Research

Modeling Land-Use Decision Behavior with Bayesian Belief Networks

Inge Aalders [1]

ABSTRACT. The ability to incorporate and manage the different drivers of land-use change in a modeling process is one of the key challenges because they are complex and are both quantitative and qualitative in nature. This paper uses Bayesian belief networks (BBN) to incorporate characteristics of land managers in the modeling process and to enhance our understanding of land-use change based on the limited and disparate sources of information. One of the two models based on spatial data represented land managers in the form of a quantitative variable, the area of individual holdings, whereas the other model included qualitative data from a survey of land managers. Random samples from the spatial data provided evidence of the relationship between the different variables, which I used to develop the BBN structure. The model was tested for four different posterior probability distributions, and results showed that the trained and learned models are better at predicting land use than the uniform and random models. The inference from the model demonstrated the constraints that biophysical characteristics impose on land managers; for older land managers without heirs, there is a higher probability of the land use being arable agriculture. The results show the benefits of incorporating a more complex notion of land managers in land-use models, and of using different empirical data sources in the modeling process. Future research should focus on incorporating more complex social processes into the modeling structure, as well as incorporating spatio-temporal dynamics in a BBN.

Key Words: *Bayesian belief networks; land cover; land use*

INTRODUCTION

Land use is the outcome of human and biophysical processes that operate in a landscape, with regard to biophysical, socioeconomic, and cultural conditions and constraints, and political context (Mather 1995, Geist and Lambin 2001, Lambin et al. 2003). The decision for land-use change is made by individual land managers based on their responses to these conditions and constraints. Land manager is defined here as the individual making the land-use decision, and can be a landowner, employee, or distant landowner, as well as tenant or crofter. Although land-use models are commonly used to enhance the understanding of complex processes, they often assume that land managers, as a group, behave uniformly and rationally (O'Callaghan et al. 1996, Lambin et al. 2000), rather than individually and emotionally, in relation to the processes that influence them and the opportunities presented by their land resource base. However, land managers' decision making can rarely be

considered uniform or rational (Burton 2004). The main challenge to modeling methodologies used for the prediction of land-use change is the development of an ability to incorporate—and manage—human behavioral effects through a complex notion of land managers (MacFarlane 1996). Decision making by individuals, groups, and societies has a profound influence on land use and land-cover dynamics (Parker et al. in press) and different methods are emerging that include decision making in land-use and land-cover change models (Briassoulis 1999). Criteria for successful incorporation of decision making into land-use and land-cover change models include a need to build on empirical evidence (Berger and Schreinemachers 2006) and, if models are to inform decision making, to be transparent.

It is recognized that over short time periods most of the land use in an area does not change. However, where change does occur, land managers are influenced by a number of factors when deciding

[1]Macaulay Institute

on the type of land use for their holding (Burton 2004), resulting in a particular land cover on their land. The natural potential of the land for the favored land use is a key factor that will determine the business choices open to the land manager. Direct evidence of the relationship between land managers' decision making, land use, and land cover is hard to obtain due to confidentiality issues. There is some evidence, based on an inventory of forestry in Grampian by Macaulay Institute, that clearly demonstrates the significance of individual land manager's decisions (Fig. 1). It shows forest land cover along the straight-lined boundaries of holdings, as opposed to diffuse boundaries of natural forest patterns, driven by biophysical factors alone. From this evidence, it can be inferred that information about land ownership (represented geographically in the form of ownership boundaries) represents the spatial unit for individual decision making and could be a valuable quantitative proxy for land managers' decision making in the prediction of land cover. Land ownership is thus used as a proxy in the modeling structures.

Individual behavior is known to be highly variable and central to decision making, due to the character of individuals themselves, and the wide range of forces that influence their decision-making behavior. For example, complex social and cultural factors lie at the root of farmer resistance to governmental woodland planting initiatives (Burton and Wilson 2000). Burton and Wilson (2000) argue that, if these behavioral factors are significant, then the value of a modeling approach based on purely rational behavior is greatly reduced. Behavioral or agent-based simulation models are being developed that deal with the decision-making process of individual land managers regarding the use of their land. This focuses on the theoretical process of decision making by individual land managers. Two examples of such an agent-based simulation model are FEARLUS, where individual behavior is examined in an artificial spatial context rather than a real world context (Polhill et al. 2002) and MABEL, which integrates agent-based modeling with geographic information systems, GIS, and Bayesian belief networks (BBNs; Alexandridis 2006, Lei et al. 2005). This paper focuses on BBNs as an alternative modeling approach, as a means to enhance our understanding of the dynamics of land-use change and to explore behavior in the decision-making process.

BAYESIAN BELIEF NETWORK MODELING

A BBN, also known as a Bayesian network or belief network, is a Bayesian methodology. It is based on multivariate probability distributions of model variables without underlying assumptions on the form of the multivariate distribution. At the same time, a BBN defines relations between variables in terms of the conditional distributions for each variable included in the network (Haas 1991), and allows reasoning under the uncertainties that are associated with these conditional distributions (Hornberger 2001, Korb and Nicholson 2004). A BBN can model situations that are characterized by inherent uncertainty, and facilitate statistical inference. There are various software packages available for modeling with BBNs, most using a graphical representation of variables and relations. The examples in this paper are generated using Netica™ Version 2.7 (Norsys, http://www.norsys.com).

The ability of a BBN to learn from new evidence for its parameters makes it possible to build a model initially with the best available information or data, and then to incorporate new and improved evidence at a later stage. This gives the opportunity to use evidence of events or variable relationships that have been observed as the best available data for building the probability distribution of the relationships in the model. The BBNs describe complex probabilistic reasoning by representing the structure of an argument in an intuitive, graphical format. Based on the evidence incorporated in the BBN and beliefs about "process," it is possible to infer (both causally and diagnostically) the state and relationships of individual variables and the impact of change. The BBNs can use limited, disparate, information sources and have the potential to incorporate qualitative information.

A variety of different modeling initiatives have explored the potential of Bayesian methods in relation to land-use and land-use change (Stassopoulou et al. 1998, Guisan and Zimmermann 2000, Prato 2000, Marcot et al. 2001, Alexandridis 2006). Some of these methodologies have developed land-use models in a purely biophysical context (Aspinall 1992, Hill et al. 1997, Tucker et al. 1997, Lineback et al. 2001), whereas others have illustrated the use of participatory modeling methods with BBNs (Bacon et al. 2002, Lynam et al. 2002, Cain et al. 2003).

Fig. 1. Forest and ownership boundaries in Grampian.
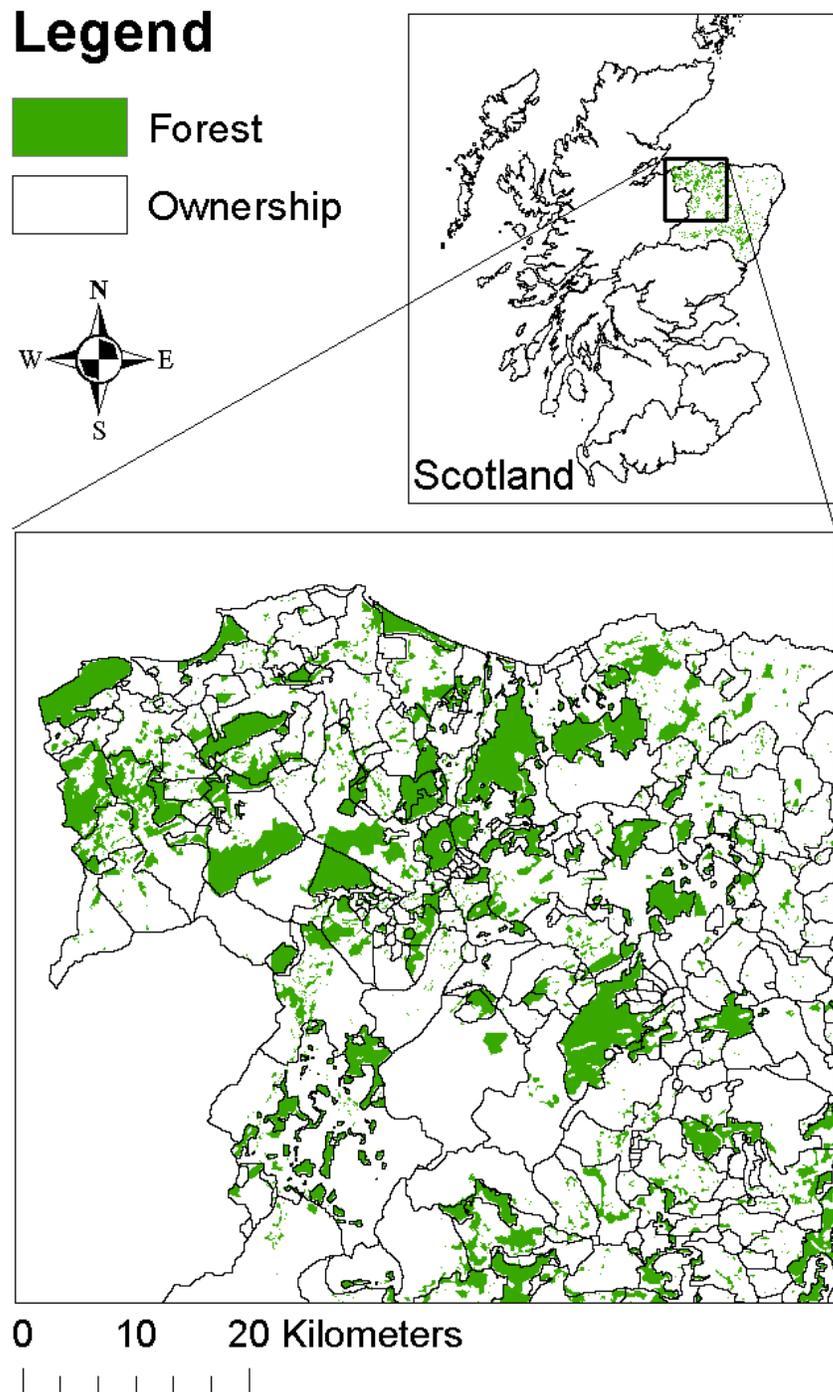
Stassopoulou et al. (1998) showed that information in a BBN flows in both forward and backward directions through the links, providing the user with the opportunity to draw both predictive and diagnostic conclusions from a network. In this way, the user can develop an advanced understanding of the process being modeled (Cain 2001). The conditional distributions in BBNs can be derived from both quantitative and qualitative information. This ability to use both types of data from different sources and to incorporate them in one model makes BBNs increasingly popular in land-use change research.

This paper uses BBNs to represent a relatively complex understanding of land managers to model and predict land-use and land-use change to enhance our understanding of land-use dynamics. The paper presents two models, one based on biophysical conditions and land-holding characteristics only, and one which expands this first model with land-manager characteristics to illustrate the additional impact of individual manager's decision making. The models in this paper adopt McConnoll and Moran's (2001) definition of land use as "...a series of operations and associated inputs on land, carried out by humans, with the intention to obtain products and/or benefits through using land resources" and land cover as "...the observed physical cover, as seen on the ground or through remote sensing, including the vegetation (natural or planted) and human constructions (buildings etc.) which cover the earth's surface."

METHODS

Study Area Selection and Description of Land Uses

Grampian region in the northeast of Scotland was selected as a study area because of its rich diversity of land cover and land uses. The area (8755 km^2) extends from the Cairngorm Mountains eastward to the North Sea coast. The northeast has been cultivated for centuries with a wide range of crops. Agriculture dominates the lowlands and much of the hill ground, about 75% of the study area is in agricultural use (crops and grass 49%, and rough grazing 26%). In recent years, Grampian has experienced change resulting from developments in agriculture, such as reform of the European Union Common Agricultural Policy (CAP) and economic

uncertainty in the oil industry caused by global market dynamics (Scottish Executive 2005). The key land-use changes observed in the region and in Scotland are caused by afforestation and urbanization (Mackey et al. 1998, Haines-Young et al. 2000). The population in Grampian of about 525 000 is strongly influenced by the local economy of the city of Aberdeen, where almost half of the population lives. Only 5% of the population are employed in the primary sectors of agriculture and fisheries (Scottish Executive 2005). The range of land-use changes driven by political and economic developments within farming systems, urbanization, and afforestation, make Grampian an ideal study area.

Model Development

For this paper, I have used existing data and information together with a logical understanding of the relationship between the variables to construct the structure of each model. I created conditional probability tables for each of the variables, based on the best available evidence, information, or knowledge. In order to test the model's ability to predict land cover, I derived separate training and test data from the available evidence. The model most successful in predicting land-cover change was selected for inference.

With Model 1, I tested the hypothesis that land cover in Grampian can be predicted based on empirical information on the land capability for agriculture (LCA) (Bibby et al. 1991) and for forestry (LCF) (Bibby et al. 1988), and the size of a holding (Wightman 1996) at a scale of 1:250 000. For this purpose, a simple BBN (Fig. 2) was developed to represent the logical relationship between these variables and the land cover of Scotland (LCS88) (Macaulay Land Use Research Institute (MLURI) 1993). The LCA integrates soils data with knowledge of climate and topography to assign areas of land into seven classes ranging from land capable of producing a very wide range of crops (Class 1) to land of very limited agricultural value (Class 7). The LCF is a similar classification, with seven classes ranging from land with excellent capability for the growth and management of tree crops to land with very limited capability for the growth and management of tree crops and urban development. The LCS88 is based on an interpretation of 1:24 000 scale black and white aerial photography. The 34 summary classes of the

original legend were simplified for this paper to six functional classes (arable land, improved grassland, rough grazing, forest, development, and other land; see Table 1).

All the spatial data (LCA, LCF, LCS88, and area) were represented in raster format, consisting of approximately 11 000 individual grid cells with a 50 x 50 m grid resolution. These data, combined in a logical overlay, represent the qualitative evidence of relationships between the variables through a joined attribute table with the four variables' values for each cell. From these data, I randomly selected training and test data, consisting of 5000 cells, with each cell representing a piece of evidence regarding the relationship between the variables. The area of ownership in the training and test data is represented as a continuous variable. However, for the purpose of the graphical representation in the BBN, this continuous variable needs to be categorized into discrete classes. Unlike the other variables in the models, which are categorical variables, area remains a continuous variable, which means that any evidence of the land capability and land cover on a holding with a particular size can be incorporated into the model. The categories of area in the graphical representation were the minimal number of classes required to distinguish different farming systems in Grampian (Table 2). This model used the area of individual holdings as spatial units to represent individual decision making, together with the two land-capability data sets, to predict land cover. I used land cover in this model as a proxy for land use because it is built on observed physical cover rather than land use as defined by McConnoll and Moran (2001).

Model 2 expanded on the spatial evidence-based Model 1 with the addition of some simple, non-spatial, factors (age, and whether or not a land manager can pass his business on to his children). In the decision-making process concerning changes in business activities, the importance of qualitative factors is based on an independent survey of land managers in the area around Huntly (Grampian) about their attitudes toward land-use change (Burton 2003). I modified the original BBN (Model 1) to include a variable of land-cover change, which linked the land-cover and land-use change variables. Land-use change provided a link between the quantitative and qualitative data by connecting land-cover change and the decision by land managers to change their land use, given an heir and their age. The reason for this construction was that

the survey data provided no insight as to the type of land use subject to change. Information about change in land cover for Grampian was limited, but data were available for other parts of Scotland, in particular, the Central Valley and the Cairngorms, where land-cover change gave an indication of the changes for Scotland as a whole. The information for the Central Valley indicated that the "rough grazing" and "other" land-cover types were the most vulnerable to change, whereas agricultural activity, forestry, and urban development were relatively stable. The results from the survey in Huntly (Burton 2003) indicated that change in Grampian had occurred in the same land-cover classes, but overall no change was estimated for approximately 80% of the land area. Therefore, this paper assumed that the drivers of change for the Central Valley of Scotland were similar to those over the same period in Grampian, and that the relationship between land cover and land-cover change derived from these data was representative for that relationship in Grampian.

In the building of Model 2 (Fig. 3), I made the simple assumption for the creation of conditional probability tables that a decision to change by a land manager means land-use change and this will result in land-cover change. I used Model 2 to infer the land manager's characteristics, based on knowledge of the land cover. In addition, I inferred the likely land cover for a particular change hypothesized in these social factors.

Populating the Model with Information

The conditional probability table of a variable in the model represented the relationship between that variable and those variables that have a link into that variable. By testing that variable based on evidence, the results provided information about the ability of the model to predict land cover and land-cover change.

With the model structure in place, the available information needs to be incorporated in the structure (training) and its reliability tested. The tests were conducted to establish whether the random training samples are representative for the study area, and whether the models after training were better at predicting land cover than the same models with random or uniform variable distributions. From the combined spatial data (LCA, LCF, LCS88, area of

Fig. 2. Model 1 (for the continuous variables, the average and the standard deviation are given below the probability distribution).
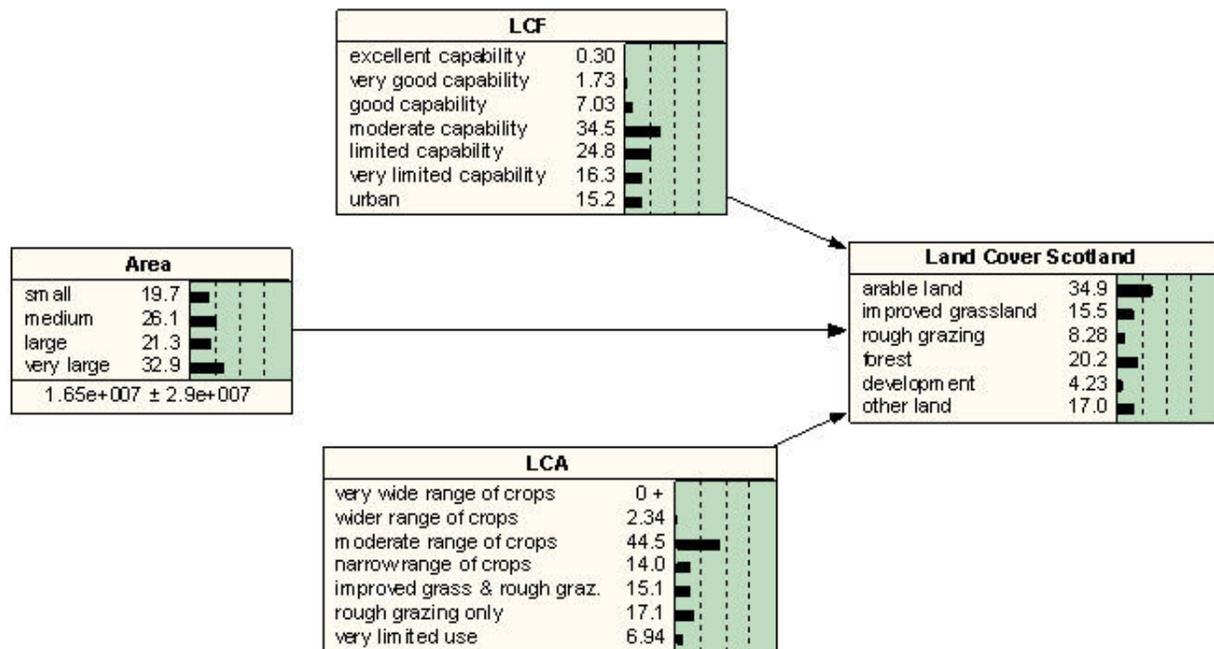
Area	Class description
small	< 0.5 ha
medium	0.5 – 100 ha
large	100-1000 ha
very large	> 1000 ha

land, and land-cover change), I selected 10 random samples of 5000 cells each representing a unique combination of values for the model variables. The samples were used as evidence of the relationship between the spatial variables. I used these samples to train the BBN structure with a uniform prior probability distribution resulting in 10 different variations of Model 1. In addition to these trained BBNs, Model 1 was built with a uniform and random probability distribution for the variables. I tested each of the resulting 12 different variations of Model 1 using the 10 test sample data. The 10 tests determined the accuracy with which the model

was able to correctly classify or predict the test sample into the appropriate classes (http://www.norsys.com/tutorials/netica/secD/tut_D2.htm). These results were collected in a confusion matrix to give an error rate, as the overall proportion of incorrect prediction of land-cover classes with respect to the probability distribution of the test evidence.

A key advantage of a BBN is that it is possible to update the conditional probability tables of the model parameters with additional data. To examine the impact of new knowledge on the accuracy of the model's predictability, in addition to the above 12

Table 1. Summary classes of LCS88.

Summary class		Original 34 classes	Original class description
1	Arable	1	Arable
2	Improved grass	2	Improved grassland
		30	Improved grassland/good rough grassland
3	Rough grazing	3	Good rough grassland
		4	Poor rough grassland
		26	Poor rough grassland/heather moorland
		27	Good rough grassland/heather moorland
		29	Good rough grassland/poor rough grassland
		31	Good rough grassland/bracken
		32	Poor rough grassland/peat
4	Forest	10	Felled woodland
		11	Recent planting
		12	Coniferous plantation
		13	Semi-natural coniferous
		14	Mixed woodland
		15	Broadleaved woodland
5	Development	22	Rural development
		23	Urban
6	Other	5	Bracken
		6	Heather moorland
		7	Peatland
		8	Montane
		9	Rocks and cliffs
		16	Scrub
		17	Freshwater
		18	Marshland
		19	Salt marshland

(con'd)

20	Dunes
21	Tidal waters
24	Missing/obscured
25	Heather moorland/peat
28	Peat/montane
33	Heather moorland/montane
34	Other mosaics

"trained" models, I built one model using all 10 training samples, and tested this "learned" model in the same way as the above models, resulting in 13 variations of the models.

The process of training and testing Model 2 was similar to Model 1 for the variables that they had in common. In addition, I trained and tested the conditional probability tables of the land-cover change variable representing the relationship between land cover and land-cover change using data for the Central Valley for the periods from 1948 to 1977 and 1977 to 1988 (Table 3). The conditional probability table of the qualitative data regarding the land manager's decision making (Table 4) was based on the expert interpretation of the survey results in Huntly (Burton 2003). With confidence in the model's predictability, I used the most successful model to infer the impact of changes in key variables on land cover and land-cover change.

Inference

After building the models through the incorporation of data and testing them as described above, I selected the best performing Models 1 and 2 of the 13 different variations, and used these to infer the behavior/response of the variables to changes in circumstances. In the process of inference, I identified the specific values of individual variables in the models as known by giving them a 100 % probability instead of the probability based on the training evidence. The BBN automatically updated the probabilities for the other variables based on that "knowledge." The probability distributions of the

various "knowledge" scenarios, including the one with no knowledge for any of the variables, were collected in tables and compared. The different "knowledge" scenarios explored in this paper provided an insight into the impact on land cover of the size of the land holding, land capability classes, and land manager's circumstances. Each of these factors could be expected to change under the influence of CAP reform and climate change, and affect future land cover in the area.

RESULTS

Test Results of Models 1 and 2

I trained Model 1 with the 10 separate random training samples, each sample consisting of 5000 cells that each represent a piece of evidence of the relationship between the model's variables. The boxplots of the posterior probability distribution of these 10 models (Fig. 4) illustrate that they were representative samples for the study area, with few outliers. Each of the variables has one dominant value.

The 10 test samples formed the basis for the results of the test of the models. The samples tested each variation of the model with a uniform, random, trained or learned probability distribution. The results show, as illustrated in the boxplot (Fig. 5), that the trained and learned models, with a mean error rate of 39% and 33%, respectively, had much better test results than the models based on a uniform and random distribution, with a mean error rate of 63% and 80%, respectively. The fact that the error rate for the learned model was better than that of the

Table 2. Area classification.

Class of area	Class description
small	<0.5 ha
medium	0.5–100 ha
large	100–1000 ha
very large	>1000 ha

trained model also suggested that the model did improve in its ability to predict land cover, through learning from additional evidence.

For Model 2, I built the conditional probability tables for the LCA, LCF, the area of holding and land cover of Scotland variables in the same way as the learned Model 1, i.e., based on the cumulative evidence of the 10 training samples, and tested the additional land-cover change relationship to land cover for its ability to predict land-cover change; the test results illustrated that the model was able to predict the land-cover change with an error rate of 12.4%.

**Inferences from Model 1 about Land Cover and Land-Cover Change**

I inferred the results of a range of different scenarios based on simple "known" values for variables from Model 1. "No knowledge" gave the probability of land cover without any particular prior knowledge regarding the variables. The other four scenarios assumed that the area of the holding, either small, medium, large, or very large, was known but with no knowledge about the other variables. Table 5 included the probability distribution of land-cover classes of these scenarios. The results showed that the land cover of small holdings was more likely to include arable land and forest, and less likely than the other type of holdings to include improved grassland, whereas medium-sized holdings were more likely to include "other land" than forest.

Table 6 similarly illustrated the probability distribution of land cover for four different types of

agricultural capability (LCA). Land in LCA Classes 2 and 3 were among the more suitable for agriculture. Therefore, it was not surprising that there was a strong probability of arable land in areas of these classes. The probability of "all findings," which was an indication of the amount of evidence from all the evidence for particular results, illustrated that Grampian had only a very small proportion of good-quality agricultural land and that most of the land was LCA Class 3, i.e., land capable of producing a moderate range of crops (Bibby et al. 1991). The results showed a shift from arable land to improved grassland and forest in areas of less favorable land. The results of the inferences (Tables 5 and 6) conformed to the existing understanding regarding the relationship between the natural resource of the land to land use, and between farm size and land use in the region. However, because land-use decision making was based on a multitude of factors, a combination of these variables for a predictive conclusion would be a more interesting and realistic inference because it was that combination that provided the land manager with the natural resource conditions for his holding.

I examined in four different scenarios the proxy for land manager decision making (area) with a particular combination of land capability conditions (Table 7). The table illustrated the probability distribution of land cover for land with the capability of producing a narrow range of crops (LCA Class 4) and suitability for forestry, which was at the more productive end of the range of classes (LCF Class 3), for each of the four classes of areas of land ownership. These results illustrated that small and medium areas of land owned with such natural

Fig. 3. Model 2 (for the continuous variables, the average and the standard deviation are given below the probability distribution).
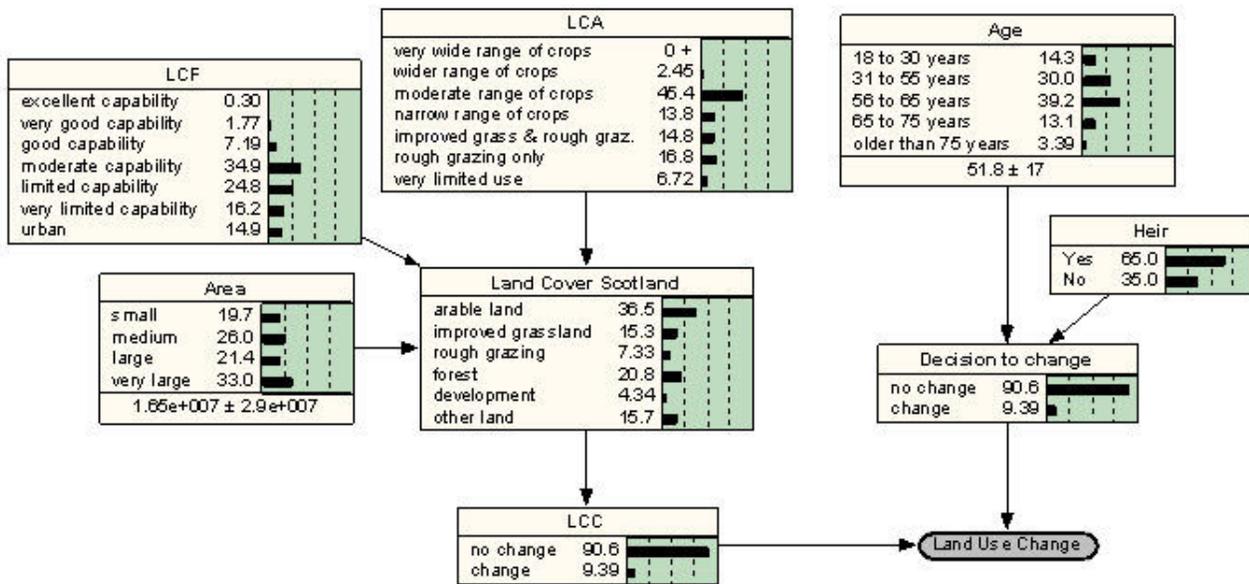
Area	Class description
small	< 0.5 ha
medium	0.5 – 100 ha
large	100-1000 ha
very large	> 1000 ha

resources had a high probability of being arable, whereas very large areas had the highest probability of being forest. For large owned areas, it was equally probable that the land cover was arable, improved grassland or forest. For this particular combination of LCA, LCF, and area the probability for rough grazing, other land and development are marginal land covers. The possible number of combinations of the values for the three variables was much larger than those for single variables. Therefore, these results, in contrast to those in Tables 5 and 6, represented only a small proportion of all findings (<0.5%). By selecting each land cover in turn, I used

the model to infer a diagnostic conclusion regarding the probability for particular conditions with that land use. In Table 8, the results showed that every land-cover type occurred on all sizes of land holding. However, there were distinct differences based on the land capability. For example, improved grassland was least likely on small farms, but most likely on farms with a capability for agriculture of LCA Class 3 (0.48) or 4 (0.25) and a capability for forestry of LCF Class 4 (0.38) or 5 (0.27). Other land could be found on holdings of similar size, but only on areas with poor land capabilities, LCA Classes 5 and above (land suited only to improved

Table 3. Conditional probabilities for land-cover change.

Land Cover Scotland	No change	Change
Arable	76.2	23.8
Improved grassland	97.7	2.3
Rough grazing	72.8	27.2
Forest	79.2	20.8
Development	71.3	28.7
Other land	89.9	10.1

grassland and rough grazing and land of very limited agricultural value). The results also showed that the probability for development on land with suitable agricultural capabilities was high (0.68) in Grampian.

The inference results clearly demonstrated the constraints that biophysical characteristics posed on land managers. Given a land capability, the area of a holding was relevant to the nature of its land use (Table 5).

Inferences from Model 2 about land cover and land-cover change

The results of inference for the prediction of land-cover change and land cover based on age groups and inheritance status are presented in Figs. 6 and 7, respectively. They show that, without an heir to the ownership of the land, the land cover is more likely to change (Fig. 6), in particular for older land managers, who will modify production to suit their physical ability in the later stages of their life. The choice of change made by older farmers is for arable land, rough grazing, and forest at the expense of improved grassland and "other land" (Fig. 7). Generally, low levels of change in land cover for those areas with an ownership inheritance suggest that inheritance supports continuity of land management, and that any changes to the system are marginal.

DISCUSSION

Based on the qualitative and quantitative data sources available for this particular paper, only the quantitative information can be spatially related and provide us with evidence of their relationship. Unfortunately it is not possible to make a spatial link between the qualitative and quantitative information, which means that the data contain a substantial level of uncertainty. This meant that otherwise very useful methodologies such as cellular automata and neural networks could not be deployed for the development of an integrated empirical model such as Model 2. Cellular automata would require more detailed information about the relationships between the variables in the model whereas, for the development of a successful neural network, at least a sound set of training data incorporating information of all the variables would have to be available. These methodologies, although useful for Model 1, cannot be used for inference. In particular, for enhancing our understanding on land-use change dynamics, a BBN's ability to infer is an added value. For example, training Model 1 leads the probability for development of 4.23 % (i.e., 0.04, Fig. 2). However, through inference (Table 8), the probability of development being located on good-quality land (LCA3) is 68% (i.e., 0.68). One could draw the conclusion that good-quality land is at risk from irreversible development, which could pose a risk to agricultural productivity; however, the probability that LCA3 is covered by development is only 0.06. This means that, although more than

Table 4. Conditional probabilities for decision to change (Model 2).

| Age groups	Heir	No change	Change
18–30 years	yes	60	40
	no	60	40
31–55 years	yes	90	10
	no	90	10
56–65 years	yes	70	30
	no	50	50
66–75 years	yes	60	40
	no	20	80
Older than 75 years	yes	60	40
	no	0	100

half the land being developed is on good-quality agricultural land, in this region that process of change does not pose a serious threat to agricultural productivity.

The availability of quantitative information allows us to test Model 1 on its reliability to model land cover. The results show that, trained with empirical evidence, the model is significantly better in predicting land cover than models with uniform or random distribution. They also show the benefit of learning from additional evidence, as illustrated by an improved error rate between the results of the model trained with one sample data set and the model trained with 10 different sample data sets. This is particularly relevant for the development of empirical models constrained by data quality and availability. More importantly, when the process or its drivers are changing, as shown by Aspinall and Hill (2000), this ability to learn can be particularly relevant. The evidence of these changes could be derived both from empirical data and computer-generated data. New evidence represents changes as modifications in the variable relationships (conditional probability table) and their probability distribution. However, despite these benefits, the integration of social factors through land ownership

in the modeling process provides no information regarding the nature or the impact of the social factors. Therefore, Model 1 is both inflexible and uninformative regarding the inference about the decision-making process of land managers. Model 2, although limited, is able to link land cover and land-cover change information in a way that allows us to infer the implications of changes within the process of land-use change on land cover. As a result, even in the rather simple example in this paper, the inference results from Model 2 show the importance of inheritance on land-use change and land-cover change.

Model 2 has been developed as an integration of the first model with non-geographic variables (age and inheritance), which aims to introduce a more complex notion of individual land managers. The results show that the inclusion of a more complex notion of land mangers improves our ability to explain and predict land-use and land-cover change. However, although all land-cover change results from a land-use change decision, not all those decisions lead to land-cover change. Therefore, more realistic models rely on improving the conditional probability table based on expert knowledge or by incorporating evidence from

Fig. 4. Boxplots of the mean posterior probability distribution for area of land.
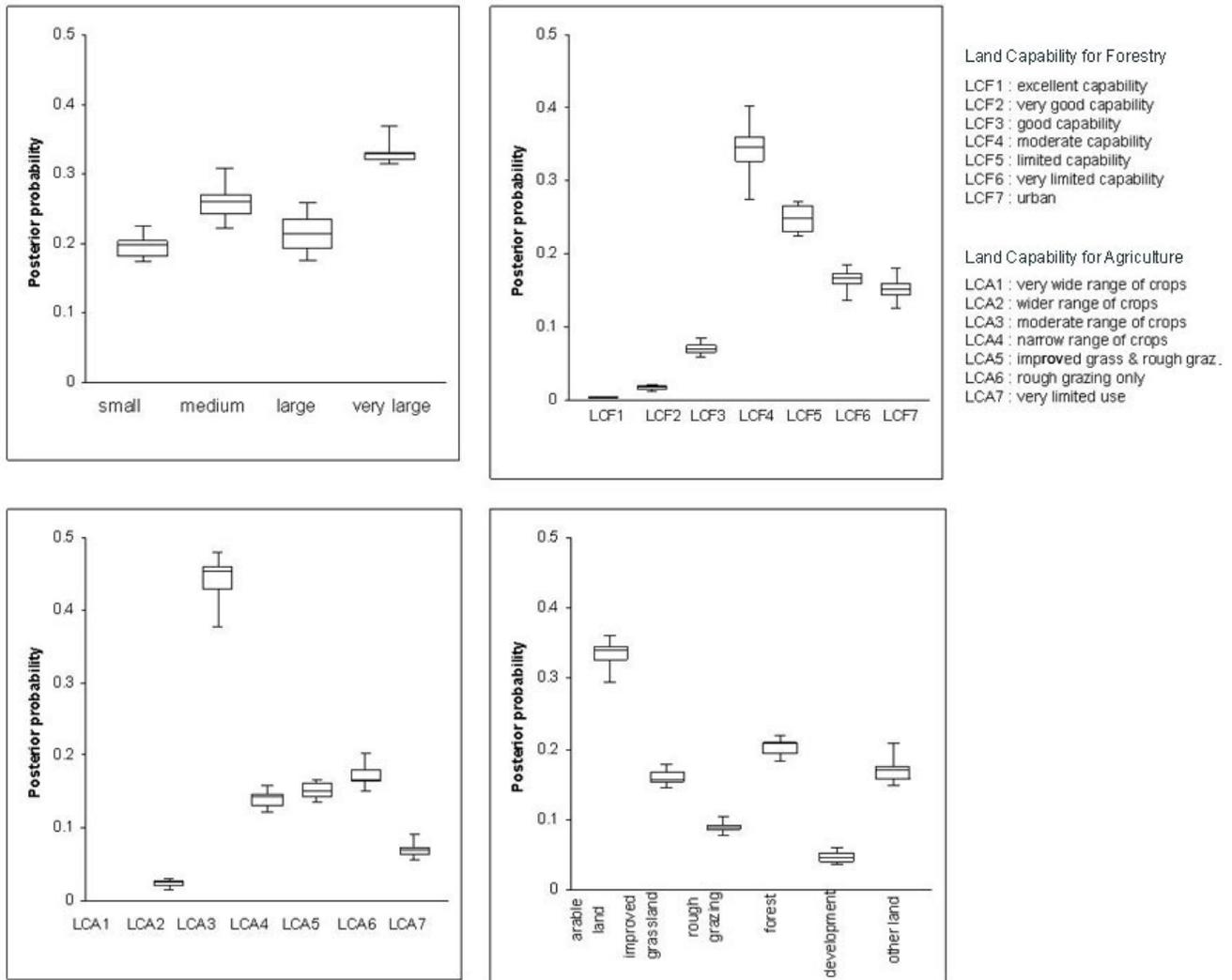
empirical or computer-generated simulation data regarding the relationship between these variables. However, this will only become meaningful when the model is developed to include more detail about the decision-making process, which may or may not lead to land-cover change. The results in this paper illustrate that the BBN is able to use different empirical data sources in a land-cover model that begins to incorporate a more complex notion of land managers than the uniform and "rational" behavior.

This inclusion and the combination of learning from new information and use of causal and diagnostic inference, can contribute to improving our understanding of the dynamics of land-use change, but this requires an expansion of the models to take account of a even more complex decision-making process, as well as expanding and improving the knowledge incorporated in the models with empirical and model simulation data.

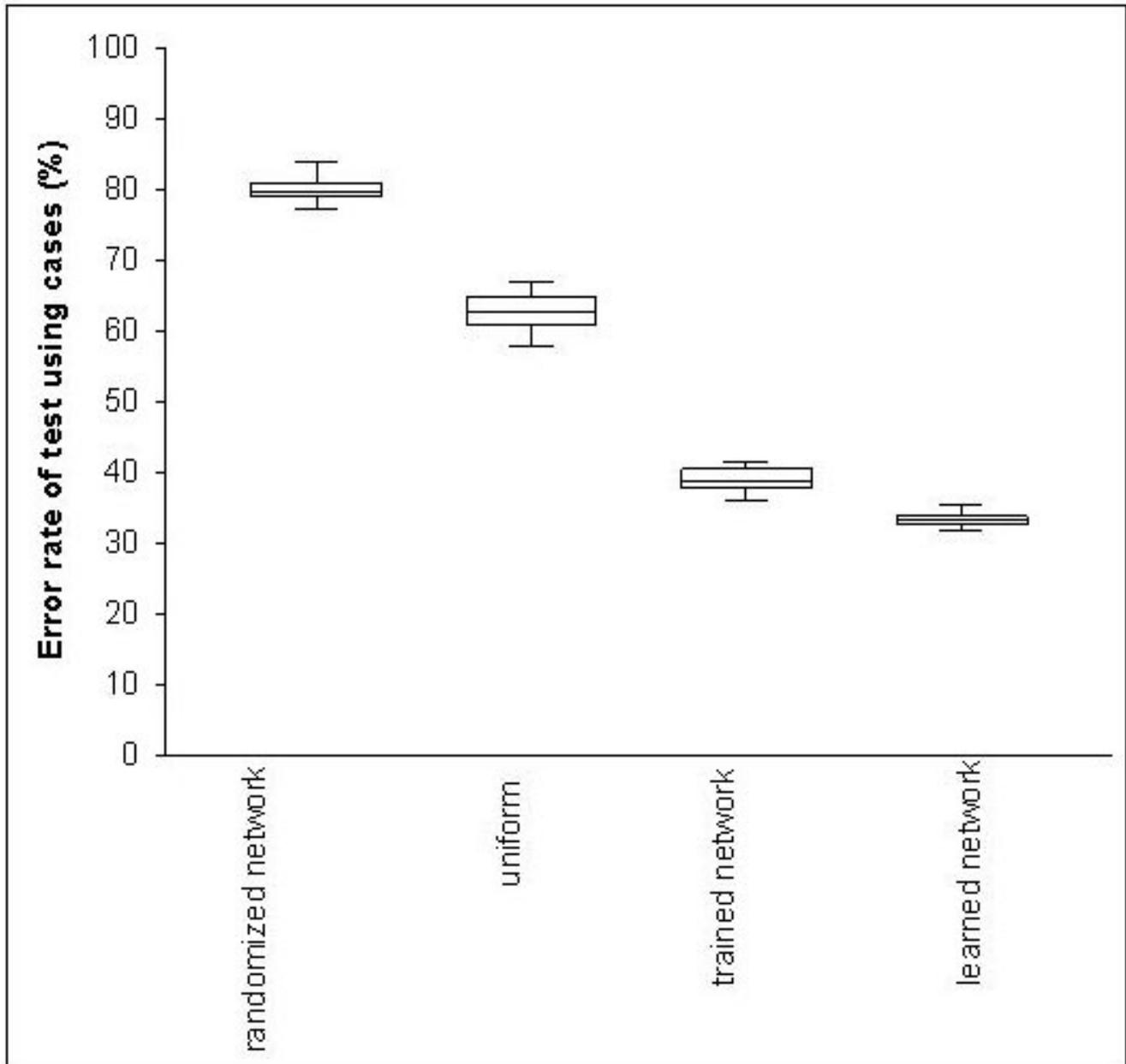**Fig. 5.** Boxplots of mean error rate of the test using cases.

Table 5. Inference results of Model 1 (area of holding).

Land cover of Scotland	No knowledge	Area = small[1]	Area = medium[2]	Area = large[3]	Area = very large[4]
Arable land	0.35	0.37	0.36	0.35	0.32
Improved grassland	0.15	0.12	0.16	0.17	0.16
Rough grazing	0.08	0.08	0.08	0.08	0.09
Forest	0.20	0.26	0.15	0.19	0.22
Developed rural	0.04	0.05	0.04	0.03	0.05
Other land	0.17	0.12	0.21	0.18	0.16

[1] Probability of all findings = 19.7%.
[2] Probability of all findings = 26.1%.
[3] Probability of all findings = 32.9%.
[4] Probability of all findings = 19.7%.

Based on available information for Model 2, the results showed that land cover is more stable when there is an heir to take over the management of a holding. The inference from Model 2 showed an increase in arable production on holdings where there are no children to follow their parents in the business. This is in line with the shift reported by Burton (2003) that can be expected, in particular, when older land managers phase out physically demanding activities such as livestock production from the farming system in favor of less demanding arable production. The models presented in this paper provide an insight into the way different social factors influence land cover in Grampian. Although land-use change only occurs on 20% of the area, the results suggest that most of that change can be related to farms where there is no prospect of inheritance. However, with the limited information available, the model is sensitive to bias and, therefore, further research that includes a more detailed representation of the decision-making process by land managers can strengthen the findings in this paper.

In general, BBNs are useful for those modeling circumstances where there are, at that stage, imperfect or disparate data and information sources that are unsuitable for other modeling methodologies. Other methodologies commonly used under those circumstances are cellular automata and neural networks. However, both these methods require evidence or understanding about the direct relationship between decision making and the biophysical environment unavailable for the creation of Model 2. This shows the advantage of a BBN's ability to use information sources that are not easily compatible, and to draw predictive and diagnostic conclusions. The key concern about using a BBN remains the risk of bias, which can affect the value of the inference drawn from it. However, during the process of inference, clear information is given about the amount of evidence supporting the outcome, as shown in Tables 5 to 7, which indicates the areas where additional evidence would strengthen the results.

CONCLUSIONS

Land-use change is driven by a complex of biophysical, social, and economic factors. For empirical modeling, this poses key challenges because the acquisition of empirical data commonly uses different methodologies, which makes

Table 6. Inference results of Model 1(LCA).

Land cover of Scotland	LCA = 2[1]	LCA = 3[2]	LCA = 4[3]	LCA = 5[4]
Arable land	0.60	0.62	0.18	0.07
Improved grassland	0.10	0.17	0.28	0.11
Rough grazing	0.07	0.04	0.12	0.14
Forest	0.08	0.07	0.26	0.40
Developed rural	0.07	0.06	0.01	0.01
Other land	0.08	0.03	0.14	0.14

[1] Probability of all findings = 2.3%.
[2] Probability of all findings = 44.5%.
[3] Probability of all findings = 14.0%.
[4] Probability of all findings = 15.0%.

integration into one model difficult. In this paper, it is shown that BBNs are able to use information from different sources that are not directly compatible to generate a model that allows us to infer implications of change both in the social and biophysical aspects of land-use dynamics. From the inference on land-use change in Grampian, the conclusion emerges that conversion of land use toward arable agriculture can be attributed to older farmers without the prospect of continuity on the farm.

The models developed for this paper are deliberately kept simple; however, even with the limited amount of detail in the land manager's decision-making process, the results demonstrate that the incorporation of a more detailed notion of land managers and a BBN's ability to infer allow us to explore the impact of different personal circumstances on the land-use decision-making process.

Although the ability to use a wide range of information sources, quantitative and qualitative, is shown in this paper, it leaves BBNs open to the criticism of bias, and to concerns about the quality of the information, which may vary within the network, with one relationship providing more reliable inference than others. However, through its ability to learn from evidence, a BBN can offer land-use change research an evolving modeling process, where its flexibility can be used to update a model and modeling structure, reflecting the emergence of new knowledge and understanding. Therefore, BBNs are most useful in the early stages of research when learning and inference from limited data can identify gaps in our understanding and knowledge, which will allow us to formulate and direct future research in land-use change Ultimately this evolving process can lead to knowledge and understanding, which would be used more effectively in other modeling methodologies.

The models in this paper are not proper dynamic models of land-use change. They represent a static but interactive representation of the process of change, which can be explored through the inference of different scenarios. The development of spatio-temporal BBNs remains part of ongoing research on the use of BBNs in land-use change modeling, as well as the development of a more detailed integration of the decision-making process and use of the advantages and strengths for integration with other modeling methodologies.

Table 8. Diagnostic inference—the probability of a land cover for particular site characteristics.

	Land cover class	Arable land	Improved gra- ssland	Rough grazing	Forest	Developed rural	Other land
Area	Small	0.21	0.15	0.18	0.25	0.25	0.14
	Medium	0.27	0.27	0.26	0.19	0.22	0.33
	Large	0.21	0.23	0.22	0.20	0.14	0.23
	Very large	0.30	0.35	0.34	0.36	0.39	0.31
LCA	1	0.00	0.00	0.00	0.00	0.00	0.00
	2	0.04	0.02	0.02	0.01	0.04	0.01
	3	0.79	0.48	0.23	0.15	0.68	0.09
	4	0.07	0.25	0.21	0.18	0.04	0.12
	5	0.03	0.11	0.25	0.30	0.02	0.24
	6	0.04	0.09	0.19	0.32	0.04	0.37
	7	0.02	0.05	0.10	0.04	0.18	0.17
LCF	1	0.00	0.00	0.00	0.00	0.00	0.00
	2	0.02	0.01	0.02	0.02	0.02	0.01
	3	0.08	0.08	0.05	0.11	0.08	0.02
	4	0.44	0.38	0.27	0.43	0.18	0.10
	5	0.27	0.27	0.25	0.26	0.13	0.21
	6	0.10	0.18	0.22	0.13	0.14	0.29
	7	0.09	0.08	0.19	0.06	0.46	0.38

Fig. 6. Probability of change/no change in land cover for a land manager's age and succession.

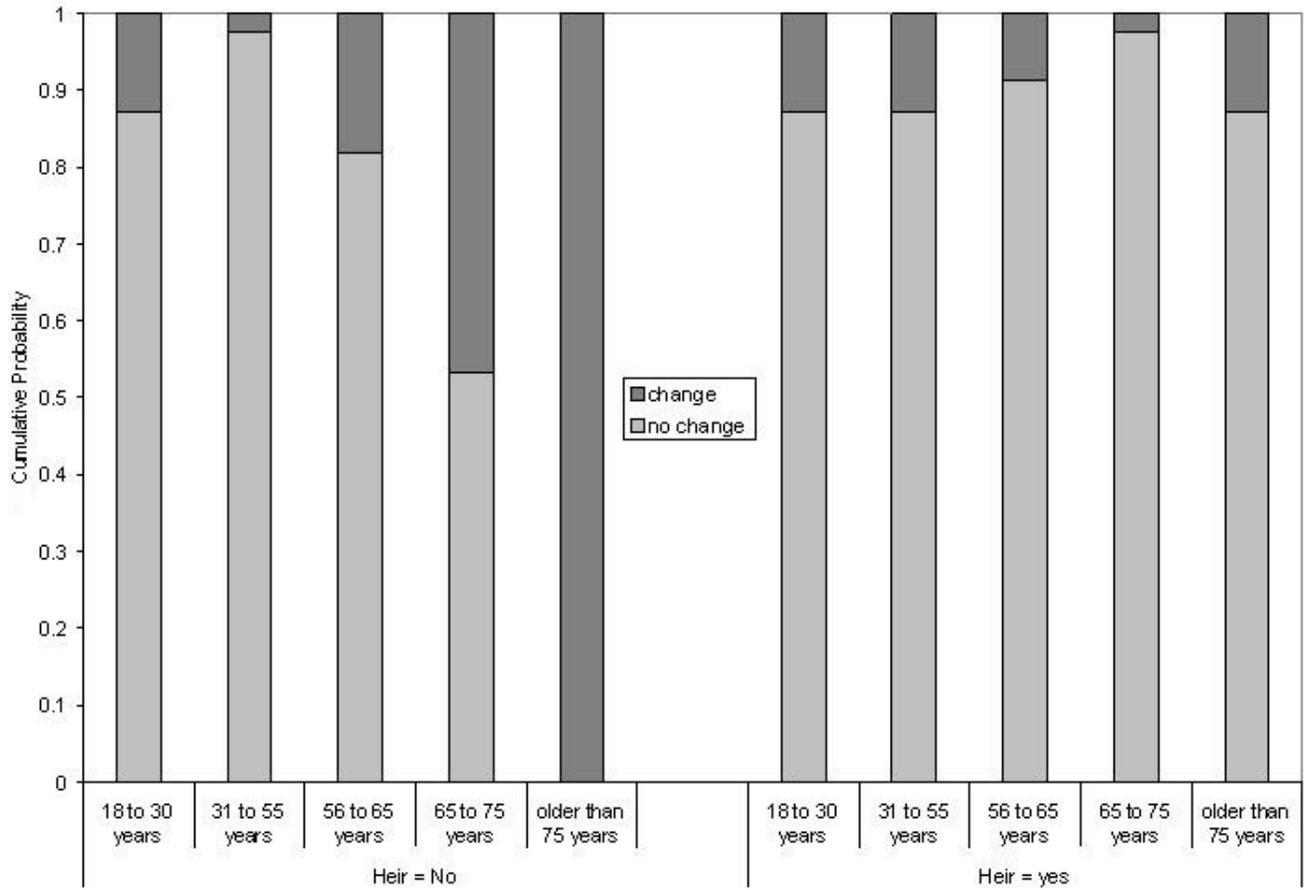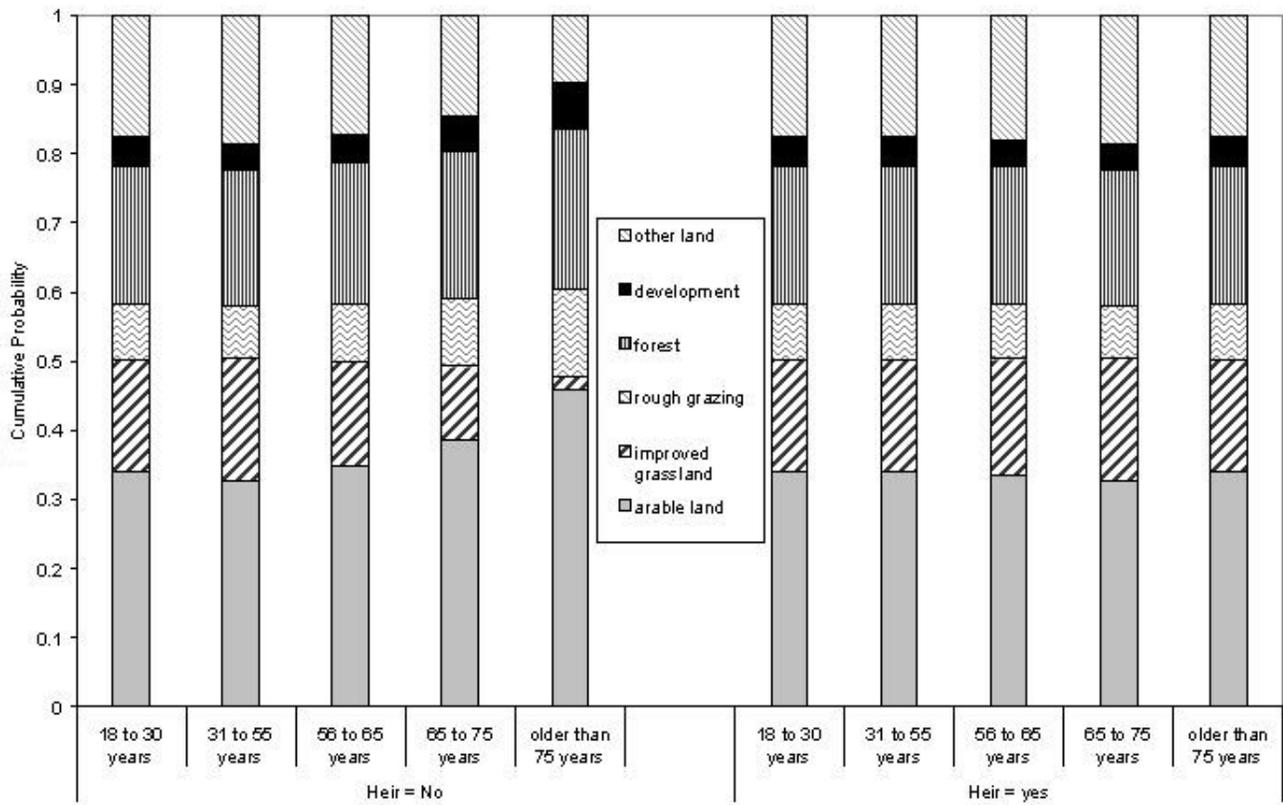**Fig. 7.** The probability of land cover for a land manager's age and succession.

LITERATURE CITED

Alexandridis, K. 2006. *Exploring complex dynamics in multi agent-based intelligent systems: theoretical and experimental approaches using the multi agent-based behavioral economic landscape (mabel) model.* Dissertation, Purdue University, West Lafayette, Indiana. ProQuest Information and Learning Company, UMI Dissertation Services (Source: DAI-B 67/09, p.4949, March 2007) 280 p.

Aspinall, R. 1992. An inductive modelling procedure based on Bayes' theorem for analysis of pattern in spatial data. *International Journal of Geographical Information Systems* **6**(2):105–121.

Aspinall, R., and M. J. Hill. 2000. Urban growth —analysis, modeling and prediction of past, present and future land-use changes in Gallatin Valley, Montana (1860–2010). *In* Proceedings of *4th International Conference on Integrating GIS and Environmental Modeling.* 2–8 September 2000, Banff, Alberta, Canada.

Bacon, P. J., J. D. Cain, and D. C. Howard. 2002. Belief network models of land manager decisions and land-use change. *Journal of Environmental Management* **65**:1–23.

Berger, T., and P. Schreinemachers. 2006. Creating agents and landscapes for multiagent systems from random samples. *Ecology and Society* **11**(2): 19. [online] URL: http://www.ecologyandsociety.org/vol11/iss2/art19/.

Bibby, J. S., H. A. Douglas, A. J. Thomasson, and J. S. Robertson. 1991. *Land capability classification for agriculture.* The Macaulay Land Use Research Institute, Aberdeen, Scotland.

Bibby, J. S., R. E. F. Heslop, and R. Harnup. 1988. *Land capability for forestry in Britain.* The Macaulay Land Use Research Institute, Aberdeen, Scotland.

Briassoulis, H. 1999. Analysis of land use change: theoretical and modeling approaches. Regional Research Institute, West Virgina University, Morgantown, West Virginia, USA.

Burton, R. J. F. 2003. Agricultural development in Scotland: the influence of the changing social structure of "family farms." *In* MLURI Annual Report 2002, The Macaulay Land Use Research Institute, Aberdeen, Scotland.

Burton, R. J. F. 2004. Seeing through the "good farmer's" eyes: towards developing an understanding of the symbolic value of utilitarian production. *Sociologia Ruralis* **44**:195–215.

Burton, R., and O. Wilson. 2000. *Farmers' resistance to woodland planting in community forest: the influence of social and cultural factors.* Occasional Paper, De Montfort University, Leicester, UK.

Cain, J. D. 2001. *Planning improvements in natural resources management—Guidelines for using Bayesian networks to support the planning and management of development programmes in the water sector and beyond.* Centre of Ecology and Hydrology, Wallingford, UK.

Cain, J. D., K. Jinapala, I. W. Makin, P. G. Somaratna, B. R. Ariyaratna, and L. R. Perera. 2003. Participatory decision support for agricultural management. A case study from Sri Lanka. *Agricultural Systems* **76**(2):457–482.

Geist, H. J., and E. F. Lambin. 2001. *What drives tropical deforestation? A meta-analysis of proximate and underlying causes of deforestation based on sub-national case study evidence.* LUCC International Project Office, University of Louvain-la-Neuve, Louvain-la-Neuve, France.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**(2–3):147–186.

Haas, T. C. 1991. A Bayesian belief network advisory system for aspen regeneration. *Forest Science* **37**(2):627–654.

Haines-Young R. H., C. J. Barr, H. I. J. Black, D. J. Briggs, R. G. H. Bunce, R. T. Clarke, A. Cooper, F. H. Dawson, L. G. Firbank, R. M. Fuller, M. T. Furse, M. K. Gillespie, R. Hill, M. Hornung, D. C. Howard, T. McCann, M. D. Morecroft, S. Petit, A. R. J. Sier, S. M. Smart, G. M. Smith, A. P. Stott, R. C. Stuart, and J. W. Watkins. 2000. *Accounting for nature: assessing habitats in the UK countryside.* DETR, London, UK. ISBN 1 85112 460 8.

Hill, M. J., R. J. Aspinall, and W. D. Willms. 1997. Knowledge-based and inductive modeling of rough fescue (*Festuca altaic, F. campestris*, and *F. halli*) distribution in Alberta. *Canadian Ecological Modeling* **103**:135–150.

Hornberger, J. 2001. Introduction to Bayesian reasoning. *International Journal of Technical Assessment of Health Care* **17**(1):9–16.

Korb, K. B., and A. E. Nicholson. 2004. *Bayesian artificial intelligence.* Chapman and Hall/CRC, London, UK.

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

Lambin, E. F., M. D. A. Rounsevell, and H. J. Geist. 2000. Are agricultural land-use models able to predict changes in land-use intensity? *Agriculture Ecosystems and the Environment* **82**:1–3.

Lei, Z., B. Pijanowski, and K. Alexandridis. 2005. Distributed modeling architecture of a Multi-Agent-based Behavioral Economic Landscape (MABEL) model. *Simulation and Modeling International* **81**(7):503–515.

Lineback, M., W. A. Marcus, R. J. Aspinall, and S. Custer., 2001. Evaluation of GIS-based terrain and wetness models for assessing landslide potential in the Boise National Forest, Idaho. *Geomorphology* **37:**149–165.

Lynam, T., F. Bousquet, C. Le Page, P. d'Aquino, O. Barreteau, F. Chinembiri, and B. Mombeshora. 2002. Adapting science to adaptive managers: spidergrams, belief models, and multi-agent systems modeling. *Conservation Ecology* **5**:24. (online) URL: http://www.ecologyandsociety.org/vol5/iss2/art24/.

Macaulay Land Use Research Institute (MLURI). 1993. *The land cover of Scotland 1988.* MLURI, Aberdeen, Scotland.

MacFarlane, R. 1996. Land-user intentions and land-use modelling. Pages 215–226 *in* D. Parker, editor. *Innovations in GIS 3: selected papers.* Taylor and Francis, London, UK.

Mackey, E. C., M. C. Shewry, and G. J. Tudor. 1998. *Land cover change in Scotland from the 1940s to the 1980s.* The natural heritage of Scotland, Scottish Natural Heritage 6. The Stationary Office, Edinburgh, Scotland.

Marcot, B. G., R. S. Holthausen, M. G. Rápale, M. M. Rowland, and M. J. Wisdom. 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. *Forest Ecology and Management* **153**:29–42.

Mather, A. S. 1995. Rural land occupancy in Scotland—resources for research. *Scottish Geographical Magazine* **111**(2):127–131.

McConnoll, W. J., and E. F. Moran. 2001. *Meeting in the middle: the challenge of meso-level integration. In* LUCC Report Series No 5, Focus 1 Office, Indiana, USA.

O'Callaghan, J. R. 1996. *Land use: the interaction of economics, ecology and hydrology.* Chapman and Hall, London, UK.

Parker, D. C., A. Hessl, and S. C. Davis. 2008. Complexity, land-use modeling, and the human dimension: fundamental challenges for mapping unknown outcome spaces. *Geoforum*: in press.

Polhill, J. G., N. M. Gotts, and A. N. R. Law. 2002. Modelling land-use change using agents in the FEARLUS project. Pages 35–40 *in* D. Parker, T. Berger, S. M. Manson, and W. J. McConnell, editors. *Agent-based models of land use and land-cover change.* LUCC Report Series No. 6, Focus Office 1, Indiana, USA.

Prato, T. 2000. Multiple attribute evaluation of landscape management. *Journal of Environmental Management* **60**(4):325–337.

Scottish Executive. 2005. Economic report on Scottish agriculture. (online) URL: http://www.scotland.gov.uk/Publications/2005/06/2290402/05450 . Scottish Executive Environment and Rural Affairs Department, Edinburgh, Scotland.

Stassopoulou, A., M. Petrou, and J. Kittler. 1998. Application of a Bayesian network in a GIS based decision making system. *International Journal of Geographical Information Sciences* **12**:23–45.

Tucker, K., S. P. Rushton, R. A. Sanderson, E. B. Martin, and J. Blaiklock. 1997. Modelling bird distributions—a combined GIS and Bayesian rule-based approach. *Landscape Ecology* **12**:77–93.

Wightman, A. 1996. *Who owns Scotland?* Canongate, Edinburgh, Scotland.