial capital, which enable the employment of strategies to survive and to attain desirable livelihood outcomes such as income, food security, well-being and sustainable use of natural resources (Carswell, 1997; Carney, 1998; DFID, 2001). Such survival strategies are intricately linked to land-use decisions, as in rural agricultural areas most of the production lines are directly dependent on land resources. Recent studies have shown that statistically causal analyses of observed data can be used to derive such livelihood typologies of agents, as well as the specific behavior with respect to land-use decisions for each human agent group (e.g. Fernandez et al., 2003; Le, 2005).

According to this discussion, this research assumes that if causal relationships exist between the biophysical environment, socio-economic characteristics of farmers and their land-use actions, farmers with different livelihood typologies living in different environmental and policy conditions will have different behavioral patterns about land-use choices. Based on this hypothesis, this chapter has two interrelated objectives:

1. To identify livelihood typologies of households, endogenous factors that differentiate such households typologies, and, based on these endogenous factors, to develop an agent categorizing procedure.
2. To determine and calibrate land-use choice models, whereby land-use behavior should be determined by the specific livelihood groups of the households.

In order to gain an overview of the living conditions and livelihood background of local farmers, first a detailed description of the socio-economic setting of the study area is given. Based on this background, the identification and categorization of livelihoods are addressed, and finally the specification of the decision-making sub-models is presented.

4.2 Socio-economic setting of the study area

4.2.1 Living conditions

The study area consists of a typical savannah parkland, with most of the land used for small-farm agriculture in the rainy season. Most of the area is covered by scattered compounds - large mud buildings - that are usually surrounded by farmland of mixed cropping of groundnuts, cereals and rice. Small grassland patches are usually scattered among the agricultural plots, serving as grazing land for the local livestock. As the area is mainly occupied by cropland and grazing plots during the rainy season, little natural vegetation is left, apart from scattered trees, which mostly have economic, medicinal or social value. Only along the river banks and in stony areas, patches of dry-savannah vegetation are left, since regular flooding and infertile soils limit the agricultural use of this land. In the dry season, small irrigated patches for tomato cultivation can be found mostly along the riverside, while the soils of the remaining area are left bare.

Field observations suggest that the living conditions vary significantly among the different households in the study area in terms of housing quality, household assets, financial means, land and labor availability, and livestock. The compound houses usually consist of several houses connected by mud walls, thereby forming a yard that is shared by all family members. Many of household activities take place in this yard, such as food preparation, cooking, eating, socializing and sleeping. The walls of the compound houses are mostly made of mud bricks, pure mud, or even cement in some cases. The roofs are usually made of corrugated iron or a combination of mud and wood, while only few of the living rooms are covered with thatch. Mostly, houses made with corrugated iron and cement were found among the better-off farmers, who were often involved in irrigation farming, whereas pure mud buildings rather represented the low-income farmer group. Many of the households owned radios or bicycles, while donkey carts, sowing machines and bullock ploughs were only found among 25 % of the households. Cars and fridges were almost completely absent in the area.

Although agriculture is the main economic activity, many households are engaged in activities such as artefacts making, wood cutting, trading, traditional medicine, and even white collar jobs such as teaching. The main sources of cash income include the sale of food crops and livestock, trading, food processing and handicrafts. Field observations suggest that better-off farmers have a tendency to derive their cash income from trading and white collar jobs, while the low-income group of farmers is more reliant on activities such as handicrafts and food processing. Some farmers could also be categorized as livestock farmers, who have a tendency to focus on cattle rearing. In general, livestock, and especially the number of cattle, turned out to be a good indicator for the household's wealth, ranging from several cattle to a few chickens. Land resources were identified to be another indicator for the household's living standard, as the amount of land varied strongly among local farmers. On average, the holdings of local households had an area of 2.4 ha, with a maximum of 22.4 ha and a minimum of 0.1 ha. Another factor describing the differences in livelihood among local households was the availability of labor. As such, households that had many children had a much lower capacity for generating income. These households also showed a different land-use behavior, as they usually focused on mixed compound cultivation, which is the common subsistence cultivation system in the area.

4.2.2 Land tenure

Understanding the land tenure system is essential for modeling the use of natural resources. Land in the study area is perceived to be a spiritual entity, which cannot be owned by an individual. The Tindaana or 'Earth Priest', usually the patrilineal descendant of the first family that settled at the place, has the spiritual authority over the land (Gyasi, 2004). The Tindaana grants usufruct rights to families or households. Each family to whom land has

Figure 4.1: Typical compound house in northern Ghana

been allocated has the prior right to cultivate the land in perpetuo (Gyasi, 2004). Although ownership rights are vested in the community, each family's access to land is secure. The inheritance of land in the study area is patrilinear, with only few women being in charge of the land in cases where the husband has died or is disabled and the male children are still of young age.

The one who first cleared a virgin piece of land 'owns' it, although ownership does not give the right to sell or lease the land (Gyasi, 2004). Although leasing of land is not allowed, some farmers tend to lend parts of their land to family members or friends, usually in exchange for small gifts or even cash.

4.2.3 Agricultural land use

Agriculture is mainly restricted to the rainy period from May to September. During the dry season, agriculture is only possible with irrigation, and about 38 % of the farmers are involved in irrigation agriculture during that season. In the following, we will describe the range of cultivation systems and the farming practices for each of the seasons separately.

Rainy season

Under a relatively low population density until the beginning of this century, the main system of farming was shifting cultivation. Nowadays, two farming systems are prevalent in the study area: The compound farming system, which is a system of mixed cropping surrounding the compound buildings, and the bush fallow system, which typically involves intercropping in out-fields operated on a rotational basis.

The bush fallow system is characterized by clearing and burning of the vegetative cover. This normally exposes the soil to erosion and leaching leading to soil infertility. While the soil fertility used to be restored by long fallow periods, the fallow periods have drastically decreased owing to population pressure (Botchie et al., 2003). The compound farms on the other hand symbolize permanent agriculture with soil fertility often maintained via household waste and animal manure. Chemical fertilizers are hardly applied in the rainy season, nor are there any soil conservation measures applied to enhance soil fertility.

Dry season

In the study area, there are two types of irrigation methods: bucket irrigation using hand-dug wells, and pump irrigation using large dugouts along the river banks or in the main river itself. Although there are also small-scale dams in the study area, these cannot be used as they are wrongly constructed. Only in the areas near Navrongo are a few small-scale dams still in use, apart from the two large-scale dams Tono and Vea, which are located west and east of the study area.

The irrigation capacity of bucket irrigation is lower than that of pump irrigation, which usually results in smaller irrigated patches for bucket irrigation. Furthermore, dugouts and wells need to be maintained almost permanently, which requires high labor input, and in many cases laborers are hired. Further expenditures for pump irrigation involve the continuous repair of the motor pumps, and the costs for oil and petrol, while for bucket irrigation only buckets and ropes are needed. The variety of crops grown during the dry season is mainly confined to local tomato varieties, either in monocultures, or in mixed cultures with small amounts of red pepper, onions or leafy vegetables. Fertilizer application is practiced by all irrigation farmers, with the main chemicals being Urea and DDT.

4.2.4 Main cropping systems

Since the spatial distribution and dynamics of land-use types is of prior concern in this study of land-use/cover change, it is necessary to obtain a valid definition of these land-use types. The difficulty in defining the main land-use types is that many crops are grown in combination with others, which leads to a high variety of land-use types if all combinations are considered. To reduce this variety in a reasonable way, different combinations of land-use types were tested for their relevance to the land-use model. To make sure that the model reflects the dynamics of land-use change in a reasonable way, those land-use types were chosen that were best represented by the livelihood background of the farming households. This way, the following main land-use types could be identified for the rainy season: The mixed compound system, mixed cultures based on groundnuts, monocultures of groundnuts, rice, monocultures of cereals and a class consisting of the minor crops soybeans (*Glycine max*) and sweet potatoes (*Ipomoea batatas*). In the dry season, where the tomato is the by far most prevalent crop, only the two land-use types monoculture of tomatoes and a mixed culture based on tomatoes could be identified.

Cropping sytems in the rainy season

The compound farm system is a permanent mixed cropping system consisting mainly of early millet, late millet, guinea corn, cowpeas and leafy vegetables. Minor crops such as tobacco and okra, which are usually grown in the inner circle of the compound, were omitted in the analysis due to their low quantities. This system is mostly located around the compound buildings, and soil fertility is regenerated by techniques traditionally involving mainly household refuse and manure from the livestock (Gyasi, 2004). This land-use type is the most widespread cultivation system, covering 48.2 % of the total cultivated area in the study area.

The monoculture system of groundnuts occupies about 7.8 % of the cultivated area. Groundnuts (*Arachis hypogaea*) are less nutrient-demanding than the other staples grown in the area and can therefore be easily cultivated on gravelly or sandy-loamy soils, which are usually not suitable for other local staples. Furthermore, there is a tendency to cultivate groundnuts on distant plots, as this crop is less labor intensive than other local crops.

Figure 4.2: Typical groundnut and millet fields in the Atankwidi catchment

In Africa, the groundnut is considered a women's crop (Kenny and Finn, 2004). This is also substantiated by an analysis of household data, showing that the percentage of women within a household is highly correlated to the percentage of area with groundnuts. Groundnuts were originally grown by women to supplement their family diet with protein (Kenny and Finn, 2004). However, groundnut production can also be a way for women to earn cash income and participate in the economy. Among rainy season crops, the groundnut is the staple most often retailed, although, in general, the disposal rate of rainy-season food crops is quite low, due to the subsistence nature of rainfed cultivation.

The mixed culture based on groundnuts is, with 29.1 %, the second most widespread cultivation system in the study area. Within this system, groundnuts are often combined with bambara beans or cowpeas, and sometimes with late millet, which helps to enhance soil fertility. Another reason for combining groundnuts with beans on distant plots is that beans are not eaten by birds and therefore do not require supervision.

In 86.7 % of the cases, the rice-based system consists of a rice monoculture. The remaining 13.7 % of mixed cultures consist in most of the cases of a combination of guinea

Figure 4.3: Typical rice fields in the Atankwidi catchment

corn and rice, and sometimes of a combination with small amounts of early millet, late millet or okro. Until recently, most of the rice cultivated was African rice (*Oryza glaberrima*), which was gradually replaced by Asian rice (*Oryza sativa*) in most parts of the study area. Rice production has increased during the last decades due to an improved access to tractors, which facilitates the field preparation on the heavy clayey-loamy soils that are usually suitable for rice cultivation. In total, rice fields cover about 6.7 % of the cultivated area.

The monoculture of cereals is, together with rice, the cultivation system with the greatest distance from the compound, with an average distance of 1 km. It consists of different combinations of Guinea corn (*Sorghum guineense*), early millet (*Milium vernale*), late millet (*Pennisetum claucum*) and sometimes maize (*Zea mays*). Guinea corn, which was originally adopted from a neighboring region, is increasingly cultivated in the study area, as it is more adapted to the reduced length of the rainy period, which is possibly a result of climate change. The small quantities of maize, which usually need chemical fertilizers to grow well, are remnants of the times before the structural adjustment program, when fertilizer was locally subsidized by the government. Cereal monocultures are usually cultivated along the riverside, where the nutrient supply is sufficient, covering about 7.4 % of the total cultivated area.

The other cropping types, covering only 0.7 % of the cultivated area, comprise monoculture of soybeans and cultures based on sweet or Irish potatoes, usually mixed with red pepper. These two cultivation types had to be combined in one land-use type, since their occurrence turned out to be too low to allocate them to two separate classes.

Cropping systems in the dry season

As the tomato is the by far most prevalent crop in the dry season (90 % of all irrigated crops are tomatoes), the only meaningful classification of land-use types in this season was a separation among monocultures of tomatoes and mixed cultures based on tomatoes. The major tomato varieties used are 'Petromech' and 'No Name', sometimes combined with onions, red pepper and leafy vegetables in a mixed culture system. These mixed systems amount to about 40 % of the irrigated area, the remaining 60 % being tomato monocultures. In general, irrigation is quite a young business in the study area. The irrigation farming in the study area only began around 16 years ago by using bucket irrigation. Nowadays, about 38 % of the farmers are involved in irrigation farming, with 35 % of them using motor pumps, and 65 % still practicing bucket irrigation. The choice of irrigation method does not seem to have an influence on the choice of land-use type.

4.3 Modeling livelihood groups

As studies suggest the importance of heterogeneity in land-use decisions (Fernandez et al., 2003), an approach to represent this heterogeneity is required. It is a common assumption that land-use decisions are related to the livelihood strategy of a farming household; thus the diversity of agents regarding land-use decisions can be achieved by a categorization of these agents into group with individual livelihood strategies. Some recent studies have shown that statistical analyses of empirical data can be used to derive such agent typologies, as well as specific behavior with respect to land use for each agent group or typology. In this chapter, the statistical methods for the derivation of agent groups as well as the range of explanatory livelihood indicators and the corresponding results are presented.

4.3.1 Identification of livelihood groups

Livelihood indicators

We applied the concept of the livelihood framework for selecting criteria that represent the livelihood structure and strategy of farming households. The livelihood framework is a concept which divides a household's resources into five different categories, called household assets. These comprise human, social, financial, natural and physical capital (Ashley and Carney, 1999; Bebbington, 1999; Campbell et al., 2001). For representing livelihood groups in a reliable way, indicators within each of these categories needed to be selected. The notable advantage of this diversified selection of indicators is that, by doing so, biased selections of grouping criteria are avoided (Campbell et al., 2001).

Based on this approach, the understanding of livelihood disparities in the study area (see section 4.2) and available studies of livelihood indicators of Ghanaian households (see Ghana Statistical Service, 2000; Ashong and Smith, 2001; Yaro, 2000), the following variables (see Table 4.1) were selected to represent the overall livelihood typology of a farming household:

1. Three variables indicating the household's human resources: household size, labor availability, and dependency ratio
2. Two variables representing the household's financial capital: total gross income and total gross income per capita
3. Three variables describing natural capital: cultivated area in the rainy season, total holdings, and total holdings per capita
4. Two variables representing physical and social capital: livestock index and number of cattle.

Apart from the above indicators, the percentages of income from the monoculture of groundnuts, the mixed culture based on groundnuts, and compound farming were included in the statistical analysis, as they directly indicate the livelihood strategy regarding land use. Field observations and statistical analysis suggest that these incomes differ significantly among households with different livelihood backgrounds. This way, households with a higher tendency to practice subsistence farming usually focus rather on compound farming, as this land-use type provides the basic staples for home consumption. On the other hand, households with a tendency towards market-based farming are more inclined to cultivate groundnuts for sale, especially in monocultures.

Statistical analyses

Based on the above selected livelihood indicators, two statistical methods were employed for the identification of agent groups, i.e. Principle Component Analysis (PCA) and k-mean Cluster Analysis (k-CA). PCA is a statistical method to condense a set of variables into a smaller set, while k-CA is a method to derive clusters of cases (in our case agent groups). We conducted PCA using all livelihood indicators (Table 4.1) to identify important indicators that differentiate household livelihood typologies. Subsequently, k-CA was applied to these condensed variables and used to identify typical household livelihood typologies.

Principle component analysis

Since the dimension of the selected set of livelihood indicators was too large for further analysis, we used the method of PCA to reduce the dimension of this variable set. This method condenses those variables that highly correlate with each other to one Principle Component, with the aim of minimizing the loss of information induced by this condensation. The Principle Components PC_j derived in such a way can be formally expressed as linear combinations of the standardized original variables:

$$PC_j = \sum_j b_{ij} \cdot X_i \tag{4.1}$$

where X_i are the standardized original variables, and the loadings b_{ij} the coefficients calculated by SPSS. The values of the coefficients are determined in such a way that the Principle Components correlate with each other at a lowest level possible. The aim of the PCA is therefore to detect components which best represent the observed coherences between the original variables.

We ran PCA with Varimax rotation and the Kaiser normalization, and the scores of extracted Principle Components were saved and standardized. Based on the values of the

Table 4.1: Livelihood indicators for categorizing farming agents

Variable	Definition
H_{size}	Size of household (number of household members)
H_{labor}	Availability of household labor (number of workers)
H_{depend}	Dependency Ratio (H_{labor} / H_{size})
$H_{gross\ inc}$	Gross annual household income (local currency)
$H_{gross\ inc\ percap}$	Gross annual household income per capita (local currency)
$H_{holdings}$	Total area of holdings (the land owned by the household (m^2))
$H_{holdings\ percap}$	Total area of holdings per capita (m^2)
$H_{cult\ rainy}$	Total area cultivated in the rainy season (m^2)
$H_{livestock}$	Livestock Index
H_{cattle}	Cattle number owned by the household
$H_{\%\ inc\ lu\ 2}$	Percentage of income from the cultivation of monocultures of groundnuts (land-use type 2) of gross income of rainy-season cultivation
$H_{\%\ inc\ lu\ 3}$	Percentage of income from the cultivation of compound farming (land-use type 3) of gross income of rainy-season cultivation
$H_{\%\ inc\ lu\ 6}$	Percentage of income from the cultivation of mixed groundnut cultures (land-use type 6) of gross income of rainy-season cultivation

weight parameters b_{ij}, we finally named the Principle Components after those initial variables that had the highest correlation to the components (Table 4.3).

K-mean cluster analysis

To derive agent groups, we used the standardized scores of the Principle Components to run k-mean Cluster Analysis. The k-means algorithm is an algorithm to cluster objects based on selected attributes into k partitions, while the objects of one partition should feature similar variable characteristics, and those of different partitions dissimilar ones. Mathematically, the objective of this algorithm is to achieve the minimization of total intra-cluster variance V, expressed as:

$$V = \sum_{i=1}^{k} \sum_{x_j \in S_i} (x_j - \mu_i)^2 \tag{4.2}$$

where $S_i, i = 1, \ldots, k$ are the k clusters (in our case agent groups), the $x_j \in S_i$ the elements of each cluster (in our case household agents), and the μ_i are the centroids or means of each cluster. Each of the x_j and μ_i has as many dimensions as the data set, i.e. one dimension for each variable. Thus, $(x_j - \mu_i)^2$ can be regarded as the distance of the agent x_j to the group centroid μ_i.

The main advantages of this algorithm are its simplicity and speed, which allows it to be run on large data sets. On the other hand, its major drawback is that it does not yield the same result with each run, since the resulting clusters depend on the initial random assignments (Bühl and Zöfel, 2000). But due to the relatively large data set, and the fact that each run resulted in the same classification, this approach seemed to be appropriate.

Results

The PCA was applied to the selected variables characterizing livelihood patterns (Table 4.1) and resulted in 7 Principle Components. The total variance explained amounted to 95 % (Table 4.2), which is quite high, meaning that only 5 % of the information was lost by the replacement of the original variables through Principle Components. In Table 4.3, the Rotated Component Matrix is presented, showing the weight parameters bij among the Principle Components and the original variables characterizing livelihood typologies, whereby values below 0.1 were omitted for a better overview.

The first Principle Component is strongly related to the variables of labor availability (b_{ij} = 0.953) and household size (b_{ij} = 0.929), and is therefore named the 'labor factor', which accounts for 25.6 % of the total variance explained. A pair correlation among these two variables showed that they are highly correlated (Pearson's R = 0.885, $p < 0.001$). The second Principle Component shows high correlations to the total area of the owned by the household (b_{ij} = 0.896), the total area owned by the household per capita (b_{ij} = 0.840), and the area cultivated in the rainy season (weight parameter = 0.755). Thus, this Principle Component was labeled the 'land factor', accounting for 15.1 % of the total variance explained. Pair correlations among these three variables were all significant ($p < 0.001$), with the Pearson's R coefficients between 0.396 and 0.631.

For the third Principle Component, the livestock index and the number of cattle were significant, showing weight parameters of 0.979 and 0.978, respectively; thus, this component

Table 4.2: Total variance explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Components	Total	% of Variance	Cumu-lative %	Total	% of Variance	Cumu-lative %	Total	% of Variance	Cumu-lative %
1	3.331	25.621	25.621	3.331	25.621	25.621	2.244	17.261	17.261
2	1.980	15.231	40.852	1.980	15.231	40.852	2.118	18.291	33.552
3	1.850	14.233	55.085	1.850	14.233	55.085	1.956	15.046	48.598
4	1.710	13.154	68.239	1.710	13.154	68.239	1.826	14.045	62.643
5	1.302	10.018	78.257	1.302	10.018	78.257	1.651	12.700	75.344
6	1.090	8.386	86.643	1.090	8.386	86.643	1.304	10.033	85.377
7	1.005	7.731	94.374	1.005	7.731	**94.374**	1.170	8.997	94.374
8	0.363	2.792	97.166						
9	0.140	1.077	98.243						
10	0.095	0.733	98.976						
11	0.059	0.455	99.432						
12	0.055	0.420	99.851						
13	0.019	0.149	100.00						

was named the 'livestock factor'. This factor accounted for 14.2 % of the total variance explained, and a pair correlation among the two representing variables showed that they are highly correlated (Pearson's R = 0.976, $p < 0.001$).

The fourth Principle Component is represented by the gross household income (b_{ij} = 0.947) and the gross household income per capita (b_{ij} = 0.931). Thus, we called this Principle Component the 'income factor', which accounted for 13.2 % of the total variance explained. Here, we again executed a crosstab analysis, resulting in a Pearson's R of 0.796 ($p < 0.001$).

The two opposing variables of the 'percentage income from monoculture of groundnuts' and the 'percentage income from mixed culture based on groundnuts' resulted in the fifth Principle component, called the 'groundnut factor'. These two variables exclude each other, because the households will either tend to use a mixed culture or a monoculture of (-0.831 and 0.960, respectively) and their Pearson's R of -0.682 ($p < 0.001$). The groundnut factor accounts for 10.0 % of the total variance explained.

The last two Principle Components are represented by only one variable each, the percentage income from compound mixed farming (b_{ij} = - 0.979), and the dependency ratio

Table 4.3: Rotated component matrix

	Principle Components						
Variables	1 - Labor Factor (25.6 %)	2 - Land Factor (15.2 %)	3 - Live-stock Factor (14.2 %)	4 - Income Factor (13.2 %)	5 - Ground-nut Factor (10.0 %)	6 - Cereal Mixed Factor (8.4 %)	7 - Depen-dency Factor (7.7 %)
H_{labor}	**0.953**		0.127				0.211
H_{size}	**0.929**	0.125	0.193				- 0.217
$H_{holdings}$	0.251	**0.896**			0.110		
$H_{holdings\ percap}$	- 0.385	**0.840**					0.251
$H_{cult\ rainy}$	0.325	**0.755**	0.157	0.175	- 0.130		
$H_{livestock}$	0.133		**0.979**				
H_{cattle}	0.148		**0.978**				
$H_{gross\ inc\ percap}$	- 0.240			**0.947**			
$H_{gross\ inc}$	0.258			**0.931**			- 0.115
$H_{\%\ inc\ lu\ 2}$					**0.960**	0.193	
$H_{\%\ inc\ lu\ 6}$					**- 0.831**	0.528	
$H_{\%\ inc\ lu\ 3}$				- 0.131		- **0.979**	
H_{depend}							**0.992**

Notes: Numbers in parentheses are percentages of total variance of the original variable set explained by the principle components.

(b_{ij} = 0.992). Here, the Principle Components are named after their original variables, the 'compound mixed factor', and the 'dependency ratio factor', explaining 8.4 and 7.7 % of the total variance respectively.

On these 7 Principle Components, the k-mean Cluster Analysis was applied to derive clusters representing the specific livelihood agent groups. The disadvantage of this method is that the number k of clusters has to be set beforehand. To solve this problem, the k-mean Cluster Analysis was run for $k = 1, \ldots, 11$, and for each run the distances of each household to the cluster centers were calculated. One household had to be omitted, as for each k this household formed a single group, which was considered as an outlier. The target was then to select the value for k that met the following two conditions: First, a low average distance to the cluster centers, and second, reasonable cluster sizes, which should be large enough to ensure statistical validity for further applications. To analyze the first condition, the cluster number k was plotted against the average distance to the cluster centers (Figure 4.4).

As visualized, the average distance to the cluster centers decreases until $k = 3$, then slightly increases, and finally decreases again from $k = 5$ upwards. Therefore, the values of $k = 3$ and $k \geq 5$ had to be considered as cluster numbers. But further analysis showed that the second condition of reasonable cluster sizes was not met anymore for values above 5. We therefore decided to set $k = 3$ for this study. Descriptive statistics then were used to check if the three clusters were meaningful (Table 4.4).

The k-CA run for $k = 3$ on the standardized scores of the Principle Components resulted in three agent groups of sizes 111, 77 and 11. In Table 4.4, for each agent group descriptive statistics of those variables are shown that best represented the Principle Components (with the highest weight parameters). In the following, a description of the characteristics of each household type is given:

Household type 1

The most conspicuous characteristic of this category of farmers is the high availability of land, ranging from 4.500 to 223.800 m^2 with a mean of 31.500 m^2. The second characteristic is the high diversity of land-use types cultivated by the households. In Figure 4.5, the percentages of the gross income from each land-use type of the total gross income of rainy-season cultivation are displayed for each farmer group. Remarkable is the difference between the three groups in the percentage of groundnut monocultures. Among farmers from the first household type, about 34.2 % of the total cultivated area is covered by groundnuts monocultures, whereas the percentages for the second and the third household type amount only to 1.6 and 3.2 %, respectively.

Apart from the relatively high land availability, the first group can be regarded as the 'middle class' of farmers, with a medium livestock index and a medium dependency ratio.

Likewise, regarding the practice of dry-season farming, this household type can be considered as the 'medium' class in comparison to the respective values of the other types, with more than 51.9 % of the farmers practicing dry-season farming. In total, this group of

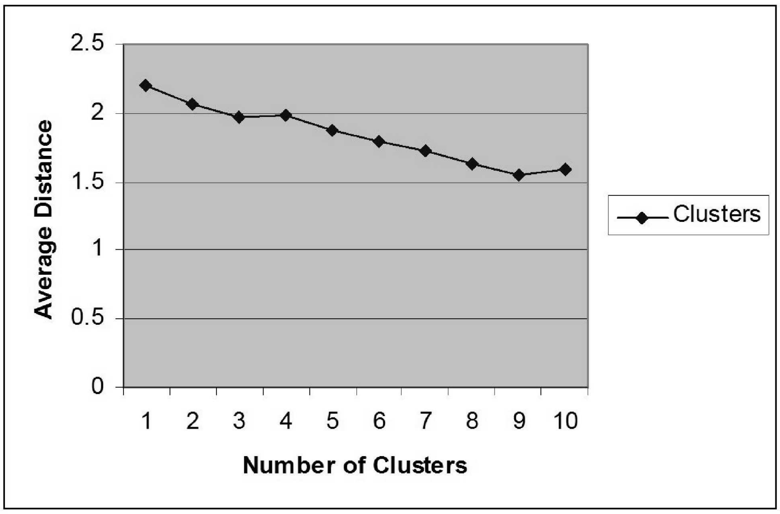


Figure 4.4: Average distances to cluster centers for k clusters

’middle class’ farmers constitutes about 38.7 % of the population.

Household type 2

This class of farmers can be considered as the poorest of all household types, with the lowest labor availability (3.724 persons per household), the lowest amount of total land holdings (18.395 m^2), income per capita (2.1 million Cedis), and the lowest livestock index. The subsistence level is the highest for this group, with an annual mean cash income of 4.9 million Cedis, compared to 9.5 and 31.5 million Cedis for the household types 1 and 3, respectively. The income proportion from mixed groundnuts and compound farming is dominant within this group, while the proportion of rice - which is considered a cash crop - is the lowest of all groups, suggesting that the level of subsistence farming is highest for this group. The fraction of households practicing dry-season farming is also quite low at 35.1 % (Figure 4.6); the majority use bucket irrigation, which is the lower-cost riverine irrigation method. In total, this household group of ’poor farmers’ makes up 55.8 % of the population.

Household type 3

Households of this group are richer than others in terms of livestock and income per capita; income ranges from 3 to 15 million Cedis per person. The ownership of land for this group

Table 4.4: Descriptive livelihood statistics

Variables	Agent Group	N	Mean	Std.Error	Minimum	Maximum	Std. Deviation
H_{labor}	1	77	7.006	0.346	2	16.0	3.041
	2	111	3.734	0.117	1	7.0	1.242
	3	11	6.090	0.709	3	9.5	2.353
$H_{holdings}$	1	77	31463	3380	4537	223800	29659
	2	111	18395	1134	1205	64078	11949
	3	11	23100	4409	4820	45042	14625
$H_{livestock}$	1	77	6872	711	368	34407	6235
	2	111	5052	690	0	56336	7267
	3	11	7441	1446	2270	16313	4795
$H_{gross\ inc\ percap}$	1	77	2165184	163638	239800	8152254	1435919
	2	111	2127251	120611	93218	6517703	1270717
	3	11	6921292	1161211	3031187	15714007	38513021
$H_{\%\ inc\ lu\ 2}$	1	77	0.342	0.037	0	0.912	0.328
	2	111	0.016	0.007	0	0.585	0.077
	3	11	0.032	0.032	0	0.353	0.106
$H_{\%\ inc\ lu\ 3}$	1	77	0.178	0.016	0.000	1.000	0.147
	2	111	0.281	0.025	0.000	1.000	0.272
	3	11	0.232	0.069	0.047	0.842	0.230
H_{depend}	1	77	0.689	0.015	0.388	1.0	0.139
	2	111	0.683	0.017	0.321	1.0	0.185
	3	11	0.705	0.046	0.444	0.9	0.154

is medium at about 23.100 m^2 per household. The pattern of gross income from rainy-season cultivation shows that households of this group focus on the cultivation of rice, with the proportion of rice being the highest among all groups (Figure 4.5). For this group, the average income from the sale of rice per household amounts to about 5.7 million Cedis, compared to only 1.2 and 0.6 million Cedis for groups 1 and 2, respectively, which indicates that rice is considered as a cash crop among farmers of this group. This further indicates that the land-use composition of this household type is more directed towards the cultivation of cash crops than subsistence crops. Furthermore, with 81.6 % of all farmers, this group is highly involved in dry-season farming, with 27.3 % practicing pump irrigation, which is the most costly local irrigation method (Figure 4.6). In total, this group of 'better-off farmers' amounts to 5.5 % of the whole population.

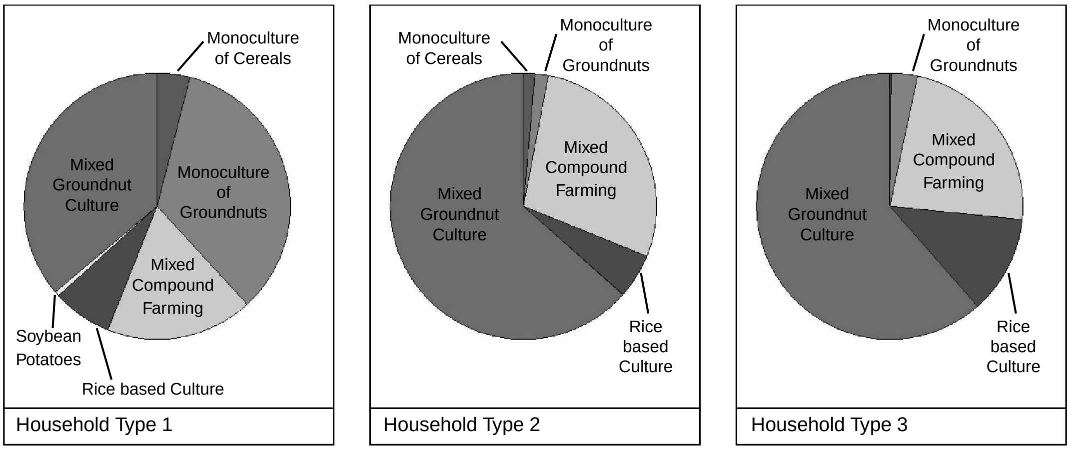


Figure 4.5: Structure of gross income from rainy-season cultivation

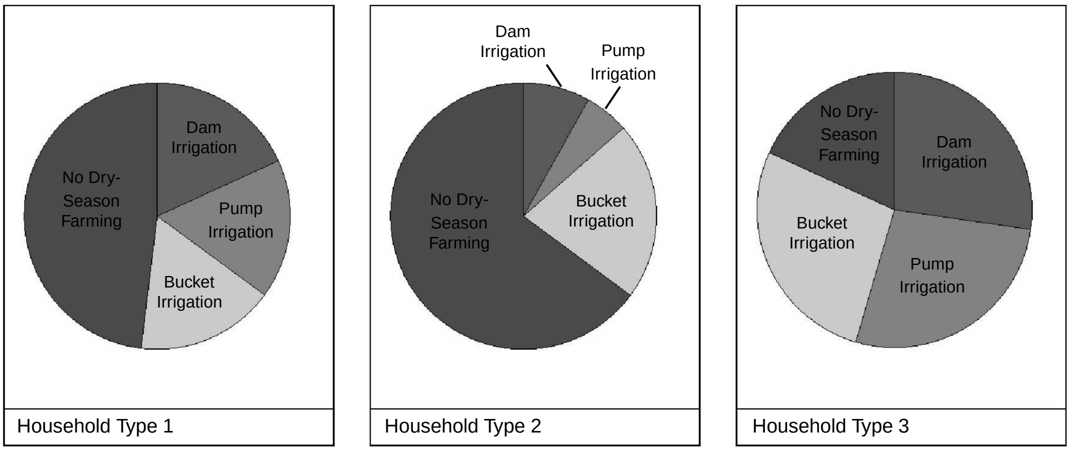


Figure 4.6: Frequency of irrigation practices of each household group

4.3.2 Agent Categorizer

The Agent Categorizer is a classifier routine (built into GH-LUDAS) to assign agents to their most similar group based on the identified grouping criteria (section 4.3.1). The most straightforward approach for classifying agents during a model run is to calculate 'distances' from each agent to each group, and assign the agent to the group with the smallest distance. There are a number of methods that can be employed to calculate such distances, including the Euclidian distance, which can be used to measure the distance between the agent's values and the mean values (of grouping criteria) of each agent group. Several methods for calculating distances were tested in a separate model, whereby the m-logit approach showed the best results, with 100 % of correct predictions. Using multinomial logistic regression, the distance of an agent A to agent group g is calculated as:

$$Dist_g = \frac{e^{\alpha_g + \sum_i \beta_{ig} V_i}}{\sum_h e^{\alpha_h + \sum_j \beta_{jh} V_j}} \tag{4.3}$$

where $Dist_g$ is the distance value of agent A to group g, V_i the values of the grouping criteria (see section 4.3.1) of agent A, and α_g and β_g the constant and preference coefficients of the grouping criteria for group g. The values for the constant as well as the preference coefficients were calculated using SPSS (Table 4.5), whereby the reference category is the third agent group. All groups as categorized by the k-mean Cluster Analysis were correctly predicted, justifying the use of this model for classification.

4.4 Modeling land-use decisions

Based on the identified livelihood groups, the main target was to develop decision-making sub-models regarding the choices among land-use types and the decisions related to irrigation farming. However, the relatively small sample size of irrigation farmers did not allow a group-wise approach for modeling the dry-season-related decisions, i.e. the decision to use irrigation and the choice of irrigation method. Instead, the preference coefficients for the m-logit models of these choices were not determined for each group separately, but for the total population. This way, unreliable results due to small sample sizes were avoided. In the following, we will describe the models of the choices among rainy- and dry-season land-use types, and the models describing the decision to do irrigation and the choice of irrigation method.

4.4.1 Modeling choices among land-use types

In this section, the models for choices among land-use types are presented, including the methodology, the specification of the range of explanatory variables, and the subsequent results. For the choice among rainy-season land-use types, an m-logit model was employed, with group-specific preference coefficients. However, regarding the choice among dry-season land-use types, a simpler approach needed to be applied, the use of which will be justified in the respective section.

Table 4.5: Parameter estimates of the m-logit model of the Agent Categorizer

	Preference Coefficients	
Variables	Group 1	Group 2
Constant	- 37.237***	128.723***
H_{labor}	- 13.604**	- 26.741*
H_{depend}	134.921*	103.340**
H_{size}	11.840*	3.735***
$H_{holdings}$	0.002	0.000
$H_{holdings\ percap}$	- 0.012*	- 0.003
$H_{cult\ rainy}$	0.000***	0.001**
$H_{livestock}$	0.000**	0.002*
H_{cattle}	2.467	0.167
$H_{gross\ inc}$	0.000**	0.000***
$H_{gross\ inc\ percap}$	0.000**	0.000**
$H_{\%\ inc\ lu\ 2}$	92.745*	- 75.260*
$H_{\%\ inc\ lu\ 3}$	- 80.803**	- 48.178***
$H_{\%\ inc\ lu\ 6}$	- 34.075*	12.853**
Model Fitting Information:	.	.
Chi-Square = 341.411, df = 26, Sig. = 0.000		
Pseudo R Square:	.	.
Cox and Snell = 0.995, Nagelkerke = 1.000, Mc Fadden = 1.000		

Specification of the variables for the m-logit model for the rainy season

Dependent variable

The dependent variable of the model is the choice of land-use type by a household in the rainy season. This categorical variable of land-use types comprises 6 land-use alternatives: monoculture of cereals, monoculture of groundnuts, mixed compound system, rice-based culture, soybean/potatoes, and mixed culture based on groundnuts (see section 4.2.4).

Explanatory variables

For the adequate modeling of land-use choice, all factors related to local household decision-making should be taken into consideration. This includes the environmental setting of the household plots, the socio-economic state, and the land-use preferences of the household (Table 4.6). The selection process of the range of variables within these three categories consisted of both intensive farmer group discussions and the supervision of the 'goodness-of-fit'

(R Square) of the m-logit model for the given variables.

Table 4.6: Range of variables for the m-logit model of rainy-season land-use choice

Variable	Definition	Data Source
Dependent Variable		
$P_{\text{land-use rainy}}$	Coded rainy-season land-use type	Interview and field observation
Characteristics of the plot user		
H_{age}	Age of the household head (in years)	Interview
H_{wives}	Number of wives of the household head (if the household head is male)	Interview
H_{depend}	Dependency ratio (number of dependants / total household members)	Interview
$H_{\text{hlds percap}}$	Total area owned by the household per capita	Interview and field measurements
H_{gender}	Sex of the household head	Interview
$H_{\text{comp head}}$	Compound head status (1 if compound head, 0 otherwise)	Interview
$H_{\text{\% lu 2 rainy}}$	Percentage of cultivated area of Monoculture of Groundnuts (land-use 2)	Interview
$H_{\text{\% lu 3 rainy}}$	Percentage of cultivated area of Mixed Compound Farming (land-use 3)	Interview
$H_{\text{\% lu 6 rainy}}$	Percentage of cultivated area of Mixed Culture of Groundnuts (land-use 6)	Interview
Environmental attributes of land plots		
P_{upslope}	Upslope contributing area	GIS-based (DEM)
P_{texture}	Soil texture (ranking scale)	Map-based calculation
$P_{\text{fertility}}$	Soil fertility (ranking scale)	Map-based calculation
$P_{\text{irr coeff}}$	Irrigation Coefficient indicating the level of irrigability (between 0 and 1)	Calculation
$P_{\text{dist user}}$	Distance of the plot to the land user (km)	Field measurement
$P_{\text{dist border}}$	Distance of the plot to the national border (km)	Field measurement

i) Environmental Variables

As our aim is to explicitly simulate the land-use decisions of local farmers, we have to understand the factors that play a role within these decisions. According to local interviews, the abundance and type of grass on a piece of land is an important indicator for the farmer whether and for which crops the soil is appropriate. Furthermore, according to traditional knowledge, soil color, texture and moisture are further indicators for the decision among the various land-use types. For instance, a grey surface and a sandy soft soil are considered to be suitable for the cultivation of groundnuts, whereas harder soils are more suitable for millet. Soil moisture should be high for rice cultivation, medium for cereals such as millet, and lowest for groundnuts.

Biophysical variables were selected to represent these soil/water conditions, which are hypothesized to be of varying importance for the different land-use types. These include soil fertility, representing the abundance of grass, upslope contributing area, irrigation coefficient, and soil texture. While the upslope contributing area approximates rather the soil moisture content caused by topography, the irrigation coefficient represents the geological component of soil moisture including factors such as groundwater level and recharge. Among topographic factors, upslope contributing area was selected, since this variable describes the relative position of a land patch, being higher for valleys and lower for mounds. This differentiation is important, as rice is preferably cultivated in local valleys, which serve as staging areas for runoff. This way, this factor can be assumed to play a role in the identification of rice plots, as the local position of the piece of land is part of the farmer's considerations. Soil texture also can be considered as an indicator of land-use choice, as the local soils suitable for the various local crops differ in the topsoil composition of particle sizes. For example, local farmers tend to cultivate groundnuts preferably on soils with a larger mean topsoil particle size, in contrast to other local staples.

Apart from such biophysical attributes, factors of spatial accessibility were hypothesized to influence land-use choice, including the distance of the plot to the compound and the distance to the national border. The distance to the compound is minimal for the land-use type of mixed compound farming, as this land-use type is always located in the immediate vicinity of the compound building. The reason is that mixed compound farming requires high inputs of animal manure, which can only be transferred over short distances. Land-use types based on groundnuts are usually located further away from the compound, as groundnuts need less attention in terms of labor and management. Another factor determining the choice of crops on distant plots is that certain crops need to be protected from livestock and birds. Local crops such as maize and cowpeas are preferably cultivated on distant plots, since their seeds not eaten by birds and therefore need less protection. Apart from the distance to the compound, the factor of distance to the national border was included in the analysis, as we noticed a spatial gradual shift in land-use patterns along the south-north axis. This difference in land-use patterns was characterized by a higher portion of cereal-based farming in the north together with a higher poverty level, indicating that the degree of subsistence farming - which is mainly based on cereal cultivation - was higher up north. This north-south gradient is, according to our field observations, caused by the remoteness of the northern part in terms of infrastructure (e.g. markets, roads), which can be explained by the close distance to the border, and by a lack of irrigation possibilities.

ii) Variables of household characteristics

The household characteristics deemed significant for land-use choice are age and gender of the household head, number of wives (if the household head is male), compound head status, dependency ratio, and total land holdings per capita. In the study area, a gradual shift among land-use types from traditional cereal farming to the cultivation of rice and groundnuts was observed during the last decades. One of the main reasons for this is that the younger generations tend to prefer cash crops such as rice and groundnuts to traditional crops; this is supported by the empirical data set, which shows a much higher percentage of such cash crops among younger farmers. To reflect this variation in land-use preferences, we hypothesized the age of the household head to be an explanatory variable for land-use choice. In a similar vein, just as there are differences among young and old farmers, there is also a difference when it comes to the gender of the household head. Female farmers usually tend to focus on the cultivation of groundnuts, since these are less labor-intensive, whereas the typical domain of male farmers is cereal farming, which requires hard work for maintenance and weeding. Therefore, we also included the gender of the household head as a hypothetical factor for land-use choice.

The dependency ratio and the number of wives of the household head both reflect

Table 4.7: Assumed effects of drivers on land-use choice

Variable	Assumed effects on land-use choice	
	Sign	Land-use type / crop type
H_{age}	(-)	Cash Crops
H_{wives}	(+)	Groundnuts
H_{depend}	(+)	Mixed Cultures
$H_{holdings\ percap}$ / 1000	(-)	Cereals
H_{gender}	(-)	Groundnuts
$H_{comp\ head}$	(+/-)	—-
$H_{\%\ lu\ 2\ rainy}$	(+)	Monoculture of Groundnuts
$H_{\%\ lu\ 3\ rainy}$	(+)	Mixed Compound Farming
$H_{\%\ lu\ 6\ rainy}$	(+)	Mixed Groundnut Culture
$P_{upslope}$ / million	(+)	Rice
$P_{texture}$	(-)	Groundnuts
$P_{fertility}$	(-)	Groundnuts
$P_{irr\ coeff}$	(+/-)	—-
$P_{dist\ user}$	(-)	Mixed Compound Farming
$P_{dist\ border}$	(-)	Cereals

the needs of the household regarding its diet. The dependency ratio reflects the number of mouths each worker feeds, thus relating to the urgency in food demands of the household (Fatoux et al., 2002). Households with a high dependency ratio could be forced to grow a larger variety of crops, since most of these would be used for home consumption. Therefore, a high dependency ratio is assumed to be an indicator for the preference of mixed cultures (e.g. mixed compound system, mixed groundnut culture). The number of wives is a similar factor explaining the urgency in food demands, but with the slight difference that each woman usually holds her own groundnut plots to feed her own family, resulting in a tendency towards groundnut cultivation.

Finally, the variable of land holdings per capita was hypothesized to be higher for the land-use types of groundnuts, since groundnuts are only a supplementary staple of the local menu. Therefore, farmers with little land might tend to focus on the main staples such as millet and Guinea corn.

iii) Land-use tendency of the household

We also have to consider that local farmers usually do not make a new land-use decision every year, but are rather inclined to maintain continuity and rely on their previous decisions.

Since such continuity cannot be reflected by the variables above, we decided to include factors explaining the general land-use tendency of the household. This land-use tendency is represented by the fractions of the land-use types of the rainy-season cultivation area from the previous year. Through the inclusion of these variables not only is the continuity in land-use decisions ensured, but also the possibility of a gradual change in these decisions, as the land-use tendencies are allowed to change over time in GH-LUDAS. Among these land-use fractions of the total cultivated area, we selected the most meaningful variables with respect to their difference among agent groups, including monoculture of groundnuts, mixed compound farming, and mixed culture of groundnuts.

Results of m-logit model of land-use choice for the rainy season

Based on these indicators, we applied an m-logit regression for the choice among land-use types for each household group separately. This resulted in group-specific preference coefficients, reflecting the overall land-use tendency of each livelihood group. In the following, we present the results as well as the goodness-of-fit for the m-logit models (for each agent group), and discuss the importance of selected significant land-use drivers.

Household Type 1

The results of the m-logit analysis of rainy-season land-use choice for household type 1 are summarized in Tables 4.8 and 4.9. The preference coefficients were calculated with respect to the land-use type mixed groundnut culture, which served as the base case. The choice of the base case did not have any influence on the calculated preference coefficients.

The chi-square test shows that the empirical m-logit model of land-use choice for this agent group is highly significant with $p = 0.000$. The Nagelkerke's Pseudo R Square of 0.541 shows that 54.1 % of the total variation in the probability of land-use choice is explained by the selected explanatory variables. Furthermore, for this agent group, 50.8 % of the choices among land-use types are correctly predicted.

Household Type 2

Using the same range of variables, an m-logit regression was also conducted for the second household group (Tables 4.10 and 4.11). The likelihood ratio test showed that the empirical

Table 4.8: Group 1: Rainy-season land-use choice: parameter estimates

Variable	Rainy-Season Land-Use Type: Mono-culture of Cereals	Mono-culture of Groundnuts	Mixed Compound Farming	Rice based Culture	. Soybean/ Potatoes
Intercept	- 21.181***	7.808***	- 0.486	- 22.465***	- 16.297***
H_{age}	0.000	- 0.003	0.006	- 0.032	- 0.016
H_{wives}	- 0.392	- 0.437	- 0.089	- 0.122	- 0.738
H_{depend}	- 0.988	- 2.210	0.002	0.544	2.028
$H_{hlds\ percap}$ / 1000	- 0.011	- 0.003	0.022	0.091	- 0.088
H_{gender}	25.067	-5.044**	0.941	23.298	16.528
$H_{comp\ head}$	0.037	- 0.391	- 0.064	- 0.352	- 0.481
$H_{\%\ lu\ 2\ rainy}$	- 2.475	1.108	0.655	- 1.159	2.222
$H_{\%\ lu\ 3\ rainy}$	- 2.144	- 0.316	0.974	- 0.631	- 0.220
$H_{\%\ lu\ 6\ rainy}$	- 4.647***	- 6.037***	- 3.711***	- 2.878	- 2.069
$P_{upslope}$ / million	0.008	0.039	- 0.023	0.022	0.005
$P_{texture}$	- 0.136*	- 0.045	0.051	- 0.207**	- 0.198
$P_{fertility}$	0.208	0.229	- 0.004	0.579	0.267
$P_{irr\ coeff}$	3.814*	- 0.713	3.697	2.630	- 3.073
$P_{dist\ user}$	0.401	- 0.381	- 1.900***	0.422	- 0.458
$P_{dist\ border}$	- 0.043	0.074	- 0.029	0.080	0.056

Model Fitting Information: Chi-Square = 194.017, df = 75, Sig. = 0.000

Pseudo R Square: Cox and Snell = 0.520, Nagelkerke = 0.541, Mc Fadden = 0.225

The reference category is: Mixed Groundnut Culture

Table 4.9: Group 1: Rainy-season land-use choice: classification table

Observed	Predicted: Mono-culture of Cereals	Mono-culture of Groundnuts	Mixed Compound Farming	Rice based Culture	Soybean/ Potatoes	Mixed Groundnut Culture	Percent Correct
Monoculture of Cereals	20	2	7	8	0	4	48.8 %
Monoculture of Groundnuts	6	19	12	3	0	5	42.2 %
Mixed Compound Farming	3	10	45	1	0	14	61.6 %
Rice based Culture	7	6	6	8	0	5	25.0 %
Soybean/ Potatoes	0	3	1	0	0	2	0 %
Mixed Groundnut Culture	3	5	16	1	0	42	62.7 %
Overall Percentage	14.8 %	17.0 %	33.0 %	8.0 %	0 %	27.3 %	50.8 %

choice model is highly significant with p = 0.000. The test for the goodness-of-fit showed that the model has an acceptably good fit, with a Nagelkerke's Pseudo R Square of 0.600. The model also has a satisfactory predictive power, as 65.5 % of the choices are correctly predicted.

Household Type 3

Because of the relatively small size of this agent group, two of the six land-use types were not found among this group, i.e. groundnut monocultures and soybean/potatoes. Out of the cases representing the remaining four land-use types 79.3 % were correctly predicted (Table 4.13).

Specification of land-use choice algorithm for the dry season

Two different dry-season land-use types were identified in the study area, namely tomato monocultures and mixed cultures based on tomatoes (section 4.2.4). The mixed tomato cultures consist on average of more than 90 % of tomatoes, with only small amounts of pepper, onions and leafy vegetables, which are mostly meant for home consumption. The decision to add such small amounts of vegetables depends on the personal taste of the farming household head, and is thus difficult to simulate. However, there are small differences in dry-season land-use choice among younger and older farmers, as well as among households with a low and a high dependency ratio. An m-logit model for land-use choice was tested with GH-LUDAS, incorporating variables such as age, number of wives, dependency ratio, as well as environmental variables, since pepper, which is the most prevalent crop after tomatoes, prefers different soil and moisture conditions. Nonetheless, this model had a low predictive power with low R Squares, which might be due to two reasons: First, the data set comprising the two land-use types was relatively small, with only 40 plots of tomato monocultures and 15 plots of mixed cultures. Second, as already mentioned above, the decision to add such small amounts of vegetables is difficult to model, as it is dependent on the personal taste of the household head and his family. For these reasons and the low predictive power of the tested m-logit model, we found that the use of such a model would not lead to reliable results, and decided to use a simpler, more robust approach.

This approach consists of the use of the mean percentages of each of the two land-

Table 4.10: Group 2: Rainy-season land-use choice: parameter estimates

Variable	Rainy-Season Land-Use Type: Mono-culture of Cereals	Mono-culture of Groundnuts	Mixed Compound Farming	Rice based Culture	. Soybean/ Potatoes
Intercept	0.251	-94.384	1.909	0.225	-489.287
H_{age}	-0.007	2.820	0.013	-0.017	0.122
H_{wives}	0.145	-37.281	- 0.193	- 0.602	0.746
H_{depend}	2.239	-185.068	-0.327	0.938	-0.438
$H_{hlds\ percap}$ / 1000	-0.052	0.304	-0.066	-0.087	0.410
H_{gender}	-0.212	425.204	-0.009	0.298	15.274
$H_{comp\ head}$	-0.142	-68.955*	-0.453	0.730	-0.097
$H_{\%\ lu\ 2\ rainy}$	2.018	434.822*	0.626	1.017	-87.211
$H_{\%\ lu\ 3\ rainy}$	-1.101	- 309.814	1.289*	- 0.135	-12.812
$H_{\%\ lu\ 6\ rainy}$	- 2.889*	- 169.786	-1.505*	-2.815**	- 8.122
$P_{upslope}$ / million	0.058	-126.312	0.630	1.048*	3.662
$P_{texture}$	-0.101	- 7.829	-0.062	- 0.132*	22.800
$P_{fertility}$	-0.148	-98.251	0.062	0.140	20.762
$P_{irr\ coeff}$	5.684**	-2785.279*	2.418	4.860**	- 146.270
$P_{dist\ user}$	0.589	47.659**	- 6.068***	0.339	-1.815
$P_{dist\ border}$	- 0.012	34.547**	- 0.047	0.075	- 0.046

Model Fitting Information: Chi-Square = 275.030, df = 75, Sig. = 0.000

Pseudo R Square: Cox and Snell = 0.559, Nagelkerke = 0.600, Mc Fadden = 0.305

The reference category is: Mixed Groundnut Culture

Table 4.11: Group 2: Rainy-season land-use choice: classification table

Observed	Predicted: Mono-culture of Cereals	Mono-culture of Groundnuts	Mixed Compound Farming	Rice based Culture	Soybean/ Potatoes	Mixed Groundnut Culture	Percent Correct
Monoculture of Cereals	8	0	4	7	0	11	26.7 %
Monoculture of Groundnuts	0	4	1	0	0	0	80.0 %
Mixed Compound Farming	0	0	93	1	0	20	81.6 %
Rice based Culture	4	1	8	19	0	19	37.3 %
Soybean/ Potatoes	0	0	1	0	1	1	33.3 %
Mixed Groundnut Culture	3	0	31	4	0	95	71.4 %
Overall Percentage	4.5 %	1.5 %	41.2 %	9.2 %	0.3 %	43.5 %	65.5 %

Table 4.12: Group 3: Rainy-season land-use choice: parameter estimates

	Rainy-Season Land-Use Type		
Variable	Mono-culture of Cereals	Mixed Compound Farming	Rice based Culture
Intercept	-44063	- 48838	- 29985
H_{age}	- 1344	- 1595	- 1070
H_{wives}	1390	1560	1052
H_{depend}	120467	142716	95436
$H_{hlds\ percap}$ / 1000	- 114	- 110	- 79
$H_{comp\ head}$	- 6119	- 7114	- 4798
$H_{\%\ lu\ 3\ rainy}$	90	178*	43
$H_{\%\ lu\ 6\ rainy}$	- 35638	- 41595	- 27489
$P_{upslope}$ (million)	- 106	- 209	- 139
$P_{texture}$	1138	1261	777
$P_{fertility}$	3689	3712	2010
$P_{irr\ coeff}$	22740**	25599	15542
$P_{dist\ user}$	- 175	- 348***	- 145
$P_{dist\ border}$	- 83	- 8	51

Model Fitting Information: Chi-Square = 124.090, df = 39, Sig. = 0.000

Pseudo R Square: Cox and Snell = 0.676, Nagelkerke = 0.714, Mc Fadden = 0.702

The reference category is: Mixed Groundnut Culture

Table 4.13: Group 3: Rainy-season land-use choice: classification table

	Predicted				
Observed	Mono-culture of Cereals	Mixed Compound Farming	Rice based Culture	Mixed Groundnut Culture	Percent Correct
Monoculture of Cereals	1	0	0	0	100.0 %
Mixed Compound Farming	0	9	0	2	81.8 %
Rice based Culture	0	1	3	0	75.0 %
Mixed Groundnut Culture	1	2	0	10	76.9 %
Overall Percentage	6.9 %	41.4 %	10.3 %	41.4 %	79.3 %

use types for each agent group. Each agent is assigned the mean percentages of the two land-use types according to the agent group he belongs to, i.e. the agent's choice among the two land-use types is determined by the corresponding probabilities of his agent group. Thus, the tendency to cultivate mixed cultures is not given by the individual agent, but is represented by the average tendency of the group members. Furthermore, as agent groups are dynamic such that an agent possibly changes his group over time, this tendency is also allowed to change during time. The mean percentages of each agent group for the cultivation of monocultures amount to 57 % for the first group, 65 % for the second, and 61 % for the third group. The algorithm for choosing a certain land-use type can be depicted as follows for an agent A:

1. If A is member of group G, set the probability to choose mixed culture P_G (which is the mean percentage of this land-use type.)
2. For a given patch, set land-use type monoculture of tomatoes.
3. Generate a random number r between 0 and 1.
4. If $r < P_G$, set land-use type mixed culture of tomatoes.

4.4.2 Modeling irrigation-related decisions

Methodology

For modeling irrigation-related decisions, we decided to use a two-fold nested m-logit model. The first m-logit model will simulate the general decision of a household agent to do dry season farming, while the second will then simulate the choice of irrigation method, if the decision of doing irrigation in the first step is positive (Figure 4.7). This two-fold nested decision is taken by each household agent in each time step of the model run after the rainy-season simulation procedures, and is independent of the group of agents.

In the following, we will describe the variables used for this nested decision-making model and give reasoning for the selection of these variables. First, we will introduce the dependent and explanatory variables of the first step of the model.

Specification of the variables of the first step of the m-logit model

Dependent variable

The dependent variable of this first step of the m-logit model is simply the choice by farming households between doing irrigation and not doing irrigation. This variable is represented in

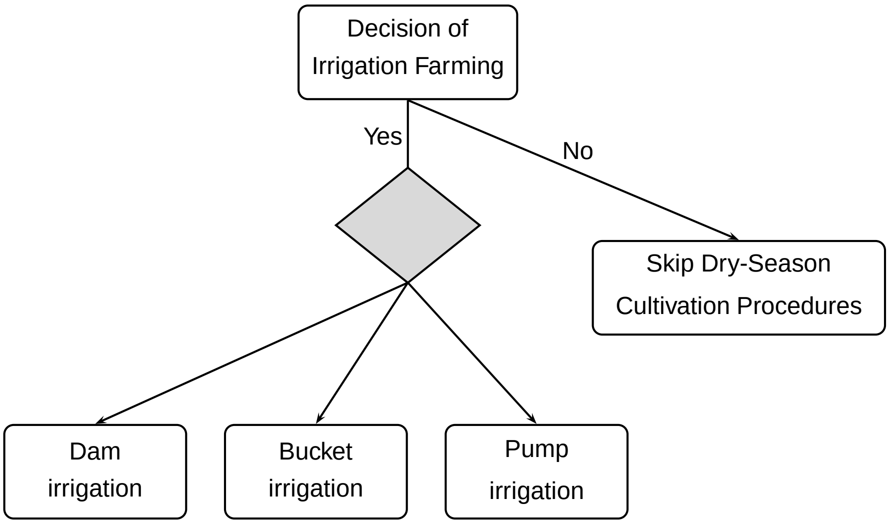


Figure 4.7: Decision tree for the nested m-logit model for irrigation decisions

the model by the dummy variable ($H_{\text{dry dummy}}$) which is 1 if the farmer is engaged in irrigation during this time step, and 0 otherwise.

Explanatory variables

In the study area, field observations and insight from the household surveys showed that most of the farming household heads who are not engaged in irrigation farming are willing to start it. Furthermore, those who are already involved in this business, would like to expand, which is due to the high profitability of this business. Only few household heads refused to get involved in irrigation farming, mostly due to old age or sickness. Thus, this decision of the household can hardly be regarded as a choice as such, but merely as a question of capability. Explanatory variables that are hypothesized to be important in the decision for dry-season farming should therefore reflect the capability of the household to practice irrigation. To reflect this overall household capability, we employed an economic approach, which defines the involvement in a business as being dependent on the availability of the four resources labor, land, capital and knowledge. However, since manpower is abundant in the dry season due to less farming activity, labor can be easily rented for irrigated cultivation, and is therefore already represented by the factor financial capital.

The factor land with respect to irrigation implies that the required piece of land should be irrigable. The access to such irrigable land is defined by local tenure rights, mean-

Table 4.14: Variables for the first step of the nested m-logit model for irrigation decisions

Variable	Definition	Data Source
Dependent Variable		
$H_{\text{dry dummy}}$	Dummy variable indicating whether the farmer is doing irrigation	Interview
Independent Variables		
$H_{\text{cash rainy}}$	Cash income from the rainy season (in Cedis)	Interview and Calculation
$H_{\text{neigh dry}}$	Percentage of immediate neighbors involved in irrigation farming	Estimation by interviewee
$H_{\text{hlds dry}}$	irrigated Area owned by the household (in m^2)	Field Measurement
$H_{\text{dist water}}$	Distance to water sources (including dams and main river) in m	Map-based Calculation
$H_{\text{perc NFA}}$	Percentage of income from non-farm activities of total annual gross income	Interview and Calculation
$H_{\text{inv strat}}$	Dummy variable indicating whether the farmer would invest in irrigation farming	Interview

ing that a single household either owns such land or can try to borrow some. Thus, two factors can be assumed to represent the access to irrigable land. First, the ownership of such land, and second, if no irrigable land is owned, a factor reflecting the chance of the household to borrow such land. In the study area, the borrowing of land is often facilitated by friendship and family relations, meaning that land is preferably granted to relatives and friends, who mostly live in the immediate neighborhood. This way, the chance of a farmer to obtain such land decreases with the distance to the irrigated area. Therefore, in order to represent the availability of land resources for irrigation with respect to land tenure, we decided to include the irrigated area owned by the household, as well as the distance of the household to water sources suitable for irrigation, which include both dams and the main river.

There are three main reasons why financial capital, the second factor, is needed: First, the purchase and application of fertilizer and other chemicals is almost inevitable for dry-season cultivation in the study area. Second, the maintenance of two of the irrigation methods hand dug wells and dugouts requires a large input of manpower, which has to be covered in many cases by rented labor. Third, other expenditures, such as the repairs and servicing of motor pumps for pump irrigation, as well as items for bucket irrigation require financial means. To represent this financial factor, we decided to use the variable of cash income of the household from the rainy season. We suppose that cash income is a better indicator for this decision than gross income, since in most cases, the transactions for the purchase of inputs and labor are made in cash. The cash income comprises the income from the sale of agricultural products and animals, as well as the income from non-farm activities such as trading, food processing and handicrafts. The use of this variable also implies that an appropriate modeling of cash income in each time step of the model is essential.

The third economic factor, which is hypothesized to be of importance when modeling the decision to do dry-season farming, is knowledge or know-how. This factor is represented by the percentage of immediate neighboring households that are involved in irrigation farming. However, this factor does not exactly reflect the transfer mechanisms of knowledge, which could also be mediated through clans or families instead of neighbors, but is nevertheless the most straightforward approach to capture this aspect as closely as possible, since the modeling of social networks was beyond the scope of this study.

Apart from these economic factors, we included a factor representing the timely fashion in which a farmer manages to start dry-season farming. In the study area, observations suggest that many farmers first get involved in non-farm activities, because these activities do not require such large cash inputs as irrigation farming. If enough financial capital is accumulated from these non-farm activities, many farmers shift to the irrigation business. In order to represent this factor, we included the percentage of non-farm activities of the total gross income (per year) as an explanatory variable in our model.

Finally, in order to capture the degree of willingness of the household head to engage in irrigation farming, we included a dummy variable - which was obtained during the socio-economic survey 2006 - that indicates whether the farmer would invest in irrigation farming if he had additional income. We call this variable the investment strategy ($H_{\text{inv strat}}$).

Results of first step of irrigation m-logit model

Based on the above variables, we calculated the preference coefficients for the m-logit model of choice between irrigation farming and no irrigation farming (Table 4.15), with the reference category being irrigation farming. All selected explanatory variables were significant at a level of $p < 0.01$, and the model had a high predictive power, with 85.4 % of the cases correctly predicted, and a Nagelkerke Pseudo R Square of 0.678 (Tables 4.15, 4.16 and 4.17).

The values of the calculated preference coefficients (Table 4.15) strongly confirm the theory of the effects of the selected variables. Thus, the lower the cash income, the lower the probability of getting involved in dry-season farming. The same is valid for the owned irrigable/irrigated area, the percentage of neighbors involved in irrigation farming, and the investment strategy, which is 1 if the farmer is wiling to invest in irrigation, and 0 otherwise. The lower all these factors are, the lower the chance of the farmer to irrigate. On the other hand, the higher the distance to water sources and the higher the percentage of income from non-farm activities, the lower is this probability.

Specification of variables of second step of m-logit model

Dependent variable

The dependent variable within the second step is the choice of irrigation method once the farmer decided to irrigate, and is represented by the household variable $H_{\text{irr method}}$. The methods are bucket irrigation, pump irrigation, and reservoir irrigation, if a dam is available.

Explanatory variables

The most significant difference among the three irrigation methods is the difference in financial requirements. Comparing pump and bucket irrigation, pump irrigation is the more profitable method, since more land can be put under cultivation, but it is also the more costly one. The maintenance of the dugout on the one hand and fuel, oil and repairs of the motor pump on the other usually cause high costs compared to the bucket method, which is usually less costly to operate. However, both types require high labor input for the maintenance of the wells and dugouts, for which labor needs to be rented in many cases, thereby increasing the input costs. Among all irrigation methods, reservoir irrigation can be regarded as the cheapest method, as the payment for use usually does not exceed the costs for the other methods. Since farmers are often forced to choose the method they can afford, we included the variable of cash income from the rainy season to represent the financial ability of the household with respect to this choice. For the m-logit model, the logarithm of this variable was selected.

Furthermore, we included three more variables in the m-logit model of choice of

Table 4.15: First step of the nested irrigation decision model: parameter estimates

No Irrigation	B	Std. Error	Wald	df	Sig.	Exp(B)	95 % Confidence Interval for Exp(B) Lower Bound	95 % Confidence Interval for Exp(B) Upper Bound
Intercept	- 0.128	0.663	0.037	1	0.847			
$H_{cash\ rainy}$	- 0.234	0.076	9.382	1	0.002	0.791	0.681	0.919
$H_{inv\ strat}$	- 2.259	0.841	7.223	1	0.007	0.104	0.020	0.542
$H_{perc\ NFA}$	0.059	0.014	17.614	1	0.000	1.061	1.032	1.091
$H_{neigh\ dry}$	- 3.680	0.798	21.290	1	0.000	0.025	0.005	0.120
$H_{dist\ dams}$	1.186	0.298	15.850	1	0.000	3.275	1.826	5.872
$H_{hlds\ dry}$ / 1000	- 0.229	0.071	10.444	1	0.001	0.795	0.692	0.914

The reference category is Irrigation

Table 4.16: First step of the nested irrigation decision model: correct predictions

Observed	Predicted: No Irrigation	Predicted: Irrigation	Predicted: Percent Correct
No Irrigation	96	15	86.5 %
irrigation	14	74	84.1 %
Overall Percentage	55.3 %	44.7 %	85.4 %

Table 4.17: First step of the nested irrigation decision model: statistics

Model Fitting Information: Chi-Square	df	Sig.	Pseudo R Square: Cox and Snell	Nagel-kerke	Mc Fadden
140.469	6	0.000	0.506	0.678	0.514

irrigation method, one representing the choice between dam and riverine irrigation, and two to separate the choice between bucket and pump irrigation. Since dam irrigation is a relatively low-cost business, the only obstacle for farmers to engage in farming along a dam is its accessibility. To represent this factor, we included the minimum distance of the farming household to dams as an explanatory variable in the model. For the choice among the two riverine irrigation methods, we selected two variables, i.e. the number of years the household has been engaged in irrigation farming, and a dummy variable indicating whether the household owns a motor pump. The number of years is a reasonable indicator, as farmers usually start their irrigation business with buckets in order to shift later to pump irrigation as soon as the necessary financial capital has been accumulated.

Table 4.18: Variables for the second step of the nested m-logit model for irrigation decisions

Variable	Definition	Data Source
Dependent Variable		
$H_{irr\ method}$	Irrigation method (dam, pump or bucket irrigation)	Interview
Independent Variables		
$H_{cash\ rainy}$	Cash income from the rainy season (in Cedis)	Interview and Calculation
$H_{dry\ years}$	Number of years the farmer is involved in irrigation farming	Interview
H_{pump}	Dummy variable indicating whether the household owns a motor pump	Interview
$H_{dist\ dams}$	Minimum distance to dams (in m)	Map-based Calculation

Results of second step of irrigation m-logit model

This model, which simulates the choice among the three irrigation alternatives, has a relatively high predictive power (Table 4.21), with a Nagelkerke R Square of 0.940, although the variables show fairly good significance levels (Table 4.19). Among the three irrigation alternatives, all cases of dam irrigation and bucket irrigation are correctly predicted, with about 76.2 % of correct predictions for the pump irrigation method. In total, 94.1 % are correctly predicted (see Table 4.20).

The results of the m-logit regression are not fully consistent with the theory of the influence of the selected variables as outlined above. In fact, cash income positively influences the choice of the more costly pump irrigation, but the pump dummy variable and the number of years the farmer is involved in irrigation farming hardly show any influence in the choice among these two riverine irrigation methods.

4.5 Summary

The assumption that differences in the livelihood background result in different land-use behavior is verified, as we have seen that the preferences for land-use types and the tendency to irrigate among livelihood groups of farmers vary strongly (see Figures 4.5 and 4.6). To

Table 4.19: Second step of the nested irrigation decision model: parameter estimates

Irrigation Method	Variables	B	Std. Error	Wald	df	Sig.	Exp(B)
Motor pump	Intercept	- 484.079	3025.502	0.026	1	0.873	
	$H_{cash\ rainy}$ (log)	27.112	203.890	0.018	1	0.894	6E+011
	$H_{dist\ dams}$	- 0.130	0.000	470405	1	0.000	0.139
	$H_{dry\ years}$	- 8.647	62.060	0.019	1	0.889	0.000
	H_{pump}	- 413.307	6970.703	0.004	1	0.953	3.18E-180
Bucket	Intercept	- 484.079	3025.502	0.026	1	0.873	
	$H_{cash\ rainy}$ (log)	26.614	203.890	0.017	1	0.895	4E+011
	$H_{dist\ dams}$	0.131	0.000	.	1	.	1.139
	$H_{dry\ years}$	- 8.457	62.060	0.019	1	0.892	0.000
	H_{pump}	- 437.689	0.000	.	1	.	8.2E-191

The reference category is irrigation

Table 4.20: Second step of the nested irrigation decision model: correct predictions

	Predicted			
Observed	Dam	Pump	Bucket	Percent Correct
Dam	25	0	0	100 %
Pump	0	16	5	76.2 %
Bucket	0	0	39	100 %
Perc.	29.4 %	18.8 %	51.8 %	94.1 %

Table 4.21: Second step of the nested irrigation decision model: statistics

Model Fitting Information			Pseudo R Square		
Chi-Square	df	Sig.	Cox and Snell	Nagel-kerke	Mc Fadden
149.595	8	0.000	0.828	0.940	0.828

derive such livelihood groups, the livelihood framework for selecting livelihood indicators was applied, followed by the application of PCA and k-CA. Based on the identified livelihood indicators, the PCA revealed seven core factors that differentiate livelihood typologies of farming households in the study area, namely land, labor, livestock, and income factors, two factors representing the preference for groundnut and compound farming, and the dependency ratio.

Based on these seven extracted components, classification using k-CA resulted in three livelihood typologies of households: the ’middle class’ (household type 1), the ’poor farmers’ (households type 2), the ’rich farmers’ (household type 3). Further land-use analyses for each household type revealed differences in patterns of land-use choice. As such, the cultivation of cash crops had a higher proportion among the rich and middle class farmers, whereas the poor farmers had a tendency to focus on subsistence crops. Moreover, there was an imbalance of irrigation practices among the identified livelihood groups, i.e. the percentage of irrigation farmers in general and pump farmers in particluar increased with the level of living/livelihood standard.

After the derivation of livelihood groups, sub-models for land-use choice were presented and calibrated, whereby the range of explanatory variables and the choice of model were justified, and the results presented. These sub-models include the choice between rainy-season and dry-season land-use types, the decision to do irrigation farming, and the choice of irrigation method. All decision models were developed on the basis of m-logit regression, apart from the choice among dry-season land-use types, as no meaningful variable set could be identified to explain choices among land-use types in this season. The preference coefficients for the m-logit model for rainy-season land-use choice were determined for each livelihood group separately, since the results of a descriptive comparison of land-use preferences among livelihood groups suggested the relevance of such a differentiation. These differences in land-use choice are reflected by the differences in the direction, magnitude and significance of the preference coefficients, which clearly show considerable heterogeneities in local land-use choice behavior. In general, households of all groups choose land-use types based on the considerations of a range of household characteristics, natural conditions and particular policy factors.

With respect to the modeling of irrigation-related decisions, a group-wise approach was considered to be unreliable due to the relatively small sample size of irrigation farmers, which did not allow any further splitting. Instead, the preference coefficients were computed for the total population, which turned out to be the more robust approach. These irrigation-related decisions were modeled as a nested m-logit model, which included the decision to do irrigation as a first step, and as a second step, the choice of irrigation method. Both environmental and household characteristics as well as policy factors were included as explanatory variables within this nested model to reflect the socio-economic as well as the environmental conditions necessary for the engagement in irrigation.

The results and structure of these land-use choice models were integrated into GH-LUDAS within the Decision Module The preference coefficients were used to compute the land-use choice probabilities/utilities, whereby each land-use option during model run is selected by an agent with its respective probability, thus allowing bounded rational decision-making behavior.

5 ECOLOGICAL DYNAMICS OF HETEROGENEOUS LANDSCAPE AGENTS

5.1 Introduction

Complex processes of land-use and land-cover change (LUCC) arise not only from the diversity of human decision-making, but also from the heterogeneous dynamics of the environment (Parker et al., 2003). Environmental drivers of land-use decisions, e.g. current land cover, topography, soil conditions, and agricultural productivity (see Chapter 4), often vary over space and time. These environmental conditions can be changed either by human interventions or by natural processes that are beyond human control (e.g. natural vegetation growth and/or climate variability). In any attempt to model environmental dynamics, it is therefore important to consider the initial spatial heterogeneity of the landscape as well as natural processes and ecological impacts driven by human agents, leading to changes in this heterogeneous pattern of the landscape.

These dynamics as well as the initial biophysical conditions should be captured and calibrated in a spatially explicit way in order to match real-world processes. According to agent-based design, a natural landscape is represented in the form of a grid of cells that are autonomous landscape agents. In order to obtain a spatially explicit representation of the processes and status of the landscape, every landscape unit needs to be endowed with internal state variables storing heterogeneous spatial data, and with internal models of relevant ecological processes, which work in response to the internal state of the landscape unit, inputs/interventions of human agents, and other global environmental factors (e.g. climate). This agent-based representation of the landscape thus treats landscape dynamics as a self-organized phenomenon, which evolves from micro-autonomous processes (Le, 2005).
Following this paradigm, two tasks were performed:

1. The identification and generation of relevant biophysical data for the initialization of the state of the landscape agents, and
2. The development and calibration of ecological sub-models, representing the temporal dynamics of landscape agents.

The first task includes the characterization of the landscape environment in a spatially explicit way, e.g. in the form of GIS raster layers using real data, including topography, accessibility, soil and land-cover classifications, and hydrological data. A land suitability analysis for irrigation is also part of this task, as a meaningful representation of the irrigability of landscape agents plays a major role in modeling irrigation-related decisions. All selected variables for this landscape characterization should be relevant to the calibration of ecological processes, or be main drivers of human decision-making regarding resource use.

The second task includes the development and calibration of biophysical sub-models, comprising productivity functions for each land-use type, a livestock dynamics sub-model, which is related to a specific forage productivity function, and a land-cover transformation model. While the former two sub-models specify yield and forage productivity, the land-cover transformation model simulates conversions among land-cover types. Since ecological dynamics of the landscape agents are the combined result of both heterogeneous natural processes (e.g. vegetation growth, erosion), and interventions of human agents (e.g. management practices), the ecological sub-models are designed to consider both natural and human drivers.

5.2 Characterization of heterogeneous landscape agents and modeling of relevant ecological processes

For a realistic representation of the landscape, both the characterization of the landscape in terms of biophysical and environmental attributes, as well as the respective ecological dynamics within this landscape have to be considered. Thus, the landscape is modeled as an aggregation of heterogenous landscape agents, each endowed with its own state variables and ecological processes. In this chapter, the landscape attributes relevant to land-use decision-making and ecological mechanisms are identified and characterized, comprising land cover, soil attributes, hydrology and topography. These attributes represent the general setup (or static condition) of the landscape as it was in 2006.

5.2.1 Landscape characterization

In this section, the basic characteristics of the landscape are presented, including a land-cover classification and the basic biophysical attributes that are of importance for the dynamics of the coupled human-environment system of land-use/cover change. These attributes are interpreted with respect to the ecosystem's primary productivity. Furthermore, the sources and methodology used to derive a spatial representation of these attributes are given.

Land-cover classification

Because land cover is clearly a key variable of MAS/LUCC models, an accurate mapping of this variable is critically important for the calibration and initialization of the simulation model (Le, 2005). An approach often used to derive main land-cover types is the analysis of satellite images via remote sensing using automatic classification methods. Such automatic classification methods extract the main land-cover types based on spectral information of the satellite image. But since some land-cover types may exhibit similar spectral properties, the accuracy of such automatic classification algorithms is often limited. Therefore, such algorithms are often used in association with other information sources to interpret the automatically derived land-cover classes, e.g. aerial photographs, a high-resolution satellite image, or ground-truth data.

An automatic classification method was conducted on the ASTER image (USGS and Japan ASTER Program, 2007), using the Unsupervised Classification procedure in ERDAS. The image was taken at the end of the rainy season when the vegetation is mature, thus showing the highest difference in spectral attributes. The Unsupervised Classification extracted 15 spectral classes, which were then interpreted using ground-truth data collected in September 2006. The ground-truth data were randomly separated into two equal sets. The first set was used to interpret the 15 spectral classes as derived by the Unsupervised Classification, while the second was used to validate the interpreted classes.

The interpretation of the 15 spectral classes resulted in 5 major land-cover types (Figure 5.1), including i) forest, ii) water, iii) cropland, iv) grassland, and v) bare land. Water covered about 0.1 % of the study area, forest about 4.3 % cropland about 63.8 %, and grassland and bare land 25.4 and 6.4 %, respectively. These values are in accordance with previous studies (e.g. Martin, 2005). The second set of ground truth data was used to validate these classes. The actual value as observed by the ground-truth survey was compared to the predicted value given by the classified land-cover map (Table 5.1). In total, 58.2 % of the land-cover classes were correctly predicted.

Since the resolution of the ASTER Image (15 m x 15 m) did not allow a correct prediction of the river network, this feature was manually digitized using the Quickbird image (DigitalGlobe, 2007), which had a higher resolution. The width of this river network was set to 30 m, which corresponds to the patch size in GH-LUDAS.

Determination of relevant soil-water attributes

Being one of the major determinants of an ecosystem's primary productivity, the inclusion of the spatial variation of the soil/water status is essential for modeling ecological processes on the landscape scale (Park and Vlek, 2002). As the determination of these spatial soil/water conditions is a complex issue, a reliable approach had to be used to represent this factor. According to agent-based modeling philosophy, the most appropriate approach to model a complex phenomenon is by identifying its basic constituent drivers. Thus, a range of parameters was chosen to explain this factor of soil/water conditions: i) two direct soil parameters to represent soil attributes, using a soil texture parameter and a soil fertility parameter, ii) several indirect indicators explaining soil formation through topographical conditions, and iii) two kinds of parameters describing water availability, representing runoff and groundwater availability, respectively. The groundwater parameters include average groundwater level as well as average groundwater recharge, while the runoff parameter is represented by a topographical wetness index, which is calculated from topographical attributes.

Table 5.1: Land-cover classification: correct predictions

	Predicted					
Observed	Forest	Cropland	Grassland	Bare Land	Total	Percentage
Forest	13	2	3	0	18	0.722222
Cropland	26	254	66	11	357	0.711485
Grassland	4	88	59	7	158	0.373418
Bare Land	0	7	16	25	48	0.520833
Total	43	351	144	43	581	0.581989

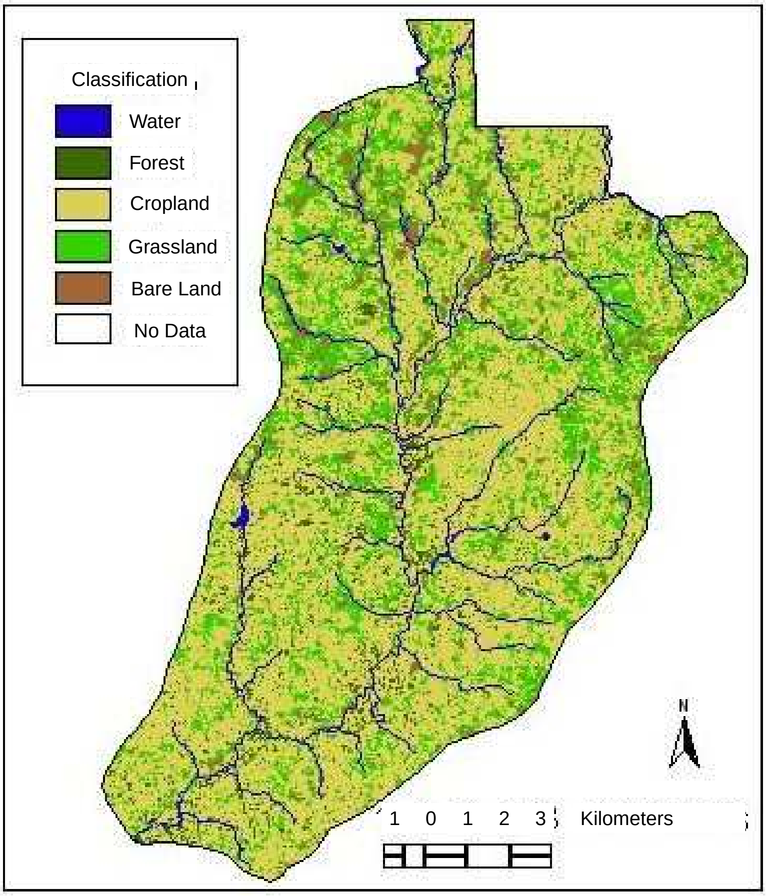


Figure 5.1: Land-cover pattern in the study area

Soil attributes

With respect to crop productivity, soil fertility is the characteristic of the soil that supports abundant plant life, being the combined effects of three major interacting components. These are the chemical, physical and biological characteristics of the soil (Soil Health, 2008). The physical and chemical characteristics of soil are far better understood than those of the biological component; therefore quite a lot is known about the desired chemical and physical status of soils. (Soil Health, 2008).

The well-known main biological conditions include the abundance of organic matter and micro-organisms, while the main chemical attributes important for plant growth comprise the abundance of and access to nutrients and minerals (Soil Health, 2008). The physical structure of the soil is the third component defining soil fertility, and includes soil texture, depth

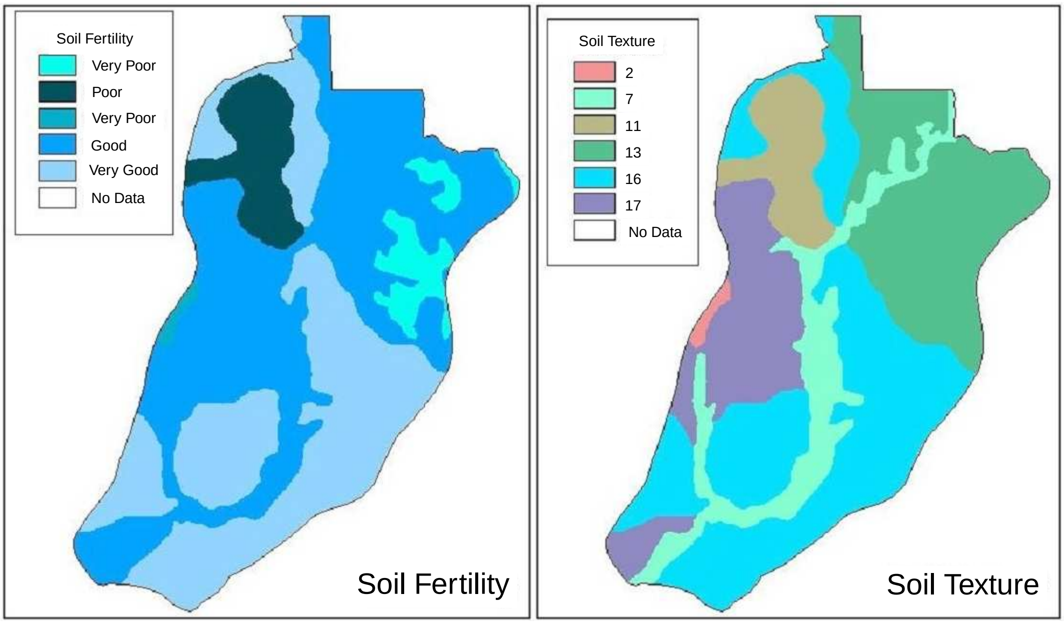


Figure 5.2: Soil fertility and soil texture classifications in the study area. Source: Adu, 1969

of topsoil, soil structure, and permeability. Since the combined effect of these attributes is a better explanatory factor for crop productivity than the sum of these single attributes, we decided to represent this factor by general soil fertility classes (as a rank from 1 to 5). Furthermore, since soil texture seemed to play a special role in the choice of land-use type and crop productivity, especially in the dry season, we decided to treat this attribute as a separate variable. Spatial data of soil texture and soil fertility were generated using soil maps and information from Adu (1969) (see section 2.5.2 for details).

Topographical factors

It is well known that the terrain regulates the flow of surface runoff and soil particles, thereby strongly determining the landscape patterns of soil and water conditions (Gessler et al., 2000). Numerous studies have shown how the shape of the land surface can affect the lateral migration and accumulation of water, sediments, and other constituents (e.g., Wilson and Gallant, 2000). These constituents, in turn, influence soil development (e.g. Kreznor et al., 1989), and exert a strong influence on the spatial and temporal distributions of light, heat, water, and mineral nutrients required by photosynthesizing plants (Wilson and Gallant, 2000).

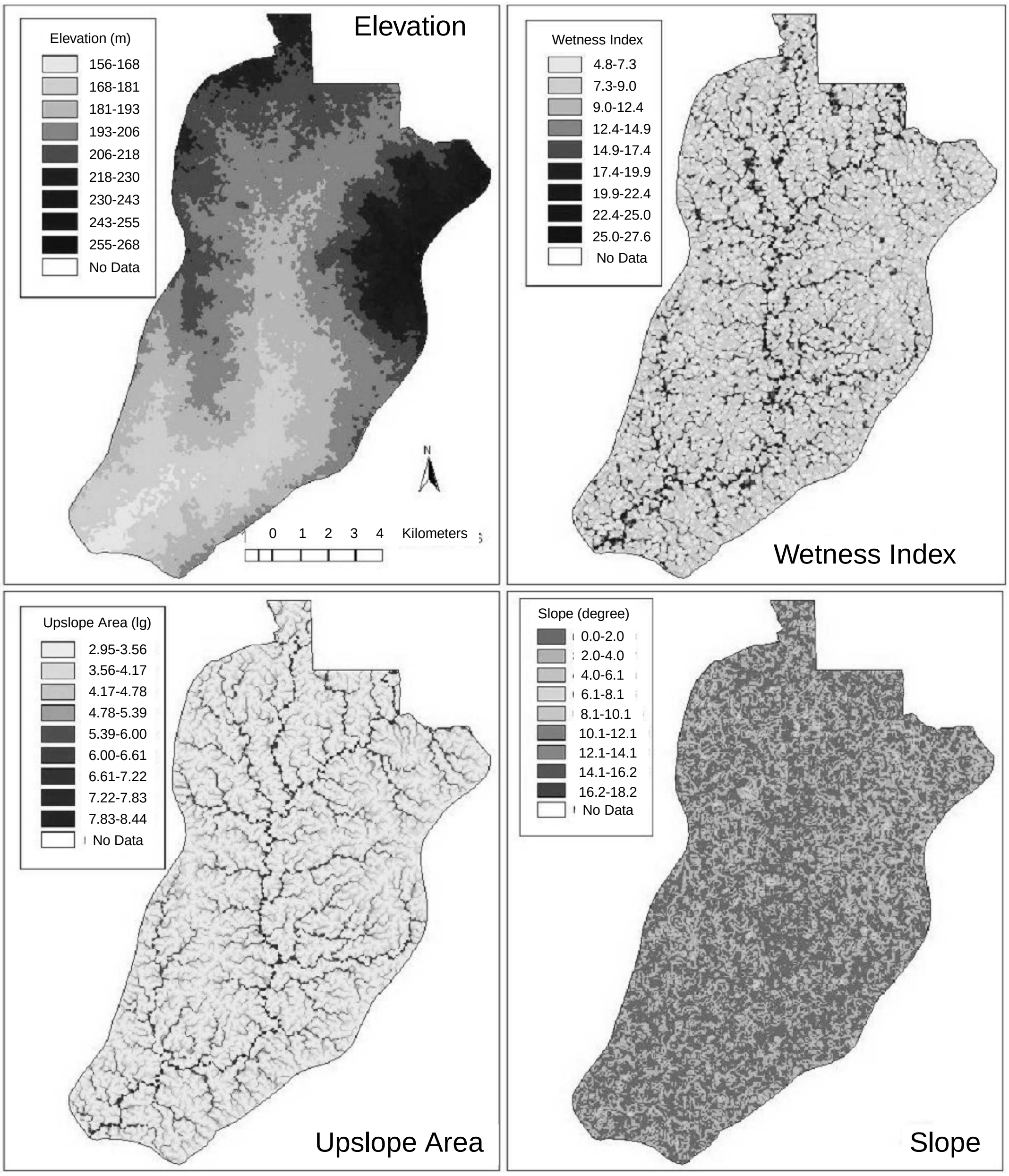


Figure 5.3: Topographic attributes of the study area

The formation of soils induced by topography refers to the concept of catena, which describes the sequence of soils along hill slopes. The catenary hypothesis is that soil development occurs in many landscapes in response to the way water moves through and over the landscape. Furthermore, terrain attributes can characterize these flow paths and, ultimately, soil attributes. Soil properties such as soil depth (Gessler et al., 2000; Park et al., 2001), pH, organic matter content and soil moisture content (Wilson and Gallant, 2000) have been shown to be dependent on terrain factors. The catena principle, together with available topographical data, has been widely used in modern soil survey techniques (e.g. Dobos, 2005; Sobieraj et al., 2004). The basic terrain factors to represent topography used in this study comprise elevation, slope degree, and upslope contributing area, which is defined as the total drainage area of the catchment above a certain point on the landscape. Furthermore, a wetness index was derived from these data, representing the spatial patterns of soil moisture content as a result of topographic surface flow, being calculated as:

$$P_{\text{wetness}} = ln\left(\frac{P_{\text{upslope}}}{tanP_{\text{slope}}}\right) \tag{5.1}$$

where P_{wetness} is the wetness index, P_{upslope} the upslope contributing area, and P_{slope} the slope gradient. The upslope contributing area (P_{upslope}) is defined as the total catchment area above a point on the landscape. For a grid cell P, P_{upslope} is computed from the grid cells from which the water flows into the cell P:

$$P_{\text{upslope}} = \frac{1}{b}\sum_{i=1}^{n} \rho_i A_i \tag{5.2}$$

where A_i is the area of grid cell P, n is the number of cells draining into the cell P, ρ_i is the weight depending on the runoff generation mechanism, and b is the contour width approximated by the cell size (Park et al., 2001). All topographical variables were calculated based on the digital elevation model by Le (2006) (see section 2.5.2).

Groundwater

The final component of the soil-water factor is represented by groundwater variables, since a wetness index alone does not describe water availability sufficiently, especially in the dry season, where rainfall plays a minimal role. Water stored from rainy-season rainfall as groundwater plays a distinct role in dry-season irrigation farming in areas where access to dams is limited. To represent this factor in an appropriate way, the following two variables were included: i) the average seasonal groundwater level, as it defines the area where groundwater can be accessed through digging, and ii) the average seasonal groundwater recharge. The latter variable has been included since it describes the water table balance of the groundwater. Spatial data on groundwater table and recharge were derived from Martin (2005) (see section 2.5.2)

Spatial accessibility

Spatial accessibility can be defined as the ease with which a target location may be reached from another location. Variables determining spatial accessibility are often key variables when modeling land-use choice, as they define the spatial variations in required patch attributes when making land-use decisions. Proxy variables that were found to play a significant role include distances to water sources (i.e. dams and the main river) and the distance to the national border. Distances to other features such as roads and local/main markets were neither statistically significant in modeling land-use choice, nor did they play a role for land-use choice according to local estimation. On the other hand, the distance of a plot to water sources such as dams or rivers can be regarded as an important proxy variable within the study area, since the decision for irrigation farming on a patch is highly dependent on this distance, as most of the irrigation activity is confined to areas along the main river and around dams.

This factor of spatial accessibility to water bodies is represented by the variable distance to water sources ($P_{\text{dist water}}$), which is calculated as the minimum distance from the considered pixel to water sources, including dams and the main river. Furthermore, the distance to the Ghana-Burkina border was another important proxy factor, as the land-use pattern varied strongly along the axis from the border in the north to the southern part of the catchment, which was the more active area with respect to irrigation farming and other activities. Due to lower soil fertility and lower water availability in the northern part, the area was less populated and farming was rather focused on subsistence crops, whereas in the densely populated southern part cash cropping was more abundant.

Features of the dams and main river were digitized using a Quickbird image, which had been taken in early 2006. The Ghana-Burkina border was extracted from national map (1:50000). Distance maps to these features were finally generated using the find distance routine in ArcView GIS 3.2. Distances to nearest dams and the main river ($P_{\text{dist water}}$) were

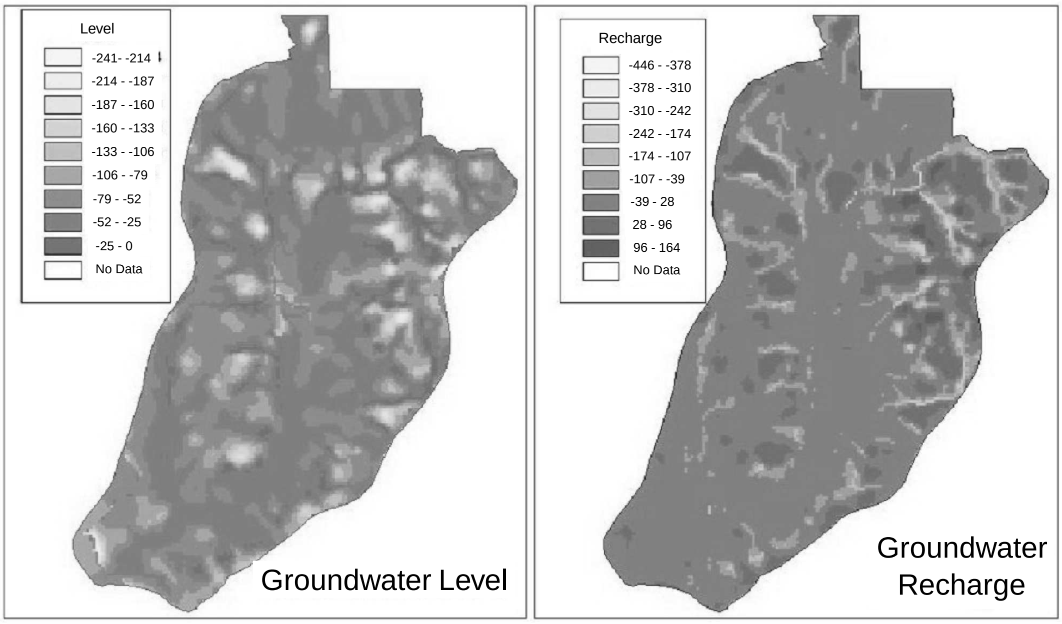


Figure 5.4: Groundwater level and recharge in the study area. Source: Martin, 2005

automatically calculated using NetLogo.

Analysis of land suitability for irrigation

In this section, a land suitability analysis with respect to irrigation will be presented for the study area. The final target is to define the irrigable area as closely as possible, as this parameter is essential in modeling the maximal extent of dry-season cultivation activities. According to the FAO Bulletin for Land Evaluation For Irrigated Agriculture (FAO, 1985), the environmental attributes explaining irrigability include topography, soil, water resources, climate, and drainage. Out of these categories, a range of parameters needed to be identified that were explanatory factors for irrigability in the study area. In the first part of this section, we will present and justify the range of selected variables. In the second part, we will present a model for the determination of irrigability based on these parameters. This model calculates an irrigation coefficient between 0 and 1 for each landscape agent, with the value of 1 indicating highest possible irrigability. Thus, a threshold between 0 and 1 for this irrigation coefficient needed to be chosen to define the final extent of the irrigable area. This threshold will be determined by analysis in the third part of this section.

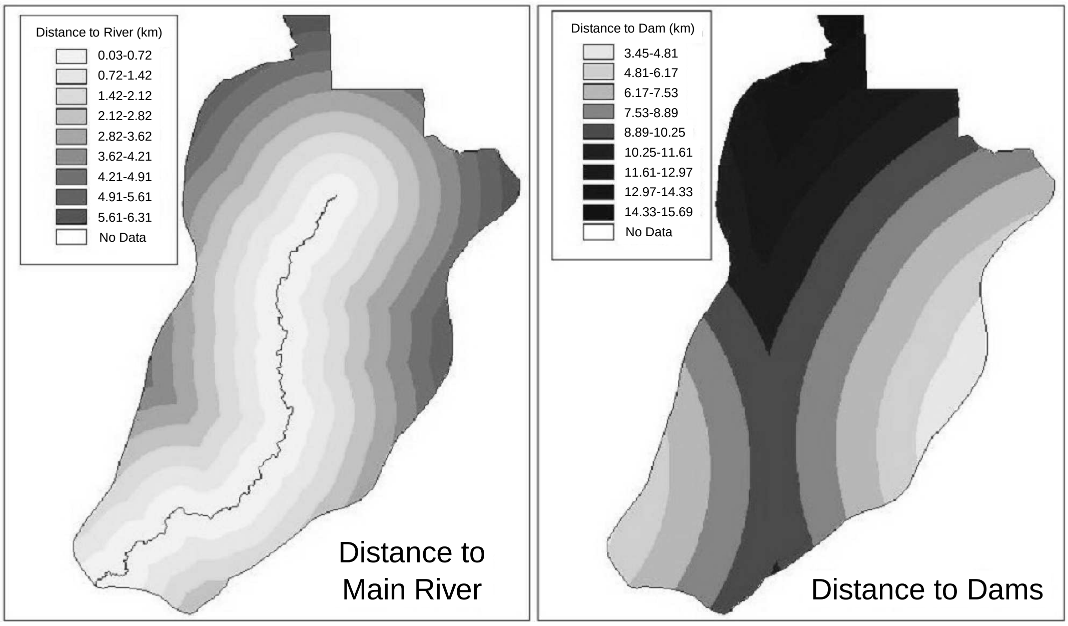


Figure 5.5: Spatial accessibility to water sources in the study area

Model of the irrigation coefficient

Range of variables

As no data were available about the extent and pattern of the irrigable area in the study area, we decided to develop a model simulating the irrigability of the landscape. According to the FAO Land Evaluation Bulletin for Irrigation (FAO, 1985), we chose a range of indicators from each explanatory category that seemed to be responsible for the pattern of the irrigated area in the study area. Factors representing climatic patterns have not been included in the analysis, due to the assumption that climate is uniform over the study area.

According to the FAO study, the topographic features influencing irrigability include slope gradient and position; the latter is defined by elevation and distance to water sources. Higher slope gradients usually limit the irrigation possibilities, but since the topography of the study area can be regarded as quite smooth, this factor should not play a role as a factor limiting irrigation. Instead, the position in relation to command area and accessibility is considered to play a decisive role, as elevation and distance of the water source often affects the irrigable land area in irrigation schemes (FAO, 1985). Thus, the distance of the patch to the main river, as well as its elevation, were included in the analysis.

Furthermore, soil attributes with respect to water-holding capacity had to be considered. Water-holding capacity is controlled primarily by soil texture and organic matter (Ball, 2001). Soils with a high percentage of silt and clay particles have a higher water-holding capacity. Furthermore, organic matter content is related to water-holding capacity in a positive way, i.e. the higher organic matter content usually results in a higher water-holding capacity because of the affinity organic matter has for water. Since data about organic matter contents were not available, we only included the parameter of soil texture in the analysis to represent irrigation-relevant soil attributes.

Third, as the component of water resources had also to be taken into account, two parameters defining groundwater availability have been included in the analysis: The average dry-season groundwater level, and the average dry-season groundwater recharge (see section 5.3.1). Furthermore, as groundwater level alone does not define the availability of water to the plant, the topographic wetness index was further included in the analysis to represent the inherent soil moisture of the soil due to topography.

Modeling the irrigation coefficient

For calculating the irrigation coefficient, first an m-logit model was developed to calculate the probability of a patch to be irrigated. The model is based on the empirical patch-based data set, including both irrigated and non-irrigated plots, together with a set of patch values of the range of explanatory variables as outlined above. Based on these empirical data, the model calculates the probability of a patch to be irrigated, with values between 0 and 1. The calculation of this probability $Prob_{irr}$ can be expressed as:

Table 5.2: Variables for explaining irrigability

Variable	Definition
$P_{elevation}$	Elevation (in m)
$P_{soil\ texture}$	Soil texture represented the rank of textural class (as a range from 1 - 21)
$P_{dist\ river}$	Distance to main river (in m)
$P_{wetness}$	Wetness Index, i.e. $\ln(P_{upslope}/ \tan P_{slope})$
P_{gwl}	Groundwater level (m below ground)
P_{gwr}	Groundwater recharge (mm/month)

$$Prob_{irr} = \alpha + \beta_1 \cdot P_{\text{elevation}} + \beta_2 \cdot P_{\text{dist river}} + \beta_3 \cdot P_{\text{soil texture}} + \\ + \beta_4 \cdot P_{\text{gwl}} + \beta_5 \cdot P_{\text{gwr}} + \beta_6 \cdot P_{\text{wetness}} \tag{5.3}$$

where α is a constant, P_i the explanatory variables, and β_i coefficients calculated by running SPSS. In Tables 5.3, 5.4 and 5.5, the results of the m-logit model are shown, under the assumption that a plot is irrigated when the probability is > 0.5. Comparing the observed to predicted variable of irrigation, among the actually irrigated patches 71.0 % are correctly predicted.

Further, we define the irrigation coefficient $P_{\text{irr coeff}}$ as the probability $Prob_{irr}$ for all patches of the landscape, i.e. using the coefficients as calculated above (Table 5.3), $P_{\text{irr coeff}}$ is calculated in the following way:

$$P_{\text{irr coeff}} = \alpha + \beta_1 \cdot P_{\text{elevation}} + \beta_2 \cdot P_{\text{dist river}} + \beta_3 \cdot P_{\text{soil texture}} + \\ + \beta_4 \cdot P_{\text{gwl}} + \beta_5 \cdot P_{\text{gwr}} + \beta_6 \cdot P_{\text{wetness}} \tag{5.4}$$

where P_i are the explanatory variables and β_i the coefficents calculated by SPSS above. This equation was used in GH-LUDAS to calculate the spatial distribution of the irrigation coefficient as defined. Naturally, all variables apart from the groundwater-related P_{gwl} and P_{gwr} variables are static, but due to the lack of a temporal hydrological groundwater model, these two variables were also considered as static.

Determination of the irrigable area

The threshold for the irrigation coefficient had to be set such that the area with values above this threshold matched the actual size of irrigable area within the catchment. The actual irrigable area can be partitioned into: the actual cultivated area during the dry-season, and ii) irrigable area not yet opened up. Thus, the size of the irrigable area can be regarded as the sum of irrigated area and irrigable area not yet developed.

To define the actually cultivated area, the irrigated area of those households that had been selected randomly from the different villages was summed up and upscaled thus

Table 5.3: Modeling irrigation of patches: parameter estimates

Not Irrigated	B	Std. Error	Wald	df	Sig.	Exp(B)	95 % Confidence Interval for Exp(B) Lower Bound	95 % Confidence Interval for Exp(B) Upper Bound
Intercept	- 8.529	4.034	4.470	1	0.034			
$P_{wetness}$	- 0.088	0.038	5.257	1	0.022	0.916	0.850	0.987
$P_{elevation}$	0.036	0.022	2.715	1	0.099	1.037	0.993	1.082
$P_{dist\ river}$	3.718	0.724	26.387	1	0.000	41.178	9.968	170.116
$P_{soil\ texture}$	0.706	0.270	6.847	1	0.009	2.026	1.194	3.439
P_{gwl}	- 0.002	0.011	0.035	1	0.852	0.998	0.976	1.020
P_{gwr}	- 0.015	0.012	1.415	1	0.234	0.985	0.962	1.010

The reference category is irrigated

Table 5.4: Modeling irrigation of patches: correct predictions

Observed	Predicted: Not Irrigated	Predicted: Irrigated	Predicted: Percent Correct
Not Irrigated	564	11	98.1 %
Irrigated	18	44	71.0 %
Overall Percentage	91.4 &	8.6 %	95.4 %

Table 5.5: Modeling irrigation of patches: statistics

Model Fitting Information: Chi-Square	df	Sig.	Pseudo R Square: Cox and Snell	Nagel-kerke	Mc Fadden
230.79	6	0.000	0.304	0.644	0.568

that it represented the total irrigated area of the whole catchment population. To determine the irrigable area not yet opened up, we followed the assumption that the maximum number of farmers involved in irrigation is only constrained by the availability of suitable land. It was observed that more farmers are inherently capable of dry-season farming than farmers actually doing it, mostly due to limitations in land availability. Therefore, the number of irrigation farmers was assumed to converge against a certain limit during time, according to the availability of irrigable land. This upper limit of farmers who can do irrigation farming is then proportional to the irrigable area. In mathematical terms, this relationship can be expressed as:

$$\frac{\text{Irrigated Area}}{\text{Irrigable Area}} = \frac{\text{Farmers doing irrigation}}{\text{upper limit of farmers doing irrigation}} \tag{5.5}$$

With help of this equation, the amount of irrigable land can be calculated if the upper limit of irrigation farmers can be determined. To derive this upper limit, the number of farmers doing irrigation from the empirical data set was plotted against time (Figure 5.6).

To approximate these data by a curve, a function had to be selected with a minimal error to the observed data. This error is usually represented by the R Square, which is the square of the correlation coefficient between observed and fitted data. To identify such a curve with maximal R Square, 150 model types were tested for their R Square using the XLfit Extension of Excel. Finally, the curve with maximum R Square (R = 0.999023) was selected, called the Richards Function (see Figure 5.6). The mathematic expression of this function is:

$$Richards(t) = fracA((1 + e^{(B-(C\cdot t))})^{(}\frac{1}{D})) \tag{5.6}$$

where A, B, C, D are constants calculated by XLfit, and t is the time. To derive the upper limit of farmers possibly doing irrigation, the limit for this function had to be determined: For $t \rightarrow \infty$, the term $e^{(B-(C\cdot t))}$ converges to 0. Thus, the limes of the function can be determined as follows:

$$lim_{x\rightarrow\infty}Richards(x) = lim_{x\rightarrow\infty}fracA((1 + e^{(B-(C\cdot x))})^{(}\frac{1}{D})) = \frac{A}{1+0} = A \tag{5.7}$$

Thus, the irrigable area can now be calculated as:

$$\text{Irrigable Area} = \frac{\text{Irrigated Area} \cdot A}{\text{Farmers doing irrigation}} \tag{5.8}$$

Based on this calculation, the irrigable area in the study area amounts to 291 ha. The threshold of the irrigation coefficient to define irrigability within the model was then set to match this number.

5.2.2 Modeling agricultural yield response

Decision-making processes in agriculture often require reliable crop response models to assess the impact of specific land management (Park and Vlek, 2002). There are two distinct

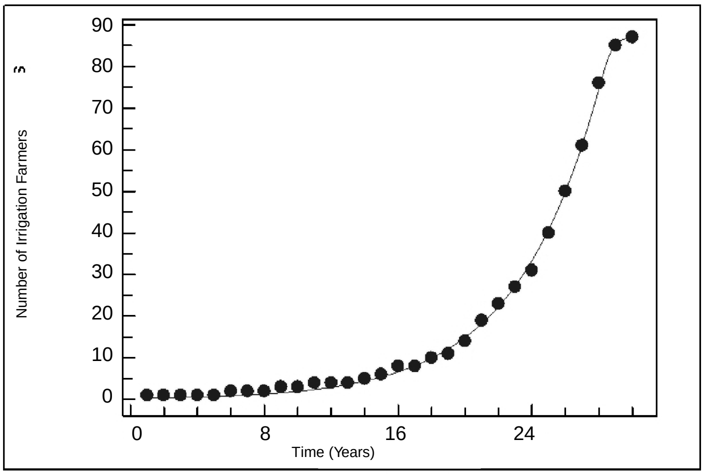


Figure 5.6: Curve estimation of the number of irrigation farmers

modeling approaches, i.e. empirical and process models, for identifying crop yield responses (Jame and Cutforth, 1996). Process-based crop growth models are built using mathematical equations to model quantitatively plant-soil-atmospheric interactions (Sinclair and Seligman, 1996; Matthews 2002). Because process models explicitly include plant physiology, agro-climatic conditions and biochemical processes, these models are supposed to be able to simulate both temporal and spatial dynamics of crop yields. Empirical models, on the other hand, attempt to determine functional relationships between crop yield and soil-land management factors using regression or correlation analysis to characterize these relationships statistically. Technologically, empirical crop growth models are relatively simple to build or develop, but these models - in contrast to process-based models - cannot take into account temporal changes in crop yields without long-term experiments (Jame and Cutforth, 1996).

While process-based models are often preferred over empirical ones in current modeling communities, empirical crop growth models still play an important role in identifying the hidden structure of crop growth processes relating to a wide range of land management options (Park and Vlek, 2002). Furthermore, process-based models require a high level of technological sophistication and calibration-verification procedures, which are limiting factors for a wider application (Sinclair and Seligman, 1996; Stephens and Middleton, 2002). The failure of many of these complex process-based crop models has, understandably, been ascribed to insufficient knowledge about the details and intricacies of the underlying physi-

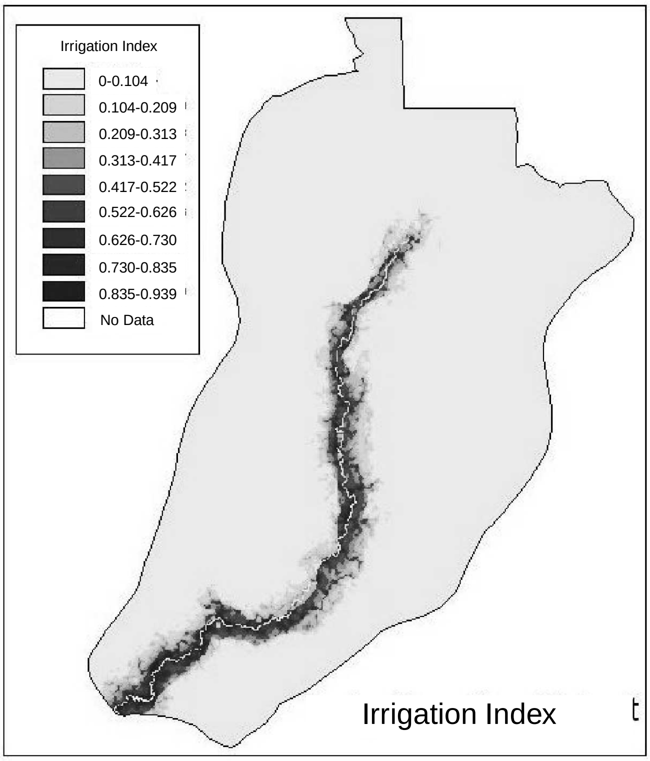


Figure 5.7: Spatial distribution of the irrigation coefficient in the study area

ological processes (Sinclair and Seligman, 1996). Naturally, these kinds of models require careful calibration and verification, which is especially problematic for developing countries, where the necessary technological and financial resources are not readily available (Stephens and Middleton, 2002). Consequently, parameterization often comes from previous research conducted in different environmental conditions or from expert opinion. The uncertainty associated with such parameterization may greatly decrease the validity of model outputs and the reliability of model application (Penning de Vries et al., 1989; Stephens and Middleton, 2002).

For this study, we selected the empirical approach to model land use productivity for three reasons. First, as our modeling scale consists of cultivation systems rather than of detailed crop varieties, it would have been unnecessarily complicated if the process-based approach had been applied. Second, as mentioned above, the calibration and verification of process-based models would require an understanding of the underlying processes and data, which are usually not available in developing countries. Third, since the main goal is the prediction of yield response rather than the understanding of the underlying processes, the approach of empirical models, which are usually more robust than process-based models, is the more straightforward one for our purposes.

Methodology

Among empirical models, three major approaches have been used to predict crop yield response in agricultural science: Linear Multiple Regression (LMR), Regression Trees (RT), and Artificial Neural Networks (ANN) (Park et al., 2005). Comparisons of the goodness-of-fit of these three approaches applied to maize yield responses in eastern Uganda can be found in Park et al. (2005). Although regression trees seem to be a quite robust model, they clearly have some drawbacks. They usually need a large data base to be reliable, as they only categorize the observed yield data according to the different explanatory factors. Furthermore, due to the use of a categorizing approach, their predictive power is low for input and yield values that lie outside the observed data range (White, 1996). Finally, the difficulty in interpreting the causal relationships is a clear drawback for the application of regression trees (Park et al., 2005). The same is valid for artificial neural networks, as these also require a large sample set and also tend to work as a black box. These latter two approaches also certainly have their strengths, but as we are not only interested in predicting crop yields, but also in interpreting the relationships between explanatory factors and yield response, we decided to apply the linear multiple regression approach, which allows such interpretations. Furthermore, the methods regression trees and artificial neural networks require a large data set, which is not given in our study, as we had to separate the yield data set into several land-use type specific samples.

The general purpose of linear multiple regression is to quantify the relationship between several independent or predictor variables and a dependent or criterion variable (in our case yield response) by using linear combinations. Furthermore, additional terms of the interactions among the predictor variables can be included in the model of crop yield response, as one might easily anticipate that soil and land management variables are highly correlated (Park et al., 2005). This way, the model can be depicted mathematically as:

$$P_{Yield} = \alpha + \sum_{i=1}^{k} \beta_i \cdot X_i + \sum_{i=1}^{k} \sum_{j>i}^{k} \beta_{ij} \cdot X_i \cdot X_j \tag{5.9}$$

where α, β_i and β_{ij} are coefficients as calculated by the linear regression procedure, and X_i the predictor variables suggested to explain crop yield response. The last term of this equation represents the interactions among the predictive variables. The great advantage of this approach is that it can take into account not only the relationships of the predictor variables with the dependent variable, but also the relationships among the multiple independent variables.

However, a purely linear relationship between predictor variables and yield is unrealistic in most cases. Instead, it is more intuitive that the yield follows a logarithmic or convergent curve in response to the explanatory variables, as there is a certain limit to agricultural output, even if input factors and biophysical suitability increase continuously. The most common approaches to generate such non-linear relationships include the use of the logarithimic, square root, and reciprocal functions (see Griffin et al., 1987). The advantage of these functions is that they still allow the use of linear regression techniques. For example, by using the logarithmic approach, linear regression tries to identify a linear relationship between the logarithm of the output variable, i.e. yield, and the logarithms of the explanatory factors. Although we have a linear relationship among the logarithmized variables, the relationship between the plain variables result in a logarithmic function. As such, the productivity function based on logarithms can be mathematically expressed as:

$$Ln(P_{Yield}) = \alpha + \sum_{i=1}^{k} \beta_i \cdot Ln(X_i) + \text{interaction factors} \tag{5.10}$$

where the interaction factors can either be products of the logarithmized or the plain variables, being $\sum_{i=1}^{k} \sum_{j>i}^{k} \beta_{ij} \cdot Ln(X_i) \cdot Ln(X_j)$ or $\sum_{i=1}^{k} \sum_{j>i}^{k} \beta_{ij} \cdot X_i \cdot X_j$, respectively. Without interactions, this function is also known as the logarithmized form of the Cobb-Douglas function, which is one of the most common functions used for predicting yield response (Griffin et al., 1987). If interactions are used, this form is known as the transcendental production function.

Accordingly, by replacing the logarithm by square roots, the square root function can mathematically be expressed as:

$$\sqrt{P_{Yield}} = \alpha + \sum_{i=1}^{k} \beta_i \cdot \sqrt{X_i} + \text{interaction factors} \tag{5.11}$$

where the interaction factors can again either be products of the square root of the variables or the plain variables, being $\sum_{i=1}^{k}\sum_{j>i}^{k}\beta_{ij}\cdot\sqrt{X_i}\cdot\sqrt{X_j}$ or $\sum_{i=1}^{k}\sum_{j>i}^{k}\beta_{ij}\cdot X_i\cdot X_j$, respectively (see Griffin et al., 1987).

Finally, the reciprocal function is expressed as:

$$\frac{1}{P_{Yield}} = \alpha + \sum_{i=1}^{k} \beta_i \cdot \frac{1}{X_i} + \text{interaction factors} \tag{5.12}$$

where the interaction factors can again either be products of the reciprocal variables or the plain variables, being $\sum_{i=1}^{k}\sum_{j>i}^{k}\beta_{ij}\cdot\frac{1}{X_i}\cdot\frac{1}{X_j}$ or $\sum_{i=1}^{k}\sum_{j>i}^{k}\beta_{ij}\cdot X_i\cdot X_j$, respectively. This type of function is usually called the modified resistance function (see Griffin et al., 1987).

General rules about which type of function to use and whether to use forms of interaction, do not exist. Rather, statistical analysis must be used to identify which functional form best fits the observed data. As such, we applied all variants of functional forms to the empirical data set in order to identify the form which best approximates the empirical yield data. The R Square, which is a common value to measure the goodness-of-fit of the respective fitted linear curve, is presented in Tables 5.7 and 5.10 for all these functional forms and for each land-use type. According to these values, we will then justify the choice of functional form.

Modeling dry-season yield response

The dependent variable of the yield response model is the total crop yield for each land-use type, but since each agricultural land-use type can include more than one crop, the harvests of crops were converted to monetary values, based on the average local prices of the year 2006.

Range of variables

Crop growth is an extremely complex process in both time and space. Changes in weather conditions influence soil moisture, root uptake and water- and temperature-related stress on plants. At the same time, different parts of the landscape experience different water availability and soil nutrient status because of pedological heterogeneity and lateral water-nutrient flows related to the shape of the terrain (Park and Vlek, 2002). Apart from that, the depletion and replenishment of soil nutrients over time and the site-specific land management (e.g labor input) lead to significant changes in crop yield. The agricultural yield of each land-use type can, therefore, be conceptually described as a function of climate conditions (C), soil/water conditions (SW), and land management practices (M):

$$P_{\text{yield-dry}} = f(C, SW, M) \tag{5.13}$$

Because of the relatively small size of the study area (about 159 km^2), is is reasonable to assume that the climate factor C is uniform over the study area. Furthermore, as no reliable data describing the relation among climate change and dry-season crop yield were available, this factor was also assumed to be constant over time.

The soil/water conditions (SW) of the patches can be approximated by the irrigation coefficient and soil fertility. The irrigation coefficient, which is calculated as a combination of soil attributes and water-related parameters, represents the factor of water availability with respect to the cultivation of irrigated crops. Soil fertility, on the other hand, represents a combination of soil-specific parameters important for crop yield. For the model of agricultural yield response, we decided to use these two coupled indices rather than a single biophysical variable, since previous studies showed that one single index alone does not always give a good representation of soil-water patterns (e.g. Western et al., 1999).

Among land management factors, labor input (in labor days) and input of agrochemicals (in Cedis) should be the prior variables for consideration, as these inputs directly

Table 5.6: Variables for predicting dry-season yield

Variable	Definition
I_{labor}	Input of labor (in labor days/m^2)
I_{chem}	Input of chemicals (Cedis/m^2)
$P_{\text{soil fertility}}$	Soil fertility (as a range from 1 to 5)
$P_{\text{irr coeff}}$	The irrigation coefficient (between 0 and 1)

influence plant growth. It is common knowledge that tomatoes respond well to fertilizer applications, especially nitrogen and phosphorus. However, the sensitivity of crop yield to these factors may be different among the two land-use types, depending on the nature of each land-use type and actual natural conditions. The instant values of labor and chemical input are determined by household agents, whose behavior is governed by the Decision Module.

Thus, the productivity function modeling dry-season yield can be formally expressed as:

$$P_{\text{yield-dry}} = f(P_{\text{irr coeff}}, P_{\text{soil fertility}}, I_{\text{chem}}, I_{\text{labor}}) \tag{5.14}$$

where $P_{\text{irr coeff}}$ is the irrigation coefficient, $P_{\text{soil fertility}}$ the soil fertility, I_{chem} the amount of agro-chemicals, and I_{labor} the total amount of labor input.

Model choice and results

Based on this range of variables, all functional forms were tested on their respective R Square for both land-use types (see Table 5.7). The logarithmic function with plain interaction terms shows the best results for both land-use types. Therefore, we selected this functional form for predicting dry-season yield based on the selected explanatory variables as described above. This way, the mathematical expression of the function is as follows:

$$Ln(P_{\text{yield-dry}}) = \alpha + \sum_{i=1}^{k} \beta_i \cdot Ln(X_i) + \sum_{i=1}^{k} \sum_{j>i}^{k} \beta_{ij} \cdot X_i \cdot X_j \tag{5.15}$$

where X_i are the explanatory variables, the β_i their respective coefficients, and α a constant, both calculated by linear regression using SPSS. The values of these coefficients indicate that many of the explanatory variables are highly correlated to yield response (Table 5.8). The basic factors $Ln(X_i)$ are significant at levels 0.1, 0.05 and 0.01. The input variables of labor and chemicals are positively related to yield response, i.e. the higher these inputs, the higher the resulting yield (although there is certainly a limit). Interesting is the fact that the irrigation coefficient is negatively related to crop yield, i.e. the higher the water availability, the lower the crop yields. The reason might be that poorly drained soils with little organic matter and high clay content, as is the case in the study area, can cause a yield decline in response to

Table 5.7: R square values for functional forms for predicting dry-season yield

	Dry-Season Land-Use Types	
	Monoculture of Tomatoes	Mixed Culture of Tomatoes
Linear		
Without Interaction Terms	0.551	0.623
With Plain Interaction Terms	0.669	0.671
Logarithmic		
Without Interaction Terms	0.452	0.782
With Plain Interaction Terms	0.661	0.967
With Log. Interaction Terms	0.554	0.842
Square Root		
Without Interaction Terms	0.570	0.665
With Plain Interaction Terms	0.615	0.865
With Sqrt. Interaction Terms	0.642	0.787
Reciprocal		
Without Interaction Terms	0.548	0.955
With Plain Interaction Terms	0.612	0.965
With Recipr. Interaction Terms	0.623	0.956

overflooding.

Modeling rainy-season yield response

Equivalent to the modeling process of dry-season yield response, in this section we will outline and justify the range of explanatory variables, the choice of model for yield prediction, and finally the results. The dependent variable of the model is land-use type specific yield response per square meter, while the yield of the single crops of each land-use type is converted to its monetary value, according to average local prices in 2006.

Range of variables

For the choice of the range of explanatory variables for rainy-season yield, we applied the same approach as for the dry season: Thus, the yield $P_{\text{yield rainy}}$ of the rainy-season land-use types can be formally expressed as a function of climate (C), soil/water conditions (SW) and management (M):

Table 5.8: Predicting dry-season yield: parameter estimates

	Dry-Season Land-Use Types	
Variables	Monoculture of Tomatoes	Mixed Culture of Tomatoes
Constant	-3.025***	56.619**
I_{labor} (log)	0.641**	1.786**
I_{chem} (log)	1.975***	1.211**
$P_{soil\ fertility}$ (log)	2.345	- 54.041**
$P_{irr\ coeff}$ (log)	- 0.593**	- 22.278*
$I_{labor} \cdot I_{chem}$	0.017	0.012*
$I_{labor} \cdot P_{soil\ fertility}$	- 7.171**	- 9.908
$I_{labor} \cdot P_{irr\ coeff}$	25.421**	28.673
$I_{chem} \cdot P_{soil\ fertility}$	- 0.001	- 0.004
$I_{chem} \cdot P_{irr\ coeff}$	- 0.003*	0.012
$P_{irr\ coeff} \cdot P_{soil\ fertility}$	0.062	6.131*
Size of training data set	46	24
Size of testing data set	15	15
R Square	0.661	0.967
RMSE	4.255	6.504
CV (RMSE)	0.0398	0.0308

$$P_{\text{yield-rainy}} = f(C, SW, M) \tag{5.16}$$

where the climate C is regarded as being constant in space, due to the relatively small size of the study area, but variable in time. Compared to the dry season, the explanatory variables representing the soil-water factor (SW) and the management factor (M) are naturally different in the rainy season, and have to be selected carefully with respect to the conditions and needs of rainy-season cultivation.

As such, the water availability required for proper plant growth in the rainy season is more dependent on rainfall than on some kind of irrigation coefficient representing groundwater availability. Parameters describing both the spatial and temporal variation in water availability due to rainfall need to be considered. The temporal variation in rainfall is represented by the annual future rainfall as simulated by the Intergovernmental Panel on Climate Change (IPCC) for the study area. The spatial variation in water availability due to rainfall is mainly due to the topographical pattern of the area, with runoff and slope gradients playing

Table 5.9: Variables for predicting rainy-season yield

Variable	Definition
I_{labor}	Input of labor (in labor days/m^2)
I_{manure}	Input of manure (Livestock Index/m^2)
$P_{soil\ fertility}$	Soil fertility (as a range from 1 to 5)
$P_{wetness}$	The wetness index, i.e. $\ln(P_{uslope}/ \tan P_{slope})$

a major role in water accumulation within the soil. In this study, we chose the topographic wetness index to represent this factor of topographical water accumulation. Furthermore, in order to consider not only the spatial variation in water availability, but also the variation in soil suitability, we included further the discrete variable of soil fertility in the analysis.

With respect to the factor of agricultural management, agricultural labor input plays a major role in successful cultivation, which includes land preparation, plowing, sowing and weeding. It is a natural assumption that an increase in these cultivation efforts has a positive impact on plant growth. Thus, the variable of total labor input, measured in labor days per square meter, was included as an explanatory management factor for crop yield response. Furthermore, the same as for the dry season, the enhancement of soil fertility through agricultural measures also plays a major role for crop yield response. In contrast to the dry season, the use of chemicals and fertilizers for rainfed cultivation in the region is minimal. Instead, animal manure is widely used to enhance soil fertility. As the exact amount of animal manure was difficult to measure, this factor is represented by the livestock index of the household divided into fractions according to the sizes of the plots that were indicated to obtain manure during the survey. The input of manure was then defined as livestock index per square meter.

Thus, the productivity function modeling dry-season yield can be expressed as:

$$P_{\text{yield-rainy}} = f(P_{\text{wetness}}, P_{\text{soil fertility}}, I_{\text{manure}}, I_{\text{labor}}, R) \tag{5.17}$$

where $P_{wetness}$ is the wetness index, $P_{soil\ fertility}$ the soil fertility, I_{manure} the input of manure, I_{labor} the input of labor, and R the annual average rainfall (in mm/m^2) as simulated by IPCC.

Model choice and results

The first step of modeling rainy-season yield response consists of the development of a spatial yield-response model based on data of the year 2006, without considering rainfall data (as these are considered to be spatially constant), while in the second step the timely fashion of crop productivity will be modeled in response to annual average rainfall. In order to select a functional form for the spatial yield model for the year 2006, the R Square for each functional form and land-use type was calculated (see Table 5.10), where the land-use type soybeans/potatoes was omitted due to its small sample size (10 plots). Instead, the yield for this land-use type was set constant at the mean crop yield level. It is obvious that the inclusion of interaction terms enhances the predictive power for all functional forms and land-use types (Table 5.10). However, there is a high variation of the R Square among the various land-use types for most of the functions, with almost all forms having one R Square below 0.2. Therefore, and in order to be consistent with the model for the dry season, we selected the functional form that had the most even distribution of R Squares among the land-use types with all values above 0.2, namely the functional function based on logarithms (see equation 5.10) with plain interaction terms, which is also called the transcendental production function.

As the input of manure I_{manure} had an empirical value of 0 for many of the cases, the logarithm could not be taken of this variable. Instead, it was embedded in the function in a linear way. Furthermore, the variable of $P_{wetness}$ was already in a logarithmic form, therefore no logarithm is taken of this variable. The results of the linear regression indicate that some of the basic variables are significant in explaining crop yield response (Table 5.11). Labor input, soil fertility, and wetness index are all positively related to crop yield for all land-use types, indicating that the higher the labor input, water availability and soil fertility, the higher the corresponding crop yield. The input of manure is also positively related to crop yield for all land-use types apart from monocultures of cereals. A reason for this negative relation could be an over-fertilization of this land-use type through manure application, as monocultures of cereals, which are usually grown along the river banks, already receive large amounts of nutrients through seasonal flooding. For further convenience, we will call the yield calculated by these factors the spatial yield ${}^{spatial}P_{yield\ rainy}$:

Table 5.10: R square of functional forms for predicting rainy-season yield

	Rainy-Season Land-Use Types				
	Mono-culture of Cereals	Mono-culture of Groundnuts	Mixed Compound Farming	Rice based Culture	Mixed Groundnut Culture
Linear					
Without Interaction Terms	0.243	0.119	0.149	0.141	0.243
With Plain Interaction Terms	0.276	0.156	0.157	0.179	0.261
Logarithmic					
Without Interaction Terms	0.373	0.169	0.158	0.250	0.315
With Plain Interaction Terms	0.456	0.215	0.220	0.264	0.321
With Log. Interaction Terms	0.579	0.228	0.191	0.272	0.318
Square Root					
Without Interaction Terms	0.392	0.155	0.188	0.213	0.287
With Plain Interaction Terms	0.413	0.170	0.215	0.223	0.302
With Sqrt. Interaction Terms	0.452	0.189	0.203	0.235	0.296
Reciprocal					
Without Interaction Terms	0.165	0.093	0.346	0.087	0.243
With Plain Interaction Terms	0.217	0.133	0.443	0.187	0.262
With Recipr. Interaction Terms	0.682	0.198	0.465	0.187	0.262

$$^{\text{spatial}}P_{\text{yield rainy}} = \text{Cobb-Douglas}(P_{\text{wetness}}, P_{\text{soil fertility}}, I_{\text{manure}}, I_{\text{labor}}) \tag{5.18}$$

In order to include the temporal effects of climate change on rainy-season crop yield, in specific changes in annual rainfall, we used a correction factor that modifies the annual crop yield as calculated by the transcendental production function. Many studies suggest a linear relationship between crop yield and rainfall (see Vossen, 1988; Sicot, 1989; Ellis and Galvin, 1994; Larsson, 1996). As such, Groten (1991) identified a relationship between crop yield (in kg/ha) for millet in Burkina Faso and annual rainfall (in mm), being expressed as:

$$Crop_Y = 0.91 \cdot R \tag{5.19}$$

where $Crop_Y$ is crop yield, and R the amount of annual rainfall. This suggests that crop yield can be generally described as being directly proportional to annual average rainfall, although

Table 5.11: Predicting rainy-season yield: parameter estimates

	Rainy-Season Land-Use Type				
Variable	Mono-culture of Cereals	Mono-culture of Groundnuts	Mixed Compound Farming	Rice based Culture	Mixed Groundnut Culture
Constant	6.534**	4.540	2.974***	5.986**	8.306***
I_{labor} (log)	0.868**	0.339	0.210	0.613***	0.575 ***
I_{manure}	- 4.317	4.830	1.769***	–.–	0.117
$P_{wetness}$	0.625***	0.39	0.214**	0.181	0.084
$P_{soil\ fertility}$ (log)	0.461	2.823	1.729***	1.924	0.643
$P_{wetness} \cdot P_{soil\ fertility}$	- 0.124**	- 0.091*	- 0.044*	- 0.055	- 0.017
$I_{labor} \cdot P_{soil\ fertility}$	13.213	1.580	3.102	- 0.042	- 2.477
$I_{manure} \cdot P_{soil\ fertility}$	1.017*	- 0.703	- 0.227**	–.–	0.181
$I_{labor} \cdot I_{manure}$	- 34.907	- 26.897	- 6.096	–.–	5.518
$I_{labor} \cdot P_{wetness}$	- 4.887*	- 0.398	1.197	0.056	1.028
$I_{manure} \cdot P_{wetness}$	0.070	- 0.153	- 0.049***	- 0.012	- 0.034
Size of training data set	51	53	160	82	167
Size of testing data set	30	30	70	45	70
R Square	0.456	0.215	0.220	0.264	0.321
RMSE	1.145	0.710	0.959	1.176	0.754
CV (RMSE)	0.228	0.099	0.176	0.188	0.108

there is certainly a limit to the positive effect of rainfall on yield. But within a reasonable range of rainfall data, this linear relationship can be regarded as valid.

Since the empirical productivity functions were derived from yield and input data of the year 2006, these functions are based on the rainfall pattern in this specific year. However, due to the linear relationship between average annual rainfall and crop yield, the effect of rainfall of year *t* in relation to the year 0 (2006) can be expressed as:

$$P_{\text{yield-rainy}} = {}^{\text{spatial}}P_{\text{yield-rainy}} \cdot \frac{R^t}{R^0} \tag{5.20}$$

where R^t is the average annual rainfall in mm for the year t, R^0 the rainfall (in mm per year) for the year 0 (base year 2006). As such, an increase in rainfall by e.g. 20 % in relation to the base year would result in an increase in yield by 20 % if all other input factors remain constant. This is in accordance with the assumption of a linear relationship as suggested by the studies as mentioned above. With the help of this equation and the transcendental production function, the yield response for a specific year *t* can be calculated.

5.2.3 Modeling livestock dynamics

The model of livestock dynamics simulates the population of livestock within the study area, being expressed by the livestock index of local households $H_{livestock}$. The model is based on the following two assumptions: The annual decrease or increase in the livestock index is randomly dependent on the livestock index of the previous year, and the total number of all livestock must be below or equal to the carrying capacity of the study area with respect to forage availability.

The first assumption of a random dependence of the livestock index of two subsequent years can be expressed as:

$$^{t+1}H_{livestock}{}^{rand} = round(^{t}H_{livestock}{}^{rand} - \sigma_{livestock} + random(2 \cdot \sigma_{livestock})) \qquad (5.21)$$

where $^{t}H_{livestock}{}^{rand}$ is the randomized livestock index at time step t, and $\sigma_{livestock}$ the standard error of the empirical data set of the livestock index. By using this equation, the livestock index in the current year lies randomly within a range of $\pm\sigma_{livestock}$ around the livestock index of the previous year. For our purposes, this random approach is the most robust and straightforward method to model variations in livestock, as the stock of animals within a household is dependent on many different factors, which are difficult to model, such as birth and death rates, diseases, sale, or the delivery of animals as gifts for funerals.

However, regardless of the small variations of the stock of animals within a household, the upper limit or carrying capacity for livestock in a specific area can be regarded as a restricting factor for the whole animal population. This carrying capacity is directly dependent on the availability of natural resources, including water and forage; we will only take into account the forage availability, as no related studies could be found to reliably model water supply of local dams (which are the main source of water for animals). Stéphenne and Lambin (2001) provide a model for determining the relationship between livestock population and biomass production under different rainfall patterns. The related equation is expressed as follows:

$$BiomPy \cdot Past_d = Liv \cdot BiomC \tag{5.22}$$

where *BiomPy* is the biomass productivity in tonnes/ha, Pastd the pastural area in ha, *Liv* the livestock population in equivalent tropical livestock unit (TLU), and *BiomC* the consumption in biomass in tonnes/TLU. TLU is a conventional stock unit of a mature zebu weighing 250 kg (Boudet, 1975). One TLU corresponds to one cattle, one horse, five asses, 10 sheep or 10 goats. Following this equation, we can calculate the number of TLU the area can sustain under normal conditions if we know *BiomPy*, $Past_d$ and *BiomC*. According to Le Houérou and Hoste (1977), biomass productivity in Sudano-Sahelian grasslands highly depends on rainfall. This is described by the following statistical relationship between dry matter biomass and rainfall, taken from ground measurements by Breman and de Wit (1983):

$$BiomPy = 0.15 + 0.00375 \cdot R \tag{5.23}$$

where R is the annual average rainfall in mm of the current year. As future scenarios of variable rainfall data are fed into the model, *BiomPy* can change over time. $Past_d$, the area in ha of pastural land is calculated from the land-cover and land-use pattern of the current year. As it is common practice that the leaves of groundnuts are dried by local farmers for animal fodder, the area of forage productivity Pastd, does not only comprise patches with the land cover grassland, but also patches covered by groundnut-based land-use types. This area comprising both grassland and groundnut cultivation is updated in each time step of the model, thus also leading to a variable outcome of biomass productivity. However, and this is the major drawback of this model, the dietary requirement per TLU (*BiomC*) is regarded as being constant at 4.6 tonnes/year (see Stéphenne and Lambin, 2001). This assumption implies that under drought conditions, the biomass consumption per livestock unit does not decline. However, related literature did not provide estimations of consumption behavior of livestock in relation to drought pressure.

According to these values, the annual carrying capacity under normal conditions can be calculated in TLU for each year. If the number of animals (in TLU) exceeds the carrying capacity, the animal population will be reduced by a factor such that the population is equal to the carrying capacity. Following this mindset, we define the annual livestock index per household as a restriction to the variable of ${}^{t+1}H_{livestock}{}^{\mathrm{rand}}$ according to the carrying capacity of TLU and the current number of total TLU:

$$ {}^{t}H_{livestock} = \begin{cases} {}^{t}H_{livestock}{}^{\mathrm{rand}} & \text{if } {}^{t}CC \geq {}^{t}TLU_{total} \\ {}^{t}H_{livestock}{}^{\mathrm{rand}} \cdot \frac{{}^{t}CC}{{}^{t}TLU_{total}} & \text{if } {}^{t}CC < {}^{t}TLU_{total} \end{cases} \tag{5.24} $$

where ${}^{t}CC$ is the carrying capacity in TLU at time step t, and ${}^{t}TLU_{total}$ the total number of TLU in the study area at time step t. ${}^{t}TLU_{total}$ is calculated as the sum of TLU per household. In order to give reasonable figures for this number of TLU per household, we decided to set this number proportional to the livestock index, which is expressed as:

$$ {}^{t+1}H_{TLU} = \frac{{}^{t+1}H_{livestock}}{{}^{t}H_{livestock}} \cdot {}^{t}H_{TLU} \tag{5.25} $$

where ${}^{t}H_{TLU}$ is the number of TLU for the household in time step t. In order to solve this equation for all ${}^{t}H_{TLU}$, the initial value of ${}^{0}H_{TLU}$ is calculated from the empirical data set. This equation ensures that the number of TLU per household in each year reflects the livestock index in the respective year.

This model of livestock dynamics has two purposes. First, it calculates household livestock numbers (livestock index) in dependence on annual rainfall and land-use behavior, and second, it provides an estimation on whether the livestock carrying capacity of the study area is reached, thus giving an indicator of the possible danger of overgrazing. Overgrazing can be defined as grazing by a number of animals exceeding the carrying capacity of a given parcel of land. Although this model assumes that the carrying capacity is never exceeded by the total number of livestock, the model indicates that overgrazing is possible if the carrying capacity is reached.

5.2.4 Land-cover transformation model

This routine models the natural changes among land-cover types in both seasons, which are beyond of human control. The range of land-cover types comprises ’forest’, ’water’, ’bare land’, ’grassland’ and ’cropland’ (section 5.3.1), whereas the land cover of grassland is absent in the dry season, where the climatic conditions are such that grass does not survive or grow. For both seasons, the land-cover types of water and forest are modeled to be stable, i.e. they do not undergo any changes, as the small patchy remnants of forest remain traditionally untouched. The task is, therefore, to analyze the changes among grassland, bare land and cropland for both seasons.

As climatic conditions in the dry season hamper cultivation and natural grass growth, most of the area is covered by bare land, apart from the small irrigated patches, which are mostly located along the river banks. Therefore, the variable of dry-season land cover does not comprise the land-cover type of grassland, and is updated in each time step $t+1$ in the following way:

$$
{}^{t+1}P_{\text{cover dry}} = \begin{cases} \mathit{forest} & \text{if } {}^{t}P_{\text{cover dry}} \text{ was forest} \\ \mathit{water} & \text{if } {}^{t}P_{\text{cover dry}} \text{ was water} \\ \mathit{cropland} & \text{if the patch is used in time step } t+1 \text{ during the} \\ & \text{dry season} \\ \mathit{bare\ land} & \text{if the patch is neither covered by forest or water,} \\ & \text{and not used in time step } t+1 \text{ during the dry season} \end{cases} \tag{5.26}
$$

In the rainy season, the land cover of bare land usually covers patches that are not fertile enough to allow cultivation or grass growth. Therefore, a conversion mechanism from bare land to other land-cover types for the rainy season was not considered. Furthermore, the modeling of conversion of grassland or cropland to bare land through erosional and other processes was beyond the scope of this study. Thus, the land-cover type of bare land was considered as being stable within the model, i.e. it does not undergo any change (like forest and water).

This way, the land-cover transformation model in general only regulates the natural conversion between grassland and cropland. This way, two directions of conversion have to be accounted for: the conversion from grassland to cropland, and the conversion from cropland to grassland. The conversion from grassland to cropland is regulated by the Decision Module, in which a procedure allows the agent to use grass patches for cultivation under certain conditions, whereas the rule for the reverse direction of conversion is dependent on natural grass growth. It is assumed that if a patch has not been used for a certain period P neither in the rainy nor in the dry season, it will be steadily covered by grass, and thus be converted to the land-cover of grassland. This period P, i.e. the number of years P a patch needs to be covered by grass, was set to 1 through discussion with local experts and farmers. This way, the update of the variable of land cover for the rainy season is expressed in the following way:

$$
{}^{t+1}P_{\text{cover rainy}} = \begin{cases} \textit{forest} & \text{if } {}^{t}P_{\text{cover rainy}} \text{ was forest} \\ \textit{water} & \text{if } {}^{t}P_{\text{cover rainy}} \text{ was water} \\ \textit{cropland} & \text{if the patch is used in time step } t+1 \text{ during the} \\ & \text{rainy season} \\ \textit{bare land} & \text{if } {}^{t}P_{\text{cover rainy}} \text{ was bare land} \\ \textit{grassland} & \text{if the patch has not been used during the last two} \\ & \text{seasons} \end{cases} \tag{5.27}
$$

5.3 Summary

This chapter gave an overview on the biophysical conditions of the study area, determined the spatial pattern of these conditions, and developed specific biophysical sub-models operating in response to these conditions, land use and socio-economic indicators. The biophysical attributes considered include land cover (for both seasons), topographic attributes (e.g. elevation), proxy variables (e.g. distance to river), soil, and groundwater data. The spatial pattern of land cover was identified for both seasons, based on an ASTER image using unsupervised classification and ground truth data collected in the study area. The methodology and sources for the development of spatial maps for local soil-water conditions were presented, including the soil attributes of soil fertility and texture, groundwater level and recharge, and the topographic attributes of elevation, slope, upslope contributing area, and wetness index. Finally, variables of spatial accessibility were determined, including distances to main river, dams and the national border in the north. Justifications for the use of these variables were given in the respective sections.

A further spatial variable that had to be determined was the variable of irrigability $P_{\text{irrigable}}$, which required the development of a specific irrigability model, as corresponding data were not available. This model is based on a land-suitability analysis approach for irrigation as provided by a related study by the FAO (FAO, 1985). In accordance with this study, a range of factors was identified to explain irrigability in the study area and, based on these factors, an irrigation coefficient between 0 and 1 was calculated for each patch by using m-logit regression, with 1 indicating highest possible irrigability. A threshold for the irrigation coefficient to define a patch as irrigable was finally determined.

The role of biophysical sub-models was then to define the productivity of the various land-use types, to regulate the population of livestock, and to determine the conversion of one land-cover type to the other. As biophysical as well as abiotic factors played a role in local crop productivity, both household (e.g. manure, fertilizer and labor input) and environmental variables (e.g. soil fertility, wetness index) were included in the models for yield response. Different functional forms for predicting crop yield were tested, and the functional form with the highest R Squares for the different land-use types was selected, being the transcendental production function. The strength of this function is its ability to represent the combined effects of explanatory variables, as it integrates interaction terms between each pair of variables. Productivity of the land-cover type of grassland was further determined in order to calculate the carrying capacity of the livestock population, which served as a restriction factor for the model of livestock dynamics. Finally, the process of land-cover conversion in the study area was analyzed, and a respective update procedure for rainy-season as well as dry-season land-cover type for each patch was developed.

6 SCENARIO ASSESSMENT OF LAND-USE/COVER AND LIVELIHOOD CHANGES IN THE ATANKWIDI CATCHMENT

6.1 Introduction

In the face of a constantly changing world, proactive land management is needed to find successful strategies for mitigating the adverse impacts of LUCC, to avoid decisions with negative externalities on the human-environment system and to enhance the sustainability of the system's functioning. As it is widely acknowledged that damage once done to the environmental system is very difficult to undo, the far-reaching consequences of land-management decisions need to be assessed before measures are taken. A useful tool for providing a knowledge base for such informed decision-making in proactive land management and planning is the simulation-based assessment of the evolution of the coupled human-environment system in response to selected policy interventions. Based on this approach, a wide range of possible future outlooks can be generated, providing a basis for informed decision-making and discussion among policy-makers.

Traditional approaches designed to simulate the complex pathway of LUCC often lack this ability to reliably project alternative pathways of the human-environment system of land-use/cover change. This is partly due to the fact that many of these approaches are only capable of projecting one timeline into the future. For instance, statistic LUCC models are only able to project a single future timeline of land-use/cover patterns, as they are mostly based on transition probabilities extracted from observed historical data. Furthermore, the range of models available to explore future otulooks has been limited due to their inadequate representation of the human-environment interrelationships. At one extreme, some LUCC models tend to ignore the explicit roles of human actors in the changing of landscapes (Huigen, 2004; Veldkamp and Verburg, 2004). The weakness of this kind of models thus lies not only in the lack of an assessment of future socio-economic indicators, but also in the limited ability to represent the direct impact of policy interventions on human land-use behavior. At the other extreme, many bio-economic models tend to treat the human influence as the main driver of LUCC, and are thus weak in assessing the role of environmental impacts on human land-use behavior (Verburg et al., 2004). These models thus often ignore the direct links between environmental conditions and land-use-related interventions, thus limiting the ability to explore future impacts of policies on the environment.

Multi-agent-based models, on the other hand, have been recognized to be well suited to exhibit the co-evolution of the human and landscape systems based on the interactions between human actors and their environment. Furthermore, the linkages of policy interventions and other external environmental or socio-economic factors to the human as well as to the landscape system can be effectively designed, as the bottom-up approach of agent-based modeling allows the modeling of the direct consequences of policy interventions on household behavior and landscape attributes. GH-LUDAS in particular, was designed to explore future outlooks of LUCC and other socio-economic indicators as a consequence of selected policy options and other external factors.

The application of simulation-based scenarios is usually seen as a useful tool to identify the variety of such possible future outlooks and to understand the consequences of selected input parameters on the performance of the system. Scenarios are accounts or synopses of projected courses of action, events or situations, and are widely used to understand different ways that future events might unfold. Unlike classical predictions, scenarios are not necessarily accurate forecasts of single future timelines drawn on past data, but multiple possible future pathways of the system evolution under a spectrum of initial conditions (Maack, 2001). The main purpose of such scenario development is thus to stimulate thinking about possible occurrences, assumptions relating these occurrences, possible opportunities and risks, and courses of action (Jarke et al., 1998). Moreover, by identifying basic trends, stakeholders can construct a series of scenarios that will help them to compensate for the usual errors in decision-making, i.e. overconfidence and tunnel vision (Schoemaker, 1995). Models that allow the simulation of user-defined scenarios of policy interventions can serve as useful decision support tools for involved stakeholders. Such tools should be user-friendly platforms in terms of their operation and the dissemination and visualization of model results, with the aim to enhance communication among the model and the user(s), and among policy-makers and other stakeholders. Multi-agent simulation models have been recognized to be able to meet these conditions, in particular the agent-based NetLogo environment in which GH-LUDAS was programmed. Visual formats of NetLogo, such as temporal calibrated maps and time-series graphs, and a user-friendly interface allow the use of GH-LUDAS as a decision support tool. Possible future scenarios the user wants to explore can be easily simulated, analyzed and communicated.

In this chapter, we present simulations of selected scenarios with GH-LUDAS, analyze the reasons for their way of performance, and communicate the corresponding results. However, to enhance communication of model results to stakeholders, the interpretation of the simulated pathways of selected scenarios should not only be grounded on the analysis of data and internal model mechanisms, but should also be vested into 'narrative storylines', which are easier to convey to local stakeholders. It is important that these storylines are consistent with data generated by the model as well as with narrative observations during field work and other related studies. The quality of scenario-based studies is dependent on the reasonability of processes involved, which can be generated by mental models in a narrative manner, or by formal models in a quantitative way. Each form has its own merits and limitations, and an efficient scenario description should therefore offer ways to integrate the narrative and quantitative traditions in a particular balance (Kemp-Benedict, 2004).
Based on this background, the objectives of this part of the study can be summarized as follows:

1. Based on the specifications of the theoretical framework (Chapter 3 to 5), to develop an operational GH-LUDAS model with the functionalities of a decision support tool to support impact assessment of selected policy options and other external factors.
2. To identify and simulate integrated scenarios of the coupled human-environment system using GH-LUDAS.
3. To provide an overview of the future pathways of these scenarios and an interpretation of these results in the form of narrative storylines based on quantitative analysis of the system functioning and field experience.

In the following, policy, climatic and demographic conditions in the study area will be described, which serve as a basis to justify the selection of external parameters of GH-LUDAS to be modified by the user. Based on this selection, the range of scenarios to be explored is presented. The subsequent section deals with the implementation of GH-LUDAS as a decision support tool, describing the mode of model operation, methodologies of output visualization and transfer, and the operation of the model interface. Finally, the scenario parameters are specified and the temporal evolution of selected relevant performance indicators of land-use/cover change and local livelihoods are analyzed and interpreted.

6.2 Selection of user-defined parameters in GH-LUDAS: Land-use policies, demography, and climate change

In the predominantly smallholder farming systems of the Upper East Region in Ghana, livelihoods are directly dependent on harvestable crop yields on a seasonal basis. The constraints to sustainable production are the dry spells during the cropping seasons, low fertility of farmlands, and farming practices that exacerbate the effects of drought and low soil fertility (CGIAR, 2000). The coping strategies resulting from these agroclimatic factors put a severe brake on investment and financial accumulation. The region's physical isolation, lack of non-agricultural investments and underdevelopment of markets result in few opportunities for economically meaningful off-farm employment or income generation (Whitehead, 2004).

The most recent agricultural polices in Ghana to tackle this problematic situation are reflected in the Accelerated Agricultural Growth and Development Strategy (AAGDS), the Food and Agricultural Sector Development Policy (FASDEP), and the Upper East Region Land Conservation and Smallholder Rehabilitation Project (LACOSREP) (IFAD, 2005). These projects broadly aim at the intensification and modernization of agriculture, income diversification, and improvement of market access. The core agricultural policies that constitute these national and regional strategies include further development of riverine irrigation, rehabilitation and construction of dams, farmer training and dissemination of new technologies, stimulation of the engagement in income-generating activities through credit, and an increased provision of infrastructure (IFAD, 2005). The promotion of irrigation through farmer education and improvement of irrigation facilities aims at improving food security in the 'lean season', and the stimulation of trade markets through increased income and demand for local products (Birner and Schiffer, 2005). Farmer training, meant to be implemented by local NGOs and the local branches of the Ministry of Food and Agriculture (MOFA), focuses on the promotion of high-yielding varieties, improvement of storage facilities, conservation measures to reduce yield losses due to soil erosion, and improved animal care. Furthermore, these organizations are also involved in the process of selecting and advising farmer groups that seek to apply for bank credits. These credits aim at financing crop production and agriculture-related small-scale enterprises, mainly targeted to women heads of households

(IFAD, 2005). Greater investment in rural infrastructure such as feeder roads and marketing facilities aim at linking remote rural areas with high production levels to agricultural markets, thus providing enhanced marketing opportunities for increasing incomes (MOFA, 2002).

However, many of these measures fail or have failed due to unefficient implementation or lack of finances in large parts of the region. Despite long years of development assistance, many communities remain poor, vulnerable and suffer from regular food shortages (Blench, 2006). In the Atankwidi catchment, the small-scale dams, which had been built to a large part in the 1970s, are silted due to misconstruction and thus can not be used for irrigation, and new dam construction projects under the new development programmes have not been implemented. MOFA, which is in charge of farmer training and education, seems to have had minimal contacts with local farmers, and their advice has not seemed to have any impact on local agricultural methods, choice of crops or livestock care. Furthermore, only 5 % of the women groups in the study area that applied for credit were finally successful, which was observed to be due to high bureaucracy levels and lack of staff on the side of MOFA.

In spite of these low levels of policy implementation, it seems there is agreement about the necessary interventions on the side of policy-makers. However, there is a high uncertainty and lack of a knowledge base about the human-environment interrelations and the policy impact on these relationships. The actual consequences on land-use and social and economic welfare of any of these measures are not well known. Scenario-based simulations could assist stakeholders in focusing their financial resources on policy measures that yield the highest returns in terms of long-term income security and equity. Therefore, with respect to the study area, we extracted those policy interventions that deserve a closer look in terms of their applicability and impact. The first strategy, the promotion of riverine irrigation farming, does not seem to be an issue in the catchment, as most of the irrigable land is already claimed. With respect to extension services, including farmer education and training either carried out by NGOs or MOFA, statistical analysis showed no impact on crop choice, agricultural techniques or input, livestock survival or crop yield. It seemed thus that even higher levels of farmer contact would not show reasonable improvements in living standards or changes in land-use or land-cover. Similarly, with respect to infrastructure, the proximity to feeder roads or marketing facilities did not significantly influence household decision-making or local marketing opportunities. The study area is already provided with a relatively extensive net of feeder roads, and market places are accessible from throughout the area on foot or by bicycle. Thus, the strategies which deserve closer attention are dam construction, as there is the ability and need among farmers to expand their irrigation business, and increased credit access, as statistics suggest a high relationship between credit provision and improvments in income levels.

However, decision-makers might not only be interested in the effects of their policies on local land-use and livelihood, but also in the future pathways caused by other factors. Reviews of the most significant changes that the region will face during the next decades comprise most importantly demographic changes and climate change. Due to climatic changes, the region experiences short and erratic rainfall, which directly affects food and livestock production (GNADO, 2000). The high population of the region is another factor that contributes to food insecurity and the poverty. Land holdings in the region are so small that food produced on one cannot sustain a family up to the next farming season (GNADO, 2000). Based on this reasoning, the following four families of scenarios were identified: i) construction of small-scale dams, ii) increased credit access, iii) population growth, and iv) rainfall changes derived from the main four IPCC climate scenarios.

Policy of rehabilitation and construction of small-scale dams

Many small-scale irrigation schemes based on earth dams and dugouts exist in the northern part of Ghana. Out of these, many were funded under World Bank projects (including the Upper Region Agricultural Development Project - URADEP) in the 1970s. The majority of small-scale structures have broken down over time due to poor maintenance and resulting siltation problems (Gyasi, 2004). Several donor agencies, government organizations and NGOs are involved in the rehabilitation of these schemes and the construction of new ones, which are to be managed by farmers. Indeed, close to 90 % of rehabilitated small schemes are successfully controlled by farmers (Dittoh, 2000). The major rehabilitation schemes in the Upper East Region have been conducted by the IFAD-funded Land Conservation and Smallholder Rehabilitation Project (LACOSREP), under which a total of 44 dams and dugouts were rehabilitated (IFAD, 2005).

The ultimate targets of the provision of communities with irrigation infrastructure include the offer of possibilities to local smallholders to engage in cultivation during the lean season, diversify their income structure, give incentives for increased marketing activity through raised cash income, and provide facilities for livestock watering and fishery (Birner and Schiffer, 2005). However, few irrigation infrastructure facilities were completed and functional on project closure, making it difficult to assess their impact properly (IFAD, 2005). A second question arises from the viewpoint of profitability, i.e. whether the obtained benefits from improved irrigation infrastructure really justify the relatively high costs of dam rehabilitation/construction, or whether other policy measures are more efficient and cost-effective. Therefore, an assessment of the long-term effects of irrigation scheme development on living standards and land-use and land-cover is of great importance.

A second policy measure with respect to the final use of small-scale dams is the application of area limitation. Our hypothesis is that the efficiency of operational dams in terms of income equity can be increased by limiting the area a farmer is allowed to irrigate around dams. This might allow more farmers to benefit from irrigation infrastructure, and reduce the number of farmers that share large parts of the irrigable areas of the scheme. Although we do not have any notice of the application of such a policy in present irrigation schemes in the Upper East Region, the investigation of the effects of this hypothetical policy measure could lead to interesting results for local stakeholders and water use authorities.

Policy of credit schemes

In an attempt to alleviate poverty and empower poor people, many NGOs and government-line agencies have been providing credit to rural women in many districts of Ghana. The essence of these credit schemes is to help the rural poor, especially women, earn a decent living through their on-going income generating activities (Ansoglenang, 2006). It was realized that women have assumed certain household responsibilities that were formerly men's gender roles, such as providing money and other material resources for housekeeping. These added responsibilities have given rural women a rare voice in household decision-making processes (Ansoglenang, 2006). Credit schemes are intended to help these women to increase their engagement in a number of income generating activities, including trade, shea-butter extraction, rice milling, pottery, local restaurant services, and alcohol brewing (Ansoglenang, 2006), and to expand these activities to small-scale enterprises. The promotion of such small-scale enterprises through credit schemes may help smallholder households to reduce risks and their dependency on agriculture through income diversification, create additional income, and stimulate marketing activity. Several case studies have emphasized the success of such credit schemes in terms of household assets, economic activity and the empowerment of women. However, rates of credit provision still remain low in the region, due to lack of staff and commitment on the side of the implementing agencies.

Population growth

Rapid population growth and low economic standards of living have had consequences for agricultural land resources in the Upper East Region (Benneh and Agyepong, 1990). Fallow lands have been reduced or eliminated, and there has been massive migration of mainly the youth to the urbanised, mining and forest areas in southern Ghana (Codjoe, 2004). The results of the agricultural land availability status (Codjoe, 2004) shows that three selected districts, namely, Bolgatanga, Bawku East and Kassena-Nankana located in the Upper East Region, would experience agricultural land shortfall in the year 2010 as a result of population growth. However, projections of annual population growth rates often lack reliable databases of past population trends and an understanding of the dynamics of migration strategies (Boadu, 2000). Although the dynamics of the single factors birth, death and migration rates are poorly understood, the observed (total) population growth rate has been estimated to 3 % in the rural Upper East Region. However, the capacity of these rural areas to sustain growing populations is limited. As land availability and reduced land productivity are considered as drivers of out-migration and ultimately as limiting factors for population growth as suggested by (Codjoe, 2004), a straightforward approach is thus to define population dynamics on the basis of the carrying capacity of the study area. A logistic function, which is defined by the annual growth rate of 3 % and the total population carrying capacity (see Chapter 3), was used in GH-LUDAS to calculate annual population increases for the study area. Based on this, the model allows the simulation of different settings of the population carrying capacity and an assessment of their consequences on local household behavior and land use.

Climate change

The Upper East Region, which is mainly a rural area with sub-humid conditions lying at the southern end of the Sahel, could be affected by climate change in terms of increased land degradation, declining agricultural productivity and changing land-use and livelihood strategies. A comparison for the region between the rainfall situation in the middle of the 20th century with the period 1970-1990 reveals a major climate deterioriation, but also that after the late 1980s the situation improved again until the 1997 drought, which was generally seen as problematic (Dietz et al., 2004). However, local farmers who were interviewed in the study by Dietz et al. (2004) saw a lot of evidence of long-term climate change, and have already been reacting to it. Changes regarding the onset of and a shortening of the rainy period has urged farmers to change the composition of their livelihood portfolios by relying more on non-agicultural sources of income, by adding more market-oriented agricultural crops (tomatoes, onions), and by changing their food production strategies to more drought-resistant varieties (Dietz et al., 2004).

It is therefore an important issue to understand the mechanisms between household decision-making and scenarios of future climate conditions, especially changes in rainfall patterns. To test household-based reactions to changed annual precipitation, we derived long-term data of annual precipitation changes for the study area from the IPCC Data Distribution Centre (www.ipcc-data.org), and linked them to functions of biomass productivity as proposed in the study by Groten (1991) (see section 5.3.3). These precipitation scenarios rely on the four basic global climate scenarios as presented by the IPCC SRES (Special Report on Emissions Scenarios), named A1, A2, B1 and B2, which cover a wide range of driving forces from demographic to social and economic developments. The annual precipitation reduction for these scenarios amounted to 2.87 mm/year for the A1, 0.36 mm/year for the A2, 2.84 mm/year for the B1, and 2.48 mm/year for the B2 scenario. Based on these values and the current average annual precipitation, the annual precipitation (mm/year) for each scenario was calculated and used for calculating forage availability (equation 5.23) and agricultural productivity of rainy-season cultivation (equation 5.20).

6.3 Developing an operational GH-LUDAS for policy decision purposes

6.3.1 Methodology

The GH-LUDAS theoretical framework (Chapter 3) was programmed in the NetLogo package 4.0.2. NetLogo, which is a freeware provided by Wilensky (1999), is a multi-agent modeling environment, which offers both a convenient language to programme agents (and their interactions) and tools to visualize and export results. The NetLogo environment consists of two main pages between which the user can switch, one reserved for the programme code, and a second, the model interface, which allows the setting of model parameters and the visualization of results. GH-LUDAS is thus a convenient platform for decision-makers, as they can easily choose among options, set parameters and view output graphs and maps on the interface page without necessarily understanding the source code. The procedures programmed in the code interface follow a schematic annual time-loop (section 3.6.2), starting with the updating of the population, followed by the routines for the dry and rainy season, and ending with the visualization and export of selected household and landscape data. These routines were verified separately as well as in combination, i.e. they were examined whether they work the way they were intended to.

The output of model simulations, which may serve as a basis for discussion among stakeholders, does not only depend on the specifications of the model routines, but also on model input data. Such input data comprise data and parameters that have been calibrated by the modeler, and external parameters that are defined by the user.

Input data of GH-LUDAS

Data that are defined by the modeler comprise calibrated input data, including spatial (GIS) data, household data, and specific parameters, mainly technical coefficients that have been extracted from quantitative analyses in case studies (Chapters 4 and 5). The household and GIS datasets were needed to initialize the coupled human-landscape, while the parameters were needed to specify various internal routines of the model. Because good-quality data are used to validate in part the MAS model, all data used by GH-LUDAS had to be calibrated and/or processed outside the model to adequately represent the reality of the coupled human-environment system. Methodologies for processing/calibrating/classifying data from different sources, organizing the household-pixel dataset, and scientific approximation of relevant data for use are discussed in detail in Chapter 4 and 5.

In contrast to these data, user-defined parameters are intended to be set externally by the user. These parameters comprise policy-related, demographic, and climate parameters, which enable the users to set their own options for scenario development. Policy-related options include the specification of the location and irrigation capacity of dams, and the annual percentage of households provided with credit, whereby an option is given to choose between a revolving credit scheme and the current (normal) scheme (see section 3.5.2). Furthermore, the growth rate and carrying capacity of the population as well as a specific IPCC climate scenario can be defined.

Ouput of GH-LUDAS

The strength of the NetLogo programming platform, and of GH-LUDAS in particular, is its provision of a set of very informative outputs. For any time step of the simulation, including two season-wise simulation steps per annum, three types of output are produced: a spatially explicit map of land-use and land-cover, graphs, and spreadsheets of predefined indicators.

Land-use and land-cover map

The land-use and land-cover map is depicted in the viewer of the NetLogo platform, and displays dry- and rainy-season land-use/cover patterns in sequence in order to reflect the real-world temporal fashion in which land-use/cover changes occur annually. With the help of the NetLogo functionality 'export-viewer', values of each pixel of the map can be exported at any of the two annual time steps in any year of simulation. Exported files of these spatially explicit maps enable experts to conduct sophisticated interpretations of the simulated land-use/cover patterns.

Digital images and graphs

A digital map interface was designed to enable the user to navigate among different landscape attributes by clicking the corresponding buttons. This allows users to visually link changes in land-use/cover to important landscape attributes such as elevation, slope, distance to river, village territories, etc. Furthermore, real-time changes in predefined indicators are visualized in graphs, e.g. average household income, percentages of land-use types within the catchment, etc. Data underlying these graphs can be exported in text files for each time step during simulation for further analysis and interpretation.

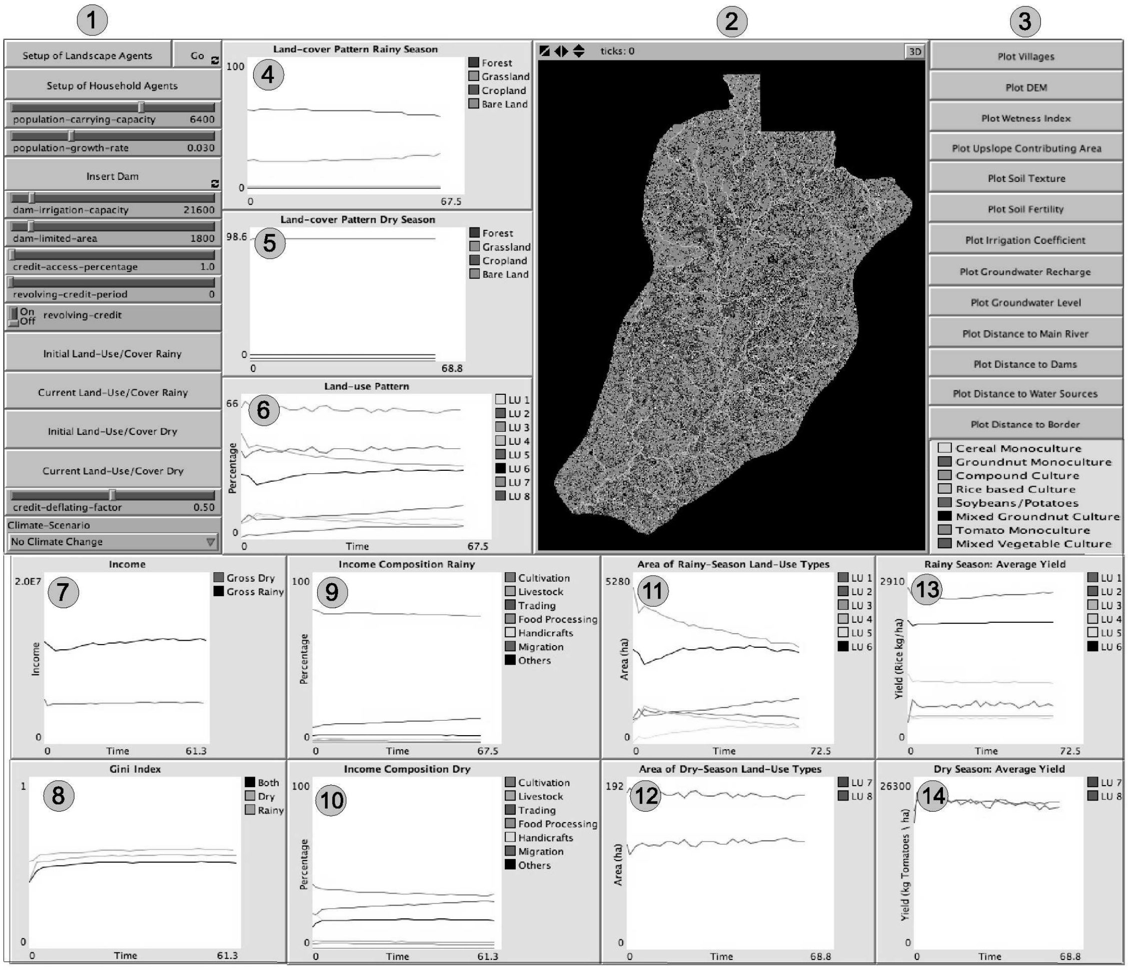
Predefined indicators

In each time step, indicators predefined by the modeler are exported to text files to serve as a basis for further analysis of the performance of the system. A wide range of indicators of households and the landscape are concurrently saved, ranging from average income from and labor input to all income-generating activities of households, over livelihood-strategy indicators and the Gini Index, to average crop yields of landscape agents, and land-use/cover performance. Adding or modifying selected indicators is an easy task if such needs arise on the side of the user.

6.3.2 Results

The user interface of the model comprise the following components: i) User's input parameters and a navigation bar for landscape attributes (parts (1) and (3) in Figure 6.1). ii) a real-time map of land use and land cover (part (2) in Figure 6.1), and iii) time-series graphs of predefined indicators of the coupled human-environment system (parts (4) to (14) in Figure 6.1). In Figures 6.2 to 6.7, the parts of the interface are depicted in detail.

By pressing the top three buttons of the input parameter bar (1), the landscape and the household agents are initialized, and the simulation of sequential annual time-loops is started. Below, parameters of population growth can be set manually by sliders, including the carrying capacity of the number of households in the catchment, and the annual growth rate. By pressing the 'draw-dam' button, dams can be inserted in the viewer by mouse click, whereby the dam's irrigation capacity needs to be defined by a slider. Below, the maximum area a dam user is allowed to cultivate can be set. Below are the credit-related settings, including a slider to regulate the annual credit access percentage of the population, a switch to choose the credit scheme and a regulator to define the timely extent of the revolving credit scheme if this option is chosen. The next four buttons allow switching among initial land-use/cover patterns in 2006 and simulated (final) patterns to enable the user to identify substantial changes visually. The last option allows the choice of rainfall scenarios, including 'No Climate Change' and the four IPCC rainfall scenarios.



1. User-defined global parameters
2. Multi-temporal map of land-use/cover
3. Navigation of different spatial attributes
4. Land-cover percentages in the rainy season
5. Land-cover percentages in the dry season
6. Percentages of land-use types
7. Average gross household income
8. Gini index of household income equity
9. Income composition (%): rainy season
10. Income composition (%): dry season
11. Area of land-use types (ha): rainy season
12. Area of land-use types (ha): dry season
13. Average yield (kg/ha): rainy season
14. Average yield (kg/ha): dry season

Figure 6.1: Model interface of GH-LUDAS

The navigation bar (3) allows the user to map major environmental attributes in the viewer, including village territories, topography-related variables (e.g. elevation, wetness index), soil attributes (e.g. texture, fertility), groundwater level and recharge, and proximity-related variables (e.g. distance to river).

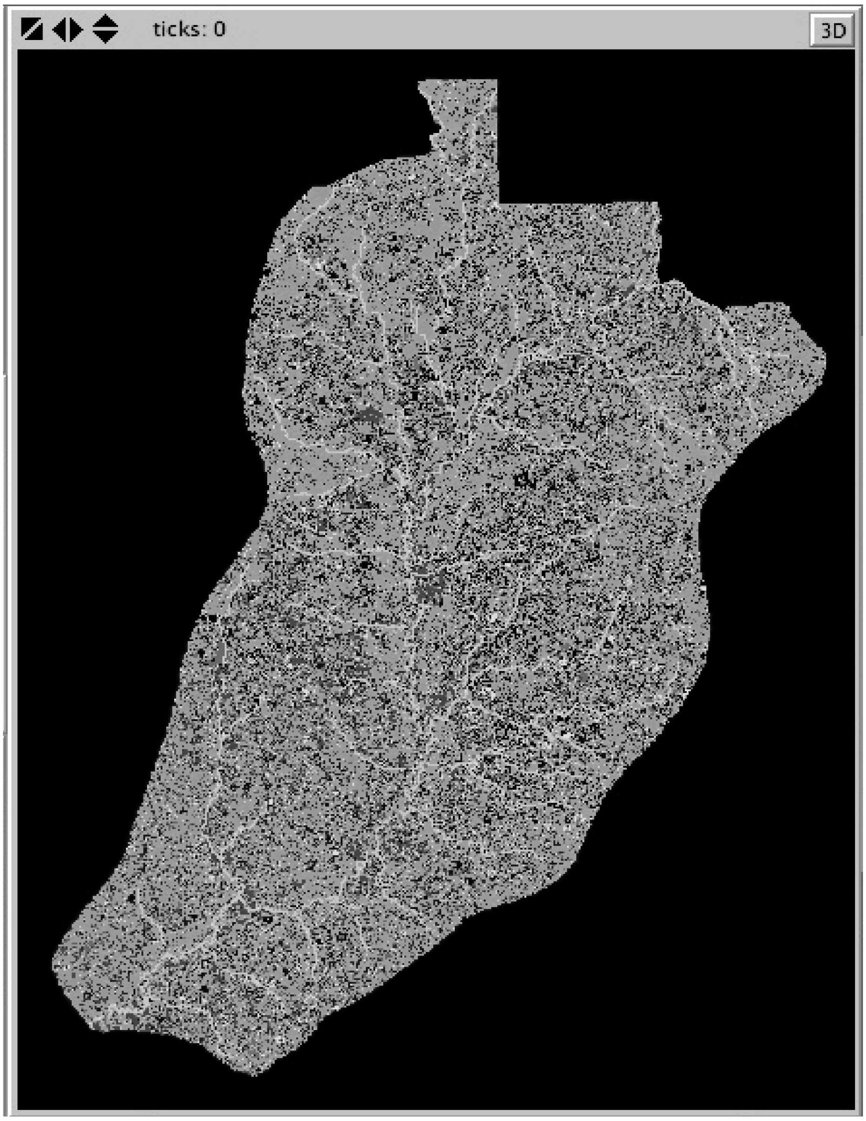


Figure 6.2: Viewer (Part 2) of model interface

The time-series graphs include two major blocks. The first block comprises graphs of indicators of the performance of the biophysical landscape system (see parts (4) - (6) and

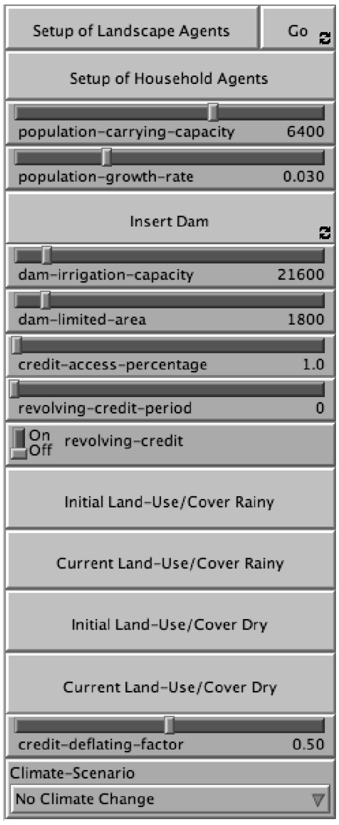


Figure 6.3: Input parameter bar (Part 1) of model interface

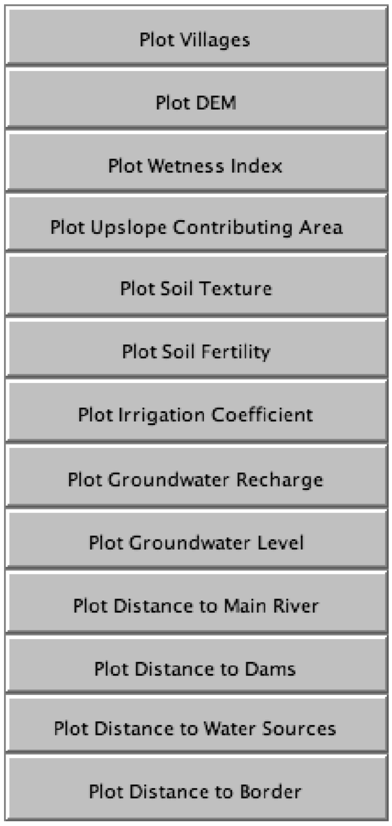


Figure 6.4: Navigation bar (Part 3) of model interface

(11) - (14)). Graphs (4) and (5) monitor changes in the coverage of the four main land-cover types for each season, while graph (6) depicts changes in the coverage of the six rainy-season land-use types (LU 1 - LU 6), and the two dry-season land-use types (LU 7 and LU 8). A legend for these land-use types is attached on the right side of the interface. While this latter graph only monitors the percental changes of land-use types of total cropland area, graphs (11) and (12) display the actual area of these land-use types in hectares.

Graphs (13) and (14) finally show the performance over time of average yield in kg rice/ha and kg tomatoes/ha for rainy-season and dry-season land-use types, respectively. Based on information of average yield and spatial extent of the single land-use types as monitored by the latter four graphs, the total crop production for the catchment can be easily calculated for each season.

The second time-series block of graphs comprises the graphs for monitoring changes

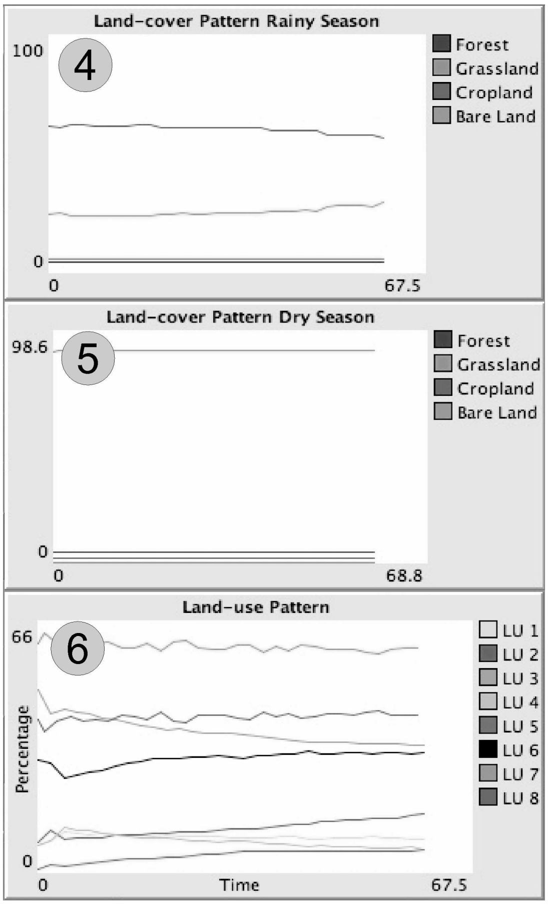


Figure 6.5: Land-use/cover graphs (Parts 4 - 6) of model interface

in the human system, parallel to the changes in the natural landscape system (see parts (7) - (10)). Graph (7) shows trends in changes in gross household income for both seasons separately, and graph (8) displays the equity of household incomes, in terms of Gini Indices of income distribution for both seasons separately and in combination. Graphs (9) and (10) show the average income structures in each season, depicting the percentages of each of the seven major income-generating activities of total gross household income. These trends allow an interpretation of changing livelihood strategies.

This user-friendly interface will allow stakeholders to test the combined consequences of selected user-defined parameters on the landscape as well as on the population level. An interaction loop may develop between the users and the model, by improving the knowledge of the effects of interventions and natural and demographic changes on the

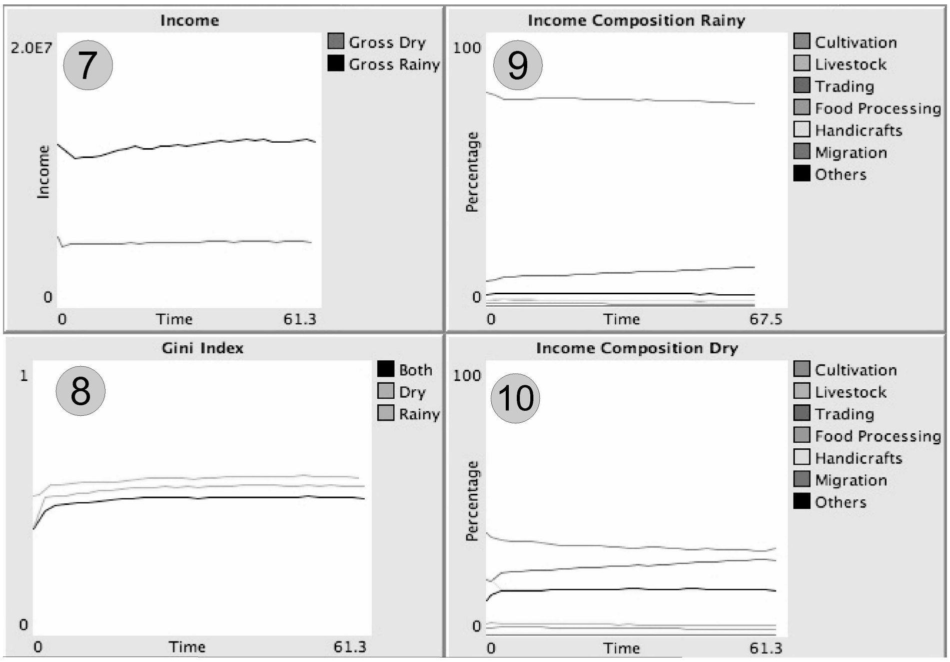


Figure 6.6: Income-related graphs (Parts 7 - 10) of model interface

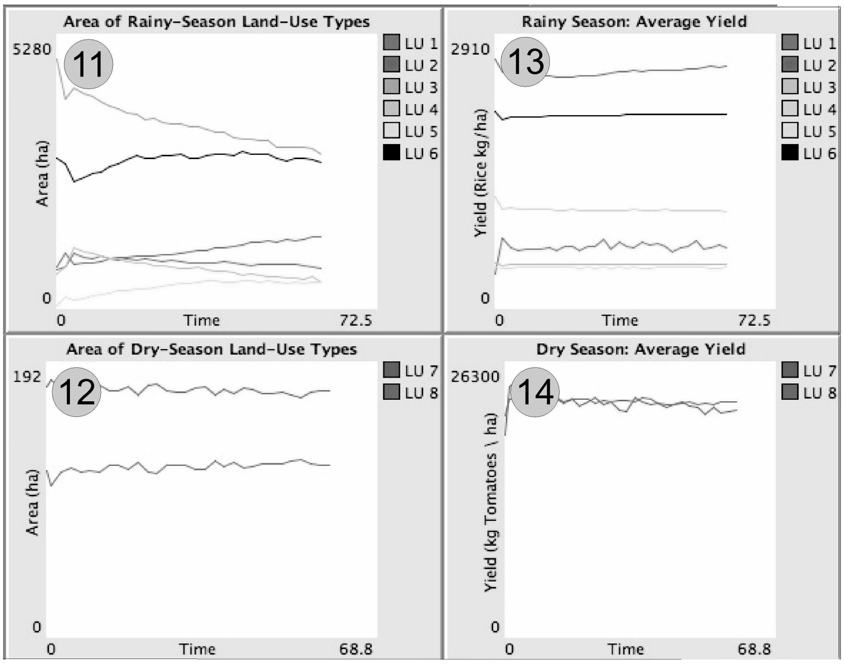


Figure 6.7: Land-use-type-related graphs (Parts 11 - 14) of model interface

coupled human-environment system. Furthermore, the interface of GH-LUDAS may enable users to develop scenarios that can be used as case studies for further analysis and interpretation.

6.4 Definition, simulation and analysis of selected scenarios

6.4.1 Methodology

According to the identified range of land-use-related factors and policies (section 6.2), the specific scenarios to be tested were systematically defined as follows:

1. The policy and global settings as in 2006 are considered the baseline scenario. This scenario assumes no rainfall or demographic changes and will be used as a baseline for evaluating the impacts of changes in land-use policies and other factors.
2. Given the baseline settings, each single policy/global factor will be shifted from the baseline to form a scenario spectrum of the considered factor. Other policy/global factors are kept the same as in the baseline scenario. Each such scenario spectrum consists of 2 to 4 single scenarios, which will enable identifying the sensitivity of this factor to socio-economic indicators and land-use/cover performance.

The different policy scenarios of each scenario spectrum are briefly described below:

Baseline scenario

The baseline scenario (S0) has the policy setting as in 2006, which is the base year of the simulation. According to statistics from the Ministry of Food and Agriculture (MOFA) in Navrongo, about 1 % of the households obtained credit every year during that time. As local dams were not operational for irrigation in the study area, no dams were inserted within this scenario. The information about past demographic statistics provided by the Ghanaian Statistical Service was too limited to serve as a basis to extrapolate future changes in local demography. Thus, for the baseline, the number of households in the catchment was assumed to remain stable. The potential consequences of an increase in household population will be separately analyzed within the demographic scenario. In a similar vein, annual rainfall was assumed to be stable in this scenario, at a level of 1100 mm per annum.

Scenarios for assessing the impacts of dam construction

Usually, variations in the location, size and number of constructed dams need to be considered in the policy of dam construction. However, as it is impractical to test all possible scenarios, we focused on variations in total available irrigable area. For this, we varied the number of dams, all having the same irrigation capacity of about 2.1 hectares, which is a reasonable value for small-scale dams in the Upper East Region. Although we are aware of the fact that the selection of the dam location underlies hydrological considerations on the side of policy-makers and contractors, we did not apply such a selection process to identify suitable locations, but assumed a random distribution of dams throughout the catchment. This procedure is justifiable, as prior simulations had shown that variations in the specific locations of these dams did not show a significant influence on the socio-economic indicators or land-use/cover at the level of the population/catchment. Following this mindset, we defined three scenarios, with a random distribution of 20 dams named the S-Dam20 scenario, of 30 dams (S-Dam30), and of 40 dams (S-Dam40). All other settings were kept the same as for the baseline scenario.

Scenarios for assessing the impacts of improved credit access

This scenario spectrum consists of different settings of annual credit access, while other parameter values are identical with those for the baseline scenario. The term annual credit access denotes the annual percentage of households that obtain credit, whereby the amount of credit is fixed to 200 000 Ghanaian Cedis (about US $ 20), which is the usual amount granted to applicants in the study area. To test the sensitivity of output values to increased ncredit access, three scenarios were defined, a percentage of 4 % (S-Cred4), 7 % (S-Cred7), and 10 % (S-Cred10). These values express a gradual change in credit coverage by a 3 % stepwise increase, based on the current value of 1 % in the baseline scenario.

Scenarios for assessing the impact of area limitation under dam construction

This scenario spectrum explores the impact of area limitation under a policy of construction of 30 dams. The scenarios include an area limitation of 900 m^2 (S-Lim900), an area limitation of 1800 m^2 (S-Lim1800), and no limitation (S-LimNo). The latter scenario has the

Table 6.1: Global-policy settings for scenario development

Scenario	Quantitative Settings: Dam Construction: Number of Dams	Dam Construction: Area Limitation (m^2)	Credit Access: Annual Credit Access (%)	Demography: Carrying Capacity (Households)	Demography: Annual. Growth Rate (%)	Rainfall Change: Scenario
Baseline Scenario (with policy settings of 2006)						
S0 (Baseline)	0	-	1 %	6400	3 %	No Change
Scenarios for exploring the impacts of dam construction						
S-Dam20	20	-	1 %	6400	3 %	No Change
S-Dam30	30	-	1 %	6400	3 %	No Change
S-Dam40	40	-	1 %	6400	3 %	No Change
Scenarios for exploring the impacts of area limitation						
S-Lim900	30	900 m^2	1 %	6400	3 %	No Change
S-Lim1800	30	1800 m^2	1 %	6400	3 %	No Change
S-LimNo	30	-	1 %	6400	3 %	No Change
Scenarios for exploring the impacts of credit access						
S-Cred4	0	-	4 %	6400	3 %	No Change
S-Cred7	0	-	7 %	6400	3 %	No Change
S-Cred10	0	-	10 %	6400	3 %	No Change
Scenarios for exploring the impacts of population growth						
S-Pop7200	0	-	1 %	7200	3 %	No Change
S-Pop8400	0	-	1 %	8400	3 %	No Change
S-Pop9600	0	-	1 %	9600	3 %	No Change
Scenarios for exploring the impacts of rainfall change						
S-ClimA1	0	-	1 %	6400	3 %	A1
S-ClimA2	0	-	1 %	6400	3 %	A2
S-ClimB1	0	-	1 %	6400	3 %	B1
S-ClimB2	0	-	1 %	6400	3 %	B2

same settings as S-Dam30, and the former two scenarios only deviate from this base scenario in their value for area limitation. Finer increments in area limitation were not possible, as the spatial resolution of GH-LUDAS is pixel of 30 m x 30 m, making up an area of 900 m^2.

Scenarios for assessing the impact of different population carrying capacities

In this scenario spectrum, the impact of increases in population sizes on socio-economic indicators and land use/cover is explored. Local population growth is simulated by the logistic S-shaped growth function, which is defined by an annual growth rate and a population carrying capacity. In all scenarios, the annual growth rate was set to 3 %, which is the current observed value in the study area, while the population carrying capacities were set to totals of 7200 (Scenario S-Pop7200), 8400 (S-Pop8400), and 9600 households (S-Pop9600).

Scenarios for assessing the impacts of reduced precipitation

This spectrum covers four single rainfall scenarios, based on the simulation of scenarios developed by the IPCC Special Report of Emissions Scenarios (SRES), namely the A1, A2, B1 and B2 storylines (see section 6.2). Data on long-term changes in annual rainfall have been derived specifically for the study area, ranging from an annual reduction in precipitation of 2.87 mm (A1), over 2.84 (B1) and 2.48 (B2) to 0.36 mm (A2). The single scenarios for simulation in GH-LUDAS were named after their original SRES name.

6.4.2 Results

Each scenario was run 5 times for 30 timesteps (years), and mean values μ and uncertainty ranges $[\mu - CI_{0.05}, \mu + CI_{0.05}]$, where $CI_{0.05}$ is the radius of the 95 % uncertainty intervall, were calculated from the generated data for each scenario. In the subsequent analyses, we will focus on those indicators that showed a significant change during time and/or showed a dependency on external (e.g. policy) settings. Changes in land cover and land use in the rainy season and their dependency on global-policy settings were analyzed, as well as mean gross household income for each season, and the Gini Index, which describes the skewness of income distribution among the population. To analyze the behavior of income classes within the local society a further single-run simulation was carried out to derive behavioral values for the high-income class, the medium-income class and the low-income class, which are separated by $0.5 \cdot$ standard deviation of annual gross income. With respect to land cover in the dry season, no changes in the composition of land-cover types could be observed, mainly as irrigated cropland in this season had almost reached its maximum spatial extent during the base year 2006. The same is valid for the choices between the two dry-season land-use types, where no significant down- or upward trends could be observed for the selected scenarios. Therefore, in the following scenario analyses, the description of indicators of dry-season land-use and land-cover patterns was omitted. Instead, another important trend regarding dry-season land-use could be observed, i.e. a change in the number of farmers practicing irrigation, and subsequently changes in the average irrigated area per (irrigating) household, which were also significantly influenced by policy interventions.

Baseline scenario

Before analysing the impacts of selected external factors and policies, the temporal peformance of the baseline scenario needs to be analyzed in order to understand the general trend of land-use/cover change and related socio-economic indicators. This baseline will then be used to compare the performance of the subsequent scenarios with that of the baseline scenario. In this baseline, the mean gross household income increased both in the dry and the rainy season (Figure 6.8 a). In total, the increase in mean annual gross household income increased from 15.9 million Ghanaian Cedis to 16.6 ± 0.05 million Cedis during the 30-year period, which was observed to be mainly due to two factors. First, the productivity per land area was increased in the rainy season, and second, an increasing portion of household labor was dedicated to the more profitable activity of trading in both seasons. In average, the percentage of income generated by trading acitivities increased from 9.9 % to 14.4 ± 0.24 % during the observed period. The higher productivity levels in the agricultural sector were not caused by a process of intensification or higher yields, but were a result of a continuing shift to more profitable crops (or land-use types).

While in 2006, 45.7 % of the land-use types consisted of groundnut- and rice-based systems, which are regarded as cash crops due to their high marketable value, the portion of these cash land-use types increased to 52.2 ± 0.38 % at the end of the 30-year period. Although this shift to more market-oriented activities (e.g. trading) and crops is the result of many interacting factors, it can be genereally said that this trend is caused by an alternation of generations, as the young generation tends to be more cash-oriented and aims at reaching a higher labor efficiency in terms of a labor-income relation. This observation also matches the impression in the field, where both old and young farmers were informally interviewed about their individual income strategy and land-use tendencies. Another factor observed during the field study was the reluctance of many young farmers to engage in hard agricultural work, as many of them preferred other less labor-intensive strategies such as seasonal migration. This observation matches the decline in total agricultural area in the rainy season (Figure 6.8 c), which decreased from 61.5 % of the total area in 2006 to 55.5 ± 0.76 % in 2036. This decline

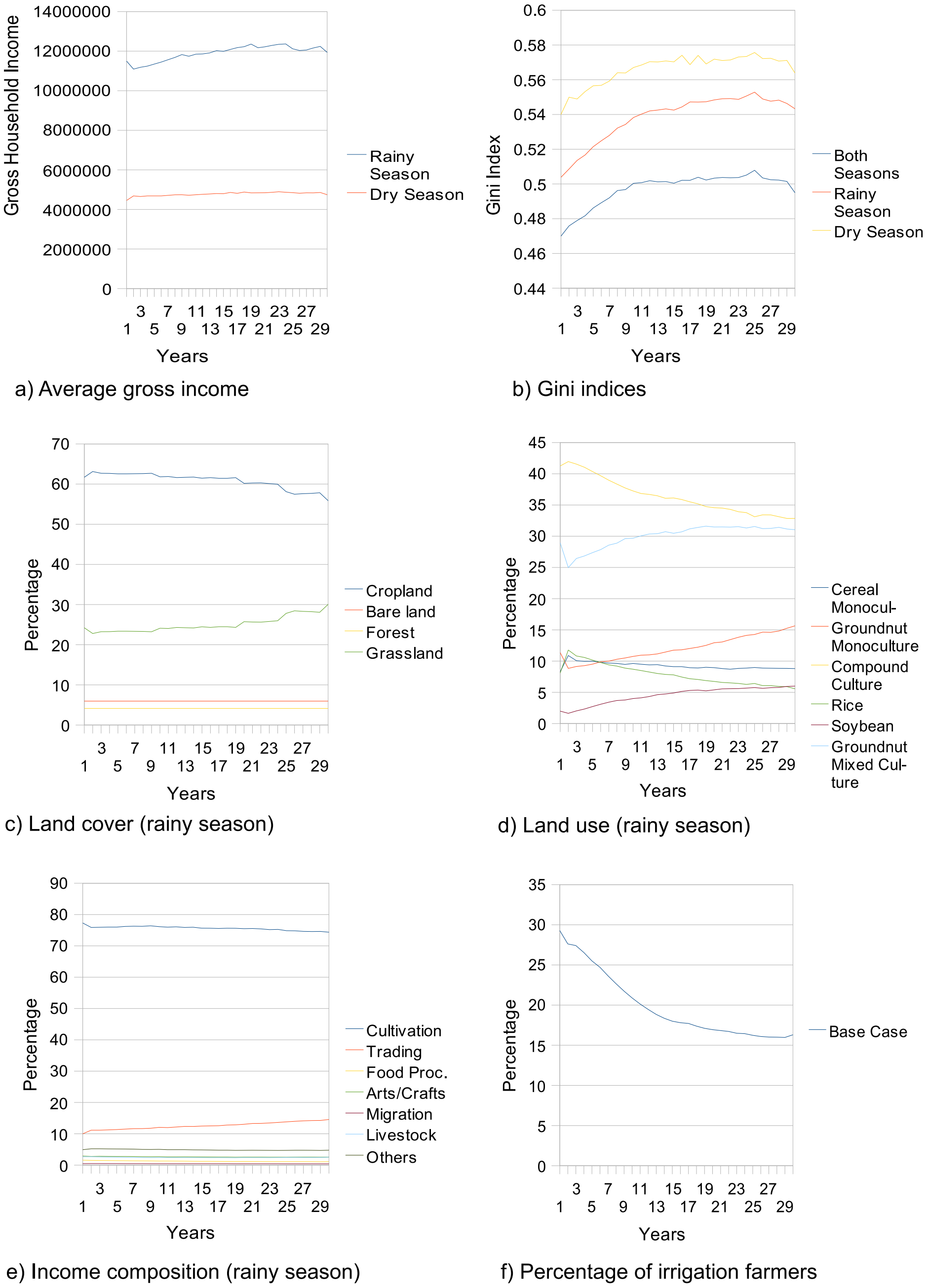


Figure 6.8: Baseline scenario: Simulated land-use/cover and socio-economic changes.
Source: Simulation results with GH-LUDAS

is supported by the fact that the household labor pool dedicated to agriculture was up to 10 % lower among the younger generation.

Although these processes lead to a general increase in average income, the equity in income distribution among the population, described by the Gini Index, seems to deteriorate (see Figure 6.8 b). During the observed period, the annual gross income of the wealthier part of the population (with income > mean income + 0.5 σ, where σ is the standard deviation) increased by 34 %, while that of the poorer part (with income < mean income - 0.5 σ) decreased by 36 %. We found that this increasing differentiation of gross income is partly due to an increasing inequity in land availability. This process is due to the fact that households falling into the lower-income class usually have many more offspring that those of the high-income class, which leads to the partitionment among many inheritents of land that is already small in size. Among high-income households, the situation is inverse. Household land is usually extensive, and its division usually does not lead to land shortages among the already few inheritants.

With respect to dry-season land use, the baseline scenario shows a decrease in the proportion of irrigation farmers of total population from 29.3 % in 2006 to 16.3 ± 0.11 % in 2036 (Figure 6.8 f). This implies that the limited irrigable area is being divided by a continuously decreasing number of farmers, which is also reflected by an increase in average irrigated area per farming household. The reason behind this process is the increasing use of pump irrigation technology, which allows the irrigation of larger areas, in comparison to the use of wells (bucket irrigation), where water has to be manually distributed. Within the 16-year old history of irrigation farming in the study area, farmers started with buckets, but in 2006, 40 % of the area was already irrigated by pump technology, which will, according to the scenario, increase to nearly 95 ± 0.14 % in 30 years.

Impacts of the policy of dam construction on land-use/cover and socio-economic status

In this scenario spectrum, the sensitivity of a dam construction policy on land-use/cover and socio-economic indicators is tested. For the dry season, average gross household income is highly sensitive to the number of constructed dams (Figure 6.9), resulting in a simulated average dry-season income of 5.72 ± 0.04 million Ghanaian Cedis in 2036 for S-Dam40, as compared to 4.74 ± 0.03 million Cedis for the baseline scenario. The additional income generated in the dry season is due to a clear shift from non-agricultural activities to irrigation farming. The analysis of simulation results reveals that in the baseline scenario only 32 ± 0.42 % of the income is generated by cultivation in the dry season, whereas in S-Dam40 this value amounts to 46.1 ± 0.14 %. This additional income does not seem to be reinvested in non-farm activities such as trading or arts/crafts, but merely in an extension of cultivation activities, especially in the dry season. Although there is a general upward trend in the involvement in such income-generating non-farm activities, the increased practice of irrigation farming does not seem to have a positive influence on this trend. Furthermore, additional income generated by irrigation farming does not seem to be invested in cash cropping during the rainy season either, as the cropping pattern is not sensitive to changes in dam numbers (Figure 6.9 e) and the uncertainty ranges overlap: According to the simulations, in 2036 about 52.6 ± 0.33 % of the cropland is used for cash crops for S-Dam40, while the figure is similar for the baseline, being 52.2 ± 0.38 %. This low effect might be due to the fact that households practicing irrigation farming usually reinvest their profit into this business, as this activity is usually more profitable than non-farm businesses such as trading or cultivation of cash crops. This behavior is in accordance with field interviews, which reveal that profit from irrigation farming is partly reinvested in irrigation, and partly used to get over the lean season. For the same reason, there does not seem to be any positive influence on income generated in the rainy season (Figure 6.9 a), which in 2036 amounts to 11.8 ± 0.196 million Cedis for S-Dam40, and 11.9 ± 0.175 million Cedis for the baseline scenario.

The Gini Index describing the equity level of income distribution is partly positively influenced by the policy of dam construction. While at the end of the first half of the simulation period the Gini Index is lower for S-Dam40 (0.486 ± 0.004) than for the baseline scenario (0.5 ± 0.004), the values seem to converge at the end of the simulation period (Figure 6.9 c). The single-run simulation to assess the local society structure reveals that for S-Dam40, the average simulated income for the low-income class (with income < mean income - 0.5 σ, where σ is the standard deviation) decreases during the simulation period by only by 25 % as compared to 34 % in the baseline scenario, while for the middle class, this value is 9 %, as compared to 14 % in the baseline. In other words, the process, which is leading to an increasingly skewed income distribution, can be slightly dampened by this policy intervention. Although it is difficult to identify the reasons for this improvement in

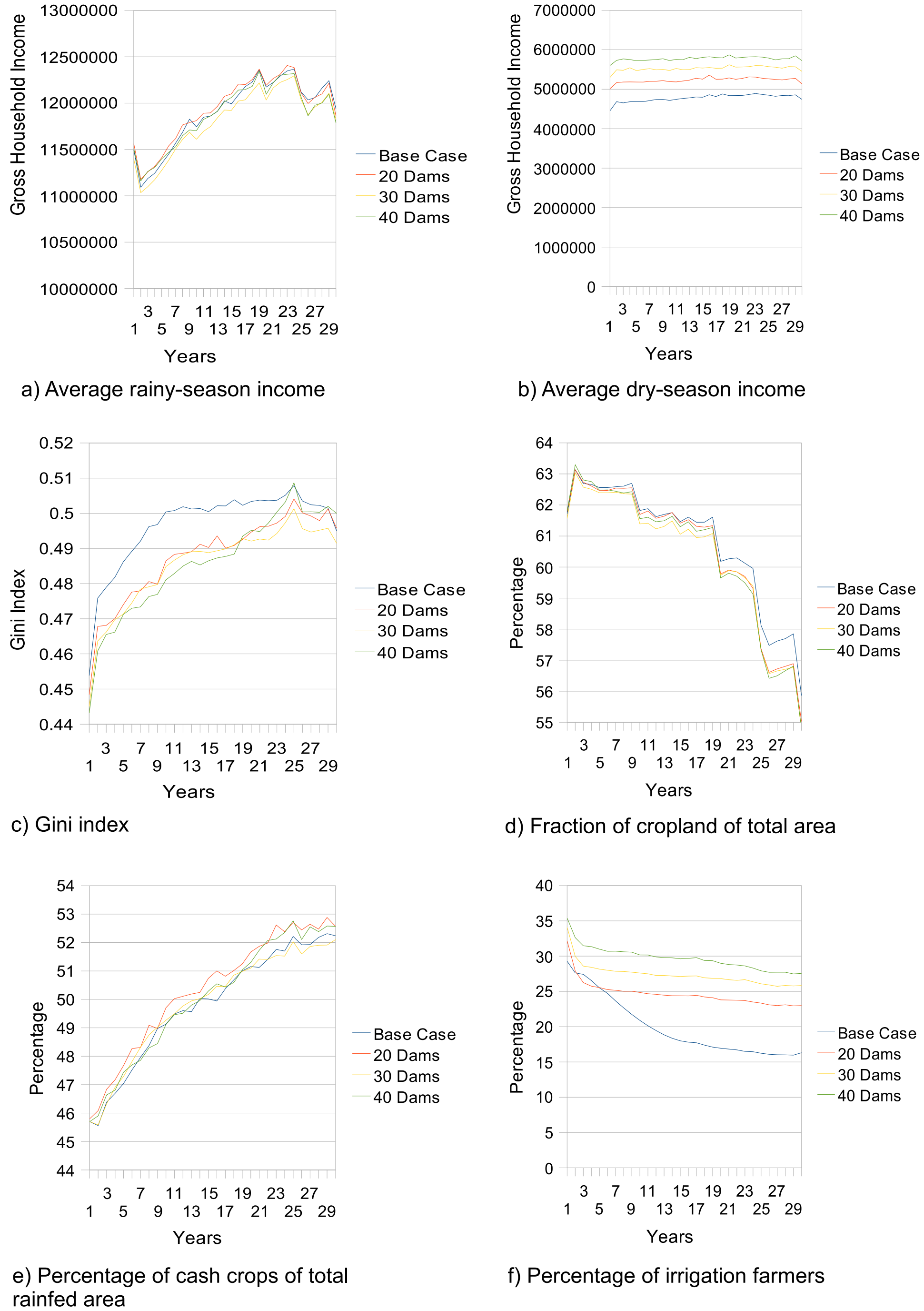


Figure 6.9: Dam construction scenarios: Simulated land-use/cover and socio-economic changes (see Tables A.1 to A.6 for means and uncertainty ranges)

income equity due to the model's complexity, two factors seem to play a major role. First, the availability of operational dams has a much higher impact on the share of irrigation farmers in the poorer class than that of the better-off class. Although the involvement of the poorer class in irrigation practices is generally low, the supply with irrigation infrastructure resulted in a 10-fold increase of the percentage of irrigation farmers among this group from 0.5 % to 5 %. The middle class experienced a 2-fold increase (from 15.5 % to 32 %), while the share of irrigation farmers among the better-off class increased only slightly. This extreme bias is caused by the fact that i) dam irrigation is generally a low-cost business that allows it to be practiced among low-income farmers, and ii) the majority of better-off farmers who share the interest in this business is already practicing it. Second, possible reason for an improvement in the income equity can be found in the correlation between irrigated land and rainfed land available to the households. The increasing involvement in the irrigation business, especially among lower and middle class farmers, seems to have an effect on their share of cultivated land in the rainy season. The increasing bias in land availability among the population as described in the baseline scenario is alleviated by the improved ability of the lower and middle class to rent additional land in the rainy season due to an improved financial situation. This trend of a higher tendency for rainy-season cultivation might also be reflected in the increased portion of cropland for the S-Dam40 scenario (Figure 6.9 d). In 2036, for the baseline scenario the percentage of cropland amounts to 55.86 ± 0.76 % while the value for S-Dam40 is 56.81 ± 1.02 %. Although the percentage of cropland seems to be slightly sensitive to the policy of dam construction, the significance is low, as the uncertainty ranges of these two scenarios overlap.

Furthermore, the policy of providing irrigation infrastructure in the form of dams can stabilize the declining trend of the share of irrigation farmers of the total population (Figure 6.9 f). In the baseline scenario, the percentage of irrigation farmers decreases from 29 % to 16 ± 0.1 % during the simulation period, in comparison to the value for S-Dam40, which is 27.6 ± 0.7 % in the final year 2036. The effect of dam construction on the percentage of irrigation farmers is significant for all four scenarios, having an average uncertainty range of ± 0.55 %. While the baseline results in larger average areas distributed among a declining number of households, this effect is clearly alleviated in the dam-based scenarios. Although irrigation farmers used to have larger cultivated areas in the beginning due to increased land availability, the upward trend is not as pronounced as in the baseline scenario. This might be due to the fact that in comparison to the baseline, a smaller portion of farmers has the financial capacity to expand their irrigation business, as there is a relatively high involvement of low-income and middle class farmers.

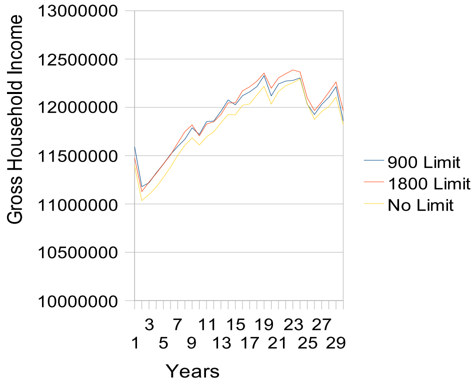
Impacts of area limitation under the policy of dam construction on land use/cover and socio-economic status

In this range of scenarios, the policy of area limitation is tested on land-use and livelihood performance. This limitation area, which is set to 900 m^2, 1800 m^2 and unlimited size, refers to the maximum area a household is allowed to irrigate in the drainage areas of local dams. The target of this secnario spectrum is to assess whether such a policy would increase the equity in income, due to an increased equity in irrigable land among local households. The Gini Index, however, gives a complex picture of this intended effect. While the Gini Index is lowest for the S-Limit900 scenario (Figure 6.10 c), which is expected to be due to an equal distribution of 900 m^2 land per household, the Gini Index for a limitation of 1800 m^2 seems to exceed that of the scenario of no area limitation. in 2036, the Gini Index for S-Limit1800 is 0.5 ± 0.008, while that for S-LimitNo is 0.491 ± 0.008 and for S-Limit900 this value is 4.87 ± 0.004. The uncertainty range of S-LimitNo overlaps with those of the other two scenarios, while S-Limit900 and S-Limit1800 have distinct uncertainty ranges. Therefore, we will only attempt to analyze the causes behind this latter diffference. However, as the system modeled is very complex, and a reliable analysis of this behavior is beyond our analytical capacities, we can only analyze the causes of this complex behavior to a limited extent. The assumption we can give is that land that is made available in the S-Limit1800 scenario seems to be mainly occupied by the high-income class. Two processes related to the difference in the Gini Index between S-Limit900 and S-Limit1800 were identified. First, the single-rune simulation shows that in 2036 the percentage of farmers belonging to the middle class is highest with 65 % for the S-Limit900 scenario, while it is lowest with 59 % for the S-Limit1800 scenario. Accordingly, the percentages of the lower and high-income classes are higher for the latter scenario, leading to an increased income gap and thus a higher Gini Index. Second, average incomes within these classes change to the disadvantage of the lower and middle income class. This process is worsened in the S-Limit1800 scenario as compared to the S-Limit900 scenario. Furthermore, there is an evident relationship between this difference in income and the difference in the allocation of irrigable (dam) areas. While the upper class claims only 28.5 % of the irrigable (dam) area in the S-Limit900 scenario, this value amounts to 32 % in the S-Limit1800 scenario, to the disadvantage of the middle class. This process seems to be in accordance with our assumption that mainly better-off farmers benfit from the implementation of an area limitation of 1800 m^2.

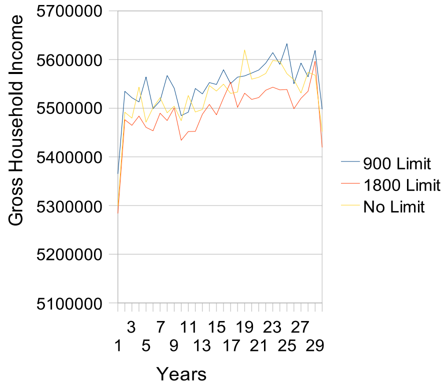
With respect to average income, the policy of area limitation does not seem to have a significant influence. Average gross income generated in the dry season in 2036 is highest for S-Limit900 (5.5 ± 0.047 million) while it is lowest for S-Limit1800 (5.42 ± 0.053 million). This dfifference, although not significant, might be caused by the fact that a higher fraction of households is involved in irrigation (30.5 ± 0.62 %), compared to 25 % in the other two scenarios). In the rainy season, households seem to compensate for their lower dry-season income in the S-Limit1800 scenario by investing in cash crops and an extension of rainfed area (see Figures 6.10 d and 6.10 e). In 2036 for S-Limit1800 55.7 ± 0.3 % of land is cropland, of which 53.1 ± 0.7 % is cultivated with cash crops, while for S-LimitNo this figures are lower, i.e. 54.8 ± 0.6 % of land is cropland, and out of these 52.1 ± 1.0 % are cultivated with cash crops. For 2036, this difference in household behavior results in a slightly higher rainy-season income for the S-Limit1800 scenario (12.0 ± 0.157 million Cedis), as compared to the S-NoLimit scenario (11.8 ± 0.256 million) (Figure 6.10 a). To summarize, a significant difference in income equity (Gini Index) can be induced by this policy, whereby effects on land use and average income are minimal and cannot be verified due to large uncertainty ranges. Possible further simulations could help to reduce this level of uncertainty.

Impacts of the policy of credit access on land-use/cover and socio-economic status

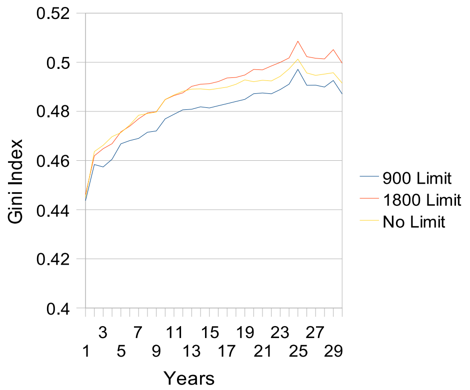
According to local data and interviews conducted in the study area, most given credits are invested in trading activities, which is the most profitable business apart from irrigation farming. Credits are usually given to women, and as trading is mainly a women's domain in contrast to the male domain of irrigation farming, women tend to start or expand their trading businesses. Since trading can be practiced throughout the year, additional income is generated in both seasons (see Figures 6.11 a and 6.11 b). For the year 2036, average rainy-season gross income amounted to 11.9 ± 0.176 million Ghanaian Cedis for the baseline scenario,



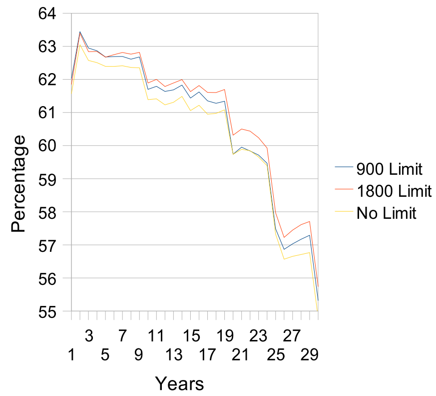
a) Average rainy-season income



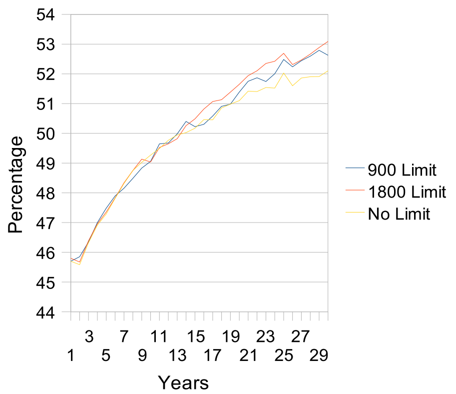
b) Average dry-season income



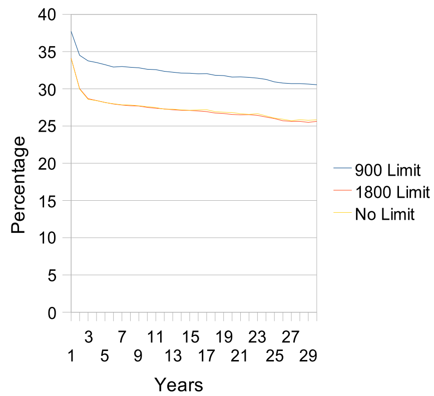
c) Gini index



d) Fraction of cropland of total area



e) Percentage of cash crops of total rainfed area



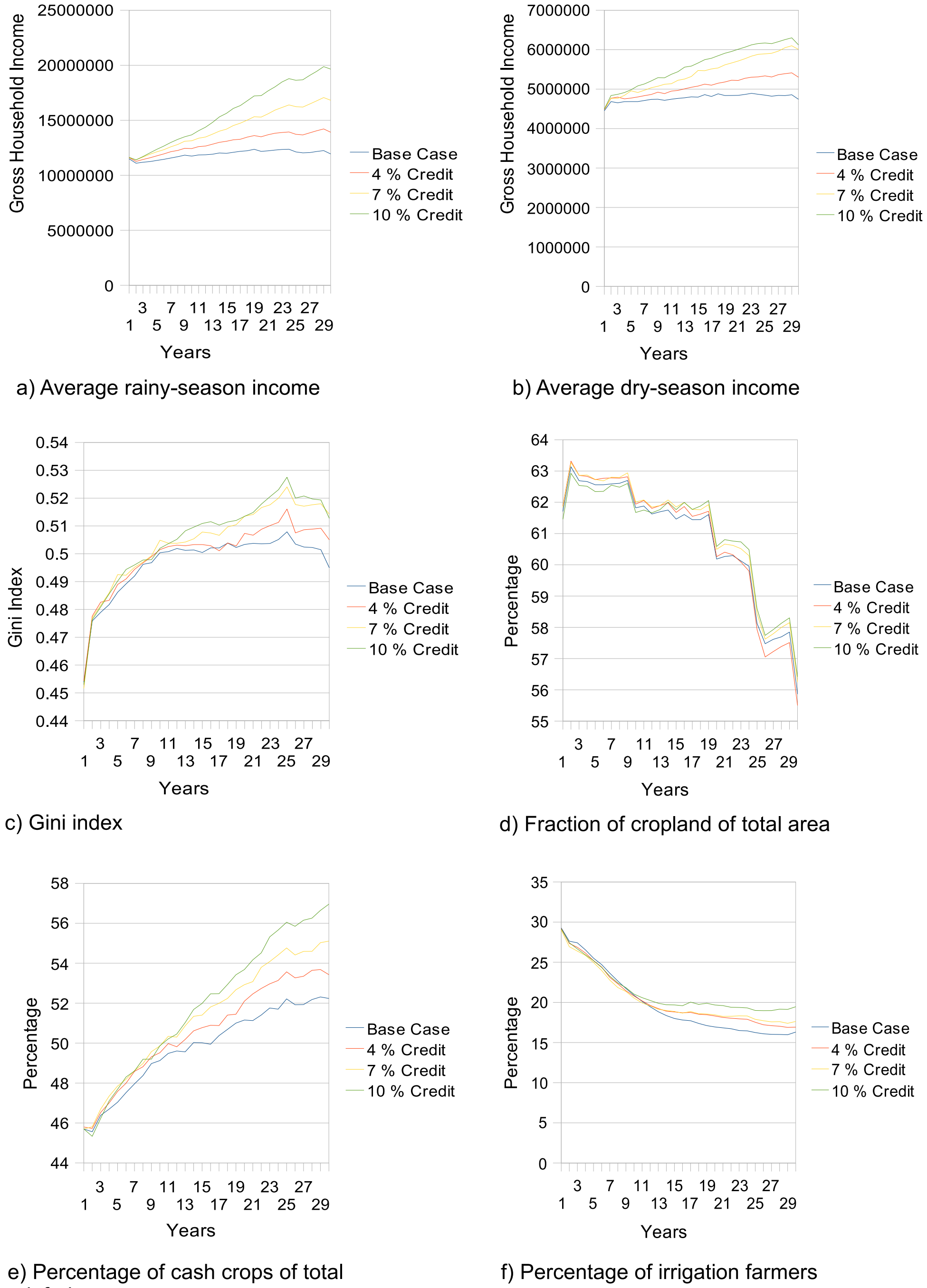
f) Percentage of irrigation farmers

Figure 6.10: Area limitation scenarios: Simulated land-use/cover and socio-economic changes (see Tables A.7 to A.12 for means and uncertainty ranges)

while for the S-Cred10 scenario this amount reached astonishing 19.6 ± 0.383 million Cedis. In the dry season, the situation is similar, but less pronounced (Figure 6.11 b). The S-Cred10 scenario led to an average gross dry-season income of 4.7 ± 0.003 million Cedis in 2036, in contrast to 6.1 ± 0.08 million in the baseline scenario. The uncertainty ranges for all four scenarios were distinct for both seasons.

The most remarkable point here is that income seems to be much more sensitive to the policy of credit access than to that of dam construction, as described above. Credit access, as the much cheaper policy intervention compared to the establishment of dam infrastructure, seems to have a much higher impact on income generation. An annual credit access percentage of 10 % would result in a maximum total annual expenditure of US $ 12.800, under the assumption that none of the credits are settled, which is an unrealistic assumption in an area where nearly 95 % of the credits are repaid. The construction of dams on the other hand would cost millions of US $, which poses the question whether such a policy is cost-effective and efficient enough to be justifiable. However, from the viewpoint of income equity, dam construction might be regarded as the more desirable intervention in terms of the equity in income distribution, as represented by the Gini Index. Improvements in credit access in the study area have the unfavorable characteristic of leading to higher income inequity (Figure 6.11 c). For 2036, the Gini Index in the baseline scenario amounts 0.495 ± 0.008 as compared to 0.513 ± 0.003 for S-Cred10. This increased inequity is reflected by an increased income gap between low-income and high-income farmers. The single-run simulation showed that in the S-Cred10 scenario the high-income class was able to more than double their average annual gross income during the simulation period, while the low-income class could increase their income by only 3 %. This skewed pattern may be caused by the increased ability of the high-income class to invest in highly profitable activities (e.g. irrigation, trading), compared to the lower class, which is usually not involved in these businesses and often reliant on low-profit activities (e.g. arts/crafts).

In the long term, correlations among income generated by non-farm activities and agriculture suggest that profit made from investments in these activities is reinvested not only in the same activities, but also in irrigation farming and cash cropping. While in 2036 for the baseline the percentage of land cultivated with cash crops in the rainy season amounted to 52.2 ± 0.38 %, the S-Cred10 scenario resulted in a significantly higher percentage (56.9



a) Average rainy-season income

b) Average dry-season income

c) Gini index

d) Fraction of cropland of total area

e) Percentage of cash crops of total rainfed area

f) Percentage of irrigation farmers

Figure 6.11: Credit access scenarios: Simulated land-use/cover and socio-economic changes (see Tables A.13 to A.18 for means and uncertainty ranges)

± 0.99 %) (Figure 6.11 e). A significant change in cropland was not observed (Figure 6.11 d), as the difference between the baseline scenario and S-Cred10 is low (0.6 %) and the uncertainty ranges overlap. The reason might be that increased credit access only seems to give an incentive to modify the cropping pattern, but not to extend farmland in general. Furthermore, the improved financial situation of local households generated by cash cropping and non-farm activities seems to be an incentive with respect to involvement in irrigation farming. The decreasing trend in the number of irrigation farmers is significantly alleviated by credit access improvement (Figure 6.11 f), as in 2036 16.3 ± 0.11 % of households are engaged in irrigation farming, while the value is as much as 19.5 ± 0.72 for S-Cred10. This difference is due to the fact that many farmers now can afford going into this business. These farmers usually prefer the low-cost alternative bucket irrigation, which is also reflected by corresponding data of irrigation technology use. In the S-Cred10 scenario, 48.2 ± 0.03 % of the irrigated area is still irrigated by buckets in 2036, while the value for the baseline is only 5 ± 0.01 %. This process of the involvement of irrigation newcomers also reduces the effect of increasing average irrigated area, as the proportion of bucket irrigation, which allows the cultivation of only small areas, is higher than in the baseline.

Impacts of rainfall change on land-use/cover and socio-economic status

In this family of scenarios, the effect of changes in rainfall is tested on system performance, where the annual changes in rainfall represent the four IPCC SRES scenarios, A1, A2, B1, and B2. A2 is the scenario with the least reduction in annual rainfall, followed by B2, A1 and A2 in this order. From the results (Figure 6.12), it is evident that the system performance changes between the B2 and A1 scenarios (e.g. Figure 6.12 a and Figure 6.14), although their annual rainfall reduction values are close. The results, although not significant in terms of uncertainty range overlap, suggest that a slight change can trigger a readjustment of the system's functioning. In the following, we will analyze the data behind this sudden system change between the B2 and the A1 scenario.

What is remarkable is the fact that average income from cultivation is lowest for B2 and highest for A1, although reduction in average annual rainfall is similar, i.e. 2.87 mm for A1 and 2.48 mm for B2. That is, the difference in annual reduction is only 0.39 mm, amounting to only 12 mm difference after 30 years, which is the simulation period. However,

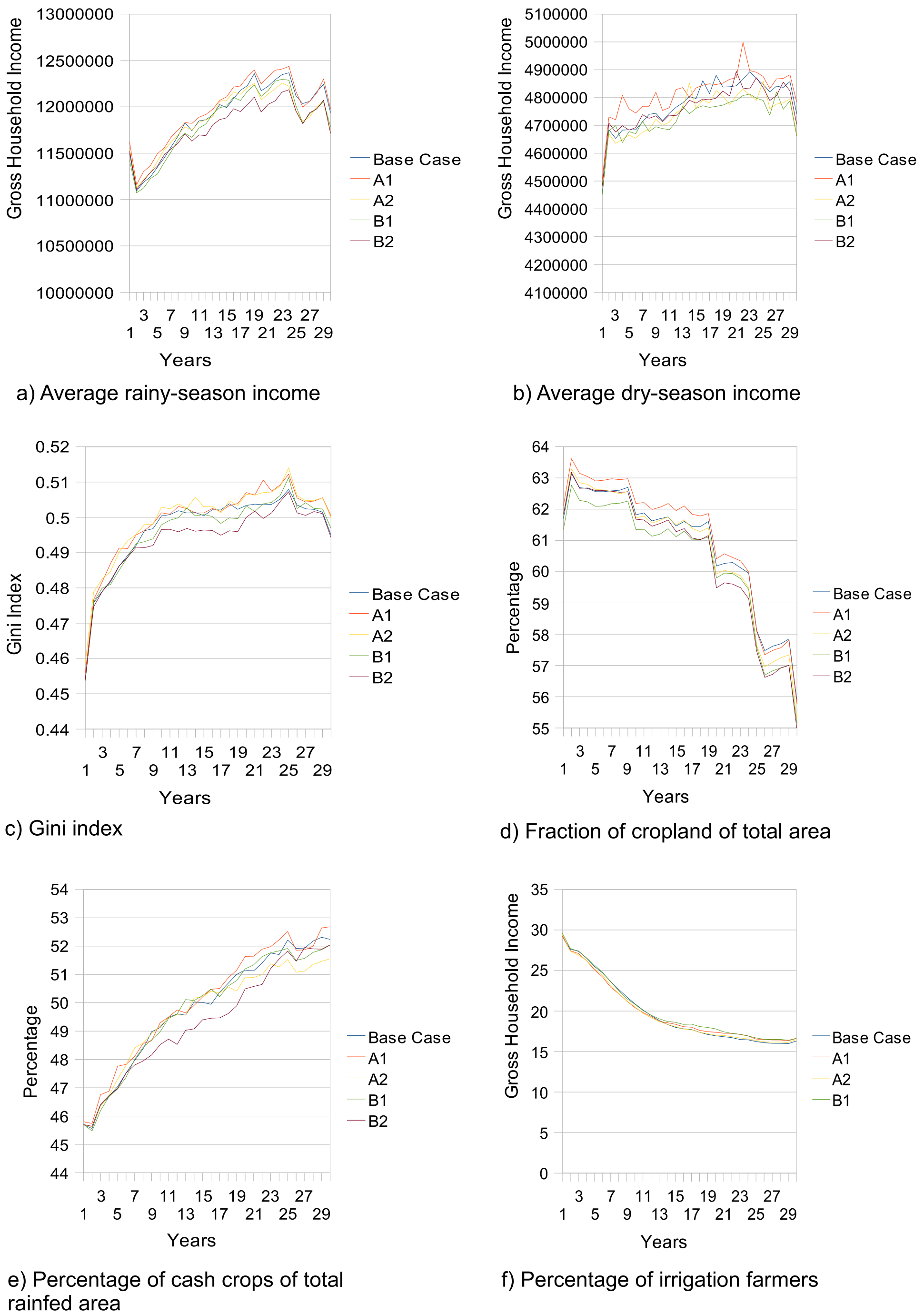


Figure 6.12: Rainfall change scenarios: Simulated land-use/cover and socio-economic changes (see Tables A.19 to A.24 for means and uncertainty ranges)

it seems that this small reduction triggers a changed system behavior. Although this change of system behavior can be hardly analyzed due to its complexity, at least the general causes for this increased cultivation-based income can be stated. In general, increased income from cultivation is due to i) higher yields, ii) the cultivation of more valuable crops, or iii) and extension of cropped area, or a combination of these. As yields are even lower for the A1 than for the B2 scenario (e.g. Figure 6.13), the income surplus must be caused by a shift to more valuable crops or an enlarged cropping area. Although not significant, results suggest that both mechanisms seem to be activated in the A1 scenario. In 2036, the percentage of cropland cultivated with cash crops is 52.7 ± 1.2 % for A1, while the value is only 52.0 ± 0.6 % for B2. Accordingly, the percentage of cropland is larger in the A1 scenario with 55.8 ± 0.93, as compared to 54.9 ± 0.92 % in the B2 scenario. This tendency might also be one of the causes for the slighlty higher average income during the rainy season, which is 12.0 ± 0.296 million Cedis for A1, as compared to 11.7 ± 0.195 million Cedis for B2 (Figure 6.12 a). As far as income equity is concerned, a slight difference in Gini Index in 2036 can be observed between the A1 scenario (0.5 ± 0.008) and the B2 scenario (0.494 ± 0.004) (Figure 6.12 c). The single-run simulation revealed that the higher Gini Index for the A1 scenario is caused by a thinning of the middle class, resulting in larger fractions of high- and low-income farmers. The more subtle reasons for this mechanism could not be revealed due to the high model complexity. Irrigation activities did not seem to be affected by decreases in rainfall (e.g. Figure 6.12 f).

Impacts of population growth on land-use/cover and socio-economic status

In this scenario spectrum, the impact of different population carrying capacities on land-use/cover and socio-economic indicators is explored. As visualized in Figures 6.15 a and 6.15 b, higher numbers of total households seem to have a negative influence on average household income, in the dry as well as in the rainy season. In the S-Pop9600 scenario, average rainy-season income amounts to 10.0 ± 0.258 million Ghanaian Cedis in 2036, while the value for the baseline scenario is higher (11.9 ± 0.176 million Cedis). The situation in the dry season is similar, with a seasonal average gross income of only 3.83 ± 0.026 million Cedis for S-Pop9600 in 2036, as compared to 4.74 ± 0.03 million Cedis for the baseline scenario. In both seasons, this declining trend is mainly related to a decline in average cultivated area,

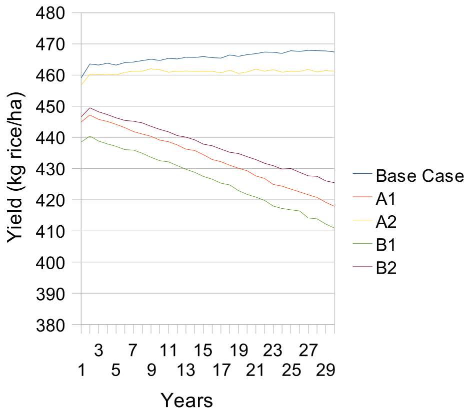


Figure 6.13: Average yield from the compound farming system

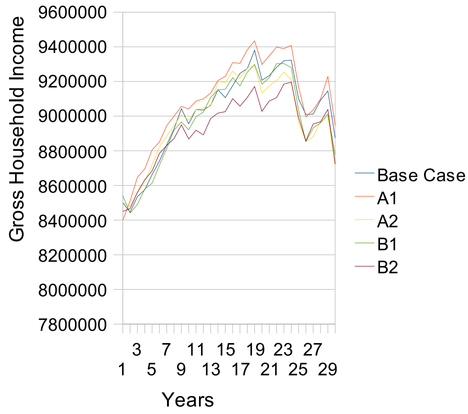
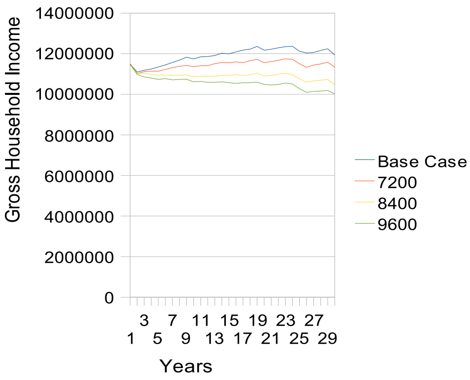


Figure 6.14: Average gross income from rainfed cultivation

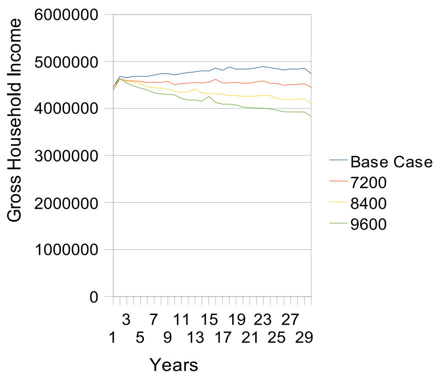
as cultivation is by far the most important contributor to household income. For the whole simulation period, the percentage of rainy-season income generated by cultivation reached as much as 77.5 and never fell below 70 %.

Average cultivated area in the rainy season showed a decline from 15400 m^2 in 2006 to 10 400 ± 89 m^2 in 2036 in the S-Pop9600 scenario, whereas in the baseline scenario this amount is only reduced to 13 900 ± 188 m^2 in 2036. Given a similar situation in the dry season, in both seasons limited available land was identified to be the main cause of this trend. While in the dry season most of the irrigable land had already been put under cultivation before 2006, arable land in the rainy season still seemed to be available, but remoteness and large distances were supposed to impede their cultivation.

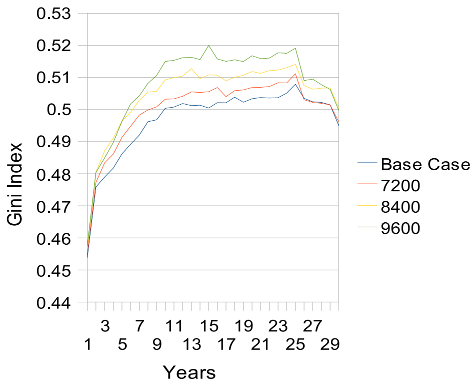
Results also suggest a higher trend of the Gini Index for the population-based scenarios in comparison to the baseline (Figure 6.15 c), being 0.495 ± 0.008 for the baseline and 0.5 ± 0.012 for S-Pop9600, although the difference between these two values is not significant due to overlapping uncertainty ranges. However, this trend is underpinned by the single-run society composition analysis among the three income classes as defined above. While in the S-Pop9600 scenario, average annual income of the high-income class experienced an increase of 16.4 % during the simulation period, the values for the middle- and lower-income class were negative. The reason for this inequity can be traced back to the increasing bias in land tenure between the low- and high-income class for the population-based scenarios. While the



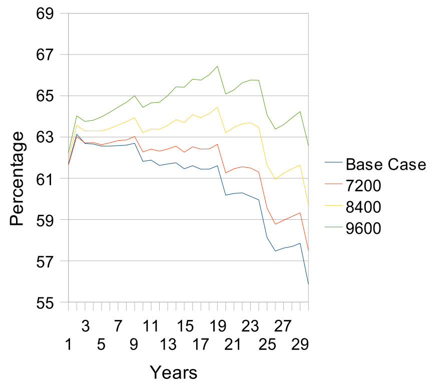
a) Average rainy-season income



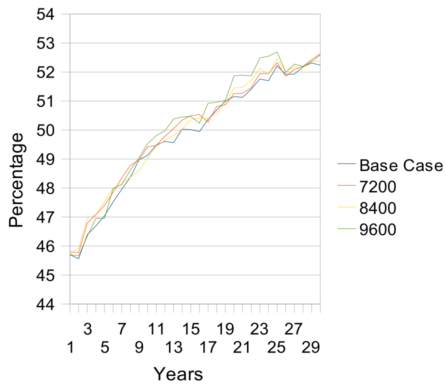
b) Average dry-season income



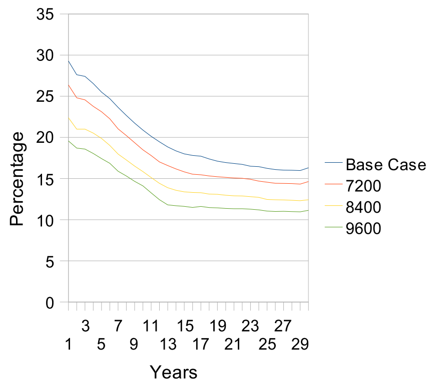
c) Gini index



d) Fraction of cropland of total area



e) Percentage of cash crops of total rainfed area



f) Percentage of irrigation farmers

Figure 6.15: Population growth scenarios: Simulated land-use/cover and socio-economic changes (see Tables A.25 to A.30 for means and uncertainty ranges)

rainfed cultivated area among high-income households remained stable in the simulation period, the area for households from the low-income class decreased dramatically. This extreme bias might be caused by the same mechanism as described in the description of the baseline scenario above. Farmers at the lower end of the income gap usually have more offspring and much less land than those at the upper end, which results in a severe fragmentation of land in this lower class, whereas the relatively abundant land of better-off farmers is usually divided among a few number of inheritants. Increasing population numbers thus amplify this mechanism.

Due to this lack of available land, the scenario spectrum showed that local households found strategies to reduce their dependency on cultivation, especially among households of the lower class. Cash cropping did not seem to be an alternative (Figure 6.15 e) as the difference of % land cultivated by cash crops did not vary significantly among the scenarios (i.e. 52.2 ± 0.38 % for the baseline scenario and 52.6 ± 0.94 % for S-Pop9600). The lack of incentives to invest in cash cropping might be caused by the low level of land availability and the fact that many households remain partly reliant on subsistence crops, leaving little land for cash cropping. As far as the trend of cropland in the S-Pop9600 scenario is concerned, the general decreasing trend in rainfed area in the baseline scenario is overlain by an upward trend in the population-based scenarios, caused by the increasing population size (Figure 6.15 d). For all four scenarios the results are significant in the sense that their uncertainty ranges are distinct.

According to the population-based scenarios, increased population numbers automatically led to a significantly lower fraction of irrigation farmers (Figure 6.15 f), being 11.4 ± 0.31 % for S-Pop9600, while being 16.3 ± 0.09 % for the baseline. This difference is mainly due to the fact that the irrigable area, which had already been almost fully reclaimed in 2006, can only sustain a limited number of households. The single-run society composition analysis for the S-Pop9600 scenario showed an extreme shift of these households to the high-income class during the simulation period, which finally cultivated 55 % of the irrigated area, but only accounting for 13 % of the population. The concentration of the irrigation business in the upper class caused by the increasing income gap also led to an accelerated spread of pump irrigation technology, as the percentage of newcomers from the lower classes remained low.

7 SUMMARY AND CONCLUSIONS

7.1 Summary

Land-use and land-cover change (LUCC) is a world-wide phenomenon, with one third to one half of the terrestrial surface already transformed by human actions (Vitousek et al., 1997). LUCC is further an integral part of global and local webs of environmental processes, being related to processes such as the hydrological cycle, climate change, land degradation and biodiversity loss. These processes may result in changes in global and local land and water resources, having immediate consequences for farming households who directly depend on the natural resource base. This interplay between human actions and the natural resource base is a vulnerable system, which is why a proactive instead of a reactive land management approach is needed to avoid damage to the ecosystem in advance. The understanding and anticipation of future land-use and land-cover change can provide a basis for such proactive land management, by trying to find strategies to mitigate future adverse impacts and possibly improve the sustainability of resource use. However, studies on LUCC processes are often challenged by their complex nature and unexpected behavior of both human and environmental drivers. The aim of this study is therefore to develop an integrated local model for a small-scale catchment in Upper East Ghana that enables policy-makers and other stakeholders to explore alternative scenarios that can improve rural livelihoods and their interplay with the environment.

The first chapter of this thesis generally clarifies how the agent-based modeling approach can be a useful tool to capture the complexity of LUCC processes, and why we used this approach for our purposes. First, the terminology of land-use and land-cover change processes is introduced, followed by an outline of typical LUCC processes (e.g. deforestation). The description of these processes give a first insight into the complex nature of LUCC processes, which is then further analyzed in detail, including the problem of scale dependencies, socio-ecological heterogeneity, interdependencies among system components, and emergent properties. The analysis shows that the complex nature of the coupled human-environment system poses great methodological challenges for LUCC modeling. To analyze the capability of current LUCC modeling approaches to capture this complexity, the most common modeling traditions are described, including a detailed analysis of their ability to represent different aspects of the complex behavior of LUCC systems. Based on this analysis and our aim to develop a small-scale decision support tool, it is finally argued that the agent-based modeling approach is the most suitable approach for our purposes.

The next chapter is dedicated to the conceptualization of the agent-based model, named GH-LUDAS (GHana - Land Use DynAmic Simulator), which will serve as a basis to project alternative pathways of LUCC into the future. First, the concepts underlying the multi-agent based approach are clarified in order to understand the further steps of model conceptualization and model implementation. Agent-based modeling in general aims at describing systems as being composed of an environment and agents located in this environment, which are endowed with automous reactive behavior templates and relations among each other. Based on this multi-agent philospshy, the conceptual framework for simulating LUCC is proposed, in which the human population and the landscape environment are represented as self-organized interactive components. The biophysical system is considered at the level of landscape agents, i.e. heterogeneous land patches with their own attributes and ecological response mechanisms with respect to environmental changes and human interventions. The human system is considered in terms of household agents, i.e. heterogeneous farm households with their own characteristics and decision-making mechanisms regarding land use. Interactions between household and landscape agents occur mainly through tenure relations and a perception-response loop. The perception-response loop involves information flows between households and patches. The information flowing from household to landscape agent reflects the decisions made by the household regarding land-use on the patch (e.g. labor input, land-use type, etc.). The information flowing from patch to household includes changes in the biophysical state (e.g. land use, land cover) and the benefits the household derives from its decisions (e.g. yield). These changes and benefits are regulated by the internal ecological response mechanisms of the single patches. Apart from the human and environmental components of the system, a third component is integrated, consisting of the external parameters regulating policy options and other macro-drivers, which directly influence system behavior through modifying household and/or patch attributes.

In Chapter 3, the theoretical specification of GH-LUDAS is outlined in detail, on the basis of the general conceptual framework previously defined. For this purpose, the GH-LUDAS framework is divided into four main modules that represent the main components of the coupled human-environment, i.e. the Human, the Decision-making, the Landscape, and the Global-policy Module. The Human Module defines specific behavioral parameters and patterns of farm households (i.e. human agents) in land-use decision-making according to typological livelihood groups. The Landscape Module characterizes individual land patches (i.e landscape agents) with multiple attributes and biophysical/natural processes representing the dynamics of crop yield, livestock and land-use/cover transitions. The Global-policy Module consists of the architecture describing how policy and other external parameters are integrated in the Human/Landscape Modules. Finally, the Decision Module, although an integral part of the Human Module, is discussed separately, due to its complicated architecture, which integrates household and environmental information into land-use decisions. This chapter provides a transparent model description, such that the internal mechanisms can be easily retraced. The speciÞcations of the model thus focused on the system architecture, describing the set of variables for each module and their interlinkages, and the system implementation, including an outline of the simulation protocol for this architecture.

The fourth chapter is dedicated to the specification and calibration of the decision-making sub-models. The choice of variables used for these sub-models needed to be based on field experience, mental models to avoid biases in variable selection, and literature describing typical variable-process relationships. To support the justification of the range of variables used, a detailed description of local living conditions and agricultural behavior is given. Based on this information and the livelihood framework proposed by Ashley and Carney (1999), meaningful indicators describing the differences in typical local livelihood typologies are identified. It is a common assumption that land-use decisions are related to the livelihood strategy of a farming household, thus the diversity of agents regarding land-use decisions can be achieved by a categorization of these agents into groups, each having an individual livelihood strategy. This categorization was carried out in two sequential steps, starting with a Principal Component Analysis (PCA) to condense the range of selected livelihood indicators into a smaller set of 'core variables', and, based on this core set, a k-mean Cluster Analysis (k-CA) was applied to derive categorical household groups.

The decision-making sub-models represent choices among discrete sets of options (e.g. choice among land-use types, choice of irrigation technology), consisting of multinomial logistic regression models, based on selected household and landscape attributes.

The coefficients for these models were calculated for each household group separately using statistical methods. The differences in coefficients represent the preferences of a particular group towards certain options of choice, thereby reflecting its general livelihood strategy. The multinomial logistic regression models were implemented in GH-LUDAS by using selected household and landscape variables and the group-wise preference coefficients, to calculate probabilities for each land-use choice option for each household. Furthermore, a routine was programmed to reallocate households to specific household groups in each time step, based on the livelihood indicators as previously specified. As the values of these indicators among households can change during time, this routine enables households to change into that group that best represents their livelihood strategy, thus changing their general land-use preferences. The methods used for this household decision-making study could capture considerable heterogeneities in land-use choice behavior, and rigorously parameterized these heterogeneities. In general, households choose land use based on the considerations of a range of personal characteristics, natural conditions of the environment, and particular policy factors. The developed model of land-use choice thus provides a basis for coupling the human and the environment system under particular policy circumstances when simulating land-use changes.

In Chapter 5, we present the calculated and derived spatial attributes of the study area, and the specified and calibrated ecological sub-models. Following a detailed description of natural and biophysical conditions of the landscape, we calibrated the heterogeneous landscape environment using GIS-based analysis and digitized maps. Because the path-dependent nature of land-use/cover changes requires careful and accurate calibration of land use/cover, current land-use/cover data were extracted from fine-resolution satellite images (ASTER), based on ground-truth points collected in the study area. Each landscpae agent was subsequently assigned a land-use/cover type based on the extracted map, representing the state in the base year 2006. Other environmental features were derived from existing databases and digitized maps such as topography, soil classes, groundwater level, and proximity variables (e.g. distance to river). All attributes were assumed to remain static over time, although a subset of them could possibly be subjected to long-term changes, such as soil attributes and groundwater level. However, due to a lack of reliable local data, it was not possible to integrate such processes in GH-LUDAS.

Furthermore, we developed ecological models that were built into the landscape agents to enable them to respond to environmental changes and human interventions. Empirical/statistical sub-models were developed and calibrated to calculate productivity levels of landscape agents. These yield functions work in response to agricultural input (e.g. labor, fertilizer), regulated by the decision-making procedure of the household agent, to biophysical attributes of the landscape agent (e.g. soil fertility), and to long-term changes in rainfall. These changes in rainfall are also integrated in the second type of ecological sub-model, i.e. the livestock dynamics sub-model. Based on a model developed by Stéphenne and Lambin (2001), the relationship between livestock population and biomass production under different rainfall patterns was established. Within this model, the calculation of biomass productivity is directly related to annual rainfall, which regulates the total population of livestock in the catchment in terms of tropical livestock units (TLU). This way, households are subjected to annual fluctuations in terms of their livestock assets, which have indirect conseqeunces on their livelihood stratgey and land-use behavior. Finally, a land-cover transition model was developed to regulate the balance between grassland and cropland. Once cropland is abandonded, it is converted to grassland after a certain period of time, which was set empirically. These sub-models, which calculate crop productivity, livestock dynamics and land-cover changes, are directly linked to the Human Module, as their results are perceived by single household agents and integrated in their decision-making routines. The interaction between decision-making and ecological reponse thus leads to an annual time-loop, which has the ability to change dynamically over time.

Summarizing, by building and calibrating sub-models for household and landscape agents, we represent the human-environment in a dynamic, adaptive and realistic manner. By defining the attributes and reactive behavior of the single entities of the coupled human-environment system of LUCC, the temporal and spatial pattern of land-use/cover change emerges from the dynamics of the interplay of the single entities. Thus, this approach does not seek to impose the nature of complexity at the top level of the system, but rather tries to let complexity emerge from the interactions of low-level entities and components. Therefore, the calibration and parameterization of agents and their reactive behavior needed to be addressed with utmost care. The range of variables and the range of most important processes involved were identified and analyzed on the basis of field experience, statistical methods and related literature. Household agents and landscape agents were parameterized based on data collected in 2006 in the study area, with the aim of representing human behavior and environmental response as realistically as possible.

The model framework (Chapters 3 to 5), was finally programmed in NetLogo, a multi-agent modeling platform, to produce the operational GH-LUDAS with functionalities of a decision support tool. The setting of external parameters allows the simulation of alternative pathways into the future, comprising parameters for dam construction, credit access, climate change, and changes in demography. While parameters of the policy-related factors of credit access and dam construction directly modify household and landscape attributes, climate change regulates the productivity of crop and biomass, thus influencing land-use behavior indirectly. Characteristics of population growth can be set by the user to define the dynamics of the number of households during time, which have indirect consequences on land and water availability for single households, triggered by increased population pressure on these resources. Through case-specific settings of these external parameters, future scenarios of land-use/cover change can be explored. Simulation outputs include a spatially explicit map of land use/cover for the catchment, graphs indicating the temporal performance of land use/cover and living standards on catchment level (e.g. average income, Gini Index), and spreadsheets of selected indicators of system performance, which can be exported to other data processing sofwares. This way, the results of selected scenarios can be compared and analyzed.

The identification, simulation and analysis of selected scenarios was thus the main focus of Chapter 6, as well as a presentation of GH-LUDAS as a decision support tool. The realtively easy handling of the model interface allows stakeholders to use GH-LUDAS as a simulation tool and a platform for communication among involved stakeholders, who do not necessarily need to understand the model code. Furthermore, integrated scenarios were developed for different (policy) options, with the purpose of identifying the range of possible future pathways triggered by policy and other external factors (policy-related purpose), and of identifying the main mechansims leading to these specific pathways of livelihood and land use (scientific purpose). First, we analyzed the environmental and policy-related conditions in the study area, and justified the selection of the range of external parameters of GH-LUDAS. With the support of this analysis and GH-LUDAS, we conducted the scenario development in a systematic and organized manner. First, we defined a baseline scenario, reflecting the policy settings as they were in 2006, and assumed no changes in climate or demography. This baseline scenario was then used to compare the pathways of other hypothetical scenarios with that of the baseline. For this, each external factor was shifted from the baseline gradually to form a scenario spectrum to assess the impact of the change in this single factor. Among others, simulation results suggest that the policy of dam construction was much less effective with respect to average annual income than that of credit provision, although it was the much more costly option in comparison to a credit scheme. Furthermore, a decline in annual rainfall seemed to trigger a shift towards cash cropping and non-farm activities, which could compensate for the losses in harvest caused by decreased precipitation.

7.2 Limitations

This first version of GH-LUDAS certainly has limitations. First, social interaction among household agents has been implemented only to a limited extent. Although neighborhood effects in the dissemination of knowledge about irrigation technology are included, other social processes are ignored, such as conflict, negotiation and competition. Competition for land resources has only been implemented indirectly through land tenure and lending, and not through direct negotiation among involved households. Such direct household-household interactions were not included, as they would require the modeling of social networks. In the study area, family networks and village affiliation play an important role in the interaction among households with respect to granting usufructuary rights on land or denying them (cases of conflict). Both cases were observed in the study area. However, the identification of realistic social networks as well as the quantification of the more qualitative benefits farmers gain from network membership is an almost impossible task. Furthermore, the networking of household agents would have meant a tremendous reduction of the computation speed of GH-LUDAS.

Second, the model cannot be transferred to other areas easily. Even within similar areas, the range of land-use and land-cover types could be different, and the decision-making and ecological sub-models needed to be area-specific. Only the basic framework of GH-LUDAS could be reused, but the range of variables and the calibration of the sub-models should undergo a detailed assessment. An accurate mapping of attributes of the biophysical landscape as well as a detailed household survey would be required, as the variables and architecture of GH-LUDAS remain rather case-specific.

The third drawback of GH-LUDAS is the assumption of static market prices. Market prices for all crops and livestock species were derived from data collected in 2006, which remain identical during the entire simulation period. However, market prices surely undergo long-term changes, due to changing global, regional and local demand-supply relations. The modeling of these processes thus would require the use and integration of global and local economic models with respect to the local goods of the study area. This integration of economic models as well as the development of local models would require intensive studies, and were beyond the scope of this thesis.

Fourth, a land suitability analysis for dam construction was not carried out, due to the limited time frame of the study. Within GH-LUDAS, the choice of the location of inserted dams is not supported by a land suitability map, but requires the knowledge of experts. A land suitability map would require in-depth knowledge of geological, pedological and hydrological data and processes, which is available only to a limited extent. Furthermore, a simulation-based analysis showed that results on catchment level were not significantly influenced by changes in dam locations, although locally, the impacts were significant.

The final drawback of the model, and maybe the most substantial, lies in the difficulty of the validation of model results. Actually, the validation of agent-based models is currently still a debated issue. While classical validation methods, e.g. sensitivity analysis and comparing simulated data with observed data, have turned out to be unsuitable for agent-based models, a number of validation strategies are proposed (see Bousquet and Le Page, 2004; Parker et al., 2003) and debated.

7.3 Recommendations

Since no model is universally appropriate, GH-LUDAS should undergo version-by-version improvements, and the first version as proposed in this study does not claim to represent the real-world human-environment system in the most realistic and fully integrated manner. However, due to the model's high flexibility, several methodological extensions regarding human decision-making and ecological processes can be easily integrated. Each of these extensions should aim at a more realistic representation of the LUCC system, although there should be a lower limit to the detailedness of the model. The selection of such extensions should thus be guided by finding a balance between a too coarse and a too detailed representation of involved processes. Furthermore, each version of GH-LUDAS, including the current one, should be validated in order to improve its credibility for decision-making support and scientific purposes. In the following, we give recommendations for methodological extensions and validation techniques for the current GH-LUDAS version.

- One of the most important processes that have not yet been integrated in the current version is the process of land degradation. Severe land degradation has been observed during the past decades in the Sudan-Savannah zone (of which the study area is part), which are the result of natural processes such as soil erosion and climate change, as well as of human-induced loss of soil fertility. Over-cultivation, over-grazing, lack of application of fertilizer and conservation measures, and reduced fallow periods have led to soil nutrient loss and decreasing agricultural production levels. Maps and models of spatial soil erosion patterns have been developed by ZEF staff, which offer a possibility to link soil erosion with land-cover change (e.g. conversion of grassland to bare land) and crop production. However, the integration of human-induced land degradation would require long-term observations in the study area, in order to establish a sub-model of the long-term consequences of human decision-making on soil productivity, and vice versa.
- The model user should be given the choice among alternative decision-making sub-models in order to explore the sensitivity of sub-model choice on model results. More research should be done on the formulization of different household decision-making strategies to examine whether particular formulizations are appropriate for particular decision-making situations. Knowledge of local decision-making processes as well as model validation should guide the selection of an appropriate decision-making architecture. In contrast to the decision-making approach of bounded rational behavior used in the current version, other approaches may reflect human behavior of local households more realistically, but also may have other shortcomings. One alternative could be the use of the BDI (Belief-Desire-Intention) architecture, which assumes that the

decisions of human agents are guided by their beliefs about other agents and their environment. However, the drawback of this approach lies in the qualitative nature of beliefs, which impedes the quantification of the internal belief structure of a household agent, and his subsequent reactions. However, the range of possible decision-making architectures is manifold, and modelers can usually select freely among them according to the mechanisms they want to focus on.

- A further challenge for the specification of the decision-making architecture is the integration of a learning mechanism. In reality, many decisions are influenced by past experiences, which serve as a basis to estimate future benefits and deliberate among options. In GH-LUDAS, currently no such mechanism is integrated. The k-nearest neighbor algorithm, which is among the simplest of machine-learning algorithms, was experimentally implemented in GH-LUDAS, resulting in a 10-fold decrease in computing speed. The integration of learning mechanisms in GH-LUDAS is thus still impeded by the computing speed of current computers, but this may change in future computer generations.

- Furthermore, as mentioned above, the economic situation in terms of market prices is assumed to remain as it was in 2006. This drawback could be compensated for by integrating at least a global model of future market price fluctuations. IFPRI's IMPACT model (International Model for Policy Analysis of Agricultural Commodities and Trade) could be used to assess future world market prices of a range of commodities until 2025. The model simulates changes in production and demand on the level of regions and single countries, which aggregate to global demand and production functions. Based on these functions, a global demand-supply balance then defines global market prices for each year until 2025. However, deviations from this global market price are often caused by a lack of infrastructure and market information, especially in developing countries, which often lead to local irregularities in commodity prices. The determination of such local price fluctuations for the study area remains a challenge for GH-LUDAS.

- A land suitability map for dam construction could be developed to support policy-makers in their decisions to find suitable locations for dams. Such a map could also support the realism of model results, as the placement of dams would follow realistic assumptions. The understanding of engineering and hydrological processes, which are required to establish such a decision-support map, could also be applied to an estimation of dam water levels due to climatic changes, which have not been considered in the current version of GH-LUDAS.

- The credibility of the model depends on how the internal structure represents the structure of the system modeled. To improve the understanding why the model was built in this way, detailed descriptions of social and environmental conditions and local agricultural behavior have been given, upon which the structure of sub-models and range of variables were grounded. Assumptions underlying the selection of variables have been clearly stated and justified. Furthermore, graphical and narrative descriptions of the model structure were given to enhance the model's lucidity and clarity, and to serve as a basis for expert assessment and comparative model-to-model studies. A documentation of GH-LUDAS will also be available as an ODD (Overview, Design concepts, and Details) protocol, which is a documentation protocol aimed to enhance the description of individual-based models and to convey the structure of the model in a unified manner. Based on this protocol, other scientists will be enabled to retrace and understand the model structure and involved variables.

- The credibility of the model should not only be enhanced by a transparent model description, but also by validation techniques such as hindcasting. With this technique, instead of simulating the future, the model is run for a past period until present, and the results are then compared to the current situation. The major drawback of this approach is the usual lack of past data necessary for model initialization. Therefore, a hindcasting approach could only be based on an approximations of past environmental and household data, which would clearly reduce the power of this method in terms of validation.

- Another validation method lies in the comparison of the results with other types of models for the same area. To validate simulated land-use/cover patterns, a statistical GIS-based model could be developed, which extrapolates observed LUCC patterns into the future. Based on classified images of several past points in time, transition probabilities among land-use and land-cover types can be calculated, which can be further used to project land-use/cover patterns into the future.

7.4 General remarks about modeling

What are the lessons that can be drawn from this modeling effort, and in particular from the use of the agent-based approach? The answer is far from clear. In general, it can be said that modeling is one of those scientific areas which experience the most criticism and distrust, especially from the non-modeling community. Especially among the social sciences, models of human behavior are often regarded as unrealistic and simplistic in their assumptions. For these reasons, what seems to be the most important challenge for a modeler is, apart from the process of model building, the justification of the assumptions he/she made about the model. With respect to the fact that many models have the reputation of claiming to be universally valid, three things have to be mentioned. First, at least in the science of land-use/cover change, it is widely acknowledged that the understanding of involved processes should rely not only on modeling efforts, but also on narrative descriptions and mental models. None of these approaches should be considered to be superior. Computer as well as mental models should be regarded as tools to improve future generations of both types of models. The second important issue is that there is the frequent misconception that models are hierarchically ordered in terms of their realism. Different types of models are built for different types of purposes in order to examine different types of problems. Each model has its limitations, and it is easy to accuse models of neglecting some part of reality. Third, models, at least in LUCC sciences, are rarely built completely objectively. The understanding of the modeler about the system functioning is absolutely necessary, but also implicates a partly subjective view of the system on the side of the modeler. Although the assumptions underlying the model should be based on objective reasoning, a trace of subjectivity can never be eliminated.

With respect to the use of the agent-based approach for studying LUCC, several aspects are important. If a realistic policy-related model is the target of a study, as it was in this case, it can be difficult to model the effects of hypothetical processes and policy interventions. As realistic agent-based models are usually based on statistical evaluation of real data, only those processes can be modeled that can be measured at the point of data acquisition. In other words, processes or reactions that do not take place in the study area (to some extent) can hardly be simulated, e.g. the adoption of a totally new crop in the future or the effects of hypothetical policy interventions. Furthermore, if an explicit representation of decision-making households is desired, multi-agent models usually need to be confined to small study areas, and the transferability of model results to other areas remains limited. Such case-specific models are often very data demanding, which results in extensive efforts of data acquisition, statistical evaluation and model programming. Some former agent-based scientists, among them LUCC scientist Couclelis (2001), doubt that the gain that can be derived from these types of models compensate for the high effort need to develop them. These major drawbacks are often ignored in agent-based LUCC studies, and this agent-based modeling is often acclaimed as a new paradigm to model LUCC. It is therefore noteworthy to mention that, the many strengths of the agent-based approach notwithstanding, modelers should be aware of the limitations in the applicability of this approach.

8 REFERENCES

Achard F, Eva HD, Stibig HJ, Mayaux P, Gallego J (2002) Determination of deforestation rates of the world's humid tropical forests. Science 297 (5583): 999-1002

Adamides ED, Tsalides P, Thanailakis A (1992) Hierarchical Cellular Automata Structures. Parallel Computing 18 (5): 517-524

Adu SV (1969) Soils of the Navrongo-Bawku Area, Upper East Region, Ghana. Soil Research Institute, Kumasi, Ghana

Ansoglenang G (2006) Rural Women and Micro-Credit Schemes: Cases from the Lawra District of Ghana. Master Thesis, Faculty of Social Science, University of Troms¿, Norway

Arthur WB, Durlaf SN, Lane D (1997) The Economy As an Evolving Complex System II. Addison-Wesley, Reading, MA

Ashley C, Carney D (1999) Sustainable livelihoods: Lessons from early experience. London: Department for International Development

Ashong K, Smith DR (2001) Livelihoods of the Poor in Ghana: A Contextual Review of Ghana-wide definitions and trends of poverty and the poor with those of peri-urban Kumasi, http://www.livelihoods.org/static/ kashongN131.html

Axtell R (2000) Why agents? On the varied motivations for agent computing in the social sciences. The Brookings Institution Center on Social and Economic Dynamics Working Paper 17, *http : //www.brookings.edu/dybdocroot/es/ dynamics/papers/agents/agents.html*

Baas NA, Emmeche C (1997) On Emergence and Explanation. Santa Fe, NM: Santa Fe Institute Publication 97-02-008

Ball AS, Pretty JN (2002) Agricultural influences on carbon emissions and sequestration. UK Organic Research 2002: Proceedings of the COR Conference, 26-28th March 2002, Aberystwyth, p 247

Ball J (2001) Soil and Water Relationships, *http : //www.medicagotruncatula.org/Ag /Soils/SoilWaterRelationships/*

Bandte H (2007) Komplexität in Organisationen: Organisationstheoretische Betrachtungen und agentenbasierte Simulation, Deutscher Universitäts-Verlag

Bebbington A (1999) Capitals and Capabilities: A Framework for Analyzing Peasant Viability, Rural Livelihoods and Poverty. World Development 27 (12), pp 2021-2044

Benneh G, Agyepong GT (1990) Land degradation in Ghana. London: Commonwealth Secretariat/University of Ghana. Pall Mall

Ben Said L, Bouron T, Drogoul A (2002) Agent-based interaction analysis of consumer behavior, proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems, Part 1, pp 184-190, Bologna, Italy (2002)

Bharati L, Rodgers C, Martin N, Plotnikova M, Intsiful J (2005). Assessing the efficiency of small reservoir irrigation systems through integrative watershed modelling. WaSiM-Workshop, Institut für Atmosphäre und Klima der ETH Zürich, Switzerland. March 17-19

Birner R, Schiffer E (2005) Water Governance Issues in the Upper East Region, Water Challenge Project Integrating Governance and Modeling, International Food and Policy Research Institute, *https : //www.uni – hohenheim.de/mas/ documents/dec*2005/*ghana_governance*

Blench R (2006) Working paper: background conditions in Upper East region, Northern Ghana, 2005, http://www.rogerblench.info/Development/ Ghana/IFAD/ LACOSREP/Blench%20UER%20working%20paper.pdf

Boadu EA (2000) Rapid Population Growth and Development in Ghana, Population Impact Project , Department of Geography & Resource Development, University of Legon, Accra, Ghana

Bonabeau E (2002) Agent-Based Modeling: Methods and Techniques for Simulating Human Systems. In: Proceedings of the National Academy of Sciences of the United States of America, Vol. 99, No. 10, Supplement 3: Arthur M. Sackler Colloquium of the National Academy of Sciences. Sackler Colloquium of Adaptive Agents, Intelligence, and Emergent Human Organization: Capturing Complexity through Agent-Based Modeling (May 14, 2002), pp 7280-7287

Booth FEM, Wickens GE (1988) Non-timber uses of selected arid zone trees and shrubs in Africa. FAO Conservation Guide 19. Food and Agriculture Organization of the United Nations. Rome. Pp 103-109

Botchie G, Seini W, Asuming-Brempong S (2003) Ghana: National Assessment of the Environmental Roles of Agriculture and the Effects of Policy. FAO Roles of Agriculture Project: Socio-economic analyses and the roles of agriculture in developing countries, f tp : //f tp.fao.org /es/ES A/Roa/pdf /2Environment/EnvironmentGhanaNA.pdf

Boudet G (1975) Manuel sur les pâturages tropicaux et les cultures fourragères. Coll Manuels et Précis Délevages, IEMVT. Ed. du Ministère de la Coopération, Paris

Bousquet F, Le Page C (2004) Multi-agent simulations and ecosystem management. Ecological modeling 176, pp 313-332

Briassoulis H (2000) Analysis of Land Use Change: Theoretical and Modeling Approaches, Regional Research Institute, West Virginia University, The Web Book of Regional Science

Brown RB (2003) Soil Texture. Fact Sheet of the Soil and Water Science Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida. *http : //edis.i f as.u f l.edu/pd f files/S S /S S* 16900. *pd f*

Campbell B, Sayer JA, Frost P, Vermeulen S, Ruiz Pèrez M, Cunningham A, Prabhu R (2001) Assessing the performance of natural resource systems. Conservation Ecology 5(2), pp 22, http : //www.consecol.org /vol5/iss2/art22

Carney D (1998) Implementing the sustainable rural livelihoods approach. In: Carney D. (ed), Sustainable Rural Livelihoods: What Contribution Can We Make? Department for International Development (DFID), London. pp. 3-26

Carswell G (1997) Agricultural intensification and rural sustainable livelihoods: A Ôthink pieceÕ. IDS Working paper 64

Case A (1991) Spatial patterns in household demand. Econometrica 59 (4), pp 953-965

Caviglia-Harris JL, Sills EO (2005) Land use and income diversification: comparing traditional and colonist populations in the Brazilian Amazon. Agricultural Economics 32 (3) , 221Ð237

CGIAR (2000) Empowering farming communities in Northern Ghana with strategic innovations and productive resources in dryland farming, CGIAR Challenge Program on Water and Food, http://www.knowledgebank. irri.org/theme1/ PN6.htm

Charney J, Stone PH (1975) Drought in the Sahara: a biogeophysical feedback mechanism. Science 187, p.434

Cioffi-Revilla C, Gotts NM (2003) Comparative analysis of agent-based social simulations: Gesim and FEARLUS models. Journal of Artificial Societies and Social Simulation 6 (4)

Codjoe SNA (2004) Population and Land Use/Cover Dynamics in the Volta River Basin of Ghana, 1960-2010 - Ecology and Development Series No.15, Cuvillier Verlag, Göttingen

Colenut RJ (1968) Building Linear Predictive Models for Urban Planning. Regional Studies 2, pp 139-143

Costanza R, Ruth M (1998) Using Dynamic Modeling to Scope Environmental Problems and Build Consensus. Environmental Management 22, pp 183-195

Couclelis H (2001) Why I no longer work with agents: A challenge for ABMs of human-environment interactions. In: Parker DC, Berger T, Manson SM (eds) Agent-Based Models of Land-Use and Land-Cover Change: Proceddings of an International Workshop October 4-7, 2001, Irvine, California, USA.

Crutchfield JP (1994) Is Anything Ever New? Considering Emergence. In: Cowan GA, Pines D, Meltzer DE (eds), Complexity: Metaphors, Models, and Reality, pp 515-131. Reading, Mass.: Addison-Wesley

DeAngelis DL, Mooij WM (2005) Individual-based modeling of ecological and evolutionary processes. Annual review of Ecology, Evolution and Systematics 36, pp 147-168

Department for International Development (DFID) (2001) Sustainable livelihoods guidance sheets, Section 2. Department for International Development, London. $http$: $//www.livelihoods.org/info/guidance_{s}heets_{p}dfs/section2.pdf$

Department of Geography and Resource Development (1992). A socio-economic survey in the Upper East Region with reference to drought and desertification control in Ghana. Final report submitted to the Environmental Protection Agency of Ghana, University of Ghana, Legon

Dietz T, Millar D, Dittoh S, Obeng F, Ofori-Sarpong E (2004) Climate and Livelihood Change in North East Ghana. In: Dietz AJ, Ruben R and Verhagen A (eds) The Impact of Climate Change on Drylands: With a Focus on West Africa, Kluwer Academic Publishers

DigitalGlobe (2007) QuickBird image for Anayare and Atankwidi catchments taken on March 06th 2007. The product was ordered and purchased by GLOWA-Volta Project. GLOWA-Volta Project Geo-database

Dittoh S (2000) Agricultural land use agreements and derives rights for granting access to farm land in Northern Ghana. Draft Report, DIFID, London

Dobos E (2005) Traditional and modern soil survey technics, 3rd European Summer School on Soil Survey 2005, $http://eusoils.jrc.it/events/SummerSchool\_2005/cd\_rom/SS2005\_Files/1DAY/Survey\_technics\_EDobos\_U.pdf$

Dorfman, R (1979) A Formula for the Gini Coefficient.The Review of Economics and Statistics 61: 146-149

Drogoul A, Corbara B, Fresneau D (1995) MANTA: New Experimental Results on the Emergence of (Artificial) Ant Societies. In: Gilbert N, Conte R (eds), Artificial Societies: the computer simulation of social life, UCL Press, London

Drogoul A, Vanbergue D, Meurisse T (2003) Multi-agent based simulation: Where are the agents? Lecture Notes in Artificial Intelligence 2581, pp 1-15

Easterling WE and Kok K (2003) Emergent properties of scale in global environmental modeling - are there any? In: Rotmans J, Rothman DS (eds) Scaling in integrated assessment, pp 263-292. The Netherlands: Swets & Zeitlinger Publishers

Egler FE (1942) Vegetation as an Object of Study. Philo. Sci. 9, pp 245-60

Ellis E (2007) Land-use and land-cover change. The encyclopedia of earth, $http://www.eoearth.org/article/Land-use\_and\_land-cover\_change$

Ellis J, Galvin KA (1994) Climate patterns and land-use practices in the dry zones of Africa. BioScience 44(5): 340-349

Enge-Rosenblatt O, Haase J, Clauß(2007) Important Characteristics of VHDL-AMS and Modelica with Respect to Model Exchange. In: Fritzson P, Cellier F, Nytsch-Geusen C (eds) Proceedings of the 1st International Workshop on Equation-Based Object-Oriented Languages and Tools, Berlin, Germany, July 30, 2007

Eoyang G (1997) Coping with Chaos: Seven Simple Tools. Cheyenne, Wyoming: Lagumo Publishing

Eoyang G, Berkas T (1998) Evaluation in a Complex Adaptive System. In: Lissak M and Gunz H (eds), Managing Complexity in Organizations. Westport, Connecticut: Quorum Books, an imprint of Greenwood Publishing Group, Inc, 1998

Epstein JM (1999) Agent-based models and generative social science. Complexity 4 (5), pp 41-60

FAO (1985) Guidelines: Land evaluation for irrigated agriculture - FAO soils bulletin 55, Soils Resources, Management and Conservation Service, FAO Land and Water Development Division, Rome, Italy

FAO (2001) Global forest resource assessment 2000 (FRA 2000): main report, FAO For. Pap. 140, FAO, Rome

FAO (2004) FAOSTAT data. FAO, Rome, http://apps.fao.org

Fatoux C, Castella JC, Zeiss M, Manh PH (2002) From rice cultivator to agroforester within a decade: The impact of Doi Moi on agricultural diversification in a mountainous commune of Choi Moi District, BacKan Province, Viet Nam. In: Castella JC and Quang DD (eds.). Doi Moi in the Mountains: Land use changes and farmers' livelihood strategies in Bac Kan Province, Viet Nam. Agricultural Publishing House. Hanoi

Ferber J (1995) Les Systèmes Multi-Agents. Vers une Intelligence Collective. InterEditions, Paris

Ferber J (1999) Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence. Addison-Wesley, Reading, MA

Fernandez LE, Brown DG, Marans RW, Nassauer JI (2003) Characterizing location preferences in an exurban population: implications for agent-based modeling. Environment and Planning B, Special Issue on Complex Systems

Fogel D, Chellapilla K, Angeline P (1999) Inductive reasoning and bounded rationality reconsidered. IEEE Transactions on Evolutionary Computation 3 2, pp 142.143

Folke C, Jansson, Larsson J, Costanza R (1997) Ecosystem appropriation by cities. Ambio 26, 167-172

Furrer R, Sain SR, Nychka D, Meehl, GA (2007) Multivariate Bayesian analysis of atmosphere-ocean general circulation models. Environmental and Ecological Statistics 14 (3), pp 249-266

Gakidou EE, Murray CJL, Frenk J (2000) Defining and measuring health inequality: an approach based on the distribution of health expectancy. Bulletin of the World health Orgaization, 2000, 78 (1)

Gessler PE, Chadwick OA, Chamran F, Althouse L, Holmes K (2000) Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Sci. Soc. Am. J. 64, pp 2046-2056

Gilbert N, Troitzsch KG (1999) Simulation for the Social Scientist. Open University Press, London, UK

Ghana Statistical Service (2000) Ghana Living Standards Survey Report of the Fourth Round (GLSS 4)

GNADO (GIA/NABIO Agroforestry Development Organization) (2000) The Ka-sena Nankana District of the Upper East Region of Ghana, GNADO, Ghana

Goldspink, C (2002) Methodological implications of complex systems approaches to sociality: Simulation as a foundation for knowledge. Journal of Artificial Societies and Social Simulation 5 (1)

Goldstein J (1999) Emergence as a construct: history and issues. In: Emergence 1(1), p. 49-72

Griffin RC, Montgomery JM, Rister ME (1987) Selecting Functional Form in Production Function Analysis. Western J. Agr. Econ.. 12(2):216-227

Groten SME (1991) Biomass monitoring for rangeland management at national level in Burkina Faso, Paper Presented at the EEC-STD3 Conference, Madrid, November 1991

Grübler A (1994) Technology. In: Meyer WB, Turner BL II (eds), Changes in Land Use and Land Cover: A Global Perspective. Univ. of Cambridge Press, pp. 287-328

Gyasi KD (2004) Determinants of Success of Collective Action on Local Commons: An Empirical Analysis of Community-based irrigation Management in Northern Ghana. Dissertation, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany

Haan de C, Steinfeld H, Blackburn H (1997) Livestock and the Environment. Finding a Balance. Report of a Study sponsored by EU, WB, DK, F, G, NL, UK, USA and FAO

Heilig GK (1994) Neglected dimensions of global land-use change: reflections and data. Population and Development Review 20 (4), 831-859

Holland JH (1998) Emergence: From Chaos to Order. Perseus Books, Reading, MA

Huberman B, Glance N (1993) Evolutionary games and computer simulations. In: Proceedings of the National Academy of Sciences Vol. 90, pp 7716-7718

Huigen MGA (2002) Spatially explicit multi-agent modeling of land-use change in the Sierra Madre, Philippines - The MameLuke Project. In Parker DC, Berger T, Manson SM (eds) Agent-Based Models of Land Use and Land Cover Change, LUCC Report Series No. 6, Indiana University

Huigen MGA (2003) Agent-based modeling in land-use and land-cover change studies. Interim Report, Laxenburg, IIASA, 35 pp

Huigen MGA (2004) First Principles of the Mamluke multi-actor modeling framework for land-use change, illustrated with a Philippines case study. Journal of Environmental Management 72, pp 5-21

IFAD (International Fund for Agricultural Development) (2005) Upper East Region Land Conservation and Smallholder Rehabilitation Project - Phase II, Interim Evaluation, *http : //www.ifad.org/evaluation/public_html/eksyst/doc/prj/region/pa/ghana/gh_lacos.pdf*

IPCC (Intergovernmental Panel on Climate Change) (2001) Special Report on Emissions Scenarios (SRES)

Irwin E, Bockstael N (2002) Interacting agents, spatial externalities, and the evolution of residential land use patterns. Journal of Economic Geography 2 (1), pp 31-54

Jame YW, Cutforth HW (1996) Crop growth models for decision support systems. Can. J. Plant Sci. 76: 9-19

Jarke M, Bui XT, Carroll JM (1998) Scenario Management: An Interdisciplinary Approach. Requirements Engineering 3 (3-4), pp 155-173

Kemp-Benedict E (2004) From narrative to number: A role for quantitative models in scenario analysis. In: Pahl-Wostl C, Schmidt S, Rizzoli AE, Jakeman AJ (eds) Complexity and integrated resources management. Transactions of the 2nd Biennial Meeting of the International Environmental modeling and Software Society. Manno, Switzerland: iEMSs

Kenny A, Finn K (2004) Groundnuts (Arachis hypogaea), Women and Development: Impact on Nutrition and Women's Roles in Western Africa, *http : //peacecorps.mtu.edu/resources/studentprojects/groundnut.html*

Kok K, de Koning F, Bouma J (2000) Possibilities of land use modeling at the supra-national scale: the case of Central America, GIS/EM4 163

Kok K (2001) Scaling the land-use system: a modeling approach with case studies for Central America (2001), PhD Thesis, Univsersity of Wageningen

Kontopoulos K (1993) The Logics of Social Structure. New York: Cambridge University Press

Kreznor WR, Olson KR, Banwart WL, Johnson DL (1989) Soil, landscape, and erosion relationships in a Northwest Illinois watershed, Soil Sci. Soc. Am. J. 53, pp. 1763Ð1771

Krider RE, Weinberg CB (1997) Spatial competition and bounded rationality: retailing at the edge of chaos. Geographical Analysis 29 (1): 16-34

Lambin EF, Geist HJ (2006) Land-Use and Land-Cover Change, Local Processes and Global Impacts, Springer 2006

Lambin EF, Geist HJ, Lepers E (2003) Dynamics of land-use and land-cover change in tropical regions. Annual review of environment and resources 28, p 205

Lambin EF, Geist HJ (2002) Global land-use and cover change: What have we learned so far? Global Change Newsletter 46, p. 27-30

Lambin EF, Turner BL, Geist HJ, Agbola SB, Angelsen A, Bruce JW, Coomes OT, Dirzo R, Fischer G, Folke C, George PS, Homewood K, Imbernon J, Leemans R, Li X, Moran EF, Mortimore M, Ramakrishnan PS, Richards JF, Skanes H, Steffen W, Stone GD, Svedin U, Veldkamp TA, Vogel C, Xu J (2001) The causes of land-use and land-cover change: moving beyond the myths. Global Environmental Change 11, p. 261-269

Lambin EF, Baulies X, Bockstael N, Fischer G, Krug T, Leemans R, Moran EF, Rindfuss RR, Sato Y, Skole D, Turner II BL, Vogel C, Nunes C, Augè JI (1999) Land-use and land-cover change implementation strategy. Stockholm, Sweden: IGBP Report No. 48

Lansing JS, Kremer JN (1993) Emergent properties of Balinese water temple networks: Coadaptation on a rugged fitness landscape. American Anthropologist 95 (1), pp 97-114

Larsson H (1996) Relationships between rainfall and sorghum, millet and sesame in the Kassala province Eastern Sudan. J. Arid Environ. 32, 211-223

Le QB (2005) Multi-Agent System for Simulating Land-use and Land-cover Change: A Theoretical Framework and its First Implementation for an Upland Watershed in the Central Coast of Vietnam. ZEF - Ecology and Development Series No. 29, Cuvillier Verlag, Göttingen

Le, QB (2006) Digital Elevation Model (DEM) at 15 and 30m resolutions for Anayare and Atankwidi catchments downscaled from USGS SRTM Elevation data (at the resolution of 92.53 m). GLOWA-Volta Project Geo-database

Le QB, Park SJ, Vlek PLG, Cremers AB (2008) Land-Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human-landscape system. I. Structure and theoretical specification. Ecological Informatics 3 (2008) pp. 135-153

Le Baron B (2001) A builder's guide to agent-based financial markets. Quantitative Finance 1 (2), pp 254-261

Lee C (1973) Models in Planning. London: Pergamon Press

Le Houérou HN, Hoste H (1977) Rangeland production and annual rainfall relations in the mediterranean basin and in the African Sahelo-Sudanian zone. J. Range Manage. 30 (3), 181Ð189

Maack JN (2001) Scenario analysis; a tool for task managers. In: Krueger R, Cassey M, Donner J, Kirch S, Maack JN (eds) Social analysis: selected tools and techniques.

Social Development Paper 36, pp 62-87. Washington D.C., USA: Social Development Department, World Bank

Mander Ü, Jongman RHG (1998) Human impact on rural landscapes in central and northern Europe. Landscape and Urban Planning 41, 3-4, 149-154

Manson SM (2001) Simplifying Complexity: A review of complexity theory. Geoforum 32 (3), pp 405-414

Manson SM (2005) Genetic Programming as an Isomorphic Analog to Bounded Rationality in Agent-Based Models. Proceedings of Geocomputation 2005. Ann Arbor, Michigan

Marks RE (2007) Validating Simulation Models: A Generic Framework and Four Applied Examples, Computational Economics 30 (3), p. 265-290

Martin N (2005) Development of a water balance for the Atankwidi catchment, West Africa - A case study of groundwater recharge in a semi-arid climate. Dissertation, Georg-August-Universität, Göttingen

Mather AS (2006) Proximate causes. In: Geist HJ (ed) Our Earth's changing land: An encyclopedia of land-use and land-cover change, vol. 2. Greenwood Press, Westport, London, pp 490-495

Matthews R (2002) Crop management. In: Matthews R and Stephens W (eds) Crop-soil simulation models, pp 69-79. Wallingford, UK: CAB International

Meyer WB, Turner BL II (eds) (1994) Changes in Land Use and Land Cover: A Global Perspective. Cambridge: Cambridge University Press

Miyao T, Kanemoto Y (1987) Urban Dynamics and Urban Externalities. Harwood Academic Publishers, New York

MOFA (Ministry of Food and Agriculture, Ghana) (2002) Food and Agriculture Sector Development Policy (FASDEP), Ministry of Food and Agriculture, Republic of Ghana

Mooney H, Neville L (2005) Consequences of Industrialized Animal Production Systems. Scientific Committee on Problems of the Environment, http://www.icsu-scope.org

Moran EF (2005) Human-environment interactions in forest ecosystems: An introduction. In: Moran EF, Ostrom E (eds) Seeing the forest and the trees: Human-environment interactions in forest ecosystems. MIT Press, Cambridge, London, pp 3-21

Otterman J (1974) Baring high-albedo soils by overgrazing: a hypothesized desertification mechanisms. Science 86, p.531

Park SJ, McSweeney K, Lowery B (2001) Prediction of soils using a process-based terrain characterization. Geoderma 103: 249-272

Park SJ, Vlek PLG (2002) Environmental correlation of three-dimensional soil spatial variability: a comparison of three adaptive techniques. Geoderma 109: 117-140

Park SJ, Hwang CS, Vlek PLG (2005) Comparison of adaptive techniques to predict crop yield response under varying soil and land management conditions, Agricultural Systems 85 (1), pp 59-81

Parker DC (2000) Edge-effect externalities: Theoretical and empirical implications of spatial heterogeneity. Ph D. diss. University of California at Davis, Davis, CA

Parker DC, Manson SM, Janssen MA, Hoffmann MJ, Deadman P (2003) Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. Annals of the Association of American Geographers 93 (2), p. 314-337

Penning de Vries FWT, Jansen DM, Berge HFM, Bakema A (1989)Simulation of ecophysiological processes of growth in several annual crops. Simulation Monographs. Pudoc/Internation Rice Research Institute, Wageningen, the Nederlands

Pietrula M, Carley K, Gasser, L (1998) Simulating Organizations. M.I.T. Press

Railsback SF, Lytinen SL, Jackson KJ (2006) Agent-based Simulation Platforms: Review and Development Recommendations. Simulation 82 (9), pp 609-623

Rand W, Zellner M, Page SE, Riolo R, Brown DG, Fernandez LE (2002) The complex interaction of agents and environments: An example in urban sprawl. In Proceedings of Agent 2002, Social Agents: Ecology, Exchange and Evolution, Chicago, IL

Ray I, Williams J (1999) Evaluation of price policy in the presence of water theft. American Journal of Agricultural Economics 81 (4), pp 928-941

Resnick, M (1995) Turtles, Termites and Traffic Jams, MIT Press, Cambridge, US

Russell SJ, Norvig P (1995) Artificial Intelligence: A Modern Approach, Prentice Hall

Schoemaker PJH (1995) Scenario Planning: A Tool for Strategic Thinking, Sloan Management Review 36 92), pp 25-40

Schultz AM (1967)The Ecosystem as a Conceptual Tool in the Management of Natural Resources. In Cieriacy-Wantrup, S. V., (Ed.) Natural Resources: Quality and Quantity, pp. 139-161, University of California Press, Berkeley

Schweitzer F, Zimmermann J (2001) Communication and Self-Organization in Complex Systems: A Basic Approach, In: Fischer M, Fröhlich J (eds), Knowledge, Complexity and Innovation Systems, Springer, Berlin

SEDAC (Socioeconomic Data and Applications Center) (2002) Land-Use and Land-Cover Change. A collaborative effort of SEDAC and the IGBP/ IHDP LUCC Project, September 2002

Shibasaki R (2003) Reviewing LUCC Models - New Challenges in Global Environmental Age, http://www-cger.nies.go.jp/lugec/Proceedings/03)Shiba-saki.pdf

Sicot M (1989) Constraintes et risques encourus par l'activité agropastorale au Sahel. Exemple de la Mare d'Oursi au Burkina Faso. In: Eldin M, Milleville P (eds) Le Risque en Agriculture. Ed. de L'ORSTOM, Coll à Travers Champs, Paris, pp 131-141

Sinclair TR, Seligman NAG (1996) Crop modeling: from infancy to maturity. Agronomy Journal 88: 698-703

Sobieraj JA, Elsenbeer H, Cameron G (2004) Scale dependency in spatial patterns of saturated hydraulic conductivity, CATENA, Vol. 55 (1), 49-77

Soil Health (2008) Soil Biological Fertility, http://www.soilhealth.com/fertility/

Soini E (2006) Livelihood, land use and environment interactions in the highlands of East Africa. Dissertation. Department of Geography, University of Helsinki

Sun Y, Cheng L (2002) A Survey on Agent-Based modeling and Equation-based modeling, http://www.cs.gsu.edu/ csclicx/Csc8350/model.pdf

Stéphenne N, Lambin EF (2001) A dynamic simulation model of land-use changes in Sudano-sahelian countries of Africa (SALU). Agriculture, Ecosystems and Environment 85 (2001) 145Ð161

Stephens W, Middleton T (2002) Why has the uptake of decision support systems been so poor? In: Matthews R and Stephens W (eds) Crop-soil simulation models. CAB International, Wallingford, UK, pp 129-147

The National Academies Press (2001) Lost Crops of Africa: Volume II: Vegetables, Development, Security, and Cooperation Policy and Global Affairs, National Reserach Council of the National Academies, The National Academies Press, Washington, D.C.

Tobias R, Hofmann C (2004) Evaluation of free Java-libraries for social-scientific agent based simulation. Journal of Artificial Societies and Social Simulation 7 (1), http://jasss.soc.surrey.ac.uk/7/1/6.html

Turing AM (1950) Computing machinery and intelligence. Mind 59, pp 433-460

Turner BL II (1989) The human causes of global environmental change. In: De Fries RS, Malone TF (eds) Global change and our common future: Papers from a forum (Committee on Global Change, National Research Council). National Academy of Sciences, Washington D.C., pp 90-99

Turner BL, Moss RH, Skole DL (1993) Relating land use and global land-cover change: a proposal for IGBP-HDP core project. IGBP Report 24, HDP Report 5. IGBP, Stockholm

Turner BL II, Skole D, Sanderson S, Fischer G, Fresco L, Leemans R (1995) Land-Use and Land-Cover Change; Science/Research Plan. IGBP Report No.35, HDP Report No.7. IGBP and HDP, Stockholm and Geneva

USGCRP (US Global Change Research Program) Land Use and Land Cover Change (2003), http://www.usgcrp.gov/usgcrp/ProgramElements/land.htm

USGS and Japan ASTER Program (2007) ASTER scene ASTL1B_071012103759071106 0029 Level 1B taken on December 10th 2007. The product was ordered and purchased by Julia Schindler. GLOWA-Volta Project Geo-database.

van der Veen A, Otter HS (2001) Land use changes in regional economic theory. Environmental Modeling and Assessment 6, pp 145-150

Veldkamp A, Lambin EF (2001) Predicting land-use change. Agriculture, Ecosystems and Environment 85, pp 1-6

Veldkamp A, Verburg PH (2004) modeling land use change and environmental impact: editorial. Journal of Environmental Management 72, pp 1-3

Verburg PH, Schot P, Dijst M, Veldkamp A (2004) Land use change modeling: current practice and research priorities, GeoJournal 61 (4), pp 309-324

Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human Domination of Earth's Ecosystems, Science 277, p. 494

Voortman RL (1998) Recent Historical Climate Change and its Effect on Land use in the Eastern Part of West Africa, Physics and Chemistry of the Earth 23 (4), pp 385-391

Vossen P (1988) An analysis of agricultural livestock and traditional crop production statistics as a function of total annual and early, mid and late season rainfall in Botswana. Agric. For. Meteorol. 42, 369-382

Weiss G (ed) (1999) Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. MIT Press

Western AW, Grayson RB, Blöschl G, Willgoose GR, McMahon TA (1999) Observed spatial organization of soil moisture and its relation to terrain indices. Water Resources Research 35(3): 797-810

Whitehead A (2004) Persistent Poverty in Upper East Ghana, Basis Brief: Collaborative Research Support Program, http://www.basis.wisc.edu/live/basbrief26.pdf

Wilensky U (1999) NetLogo. http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, Illinois

Williams M (2003) Deforesting the Earth: From prehistory to global crisis. University of Chicago Press, Chicago, p.715

Wilson JP, Gallant JC (2000) Terrain Analysis - Principles and Applications. John Wiley and Sons

Wooldridge M (2002) An Introduction to Multi-Agent Systems, John Wiley & Sons, Chichester, England

Yaro JA (2000) Poverty and Land Degradation: A Case Study of the Gia-Kajelo Community, Kassena-Nankana District of Northern Ghana (MPhil. Thesis) Department of Human Geography, University of Oslo, Oslo, Norway

Zeng N, Neelin JD, Lau KM, Tucker CJ (1999) Enhancement of interdecadal climate variability in the Sahel by vegetation interaction. Science 286, pp 1537-1540

ACKNOWLEDGEMENTS

First and foremost, I would like to express my great gratitude and appreciation to Professor Dr. Paul Vlek for his trust in my abilities, his excellent supervision, and his moral support during the conduction of this thesis. Furthermore, I want to express my strong gratitude to three people without whom this doctoral thesis would not have been possible. First, I am deeply indebted to Dr. Quang Bao Le, my tutor, who assisted me in all questions and problems I had during the beginning of the doctoral work to its end. Especially, I benefited from his experience and skills of preparing the thesis, his theoretical and practical knowledge in modeling issues, and his constant support and moral encouragement. The second person I owe many thanks to is Jacob Afeliga, who greatly assisted me in preparing and conducting my field work in the study area, who shared with me his outstanding knowledge of local conditions, and who showed a remarkable persistence and patience as he went with me through this laborious time. Last but not least, I want to give my heartful thanks to my husband Jean, who showed great interest in my research and who supported me with his love and encouragment during all hard times.

My sincere thanks are given to the Robert Bosch Foundation for giving me the financial support for this doctoral study, in the form of a scholarship and a generous field research budget. Furthermore, I would like to express my sincere gratitude to Dr. Wolfram Laube and Dr. Wilson Agyare for their advisory support during field work. I also want to thank heartly my field workers Ben, James, Isaac, Lambert and Callistus, who did a great job during the field surveys. This also applies to the four GLOWA drivers, whose knowledge of the study area and towns throughout Ghana was of great help. My sincere thanks are extended to the farmers of the Atankwidi catchment for their great hospitality and patience, which made field work quite enjoyable.

I am very grateful to the staff of the International Water Management Institute (IWMI) in Accra for offering me the possibility to conduct my studies in Accra and providing me with important contact addresses. My special thanks go to the GLOWA coordinator, Dr. Boubacar Barry, who supported me with his constant encouragement, technical devices and financial means, and to Daniel Ofori, IWMI's administrative assistant, for his help in accounting my expenditures.

Further, I want to express my gratitude to members of other Ghanaian institutions, including the Ghanaian Statistical Service in Accra and Bolgatanga, the Mininstry of Food and Agriculture (MOFA) in Navrongo, the Forestry Department and the Health Research Center in Navrongo, the Kwame Nkrumah University of Science and Technology in Kumasi, and the University of Ghana, Legon, Accra, especially the Centre for Remote Sensing and Geographic Information Services (CERSGIS).

I am very grateful to all ZEF scientific staff that have advised me with practiical and scientific issues. Further, I owe thanks to Dr. Guenther Manske, Rosemarie Zabel, and Karin Hagedorn-Mensah for their help with many administrative matters. I also thank Margaret Jend for editing the language of this dissertation, and my colleagues of the batch 2005 at ZEF for their companionship, which often helped to alleviate life as a doctoral student.

Finally, I give my heartful thanks to my mother, Barbara Schindler, and my brothers Thomas and Michael Schindler, for their familial love and moral support during my field work in Ghana and during the thesis writing process.

Ecology and Development Series

Band 1
Fatondji D (2002) Organic amendment decomposition, nutrient release and nutrient uptake by millet (*Pennisetum glaucum* (L.) R. Br.) in a traditional land rehabilitation technique (zai) in the Sahel. ISBN 3-89873-476-6 (Cuvillier Verlag, Göttingen)

Band 2
Relos de Macale MA (2002) The role of *Azolla* cover in improving the nitrogen use efficiency of lowland rice. ISBN 3-89873-468-4 (Cuvillier Verlag, Göttingen)

Band 3
Widyastuti R (2002) Soil fauna in rainfed paddy field ecosystems: their role in organic matter decomposition and nitrogen mineralization. ISBN 3-89873-571-0 (Cuvillier Verlag, Göttingen)

Band 4
Kaizzi KC (2002) The potential benefit of green manures and inorganic fertilizers in cereal production on contrasting soils in eastern Uganda. ISBN 3-89873-593-1 (Cuvillier Verlag, Göttingen)

Band 5
Sawhney P (2003) People-park interaction: A case of Bandhavgarh National Park, India. ISBN 3-89873-641-5 (Cuvillier Verlag, Göttingen)

Band 6
Gichora M (2003) Towards realization of Kenya's full beekeeping potential: A case study of Baringo District. ISBN 3-89873-675-X (Cuvillier Verlag, Göttingen)

Band 7
Li Z (2003) Diversity and ecology of mountain bamboos in the Shennongjia National Nature Reserve of Central China: Implications for resource management and biodiversity conservation. ISBN 3-89873-727-6 (Cuvillier Verlag, Göttingen)

Band 8
Hafeez MM (2003) Water accounting and productivity at different spatial scales in a rice irrigation system: A remote sensing approach. ISBN 3-89873-769-1 (Cuvillier Verlag, Göttingen)

Band 9
Gehring C (2003) The role of biological nitrogen fixation in secondary and primary forests of Central Amazonia. ISBN 3-89873-843-4 (Cuvillier Verlag, Göttingen)

Band 10
Tadesse Woldemariam Gole (2004) Vegetation of the Yayu forest in SW Ethiopia: impacts of human use and implications for *in situ* conservation of wild *Coffea arabica* L. populations. ISBN 3-98973-866-3 (Cuvillier Verlag, Göttingen)

Band 11
Le HTT (2003) Role of *Pythium aphanidermatum* (Edson) Fitzp. in tomato sudden death in the tropics with emphasis on integrated disease management. ISBN 3-89873-872-8 (Cuvillier Verlag, Göttingen)

Band 12
Win KK (2003) Plot-specific N fertilizer management for improved N-use efficiency in rice-based systems of Bangladesh. ISBN 3-89873-885-X (Cuvillier Verlag, Göttingen)

Band 13
Tuladhar JK (2003) The effect of *Azolla* on nitrogen use efficiency in rice-wheat rotations of Nepal. ISBN 3-89873-974-0 (Cuvillier Verlag, Göttingen)

Band 14
Braimoh AK (2004) Modeling land-use change in the Volta Basin of Ghana. ISBN 3-86537-044-6 (Cuvillier Verlag, Göttingen)

Band 15
Codjoe SNA (2004) Population and land use/cover dynamics in the Volta River Basin of Ghana, 1960-2010. ISBN 3-86537-038-1 (Cuvillier Verlag, Göttingen)

Band 16
Duadze SEK (2004) Land use and land cover study of the savannah ecosystem in the Upper West Region (Ghana) using remote sensing. ISBN 3-86537-041-1 (Cuvillier Verlag, Göttingen)

Band 17
Agyare WA (2004) Soil characterization and modeling of spatial distribution of saturated hydraulic conductivity at two sites in the Volta Basin of Ghana. ISBN 3-86537-100-0 (Cuvillier Verlag, Göttingen)

Band 18
Ajayi AE (2004) Surface runoff and infiltration processes in the Volta Basin, West Africa: Observation and modeling. ISBN 3-86537-097-7 (Cuvillier Verlag, Göttingen)

Band 19
Garcia M (2004) Effects of pesticides on soil fauna: Development of ecotoxicological test methods for tropical regions. ISBN 3-86537-106-X (Cuvillier Verlag, Göttingen)

Band 20
Intsiful JD (2004) Upscaling of land surface parameters through inverse-SVAT modeling. ISBN 3-86537-174-4 (Cuvillier Verlag, Göttingen)

Band 21
Wickel B (2004) Water and nutrient dynamics of a humid tropical agricultural watershed in Eastern Amazonia. ISBN 3-86537-187-6 (Cuvillier Verlag, Göttingen)

Band 22
Oguntunde PG (2004) Evapotranspiration and complimentarity relations in the water balance of the Volta Basin: Field measurements and GIS-based regional estimates. ISBN 3-86537-225-2 (Cuvillier Verlag, Göttingen)

Band 23
Ibrakhimov M (2004) Spatial and temporal dynamics of groundwater table and salinity in Khorezm (Aral Sea Basin), Uzbekistan. ISBN 3-86537-325-9 (Cuvillier Verlag, Göttingen)

Band 24
Rücker GR (2005) Spatial variability of soils on national and hillslope scale in Uganda. ISBN 3-86537-344-5 (Cuvillier Verlag, Göttingen)

Band 25
Bagamsah TT (2005) The impact of bushfire on carbon and nutrient stocks as well as albedo in the savanna of northern Ghana. ISBN 3-86537-406-9 (Cuvillier Verlag, Göttingen)

Band 26
Mekuria Argaw Denboba (2005) Forest conversion – soil degradation – farmers' perception nexus: Implications for sustainable land use in the southwest of Ethiopia. ISBN 3-86537-444-1 (Cuvillier Verlag, Göttingen)

Band 27
Anhar A (2005) The role of biological nitrogen fixation in the cacao agroforestry system in Central Sulawesi Indonesia. ISBN 3-86537-473-5 (Cuvillier Verlag, Göttingen)

Band 28
Syahrinudin (2005) The potential of oil palm and forest plantations for carbon sequestration on degraded land in Indonesia. ISBN 3-86537-481-6 (Cuvillier Verlag, Göttingen)

Band 29
Le QB (2005) Multi-agent system for simulation of land-use and land cover change: A theoretical framework and its first implementation for an upland watershed in the Central Coast of Vietnam. ISBN 3-86537-627-4 (Cuvillier Verlag, Göttingen)

Band 30
Lulseged Tamene Desta (2005) Reservoir siltation in Ethiopia: Causes, source areas, and management options. ISBN 3-86537-638-X (Cuvillier Verlag, Göttingen)

Band 31
Howard NK (2005) Multiscale analysis of landscape data sets from northern Ghana: Wavelets and pattern metrics. ISBN 3-86537-653-3 (Cuvillier Verlag, Göttingen)

Band 32
Akramkhanov A (2005) The spatial distribution of soil salinity: Detection and prediction. ISBN 3-86537-654-1 (Cuvillier Verlag, Göttingen)

Band 33
Puig CJ (2005) Carbon sequestration potential of land-cover types in the agricultural landscape of eastern Amazonia, Brazil. 3-86537-694-0 (Cuvillier Verlag, Göttingen)

Band 34
Amisigo BA (2005) Modelling riverflow in the Volta Basin of West Africa: A data-driven framework. ISBN 3-86537-701-7 (Cuvillier Verlag, Göttingen)

Band 35
Tipraqsa P (2006) Opportunities and constraints of integrated farming system in Northeast Thailand. A case study of the Huai Nong Ian catchment, Khon Kaen province. ISBN 3-86537-804-8 (Cuvillier Verlag, Göttingen)

Band 36
Compaoré H (2006) The impact of savannah vegetation on the spatial and temporal variation of the actual evapotranspiration in the Volta Basin, Navrongo, Upper East Ghana. ISBN 3-86537-805-6 (Cuvillier Verlag, Göttingen)

Band 37
Forkutsa I (2007) Modeling water and salt dynamics under irrigated cotton with shallow groundwater in the Khorezm region of Uzbekistan. ISBN 3-86537-806-4 (Cuvillier Verlag, Göttingen)

Band 38
Feyera Senbeta Wakjira (2006) Biodiversity and ecology of Afromontane rainforests with wild *Coffea arabica* L. populations in Ethiopia. ISBN 3-86537-807-2 (Cuvillier Verlag, Göttingen)

Band 39
Khamzina A (2006) The assessment of tree species and irrigation techniques for afforestation of degraded agricultural landscapes in Khorezm, Uzbekistan, Aral Sea Basin. ISBN 3-86537-852-8 (Cuvillier Verlag, Göttingen)

Band 40
Bagayoko F (2006) Impact of land-use intensity on evaporation and surface runoff: Processes and parameters for eastern Burkina Faso, West Africa. ISBN 3-86537-843-9 (Cuvillier Verlag, Göttingen)

Band 41
Martin N (2006) Development of a water balance for the Atankwidi catchment, West Africa – A case study of groundwater recharge in a semi-arid climate. ISBN 3-86537-885-4 (Cuvillier Verlag, Göttingen)

Band 42
Moser P (2006) Regeneration and utilization of *Faidherbia albida* and *Acacia erioloba* along ephemeral rivers of Namibia. ISBN 3-86537-948-6 (Cuvillier Verlag, Göttingen)

Band 43
Herbst S (2006) Water, sanitation, hygiene and diarrheal diseases in the Aral Sea area (Khorezm, Uzbekistan). ISBN 3-86727-010-4 (Cuvillier Verlag, Göttingen)

Band 44
Kassahun Tesfaye Geletu (2006) Genetic diversity of wild *Coffea arabica* populations in Ethiopia as a contribution to conservation and use planning. ISBN 10: 3-86727-986-1, ISBN 13: 978-3-86727-986-4 (Cuvillier Verlag, Göttingen)

Band 45
Odeny DA (2006) Microsatellite development and application in pigeonpea (*Cajanus cajan* (L.) Millspaugh). ISBN 10: 3-86537-991-8, ISBN 13: 978-3-86727-991-8 (Cuvillier Verlag, Göttingen)

Band 46
Taye Kufa Obso (2006) Ecophysiological diversity of wild Arabica coffee populations in Ethiopia: Growth, water relations and hydraulic characteristics along a climatic gradient. ISBN 10: 3-86727-990-X, ISBN 13: 978-3-86727-990-1 (Cuvillier Verlag, Göttingen)

Band 47
Schmitt CB (2006) Montane rainforest with wild *Coffea arabica* in the Bonga region (SW Ethiopia): plant diversity, wild coffee management and implications for conservation. ISBN 10: 3-86727-043-0, ISBN 13: 978-3-86727-043-4 (Cuvillier Verlag, Göttingen)

Band 48
Thenya T (2006) Analysis of macrophyte biomass productivity, utilization and its impact on various eco-types of Yala Swamp, Lake Victoria Basin, Kenya. ISBN 10: 3-86727-040-6, ISBN 13: 978-3-86727-040-3 (Cuvillier Verlag, Göttingen)

Band 49
Lopez RC (2007) Prevention of *Imperata* invasion in upland cultivated fields in the montane forest of Central Sulawesi, Indonesia. ISBN 978-3-86727-125-7 (Cuvillier Verlag, Göttingen)

Band 50
Almanza Fandiño MT (2007) Management of *Bombus atratus* bumblebees to pollinate Lulo (*Solanum quitoense* L), a native fruit from the Andes of Colombia. ISBN 978-3-86727-187-5 (Cuvillier Verlag, Göttingen)

Band 51
Kpongor DS (2007) Spatially explicit modeling of sorghum (*Sorghum bicolor* (L.) Moench) production on complex terrain of a semi-arid region in Ghana using APSIM. ISBN 978-3-940124-00-5

Band 52
Capote Fuentes RT (2007) Resilience of mangroves on the south coast of Havana province, Cuba. ISBN 978-3-940124-01-2

Band 53
Nyarko BK (2007) Floodplain wetland-river flow synergy in the White Volta River basin, Ghana. ISBN 978-3-940124-02-9

Band 54
Kasina JM (2007) Bee pollinators and economic importance of pollination in crop production: Case of Kakamega, Western Kenya. ISBN 978-3-940124-03-6

Band 55
Sandwidi WJP (2007) Groundwater potential to supply population demand within the Kompienga dam basin in Burkina Faso. ISBN 978-3-940124-04-3

Band 56
Nhamo N (2007) The contribution of different fauna communities to improved soil health: A case of Zimbabwean soils under conservation agriculture. ISBN 978-3-940124-06-7

Band 57
Goh HC (2008) Sustainable tourism and the influence of privatization in protected area management. A case of Kinabalu Park, Malaysia. ISBN 978-3-940124-07-4

Band 58
Tia L (2008) Modeling of vegetation dynamics and its contribution to the water balance in semi-arid lands of West Africa. ISBN 978-3-940124-08-1

Band 59
Mdemu M (2008) Water productivity in medium and small reservoirs in the Upper East Region (UER) of Ghana. ISBN 978-3-940124-09-8

Band 60
Beining A (2008) Ecophysiological diversity of wild *Coffea arabica* populations in Ethiopia: Drought adaptation mechanisms. ISBN 978-3-940124-10-4

Band 61
Scheer C (2008) Biosphere-atmosphere-exchange of C and N trace gases and microbial N turnover processes in irrigated agricultural systems of the Aral Sea Basin, Uzbekistan. ISBN 978-3-940124-11-1

Band 62
Obuobie E (2008) Estimation of groundwater recharge in the context of future climate change in the White Volta River Basin, West Africa. ISBN 978-3-940124-12-8

Band 63
Friesen JC (2008) Regional vegetation water effects on satellite soil moisture estimations for West Africa. ISBN 978-3-940124-13-5

Band 64
Ayalew Belay E (2009) Growing lake with growing problems: integrated hydrogeological investigation on Lake Beseka, Ethiopia. ISBN 978-3-940124-15-9

Band 65
Callo-Concha D (2009) An approach to environmental services assessment: functional biodiversity in tropical agroforestry systems (The case of Tomé-Açú, Northern Brazil). ISBN 978-3-940124-17-3

Band 66
Maguta J (2009) Conservation tillage in Kenya: the biophysical processes affecting its effectiveness. ISBN 978-3-940124-17-3

Band 67
Mensah A (2009) The influence of land-use activities on nutrient inputs into upland catchment streams, Ghana. 978-3-940124-18-0