Linking processes and pattern of land use change

Overmars, K.P.

3

Comparing inductive and deductive modelling of land use decisions: Principles, a model and an illustration from the Philippines

Abstract

Understanding the causes of land use change is of great importance for issues of tropical deforestation, agricultural development and biodiversity conservation. Many quantitative studies, therefore, aim to link land use change to its causal 'driving forces'. The epistemology of virtually all these studies is inductive, searching for correlations within relatively large, sometimes spatially explicit, datasets. This can be sound science but we here aim to exemplify that there is also scope for more deductive approaches that test a pre-defined explanatory theory. The chapter first introduces the principles and merits of inductive and more deductive types of land use modelling. It then presents one integrated causal model that is subsequently specified to predict land use in an area in northeastern Philippines in a deductive manner, and tested against the observed land use in that area. The same set of land use data is also used in an inductive (multinomial regression) approach.

With a goodness-of-prediction of 70 percent of the deductive model and a goodness-of-fit of 77 percent of the inductive model, both perform almost equally well, statistically. Because the deductive model explicitly contains not only the causal factors but also the causal mechanisms that explain land use, the deductive model then provides a more truly causal, as well as more theory-connected, understanding of land use. This provides land use scholarship with an invitation to add some more deductive (theory-driven and theory-building) daring to its methodological repertoire.

Based on: Overmars, K.P., De Groot, W.T., Huigen, M.G.A. 2006. Comparing inductive and deductive modelling of land use decisions: Principles, a model and an illustration from the Philippines. Human Ecology (Accepted).

3.1 Introduction

The face of the earth is rapidly changing, with great consequences for rural livelihoods, biodiversity conservation, urban quality of life and the global climate. Understanding la use change is therefore a matter of obvious import and urgency, reflected, inter alia, in LUCC, the joint land use program of the International Geosphere-Biosphere Programme (IGBP) and the International Human Dimensions Programme on Global Environmental Change (IHDP) (Turne *et al.*, 1995).
Land use change is the result of the often complex interplay of underlying causal factor usually referred to as 'driving forces', that may vary across scales and organizational levels, that may work directly or through longer causal routes and that may be associate with quite different societal and scientific realms, such as markets, policies, demograph culture and biophysical factors. How can such complexity be handled scientifically? On approach is to focus on only one or a few factors, and accept that explanations can only partial. Generally, however, land use scientists desire to do a more 'integrated' (multi-fa tor) analysis. As shown, for instance, in the overviews of Walke *et al.* (2002) and Verburg *et al*. (2004d), the great majority of the present-day blooming of quantitative integrated (multi-factor) studies of land use change follows an inductive approach, sometimes guic by theory but without testing the theory as such. In the present chapter, we make a case the present state of the art allows to perform integrated research and yet use a more ded tive epistemology, and that this option, in interaction with inductive work, will enhance causal insight and cumulative scientific progress in land use science. We aim to strengtl our case by showing and discussing the performance of a deductive and an inductive ap proach, applied parallel to each other to explain the land use in a single example regic
The chapter is organized as follows. The following section discusses the principles and merits of inductive and deductive approaches to land use science. Since deductive work requires a theoretical model to be tested, the third section is devoted to the structure of the model for our case study. The material and methods section then introduces the stuc area and the data gathering methods. The fifth section formalizes the deductive model for our example region and the sixth quantifies the model. The seventh section describe the results of the deductive model as well as those of a multinomial regression model, which was used for the inductive modelling exercise. The ensuing discussion shows the value of the deductive modelling approach within a range of approaches from inductive to deductive

3.2 Inductive versus deductive modelling

3.2.1 Deductive and inductive epistemology in integrated land use explanations

For most of us, the "empirical cycle" must have been the first concept taught in lectures about how science proceeds. First there is a theory; then a concrete predictive hypothesi is deducted from that theory. Then this hypothesis is tested in the real world and with tl result, the theory is either falsified or strengthened. This, in short, is deductive epistemc ogy. Contrasting with this approach, inductive methodology works the other way arour It begins with observations of reality and then tries to find regularities in these data. Th regularity is then declared to be a general pattern (a model, a theory). This claim can be based, for instance, on the randomness of the sampling that was used. Statistical work ca be part of both approaches; statistical testing is a characteristic part of deductive methoc ogy, while multiple regression is often used for inductive approach

The present chapter does not find fault in this basic epistemological scheme. We do, ho ever, think that for a proper understanding of how land use science proceeds in practice is necessary to define a number of methodological positions that lie in-between the ded tive and inductive extremes. In order to arrive there, it helps to first specify what exactl 'a model' or 'a theory', especially in terms of what may be called the degree of specifica of that model or theory. A first case is that a researcher has no model or theory at all. Obviously then, the only methodology available is extreme induction, or data mining as it will be called below, in which the researcher attacks large datasets, basically 'correlati everything with everything else' in order to see if any patterns may be found. A second is that the researcher has a notion of what factors may be relevant for the explanation of what he seeks to explain. This may be called a weak model. The researcher may then sti fall back on data mining approaches, but he may also concentrate on the candidate facto to see if these indeed play a role as was hypothesized. A third case is when the research avails of what may be called a strong or structured model, that not only states what fact are important but also how they are important. According to Elster (1989), it is only ther that true explanation comes within reach, because true explanation requires insight not only in the factors but also in the mechanisms. An example of a structured model is the well-known law of Liebig on plant growth that specifies not only that nutrients like nitr gen and phosphorus are important but also how they are, with the plant responding onl to the nutrient that is 'in the minimum'. In formula: *plant growth = a + b*MIN [phosphorus, c*nitrogen]*. Note here how different this formula is from the usual structure of multiple regression, which would be: *plant growth = a + b*phosphorus + c*nitrogen*. In the latter for- mula, adding more nitrogen would always result in more plant growth even if phospho would be at zero. In the first formula (and in reality), the plant does not respond at all. I the case of the researcher availing of a strong model, he can try to induce the parameters his particular case in the model's structure. He may also fall back on a traditional multip regression with the nutrients arranged in the additive structure, or even on blind data m ing. See for instance De Groot *et al.* (1987) for an example of induction using both Liebig' structure and traditional regression on plant growth. A final case is when the researche avails of a fully quantified model, e.g. Liebig's structure with the parameters a, b and c specified. It is only then that true prediction, hence true deduction, is possib

We can now come to see the deduction/induction pair of terms as defining a gliding sca between two extremes. On the one hand, there is extreme deduction of the Popperian ki (Popper, 1963), in which the empirical cycle is followed strictly and theory falsification rather than verification is seen as the key to progress. On the other hand, there is extren induction, in which the researcher aims to find patterns in large datasets without any theoretical guide. Both extremes have their advantages in some cases, e.g. if very strong theories are available, or if no theory at all is as yet defined, respectively. Both have stro disadvantages too, however. In the social and economic sciences, extreme deductivism would lead to an endless rejection of theories because simply none of them is able to gra the full complexity of the system described. Extreme inductivism, on the other hand, lea to an immense amount of correlations that cannot be interpreted as causes and never ac- cumulate into a coherent theory

One response to this dilemma is the validation of inductive models following the sugges
tion made (in passing) in many statistical textbooks to split one's data in half, use the fi r
half for a free induction of any kind, and then use the other half to test the induced mod
An example is in Nelson *et al*. (2004), who use 1/25th of their large and spatially explicit
dataset to induce their explanatory land use model and then use the model to predict la
use over the whole map. A more radical way out of the dilemma has been suggested by
Brox (1990), discussing the epistemological status of 'grand theories', in his example the
common property theory applied to fisheries. Brox' solution is that we forget about the
empirical claim of such theories at all but rather regard them as analytical tools. Using tl
theory we may discover which part of reality behaves according to the theory (which is
interesting), and which part does not (which is interesting toc
In most research practice, researchers find a less daunting solution by seeking or simply
adopting a position, usually implicit and led by disciplinary traditions, somewhere on tl
continuum between extreme induction and extreme deduction. For the present chapter
and including the two extremes, we may define six of these positions. We concentrate h
mainly on quantitative work

1. *'Extreme induction'*. This is the extreme of data mining, "knowledge discovery in databases" (Liao, 2003)[1]
2. *'Unstructured factors induction* Under this term we subsume all research approache that apply a broad conceptual framework of some kind, usually derived from com mon sense or literature overview, in order to specify a usually long list of factors (roads density, slope, off-farm income, tenure security, distance to recreation sites, household composition and so on, often each with several variants of further spec fication and measurement) that are candidate to help explain land use or land use change. (Alternatively, some kind of theory may be invoked as well, e.g. as Nels *et al*. (2004) do, saying that land users choose for the most profitable land use, but th these theories are in fact only serving as a broad conceptual framework.) Often, th factors are proxies of the actual factors that influence the land use process, since th processes themselves are not specified. The studies then leave it to the procedures statistical inference to find the correlations between these varial[2] Characteristically these studies do not end with a discussion of theoretical perspectives but only wit a discussion of the significance of correlation coefficients and suchlike in the spec case studied. Many land use change studies fall into this category (e.g. Geoghega *et al*., 2001; Serneels and Lambin, 2001; Overmars and Verburg, 2005 (Chapter 2
3. *'Theory-guided factors induction'*. This term denotes all studies that take an explicit theory of land use change as point of departure to critically specify a theory-connected (and usually shorter) list of explanatory variables. Strictly speaking, this list is still unstructured; it is only a list, after all, without specification of how the

---

[1] *In quantitative research, this extreme is often, and understandably, seen as something to be done onl very sparsely. In qualitative research, remarkably, extreme induction is often seen as the ideal basis fo 'grounded' theory building (Glaser and Strauss, 1967), allowing respondents to speak in their own voice and analysing their visions without any preset notions of the researcher. Great progress has been achieved this way, e.g. the famous discovery of the 'ethics of care' (Gilligan, 1982).*

[2] *See Overmars and Verburg (2005) for a factors-led inductive study on the research area in the Philip pines and Geist and Lambin (2002) for an inductive meta-analysis of 152 studies on tropical deforesta tion.*

variables are supposed to interact. On the other hand, the variables are not simply 'candidates' that are dropped if they do not contribute to the explanation. If they do not contribute, something is 'wrong' with the theory or its interpretation, which needs to be discussed. One quantitative example is in Perz and Walker (2002), focu ing on secondary forest growth in Amazonia in connection with Chayanovian theo Another example is by Rudel and Roper (1997) who arrived at their "frontier mode and "immiserization model" of tropical deforestation by a careful construction, examination and re-examination of a relatively small dataset rather than by blind for applied to a large one. Interesting results have also been reached in a more qualitat manner, exemplifi ed by Ostrom (1990) who arrived at her well-known conditions successful common property management by a stepwise induction of case studies. Characteristic for all studies of theory-guided induction is that the relevance of the results is wider then those of type 1 and type 2 studies. Guided by theory, inductic can become theory building

4. *'Imposed theory structure'* The next rung on the induction/deduction ladder is forme by studies that impose not only theory-guided factors but also a theory-guided stru ture (the 'behavioural statements', as Walker (2004) says) on reality before multiple regression is applied in order to induce the parameters within that structure. If our theory would be, for example, that people only choose for a land use type to the extent that this land use type is both culturally appropriate and profi table, our mo structure would look like Liebig's law, e.g. that the land use depends o $ß_0 + ß_1{*}MIN\ [ß_2{*}CULT, PROF$. In the same vein Tadepally (1999) stated that in order to rehabilit their village-level irrigation systems, villages should avail of both the capacity (spe fi ed by Tadepally as collective social capital) and the motivation to do so (specifi e by Tadepally as low rehabilitation cost), and found a strong relationship between these two variables and the success of NGO intervention for rehabilitation, with an imposed structure o $SUCCESS = ß_0 + ß_1{*}MIN\ [ß_2{*}CAP, MOT$. (It is interesting to note that the 'imposed structure' approach can also be used in a more qualitative style. We then use a theory to 'tell the story' of a specifi c case of land use change as do, f instance, Walker and Solecki (2004) and De Groot (1999) who apply dynamic versi of Thünian theory to tell the land use history of the Everglades and of the Cagayan Valley in the Philippines, respectively. If the story is good, or at least signifi cantly more insightful than others, this is a test that reality indeed works as the theory prescribes. This test will always remain soft , however, since qualitative theories ar stories will always be quite malleable in the hands of good storytellers
5. *'Imposed theory'* A purely deductive approach is reached when a land use theory is specifi ed for a real world case in terms of both structure and parameters, and the l use thus predicted is tested against real land use. As an example, in the case study presented in this chapter we will develop a theory-based model structure, quantify and then test it on a dataset from the Philippines
6. *'Extreme deduction'* We keep the 'Popperian' extreme separate here because in step 5, the model and the data gathering are not geared towards falsifi cation and neith need theories to be dropped if they do not work adequately ye

A few technical remarks are in order here. First, induction, deduction and the continuum between them, even though central tenets of epistemology, do not cover the full spectru of scientifi c methodology. Creative inference ('abduction') and the heuristic concepts of 'event ecology' (Vayda and Walters, 1999) are cases in point. Second, we may note that t six rungs of the induction/deduction ladder are naturally not the only possible ones. Researchers may also find intermediate and mixed positions, or work sequentially, with m or less extreme induction generating patterns that may be later used for a more deducti approach or, the other way around, starting from a theory. We do not go into these issue here, however, and regard the listing as good enough to indicate what we mean when saying that land use studies could or should become 'more deductive

3.2.2 Could land use explanation studies become more deductive?

Overlooking the field of explanatory land use studies, we find a quite skewed distribut over the induction-to-deduction axis. Examples abound of unstructured factors inductic Theory-guided factors induction is present in much smaller numbers. Imposing of theor structure is virtually non-existent. This may have historical and cultural backgrounds. T begin with, strong theories that may be tested are simply not massively present in any young science field. Furthermore, theories and deduction are not really *en vogue* in post-modern times (they are top-down, they turn a blind eye to the multiple complexities anc voices of social realities, etc.). And finally, the attraction that land use studies appear tc have had to econometrists and GIS-based geographic data technology may have had a flipside too, namely to block growth towards more deductive, theory-guided wo
In our opinion, explanatory land use studies could become more deductive. We do have land use theories to use and test, if only simple. Examples are Neo-Malthusian theory speaking about poverty traps, neo-Boserupian theory speaking about the positive effect of population on land use sustainability and neo-Thünian theory speaking about movin frontiers and urban markets (coupled as in De Groot (1999), or decoupled as in Walker (2004)). And we have more general explanatory theories waiting to be applied and teste on land use situations, such as rational choice theory, cultural theory, theories of collecti action and common property management, and so on. [3] Furthermore, much knowledge has accumulated and great datasets have been collected – knowledge and datasets that c not need to be used one-way but may also be re-used in more deductive manners. In the present chapter, our example shows that nothing difficult is at stake here, conceptua

3.2.3 Should land use explanation studies become more deductive?

Should explanatory land use studies try to move up a bit on the deduction ladder? Ther two main advantages of using deductive methods. First, deduction yields the intrinsical better proof of causalit *i.e.* true explanation. Let us take Nelso *et al.* (2004) as an example Their causal model is that on each site, the most profitable crop is grown. However, this is not tested as such because, as Nelson *et al*. say, profitability is not measurable. Insteac factors such as land slope are used as independent variables. Slopes, obviously, do have influence on profitability but they causally underlie many other values too. If, say, mai found to be associated with medium slopes, would that be because of its relative profita ity there? It could also be that traditions do not allow maize elsewhere, or because of ris

[3] *We disagree here with Couclelis (in Parker et al., 2002: 6), who rightfully states that predictive mode should be structurally appropriate, but then adds that the process theories (i.e. theory specifying causa mechanisms of these model structures) are simply not available in the land use fiel*

aversion, or because maize from these soils tastes best. The slopes/maize correlation doe not establish any causality. If, however, Nelson *et al*. would have *calculated* the spatially explicit profi tability of crops (based on prices, distance to road, slope etc.), then would have set the model to predict crop distribution on the basis of highest profi tability and then would have found the model predicting maize on medium slopes, a strong proof o causality would have been delivered. In the words of Elster (1989), this is because not or the causal factor but also the causal mechanism (in this case, profi tability) has been test The second benefi t of a more deductive approach is that it bett er facilitates the accumu tion of insight on the level of the discipline as a whole. Referring back again to the exam of Nelson *et al.* (2004), the outcomes of type 2 studies are basically the strengths-of-correl tion between land use and the usually long list of independent variables such as slopes, prices, household composition variables and so on. Conclusions then necessarily tend tc remain largely stuck on that level, e.g. that maize tends to converge on certain slopes in case, or that the number of children does not appear to have an infl uence in that case. Ir order to reach some degree of generalization, such studies then have to wait until enoug of them have accumulated to themselves become data in a meta-analysis such as that of Geist and Lambin (2002) who, characteristically for an inductive approach in the meta-analysis of inductive studies, come up with a generalized and regionally patt erned listi of proximate factors and underlying driving forces of tropical deforestation. Obviously useful as this may be, more progress would be made if not only the incidental meta-ana but also the researchers themselves, in their own studies, would be able to participate in a permanent intertrade of generalization. This can be achieved if these studies would be more deductive, *i.e.* more theory-led. That way, all land use scientists could contribute tc progress around a relatively limited number of theoretical themes in stead of only addir more detail about a very large number of empirical variables. Theory-led work, feeding back into theory, leads to theory building

In all this, we assume that empirically based theories are good to have. In other words, v assume that land use scientists do not become addicted to theories, especially their own, to a degree that theories begin to block entry for the surprises of reality (Vayda, 1983) or become objects of counterproductive controversy (Brox, 1990

3.2.4 Model choice and levels-of-deductivism of this chapter

As said, the objective of the present chapter is to expound and illustrate the deduction/ir duction dimension for integrated land use explanation. For the deductive part, we have therefore chosen to test a broad model that is able to take up all factors that should be cc prised in a truly integrated approach, hence including cultural, economic and biophysic data. It does have to be a model, however, and not some underspecifi ed agglomerate. F this deductive 'core structure', we have chosen for the Action-in-Context framework of De Groot (1992), which may be characterized as broad rational choice. For the inductive approach a multinomial logistic regression model is applied. As for the positions on the deduction ladder, we have chosen to compare an 'unstructured factors induction' (rung with a fully deductive approach, 'imposed theory' (rung 5), hence a true test without an subsequent fi tt ing on the dataset. (In the remainder of this chapter these two approach are referred to as the inductive and deductive approach or model, respectively). Logical too, we put all emphasis on the comparison and not on the cultural or land use intricaci of the study area

3.3 The Action-in-context framework and decision model

Action-in-Context (AiC) (De Groot, 1992; Verburg *et al*., 2003) is a framework designed for the explanation of human actions, especially in the environmental field. Based on the concept of progressive contextualization (Vayda, 1983), the idea of AiC is to start ou from the action to be explained, then identify the (individual or collective) actors directl causing this action, then identify the range of options available to these 'primary' actors and the motivations attached to these options, and then identify other ('secondary') act and factors influencing these options and motivations, thereby putting the action in its relevant causal context without *a priori* bias towards any scientific discipline (Vayda an Walters, 1999). With that, AiC is a fully actor-based framework, which is a logical choice explanatory work because actors, not systems, are the social entities that cause change d rectly[4] AiC may be used as a framework to guide the research process, but can also be u as a template for models. These models can be, for example, detailed multi-agent model that model individual agents (an example is in Huigen, 2004), or models that explain th choices of a smaller number of large actor categories. The latter is of course much simpl to implement and the way we will proceed in this stud

Action-in-Context has four interconnected components. (1) The first is an often repeated "core element", comprising of the action, the actor, his options and his motivations. In Elster (1989), the latter two are called "opportunities" and "desires" but the structure is of the same simplicity: in order to act, people must have both the capacity and the will to do so. The other components of AiC are elaborations of the core element. (2) The "actors field" is an aspect of AiC that is, to our knowledge, unique in the social sciences. It describes the chains of social influence (causality, power) that run from the primary actors outward to other actors. Such chains may run, for instance, from farmers NGOs, big landowners, traders, government agencies and the World Bank; an example in Verburg *et al.* (2003). The method of constructing actors fields is by posing the questio what actions (hence what actors) have an influence on the options and/or motivations o the primary actors. The secondary actors thus identified have their own options and mo tivations for these actions, which then may lead to the identification of tertiary actors, a so on. Moving from primary to secondary and further actors in AiC is the actor-based w of moving from proximate factors to underlying drivers *sensu* Geist and Lambin (2002). (3) The next component in AiC may be mixed freely with the preceding one and consist of a "deeper analysis" of the options and motivations of selected actors, distinguishing, *inter alia*, between elements of knowledge, resources, economic merit and culture. Figure 3.1 is AiC's broadly rational decision model designed to support this step, which will be discussed in some detail below. The deeper analysis is a second way to connect proxima factors to underlying culture and structure. (4) The final component of AiC is called the "actor model", which defines how the actor evaluates the options and motivations to co to his decision. In qualitative research, the actor model can often remain implicit. In suc cases, the researcher 'puts himself in the place of the actor' (Vayda, 1983) and trusts that his audience can do the same, thus understanding the logic of the actor's choice without

---

[4] *Also Blaikie (1985) has this basic notion of explanation by putting actions of actors in context, but h contexts are conceptualised as systems rather than other actors. In AiC, explanations may reach up to the global level but this level is then still present as actors, e.g. the IMF in its own global 'life-world' and with its own options and motivations to act.*

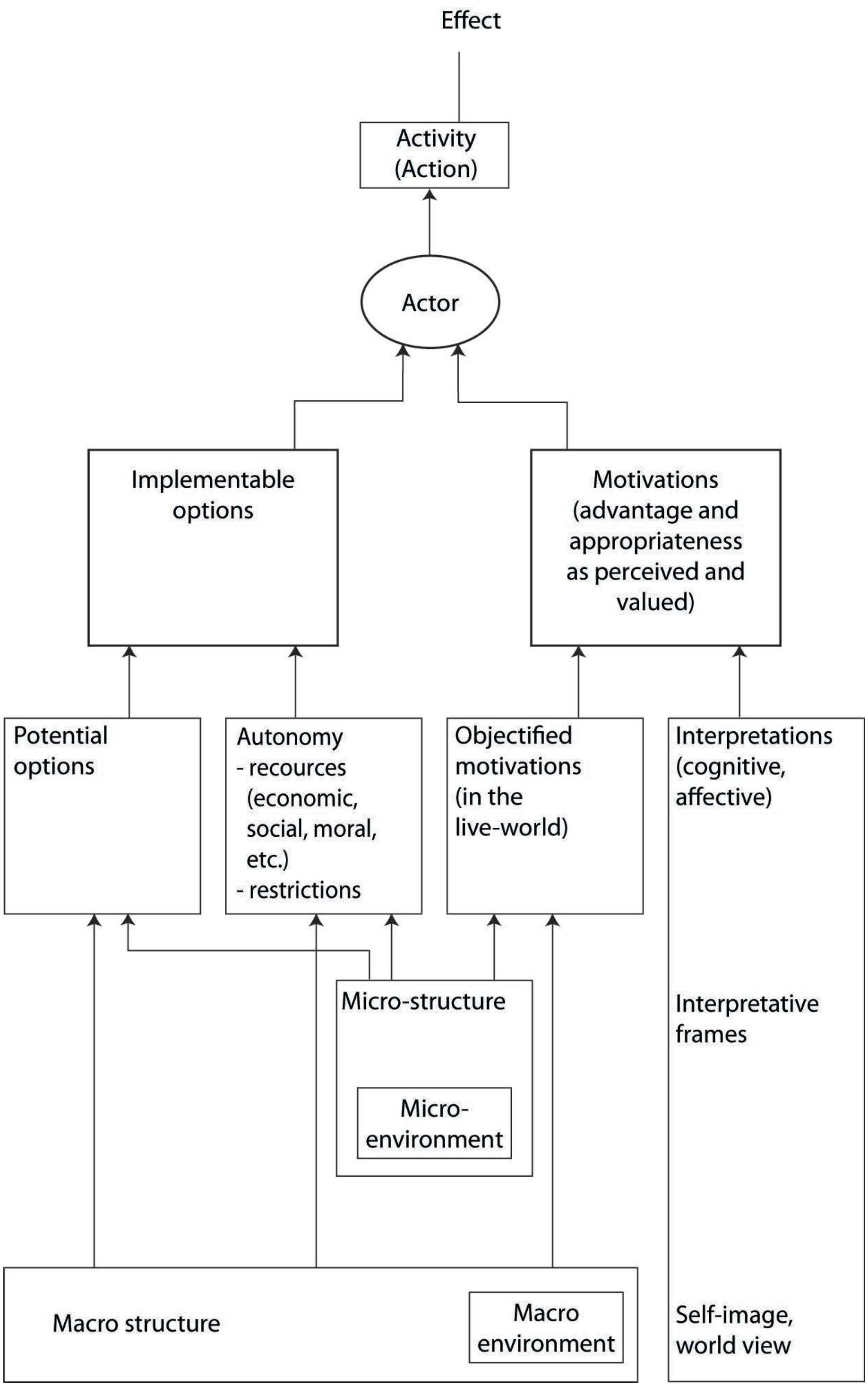

Figure 3.1: The decision model structure of AiC

a formalized model being needed. Another actor model is rational choice, which is to sa that the actor chooses for the option of maximum merit – or utility, or profi t, in which tl defi nition of these terms marks the diff erence between narrow and broad rational choi Broad rational choice may also be formalized in the shape of a multi-criteria table, as do by Hobbes *et al.* (n.d.). Specifi c for AiC is that it also off ers an actor model for cases whe a deeper refl ection on the logics of human decision-making is warranted. This actor mo distinguishes between three 'moral domains' of reasoning: *homo economicus* (*i.e.* rational choice), *homo honoris* and *homo communalis* (or 'ethics of care'); more detail is in De Groot (1992).[5]

Action-in-Context may be applied in many forms, in full or partially, formal or informal as heuristic tool for guiding fi eldwork or as a static model. See, for instance, De Groot a Kamminga (1995), Van den Top (1998) and Cleuren (2001) for qualitative applications or tropical deforestation. In the present chapter, we will especially use a quantifi ed versior of the decision model of the deeper analysis. The decision model (Figure 3.1) will be use as a structured model describing land use decision-making, which will be quantifi ed an tested in full. In this respect the approach is deductive since the model and its paramete are determined using the pre-defi ned decision-making structure and the prediction is derived from this model, aft er which the result is tested against observations. Therefore the decision model, depicted in Figure 3.1, is described in detail belo

In Figure 3.1, all arrows stand for causal relations. The top layers of the fi gure only repe the core element of AiC, be it that the options are now specifi ed as "implementable" op tions and that the motivations are the options' merits ("advantage and appropriateness" "as perceived and valued" by the actor. These specifi cations facilitate the defi nitions or the next lower level, which is the one of most interest here. At this level the implementa options are seen to result from the actor's "potential options" and "autonomy". Potentia options are defi ned as everything the actor could do if he were infi nitely autonomous (powerful). Basically, they are all options that the actor knows to be possible. In land use issues, the typical role of agronomic research and extension is to bring more potential o tions to farmers (the former are then secondary actors). This is not to say, of course, that these options will also be implemented*i.e.* become an action in Figure 3.1); farmers shoul also have the capacity (autonomy) to be able to implement them, as well as the motiva- tion. "Autonomy" is capacity-to-implement, and is defi ned as the sum of all resources t actor can access (economic capital, private social capital, cultural capital, entitlements to common goods, etc.), taking into account possible restrictions (e.g. zoning regulations). Potential options and autonomy together determine the implementable options. Figure 3.1 does not specify the structure of this joint causality but we may note that it can not be some simple form of addition; just adding potential options does not automatically add to implementable options (let alone change actions), and neither does just adding to the actor's autonomy ('empowerment'). The case study of this chapter shows one way o modelling this

---

[5] *In land use studies, it seems logical to assume that many choices will have a rational choice character Deeper refl ections may sometimes be needed, however, for instance to explain why people may stop planting trees once they are off ered a fi nancial incentive. It may be that planting trees was done in th moral domain of ethics of care or ethics of honour (we do it for each other, we do it for the children, we do it for the pride of the village), but fl ipped into the domain of homo economicus reasoning, triggered by the sudden association with monetary gain. And then of course, we do not plant trees for so litt le money.*

Motivations are the merits of the options. In Figure 3.1, the motivations "as perceived ar valued" are separated into "objectifi ed motivations" and their "interpretations". Object motivations are all those that may easily be quantifi ed, such as economic cost and benet time expenditure, risk probabilities, caloric value of food and so on – in short, all these fa tors that micro-economists and farming system analysts feel at home with. Interpretatio on the other hand, are all those factors that give weight, coherence, shape and colour to objectifi ed motivations. Note that this way the interpretations are set as somehow multi ers of the objectifi ed motivations rather than a 'fi lter' between actor and reality; psycho and culture add life to the actor, so to speak. Deeper down in the fi gure (but without ca arrows, indicating that the relationship is diffi cult to quantify) these interpretations are supposed to rise out of broader "interpretative frames" and "self-image / worldview". C example is the actor's image of what it is to be a good farmer (Zuiderwij k, 199
In Figure 3.1 furthermore, the third-layer elements are supposed to arise out of the actor micro-structure (defi ned as all structures, social and physical, where the actor makes a c ference) and macro-structure. Since these relationships do not play a role in our quantifi model, we do not go into them here
Overall, Figure 3.1 is obviously not something special as is AiC's actors fi eld but rather designed as the reverse. It aims to overarch and coherently integrate all elements of broa rational choice theory, including cultural elements, the 'capitals' of Bebbington (1999) ar so on, and remain close to the models of social psychology (albeit dropping the cumber-some intervening variable of 'att itudes'). Roughly then, many disciplinary focal points a included in the model: the options of agronomy and forestry, the objectifi ed motivation economics, the culture of anthropology, the capitals (autonomy) of development studies the environment of geographers, and so on. Thus, the model facilitates explanatory wor without preoccupation towards any specifi c disciplin

3.4 Material and methods

3.4.1 Study area

The study area is situated in Cagayan Valley in the northeastern part of the island Luzon the Philippines (Figure 3.2). The study area includes 16 villages *barangays*) in the munici-pality of San Mariano, in the province of Isabela, and comprises approximately 260 km2 It is situated between the town of San Mariano in the west and the forested mountains o the Sierra Madre in the east
The population is approximately 16,500 persons (about 3,150 households) of various eth groups, among whom the Ilocano, Ibanag and Ifugao, who are all migrants or descende of migrants that came to the area from the 1900s onwards, and the Kalinga and Agta, who are the indigenous inhabitants. Before immigration started, the area was completel forested with tropical lowland forest. At present, the study area shows a clear land use gradient ranging from intensive agriculture, with wet rice and yellow corn, near San Ma ano via a scatt ered patt ern of wet rice, yellow corn, banana, grasses, and (fruit) trees in foothills to residual and primary forest in the eastern par

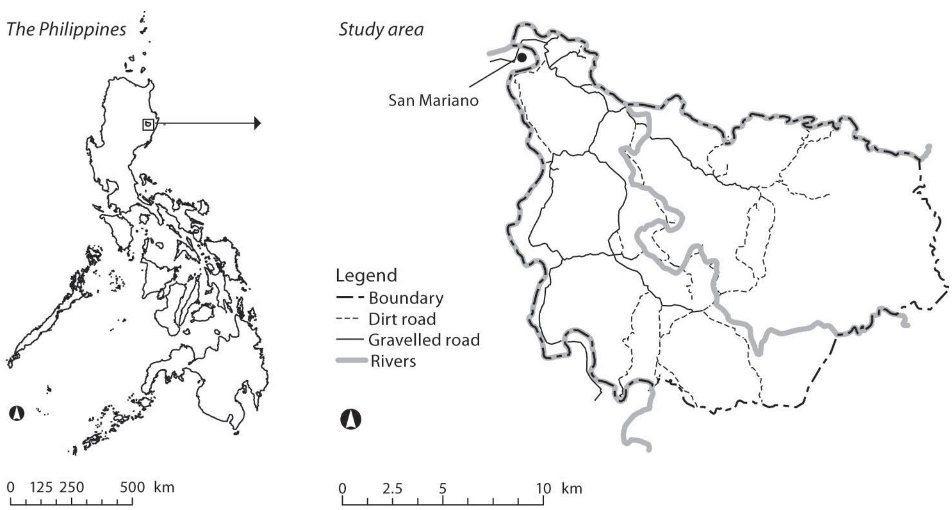

Figure 3.2: Location of the study area in the Philippines (left) and topographic features of the area (right)

3.4.2 Data collection

To collect the household-level data we conducted an interview campaign between June and November 2002 in 13 of the 16 villages, using a structured questionnaire to collect d on land use and its explanatory variables. Selection of households to be interviewed wa based on systematic random sampling using population data available per village. In all villages every twentieth household was selected through systematic random sampling. total of 151 households were interviewec

The household questionnaire was structured in a nested hierarchy (see also Overmars a Verburg, 2005 (Chapter 2)) with the household level at the top and the plot level and the fi eld level underneath. The plot is defi ned as a piece of land owned or used by the hou hold, and a fi eld is a specifi c part of the plot used for one of the land use types. On ave a household owns 2.07 plots, with 1.15 diff erent crops per plot. Variables were collectec on their appropriate levels, e.g. soil characteristics at the fi eld level, travel time to plot a the plot level and ethnicity at the household level. Records containing missing values ar households without any land were excluded from the dataset. In total 114 households w included in the analysis. These households are the managers of 236 plots with 272 fi elds This data was used to fi t the multinomial regression model of the inductive approach a to validate both the inductive and the deductive model. Table 3.1 shows a selection of th data set consisting of those variables that turned out to be relevant in this case stuc

Besides the questionnaires, semi-structured interviews were held with farmers and key-actors like heads of the villages and the elderly. If possible, these interviews were held with a group of people to enable discussion and verifi cation. These interviews dea *inter alia*, with the motivations of the people to grow one crop or another. People were asked, for instance, to compare the diff erent options for a specifi c fi eld or to compare crops in general. People were also asked hypothetical questions about what they would change (

Table 3.1: Description and descriptive statistics of the variables of the household survey (n = 272)

Variable name	Description	Min.	Max.	Mean	St. dev.	Incl. in ded. model	Incl. in ind. model
Dependent variables							
Yellow corn	1 if yellow corn, 0 otherwise	0	1	0.58		Y	N
Wet rice	1 if wet rice, 0 otherwise	0	1	0.13		Y	N
Banana	1 if banana, 0 otherwise	0	1	0.24		Y	N
Fruit trees	1 if fruit trees, 0 otherwise	0	1	0.05		Y	N
Land use	Yellow corn (4), wet rice (1), banana (2) and fruit (3) (nominal)	1	4			N	Y
Independent variables							
Slope1	1 if slope category is fl at, 0 otherwise	0	1	0.38		Y	N
Slope2	1 if slope category is fl at to rolling/moderate, 0 otherwise	0	1	0.23		Y	N
Slope3	1 if slope category is rolling/moderate, 0 otherwise	0	1	0.28		Y	N
Slope4	1 if slope category is rolling/'moderate to steep/hilly, 0 otherwise	0	1	0.08		Y	N
Slope5	1 if slope category is steep/hilly, 0 otherwise	0	1	0.03		Y	N
Slope	1 if fl at to 5 if steep (ordinal	1	5			N	Y
Ethnicity Ifugao	1 if male household head is Ifugao, 0 otherwise	0	1	0.10		Y	Y
Ethnicity Ibanag	1 if male household head is Ibanag, 0 otherwise	0	1	0.30		Y	N
Ethnicity Ilocano	1 if male household head is Ilocano (or Tagalog speaking), 0 otherwise	0	1	0.55		Y	Y
Ethnicity Kalinga	1 if male household head is Kalinga, 0 otherwise	0	1	0.03		Y	N
Ethnicity Other	1 if male household head is other than Ifugao, Ibanag, Ilocano or Kalinga	0	1	0.02		Y	N
Municipality of origin	1 if male or female head is born in San Mariano, 2 if both, 0 otherwise	0	2	1.17		Y	Y
Creek	1 if there is a creek or spring trough or bordering the plot, 0 otherwise	0	1	0.58		Y	Y
Plot distance	Minutes walking to the plot (min)	0	240	23.81	33.50	Y	Y
Transportation cost	Cost to transport a bag of corn from the house to San Mariano (pesos)	7	45	23.85	12.49	Y	Y
Average age	Average age of household heads (years)	21	78	42.71	13.07	Y	Y
Tenure tenant	1 if the household is tenant of the plot, 0 otherwise	0	1	0.19		Y	Y
Tenure SIFMA	1 if the plot is has SIFMA tenure, 0 otherwise	0	1	0.06		Y	Y

not) in their land use practices if certain conditions would change. These semi-structure interviews were guided by the AiC framework, which was used for progressive context alization (Vayda, 1983) in this stage. The information from these interviews was used to qualitatively describe the processes in the area as well as to quantify the decision model and to determine the calculi to relate the elements of this mode

3.4.3 Analysis

The inductive (regression) mode

For the inductive approach to predict the land use on a fi eld we applied multinomial lo regression, which is regularly used in land use analysis (e.g. Nels *et al.*, 2001; Müller an Zeller, 2002). Multinomial logistic regression extends the possibilities of logistic regressi by allowing for more than two categories in the dependent variable. In this case four lan use categories are included. The parameter estimates are calculated simultaneously and the probabilities of the diff erent land use categories add up to one. The fi nal prediction the land use with the highest probability. In the multinomial model, the estimated parameters are to be interpreted in relation to one of the categories of the dependent variabl which serves as reference category. Under the assumption that all error terms are mutu independent and have a log Weibull distribution the multinomial logistic regression mo can be expressed as follows

$$P(y_i = j) = \frac{\exp(x'_{ij}\beta)}{1+\exp(x'_{i2}\beta)+\ldots+\exp(x'_{iM}\beta)}, j = 1,2,\ldots,M \qquad (3.1)$$

where y is the dependent variable (land use typ j are the alternatives o M land use types i is the i^{th} fi eld x are the explanatory variables, an β is a vector of regression coeffi cient In this equation $(x'_{i1}\beta)$ is normalized and set to zero, (exp(0)=1), so in this case alternativ (yellow corn) is used as the reference category (Verbeek, 200

The probability ratio (odds ratio) for the normalized model is given b

$$\frac{P(y_i = 2)}{P(y_i = 1)} = \exp(x'_{i2}\beta) \qquad (3.2)$$

This probability ratio is used to interpret the eff ect of a change of the values of one of the explanatory variables. Equation 3.2 shows that the probability ratio is only depende on the reference alternative and the alternative under study and does not depend on the nature of any of the other alternatives (Verbeek, 2000). An increase in the probability of a land use relative to the reference land use may have no signifi cance on the probability when compared to other possible land uses (Nelso *et al.*, 2001).

In the multinomial regression the following variables were taken into account: slope, et nicity, municipality of origin, presence of creek, plot distance, transportation cost, avera age of the household heads and tenure (see also Table 3.1). The approach follows the ru 2 level of the deduction/inductive ladder presented previously: selection of the factors w inspired by several (disciplinary) land use theories, prior studies and fi eld knowledge.[6]

---

[6] *The original 'rung 2' study upon which the multinomial regression model in this chapter was based i described in Overmars and Verburg (2005) (Chapter 2)*

These theories are not tested as such, but their factors are used in a regression analys
The variable slope was taken into account as if it were a continuous variable with a ratic
scale (fl at slopes were given the value 1 up to steeps slopes with the value 5). Doing this
we suggest a linear relation between the slope categories and the land use types. Includ-
ing four of the slope categories as nominal variables, as we do in the deductive model,
was not possible because, due to (quasi) complete separation, the maximum likelihood c
that multinomial regression model was impossible to calculate. For the same reason the
variable ethnicity Ibanag male was not included in the regression analysis. All variables
were tested for multi-collinearity by regressing each of the independent variables upon
remaining independent variables. This test revealed no multi-collinearit

The causal model for the deductive (AiC) approac
The causal model we applied is a quantifi cation of the decision model of the Action-in-C
text framework (the deeper analysis[7] As actor model we use broad rational choice, sayir
that the actor chooses for the implementable option of maximum merit. In the next secti
a qualitative description of the case is provided following the structure of the framewor
A detailed description of the actions is given and the primary actors are identifi ed and
described. Then, the potential options and autonomy are described in detail, followed b
the objectifi ed motivations and interpretations. The subsequent section is devoted to the
quantifi cation of this causal decision model. This model was used to predict the probab
of the occurrence of the land use types. In explaining and quantifying the model we use
the same set of explanatory factors as in the multinomial model, together with some ad-
ditional constants like, for example, maximum benefi t and investmen

Comparison
Since the inductive (regression) model and the deductive (AiC) model describe the land
use system in the same area and use the same variables, the performance of the models c
easily be compared. The performance was calculated using cross-tables (also called pre-
diction matrix). The outcomes are a goodness-of-fi t and a goodness-of-prediction for the
inductive and deductive model, respectively. The cross-tables show the observed catego
against the modelled category of land use. Subsequently, the diagonal of the table conta
the correct predictions. Besides the number of correctly predicted fi elds also the relative
quantity of the predicted categories will be taken into account while assessing the mode
performance. In this application we consider not only the total score of good predictions
to be important, but we also want that the correct predictions do not have an extreme bi
towards only a few of the categories

3.5 Qualitative description of the deductive model

3.5.1 Actors

Ninety percent of the households in the study area have a piece of land to cultivate. Farı
ing is the main source of income of 80 percent of the households and the second source c

[7] *The actors fi eld of land use choices has not been part of the model. The actors fi eld explaining yellow corn production in the area and the actors fi eld of forest migration are described by Van den Top (1998).*

income for 10 percent of the households. Besides this, most people work also as a labour for other farmers, which provides them with additional income. The actors considered in the analysis are those households that have control over a piece of land that they can possibly cultivate. They will be called farmers in the remainder of the chapte

3.5.2 Actions

The analysis focuses on the decision-making on agricultural land use types. However, tl possibilities of making a living in the area are broader than agriculture alone. Besides op tions that do not involve land (e.g. carpenter, storekeeper, driver), the people in the area also have other land use options such as small-scale logging and collecting non-timber forest products. Both these option categories are not considered in this study. Hence, the research question is why farmers cultivate a certain crop at a certain location. The area that the farmer cultivates per crop is not subject of analysis. This study is restricted to th explanation of the occurrence of agricultural land use types on existing fiel

3.5.3 Potential options

To construct a list of potential land use options we could include, for example, all crops grown in the region over the last 30 years. Based on data from 1971 onwards, this would include tobacco, peanut, white corn, vegetables and several other crops, besides the four most important crops at present, which are wet rice, yellow corn, banana and fruit trees. Currently, these land use types account for 92 percent of the observed fields. Consideri our research objective, which is a methodological comparison rather than location speci detail, we chose to include only these four land use types. Methodologically, it is import to know that these four potential options are not all of them implementable options at al locations, as we will see. Some more detail on these four crops is supplied belo

Yellow corn is the most important cash crop in the area. To cultivate yellow corn, the farmers use hybrid seeds, often apply large quantities of fertilizer and most of them use pesticides. Most farmers get these agrochemical inputs, which are indispensable to culti- vate yellow corn, on credit from traders. Often, these traders also provide the farmers w consumptive credit. The farmers are obliged to sell the harvest to the same traders, who charge a high interest rate, to pay back their debts. This reduces the farmers' freedom to the highest price on the market. Due to the risks inherent to corn production and due to credit system farmers end up in a strong debt bondage with the traders. Yet, many farm continue to seek the patronage relationship with the traders because corn is in fact is the only access to credit and the traders may also help out in times of need (Van den Top, 19

Yellow corn is cultivated twice per year. The main risks for yellow corn, as reported by t farmers, are pests like rats, insects and birds and climatic conditions like droughts, floo and typhoons

The cultivation of wet rice is for consumption by the household and is only marketed if there is a rare surplus. Water sources in the area, necessary to cultivate wet rice, are rain natural irrigation by creeks and springs, and a large irrigation scheme. Most farmers tra plant the rice, though some use the system of direct seeding (broadcasting), which is les time-consuming, but requires some skills. Fertilizer and pesticides are used, but less tha yellow corn production. If sufficient water is available for irrigation two crops are culti per year. The most reported risk is crop damage by rats, insects, birds and sna

Banana is largely a cash crop, but the input level of fertilizers and pesticides is low. The most important varieties that are used in the area are Damilig, which is a cooking banan for industrial use (banana chips and banana ketchup), and Lakatan and Turdan, which are dessert bananas. These three have different characteristics regarding spacing, harve and price. Damilig is normally sold at a lower price, but the plantation has to be renewe less often and is more resistant to diseases than Turdan and Lakatan. The time between planting and the first harvest is 1 to 1.5 years (Sterken, 2004). The main risks for banana typhoons, Banana Bunchy Top Virus and wild pigs. Newly planted banana fields are of intercropped with yellow corn or upland rice for the first one or two years. In some area bananas are cultivated in rows between field

Fruit trees are not grown on a large scale in the area. Recently, a number of farmers part pated in a program called SIFMA (Socialized Industrial Forest Management Agreement which awarded them with 25 years of tenure rights provided that they plant a certain area with (fruit) trees (mainly mango, citrus and coconut), which were provided for free by an NGO (General, 1999). This land use type was included in the analysis, because it might become an important land use type in the future. However, marketing of fruit is still underdeveloped in the area and also maintenance of the plantations is often lacking which prevents fruit tree plantations from being successful so far. Fruit trees are often intercropped with yellow corn, especially when the fruit trees are small and not bearing fruits yet. Others plant the fruit trees on the borders of their fields. Only a few farmers have fruit trees as their only crop (Klein, 2003). The most important risks for fruit trees a typhoons and fire

3.5.4 Autonomy

The level of autonomy determines if a potential option can be implemented or not. Autonomy consists of two elements: resources and (absence of) restrictions. The autonomy the farmer is restricted by the variables tenancy, creek and ethnicity. If the farmer is a te of the land he cultivates, the landowner often decides what the tenant should cultivate, which is most often yellow corn. So, the tenant cannot make an autonomous decision ab what to cultivate. For the cultivation of wet rice two restrictions were added: presence o a creek and the farmers' ethnicity. Wet rice cultivation is restricted to Ifugao and Ilocanc because, generally spoken, they have better skills and knowledge in constructing rice fi and rice terraces and cultivating wet rice. Ibanag people, who are the original lowlander of the Cagayan valley, have a long tradition in corn cultivation. Formerly, they produce white corn as staple food because growing white corn could be combined with tobacco, which was an important crop in the region during the Spanish time (Van den Top, 1998) The assumption is that many Ibanag farmers do not know (or know to a lesser extent) h to cultivate wet rice because it was not part of their tradition (Romero, pers. comm.). Th presence of a creek on or near the plot is important for the cultivation of wet rice, since it needs a water source. The source of water could be a pump or an artificial irrigation system, though in most cases this is a small river or stream that is diverted towards the r field. This stream should be close to the rice field. So, a creek nearby is considered to be prerequisite to cultivate wet rice

The other element of autonomy is the resources of a farmer. In this case study, resources are considered to be necessary to do initial investments to start a new land use type, like clearing a forested area for corn cultivation or constructing a rice terrace. If the resource are suffi cient to do the investment the land use type is an implementable option. So, th initial investments function as a threshold. They are built up of two components: basic investments and, for rice only, additional investments dependent on slope. In our mode the resources are composed of the 'level of assistance', the possibility to obtain credit to buy inputs for a crop and participation in the SIFMA program, which together should b suffi cient to do the initial investments for a specifi c land use ty

The resource 'assistance' is composed of the factor municipality of origin and a factor proportional to the average age of the household heads. The municipality of origin of th household heads is considered to be indicative for the size of a household's social netwc (roughly: social capital). The assumption is that people who are born in San Mariano ha more relatives and friends nearby than people coming from outside the municipality. Tl social network is necessary for farmers to organize a group of people to do the work at relatively low costs. In many places in the area it is a custom to help one another by wor ing in a large group to do the larger jobs like cleaning, planting and harvesting (Mooner 2002). Ifugao were considered to have assistance from relatives even when they are not from San Mariano, because oft en they migrate aft er invitation of relatives or friends an cluster together. Also a higher age is considered to be indicative for a larger network to organize labour (children, relatives)

Another way to meet the necessary investments is to borrow money. In the research are credit is almost exclusively provided for yellow corn. Other sources of capital to make investments for other crops are hardly available, which actually restricts farmers in thei options.

The last resource is participation in the SIFMA program, which provides tenural securit and assistance in starting an agro-forestry plantation and therefore applies to the land u type fruit trees. In the study area, land titles can only be obtained for the so-called A&D ienable and disposable) lands, which are the fl at areas. Sloping lands are classifi ed as fo and owned by the state and for these lands no offi cial titles can be acquired. Neverthele people cultivate these state-owned forest lands. Governmental as well as non-governme organizations encourage farmers on these lands to invest in agro-forestry systems, whic are considered to be more sustainable than arable farming. However, insecure property rights hamper the development of these agricultural systems because they require high investments and need a long time to become profi table (e.g. tree planting and conserva measures). Farmers do not have the money to invest and they are not sure if they can sti use the land at the time the crops become profi table. Therefore, the SIFMA program all farmers to apply for a 'stewardship contract' for 25 years while committ ing themselves to a more sustainable way of farming. Farmers that were awarded a SIFMA contract can receive free fruit tree seedlings to be planted on their SIFMA lot, covering a part of the h initial investment costs

3.5.5 Objectifi ed motivations

Motivations are composed of objectifi ed motivations and interpretations. In this study t objectifi ed motivations are considered to be the net economic benefi t from one hectare of a land use type at the moment the product is sold in San Mariano (in case of yellow corn, banana and fruit trees) or consumed (in case of rice). The net benefi t is defi ned as the maximum benefi t under ideal climatological (no extraordinary droughts or typhoor and biophysical conditions (fl at area with a good soil) for an average price, multiplied b

a yield-reducing factor depending on slope (for yellow corn) and a yield-reducing factor
depending on risks lowered with the transport cost. The maximum benefit is considered
be the same at all locations in the study area
Steeper slopes will decrease the objectified motivation towards corn because the costs a
higher and the yield is lower. Ploughing is more difficult or impossible on steeper slope
which increases the costs spent on planting the corn. On steep slopes, seeds and fertilize
are washed away during heavy rains. This will reduce the yield of such a field in compa
son with flat fields. So, on steeper slopes the cultivation of corn will cost more in effort
time and will yield less because of the poorer productivity of the plo
Bananas can grow in every landscape position, unless soil drainage is very bad (Valmay
et al., 1990). Many of the drawbacks that corn has on steep slopes do not apply to banana
Banana cultivation does not involve tillage, so ploughing is not required. Bananas are re
newed only once every 5-15 years. The productivity of banana is the same on steep slope
and flat areas. So, slope does not influence the motivation towards growing banana. On
contrary, many farmers plant bananas to prevent soil erosion on steep slope
Transportation cost is the cost to transport the product from a farmer's home to the mar
In this study transportation costs apply to yellow corn, banana and fruits. Rice is used fo
household consumption or sold in the neighbourhood. Additionally, the distance from t
plot to the residence of the farmer is taken into account. If a plot is far from the farmer's
lage, the farmer needs to invest more effort and time in cultivating a crop on that plot. T
effect will be most prevalent with yellow corn and wet rice, which need to be frequente
by the farmer more often than other crops like banana. Moreover, fields that are far aw
have more risk to be damaged by fire, water buffalos or peop

3.5.6 Interpretations

The objectified motivations are adjusted to the interpretation of the individual land man
ager. In this model interpretations are simplified to crop preferences of the different eth
groups. The traditions and cultural values of the ethnic groups are different for the crop
considered. These traditions make that people feel at ease with growing certain crop or
that they are proud to have it. As said before, Ifugaos and Ilocanos have a tradition in w
rice cultivation whereas the Ibanag have a tradition in corn cultivation and not in wet ric
cultivation. This is reflected in their preference for corn and rice. The preference for ban
and fruit trees seems to be the same for all ethnic groups. The objectified motivations ar
combined with the interpretations to become the motivations "as perceived and valued"
This may cause people of different ethnicity to choose a different land use option even
the objectified motivations are the same for both ethnic grou

3.6 Quantifying the deductive model

Based on the fieldwork and the qualitative analysis in the previous section, which is de-
rived from this fieldwork, the formal model with the structure of the deeper analysis of
Figure 3.1, is quantified as follows. The core of the model is that the predicted land use
is the implementable land use option with the highest motivation (Equation 3.3). Startin
with the options side of the model, Equation 3.4 shows that the implementable options a
composed of potential options and autonomy. The potential options are yellow corn, we

rice, banana and fruit trees. Autonomy (Equation 3.5) is determined by restrictions and resources. If a restriction is 1 or the resources are 0, the autonomy is 0 and the potential option cannot be implemented

Action = f(implementable options, motivations) (3.3)
*Implementable options = potential options * autonomy* (3.4)
*Autonomy = (1 - restrictions) * resources* (3.5)
Restrictions = f(T_TENANT [a]*, CREEK, ETHNICITY, CROP) (3.6)*
Resources = IF(assistance + credit + tenure_SIFMA – investments>=0),
resources = 1, else 0 (3.7)
Assistance = (MUNICPALITY_ORG + (AGE/34))/3 (3.8)
Credit = f(CROP) (3.9)
Tenure_SIFMA = f(TENURE SIFMA, CROP) (3.10)
Investment = inv_basic + inv_slope (3.11)
inv_basic = f(CROP) (3.12)
inv_slope = f(SLOPE, CROP) (3.13)

[a] *Variables in the equations are written in capit*

As described in the qualitative model description the restrictions in this study are a func of tenancy, creek, ethnicity and crop (Equation 3.6). If the land manager is a tenant we o consider yellow corn to be an option. So, if the variable tenant is 1, all land use types exc corn were given value 1 (Table 3.2). In the Equations 3.4 and 3.5 this leads to an autonon of zero and therefore to a zero for the implementable options calculation, meaning the la use type is no option. If the field is not cultivated by a tenant (tenure tenant = 0) all opti are possible. Wet rice is only possible if a creek is nearby and if the field is cultivated by farmers of the ethnicity Ilocano or Ifugao. These restrictions are summarized in Table 3. Calculations run similar to the example above. These relations are intuitively determine based on field experience and the interviews and are not fitted in any w
In the model the resources assistance, credit and tenure SIFMA should cover the investments for a land use type to make this land use implementable (Equation 3.7). The assist depends on municipality of origin and average age of the household heads and is specif in Equation 3.8. (For Ifugao the value of the factor municipality of origin was set on 2 ev if they are not born in the municipality of San Mariano). The equation is formulated in such a way that the result is centred around one for a specific age (34 yrs.). This specific parameter was optimised, since no clear theoretical idea was available to determine the influence of age. As explained in the previous section credit is 1 (possible) for yellow co and 0 (not possible) for the other crops. The resource due to the assistance by the NGO in the SIFMA areas is 1 for fruit trees (Table 3.2). The investments consist of two parts: basic investments and investments due to slope (Equation 3.11). The basic investments a defined as the basic investments necessary to start a new field for a specific land use ty The values of the basic investments (Table 3.2) are relative to the initial investments for yellow corn, which were set on 1. This relation was estimated by the authors based on fi experience. The relation between slope and the investment necessary to build a rice terra (Table 3.3) was estimated according to the amount of labour necessary to build a terrace (Romero, pers. comm.), which was calculated as an average from field observations. Th extra investment due to slope was set on 1 for the terraces on slope category 3 and the o categories were estimated calculated to this valu

Table 3.2: Factors that determine autonomy through restrictions and resources (Values in the tables ar used in the model)

Variable	Yellow corn	Wet rice	Banana	Fruit trees
Restrictions				
Tenure tenant = 1	0	1	1	1
Tenure tenant = 0	0	0	0	0
Creek = 1	0	0	0	0
Creek = 0	0	1	0	0
Eth. Ilocano and Ifugao	0	0	0	0
Eth. Ibanag, Kalinga and Other	0	1	0	0
Resources				
Credit	1	0	0	0
Tenure SIFMA = 1	0	0	0	1
Tenure SIFMA = 0	0	0	0	0
Investments	1	1.2	0.3	1.5

The result of the model structure and the parameters is that corn is possible for all farmers because investments can be covered by credits, banana is also possible for all farmers because the initial investments are low and that fruit trees is possible for people that hav a SIFMA lot. Initially, the calculation resulted in no possibilities for wet rice, because of too high initial investments. Since rice does occur in the area this rule was relaxed a litt l This can be justifi ed by the fact that rice fi elds are usually smaller than a hectare and th calculation is per hectare and therefore initial investments are smaller in reality than the calculated investments

Table 3.3: Calculation of investment term for the construction of rice terraces

Slope category	Days labour per ha*	Investment term for rice
Slope1	52	0
Slope2	292	0.36
Slope3	716	1
Slope4	2209	3.25
Slope5		4.33**

** Source: Romero (pers. comm.) (n = 28)*

*** Estimated by the authors*

The right branch of the AiC model (Figure 3.1) deals with the motivations. The motivati (as interpreted) consist of objectifi ed motivations multiplied with a factor for the preferences (Equation 3.14). In this case the objectifi ed motivations are expressed in Philippin Pesos and consist of the maximum benefi t, a slope factor, a risk factor and transportatio cost. The maximum benefi t is expressed in Table 3.4. These values stem from average yields reported in interviews, except from the maximum benefi t for fruit trees, which w calculated by Klein (2003). For yellow corn the maximum benefi t is multiplied by a yiel factor depending on slope (Table 3.5) and an average yield reducing factor depending o estimated risks for all land use types (Table 3.6). The former were derived from reporte yields on fi elds with diff erent slopes and the latt er was derived from interpretations o damage reports in the interviews. This risks table does not include the regular pest and diseases, because these are incorporated in the estimated yields. The high typhoon risk for banana is related to the fact that the banana is not productive for 1 to 1.5 years aft er typhoon, while other crops can be replanted and productive several months aft er destru tion. The transportation costs are computed according to Verbur *et al.* (2004a). Travelling distance to the plot (variable 'plot distance') was translated into monetary costs. For wet rice transportation costs were only based on costs from the residence to the fi eld, since t product is not marketed, and for the other crops the calculation is a combination of costs from fi eld to residence and from residence to the town of San Mariano. The preferences (Table 3.7) based on ethnicity were quantifi ed by the authors based on qualitative descr tions by the farmers. It may be noted that in this model the eff ect of the preference for w rice cultivation is cancelled out by the much higher net benefi t of wet rice compared to other crops, so diff erences in preference do not change the prediction of ri

*Motivations = objectifi ed motivations * preferences* *(3.14)*

*Objectifi ed motivations (net benefi t) = max_benefi t * slope_fact * (1-risk) - tr* *(3.15)*

Max_benefi t = f(crop) *(3.16)*

Slope_fact = f(slope, crop) *(3.17)*

Risk = f(crop) *(3.18)*

Tr_costs = f(tr_cost, plot_distance, crop) *(3.19)*

Preferences = f(ethnicity, crop) *(3.20)*

The objectifi ed motivations with the interpretations combine into the motivations (as perceived and valued) for each fi eld for all four crops. These motivations (as perceived and valued) are summarized in Table 3.8. Cultivating wet rice is by far the most profi ta followed by fruit trees. The benefi ts from corn and banana are very simil

Table 3.4: Maximum benefi t (in Ph. Pesos, calculated from fi eld data) per land use ty

Crop	Max_benefi
Yellow corn	22435*
Wet rice	42000*
Banana	21213*
Fruit trees	32230**

** Source: fi eld dat*

*** Source: Klein (2003)*

Table 3.5: Calculation of slope factor for yellow corn

Slope category	Average yield (kg/ha)	Slope_factor
Slope1	3581	1.00
Slope2	3829	1.07
Slope3	3070	0.86
Slope4	no data	0.50*
Slope5	no data	0.20*

** Estimated by the authors; other data based on fi eld observations (n = 37*

Table 3.6: Risk factors of crops

Risk/Crop	Yellow corn	Wet rice	Banana	Fruit trees
Typhoon	0.10	0.05	0.20	0.10
Drought	0.10	0.05	0.02	0.02
BBTV	0.00	0.00	0.10	0.00
Risk total	0.20	0.10	0.32	0.12

N.B. All numbers are estimated by the authors

Table 3.7: Preference factors based on ethnicity

Ethnicity/crop	Yellow corn	Wet rice	Banana	Fruit trees
Ifugao	0.9	1.2	1	1
Ibanag	1.2	0.9	1	1
Ilocano	1	1.2	1	1
Kalinga	1	1	1	1
Other	1	1	1	1

N.B. All numbers are estimated by the authors

Table 3.8: Summary of the motivational value for all fields per land use typ

Land use	Average (Peso/ha)	St.dev
Yellow corn	14239	3834
Wet rice	40694	5181
Banana	12474	672
Fruit trees	23313	1739

3.7 Model Results

3.7.1 The inductive (multinomial regression) model

The inductive model (Table 3.9) shows the estimated parameters of wet rice, banana anc fruit trees in relation to yellow corn, which is the reference category. The estimated coefficients should be interpreted relative to this category. For example, one unit increase in the explanatory variable creek will increase the $ln(P_{wet\ rice} / P_{yellow\ corn})$ with 1.988. Positive coefficients result in an increase of the probability relative to the reference category anc negative coefficients in a decrease. In multinomial regression analysis the interpretatio of the estimated coefficients is not completely straightforward, because the coefficient only tell us the relation between one land use category and the reference category. This complicates direct comparison of the inductive model with the deductive mod
The cross-tabulation (Table 3.10A) shows the number of observations that is modelled c rectly (the bold diagonal figures) and if not, in which category. The right column shows percentage of the observations that was fitted right. Especially yellow corn was fitted well (91 percent), banana was fitted reasonably well (66 percent) and wet rice (50 perce and fruit trees (43 percent) were fitted somewhat weakly. In total, the multinomial regr

sion model fi tt ed a total of 209 out of 272 (77 percent) observations correctly. A test was performed to what extent the observed and modelled land use distributions are alike. Tl Chi-square statistic of this test is signifi cant at the 0.0001 level. The kappa statistic, whic indicates the proportion of agreement aft er chance has been excluded, is 0.5.

Table 3.9: The multinomial regression model

	Wet rice		Banana		Fruit trees	
Variables	**b**	**s.e.**	**b**	**s.e.**	**b**	**s.e.**
Intercept	-3.182	1.764	-9.936***	1.958	-11.420***	3.215
Slope	-1.302**	0.408	2.224***	0.333	1.628***	0.489
Ethnicity Ifugao male	2.631*	1.073	-0.295	1.243	-1.588	1.661
Ethnicity Ilocano male	1.678*	0.705	0.380	0.509	-0.131	0.965
Municipality of origin	-0.668	0.359	-0.097	0.342	-0.402	0.601
Creek	1.988***	0.554	0.013	0.505	0.502	1.001
Plot distance	-0.008	0.014	0.008	0.007	0.016*	0.008
Transportation cost	0.051*	0.023	0.065**	0.022	0.050	0.038
Average age	0.011	0.021	0.037	0.021	0.067	0.036
Tenure tenant	0.084	0.594	-0.921	0.678	0.159	1.261
Tenure SIFMA	-0.908	1.363	0.749	1.297	3.931*	1.661

** p<0.05, ** p<0.01, *** p<0.001*

Table 3.10: Observed land use vs. modelled land use of the inductive (regression) model (A) and deductive (AiC) model (B)

A.	Predicted land use type					
Inductive (regression) model	**Wet rice**	**Banana**	**Fruit trees**	**Yellow corn**	**Total**	**% Correct**
Observed land use type						
Wet rice	**18**	1	1	16	36	50.0
Banana	0	**42**	3	19	64	65.6
Fruit trees	0	5	**6**	3	14	42.9
Yellow corn	4	11	0	**143**	158	90.5
Total	22	59	10	181	272	76.8

B.	Predicted land use type					
Deductive (AiC) model	**Wet rice**	**Banana**	**Fruit trees**	**Yellow corn**	**Total**	**% Correct**
Observed land use type						
Wet rice	**21**	1	1	13	36	58.3
Banana	2	**31**	5	26	64	48.4
Fruit trees	1	3	**5**	5	14	35.7
Yellow corn	18	5	1	**134**	158	84.8
Total	42	40	12	178	272	70.2

3.7.2 The deductive (AiC) model

The results of the deductive model (Table 3.10B) are largely the same as the results of th inductive model. Wet rice is predicted better than in the multinomial model and for the other land use types the deductive model performed slightly less. The model was able to predict 70 percent of the occurring land uses of a dataset of 272 fields. The Chi-square statistic is significant at the 0.0001 level. For this model the kappa statistic is 0.471. The kappa statistics of the two models are not significantly differe $p < 0.05$) (Couto, 2003). So based on the kappa statistic it cannot be shown that the inductive model performs bette than the deductive model

3.8 Discussion and Conclusions

Following the objectives of the chapter, this section will discuss some of the case study outcomes, but pays special attention to the differences between inductive and deductiv research approaches and especially those presented in this stud

3.8.1 Factors of land use change

The AiC framework is designed to incorporate relevant factors from all scientific discip in a balanced manner. Using the deeper analysis of the AiC framework as a template fo the deductive model, we were able to incorporate variables from various different disci plines, including geographic (e.g. slope, presence of creek), economic (e.g. investments, net benefit), social (age, municipality of origin), anthropologic (ethnicity), and policy (th tenural instrument SIFMA). The same factors are incorporated in the inductive model a in that respect both models are equally multi-disciplinary, 'integrated' models. The facto comprise a good many of those listed in the recent overviews of driving factors by Geist Lambin (2002) and Lambir *et al.* (2003), even though we have focused only on simple cro choices. Since we have not compared land use in two or more points in time, our factors explanatory factors rather than dynamic 'drivers' of land us change, formally. Predictio of the effect of incremental changes in factors may be derived from both models, howev (as in Nelson *et al.* 2001, for example). In the sense of factors and predictions, therefore, t present study is comparable to mainstream land use studie

3.8.2 Field-level conclusions

The inductive approach (of type 2 in terms of the deductive/inductive ladder) has been t fit all factors to the actual land use in a multinomial regression, thus generating a struct of land use depending or $ln(P_{category\ M} / P_{reference\ category}) = ß_1x_1 + ß_2x_2 + \ldots + ß_nx_n$. The deductive ap proach (of type 5 of the deductive/inductive ladder) has been to use the factors as eleme in a model of the decisions of the land users, to quantify the model on the basis of field knowledge and then test the whole causal structure against the actual land use data. Th results of the two approaches look much alike, on the surface. The inductive model fitt percent of the observations correctly, and the deductive model predicted 70 percent of t observations correctly. The performance on yellow corn was good in both models (90 ar 85 percent, respectively). Both models overestimated the total amount of yellow corn an underestimated the amount of banana fields. In both models some of the observed bana fields were classified as yellow corn. This is possibly due to the fact that the maximum benefit of both crops is almost the same and both crops have hardly any restrictions (co supported with credit, banana has low initial investments). Therefore, small imperfectio in the model can cause the prediction to go wrong. Another cause of poor prediction of banana could be the existence of a time lag between changes in context and changes in t land use observed. Most bananas are cultivated for 5 to 15 years and may still be standir even when benefits are low. The predictive capability for wet rice was low for both moc The deductive model performed a little better than the inductive model. In the deducti model, the occurrence of wet rice is completely determined by the restrictions (absence c a creek, 'non-rice ethnicity' and tenancy) and these restrictions may be too rigid. In both models the prediction of fruit trees is weak. This is due to the fact that only five percent the observations are fruit trees and that the mechanisms behind the planting fruit trees i this area are not completely understood yet

3.8.3 Differences between inductive and deductive approaches

The real and important differences between the two models lie on a deeper level. As sai the second section, deductive approaches, because they start out from theory and maint theoretical structures, better feed back into theory development than inductive studies tend to do. The present study, for instance, is a true verification of broad rational choice theory expressed in the structure of Action-in-Context's decision model. The Popperian critique here would be, of course, that this is no surprise because one should aim to veri *un*likely structures, or to falsify the likely ones. Nevertheless, an entry to the theory leve it is and once there, progress may be pursued in many directions, including the testing of less likely actor models or spatially explicit neo-Thünian theory of moving land use frontiers (e.g. De Groot, 2003). The model structure may also be expanded easily to inclu multi-level causal influences on the region's land use, for instance through AiC's actors field, see Verburg *et al.* (2003). Using the actors field, we arrive at a multi-agent modellir of land use

A second advantage of deductive work is that, as it tests full causal structures rather tha separate causal factors, a much better grip on causality is established. Two specific aspe of this characteristic stand out from the present study. (1) Any inductive model, working it does from the data 'upward', can only fit for variables that vary across the dataset. Th influence of all factors that are constant across the dataset, such as in our case the marke price of corn, end up, implicitly, in the intercept (β_0). Therefore, it is less straightforward t predict the amount of land use change for a change in one of the factors accumulated in the intercept. A deductive model, however, allows the inclusion of all factors assumed t have causal influence (hence relevant for explanations as well as policy-oriented predic tions). In the deductive model, for instance, changes from corn to another land use coulc be predicted if the corn price, and with that its net benefit, would fall. (2) For the same reason of testing full structure rather than factors, deductive models are able to handle new phenomena, assuming that they do not alter the model structure. In the study area, for instance, cassava may be an alternative cash crop. The inductive model cannot handl this, because cassava is new, hence absent from the dataset and therewith from the mod In order to make a prediction of the region's response to cassava by way of the deductiv model we do need general cassava production data such as its price, productivity, positi in cultural preferences, accompanying credit scheme and so on, but once we have these, prediction is produced and various policy scenarios may be studied. (The predictions m turn out to be wrong, of course, but that is a normal issue of model quality

3.8.4 Reaping the benefi ts of combining inductive and deductive approaches

Research programs oft en tend toward a certain development on the deduction/inductio ladder. For example, starting out with a more or less extreme data mining (rung 1), the selection and shaping of causal factors may become more led by insight (rung 2) and researchers may end up in studies more consciously and fully in discussion with theory (rung 3). In fact, quite a number of inductively acquired insights into our fi eld research region, gathered by previous studies, informal discussions and visits, interviews and observations, underlie our own deductive model. In other words, we sojourned long on rungs 1 and 2 before our deductive jump to rung 5 in the present chapter. For the sake of clear-cut illustration we refrained here from what would in fact have been the natural follow-up of our strictly deductive approach, namely, to bett er calibrate and fi t our mo parameters on reality, *i.e.* move one level down to rung 4, searching to reach a bett er fi t than the 70 percent of the strictly deductive model. Aft er that, we could even have begu to adapt our whole model structure in discussion with theory and fi eld realities, thereby arriving at rung 3

Overall, then, the most eff ective way to reap the benefi ts of more deductive work does seem to rigidly 'go deductive' and stay there. Rather, the message should be that researc will profi t most from a consciousness of the whole spectrum between the inductive and deductive extremes, an awareness of the advantages of deductive approaches versus the currently dominant inductive research routines, and then seek the most fertile sequence and interactions between inductive and deductive work. This then is the invitation mear to emanate from the present chapter to the scientifi c community of land use change research.