A Heuristic Approach to Modelling Thinnings

Stefan Daume and Dave Robertson

Daume, S. & Robertson, D. 2000. A heuristic approach to modelling thinnings. Silva Fennica 34(3): 237–249.

Thinnings play an important role in guiding forest development and are considered by many to be the most important influence on forests in Central Europe. Due to their importance, thinning models are a major part of any forest growth model for managed forests. Existing thinning model approaches have a number of problems associated with structure and model development that weaken their reliability and accuracy.

To overcome some of these problems this paper proposes a heuristic approach to modelling thinnings, where the focus is on distance-dependent, single-tree models. This alternative approach tries to capture the information, strategies and deductive processes likely to be employed by a forester deciding on the removal of individual trees in a stand. Use of heuristics to represent thinning knowledge simplifies the construction and refinement of a thinning model and increases its plausibility. The representation of thinning heuristics in Prolog – a programming language based on formal logic – is a straightforward process without losing expressiveness of the original heuristics. Limited tests of the model implemented in Prolog indicate that the proposed model outperforms its competitors.

Keywords thinning, heuristic model, rule-based, knowledge-based
Authors' address The University of Edinburgh, Institute for Representation and Reasoning, Division of Informatics, 80 South Bridge, Edinburgh EH1 1HN, United Kingdom
Fax +44 131 650 6513 **E-mail** stefand@dai.ed.ac.uk
Received 21 June 1999 **Accepted** 15 June 2000

1 Introduction

Thinnings are used by foresters to direct forest development. In Central Europe where only remnants of natural forests are left, thinnings can be considered as the most important influence on forest development. As such they are an indispensable part of any forest growth model that attempts to model the development of a managed forest. In addition, thinning models can also be a source of information guiding the selection of silvicultural strategies (Gadow and Hui 1998).

The foremost purpose of a thinning model is to predict the outcome of a thinning, whether to model a component of forest growth or to determine the immediate result of removing trees from a stand. Construction of a thinning model is an empirical process. The model has to serve as an explanation for available observations – in our case an actual thinning. Given satisfactory performance of a model with respect to our observations, we assume that the model is suitable for thinning predictions in similar stands.

1.1 Problems with Existing Thinning Models

Numerous examples for thinning models exist covering a variety of different approaches and levels of detail with respect to the tree information they incorporate. Existing models range from examples focusing on the whole stand and using a single attribute to more elaborate single-tree models, incorporating multiple tree attributes. Examples for the former are *dbh distribution* (dbh = diameter at breast height) based models (see for example Murray and Gadow 1991, Lemm 1991, Nagel 1996), while *rank order* or *probalistic models* are examples for models at the latter end of the spectrum (see for example Vanclay 1989, Kahn 1995, Daume et al. 1998, Albert 1999).

As different as all these approaches along the spectrum might be, they are similar in the sense that they focus on the modelling and accurate prediction of a thinning result. With respect to its granularity and level of detail the approach we are going to present fits well into the latter end of the mentioned spectrum, while it is different from all approaches in that its main focus is not on the replication of a thinning result, but a thinning process.

We claim that focusing on the replication of a thinning process rather than a thinning result influences the accuracy of a model's predictions and we argue that an approach which tries to explain how a thinning was conducted can lead to more reliable thinning models. Furthermore, many existing models seem to make implicit assumptions which weaken their reliability considerably. In the following we discuss the conceptual shortcomings of existing model approaches in more detail before presenting our heuristic thinning model. Fig. 1 reflects our view of the different components of a thinning model and the process of constructing and refining it. We will employ this structure for the discussion of problems related to existing thinning models by referring to the *input* of thinning models, the *algorithm* employed to combine these inputs in order to reach a decision for tree removal and finally the *evaluation of the performance* of the model with respect to observation of actual thinning practice. This structure is replicated in the presentation of our heuristic approach in Section 2.

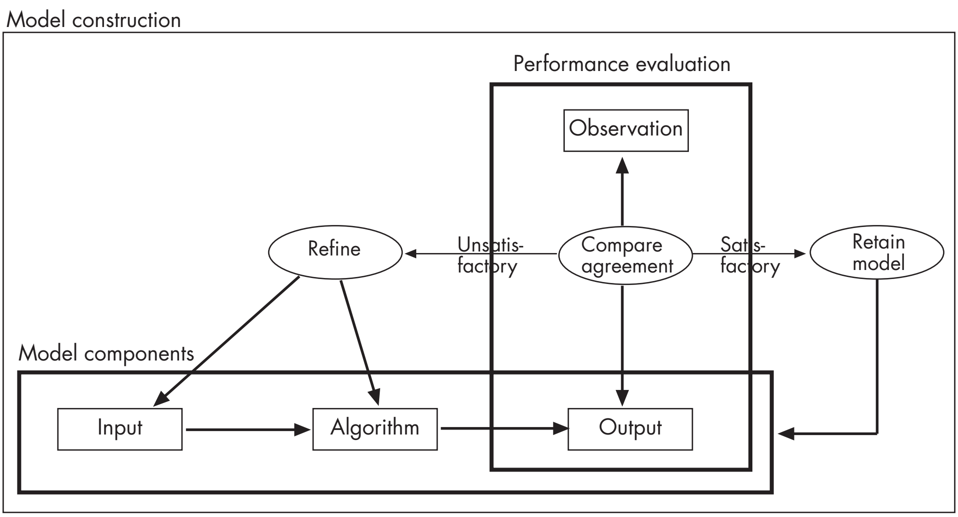

Fig. 1. Components and construction of a thinning model.

1.1.1 Input

Problems arising from the input to a thinning model have to do with the number and the type of inputs used. The most important form of input to a thinning model are the attributes describing the trees in a stand. We can distinguish between two types of attributes: *absolute single tree attributes* (i.e. species or dbh) and *relational attributes* describing the tree in relation to other trees (i.e. distance to another tree).

Those attributes can either be provided initially or inferred on the basis of the initial data. Relational attributes will in general be inferred on the basis of absolute single tree attributes. But it is also possible to infer absolute single tree attributes. As a simple example we might consider an attribute which assigns *dbh classes* as high, medium or low, but our measurements are numerical dbh values. In this case we need to infer the attribute *dbh class* from the numerical dbh measurements. In order to infer attributes we need to know how this has to be done. Procedures defining how to compute new attributes also have to be considered as input to the model.

The attributes that are used as model input are assumed to provide a sufficient explanation for the removal of a tree. Thus, we assume that the selected attributes are the most important factors influencing the thinning decision. Incorrectly chosen inputs will lead to inaccurate thinning models. Three forms of incorrect choice can be identified:

Too few attributes are used. The dbh distribution is an example for a model that uses only one attribute to explain the removal of trees. Although examples exist for which one attribute is the most important and exclusive property of a tree that qualifies it for removal this will only be true for certain types of thinnings. Given the heterogeneity and diversity of forests it is likely that more than one attribute may be required to explain removal decisions.

Wrong attributes are used. Given that the observation our model should explain is a thinning carried out by a forester, the model should use attributes that can be easily accessed, estimated or inferred by the forest expert. Competition indices of varying complexity are a typical counterexample for that (see Hegyi 1974, Biging and Dobbertin 1992, Kahn 1995 for examples). They usually require computation that is unlikely to be carried out by a forester and thus the parameter itself will never take part in the decision of the forester. Using these kinds of parameters might give an alternative approach to explaining a thinning, but will make it difficult or impossible to validate a thinning model with respect to a forester's decision.

In order to use the right attributes the type of thinning that should be modelled has to be considered as well. A thinning that targets trees that compete with trees whose growth conditions should be improved is unlikely to be explained on the basis of absolute tree attributes only, but will need certain relational tree attributes (see above). This excludes for example the dbh distribution approach as an explanation for a large portion of thinnings and could account for the poor performance of thinning models in many other cases.

Unsuitable inference of attributes. When attributes are computed or inferred on the basis of initial tree and stand information they could be wrong or at least misrelate information. This can be illustrated using the example of the relative dbh as used in a thinning model by Kahn (1995). The relative dbh is an elegant way of characterising a tree's dbh with respect to all other stand members and is defined as the relative cumulative frequency of a dbh in the stand. Using this parameter as a criterium for tree removal makes the implicit assumption that the dbh of a certain tree in relation to the rest of the stand has an influence on the thinning decision. Because of the variability of a forest it is however likely that a decision in one section of the stand might be completely independent from the rest of the stand. Thus, the value of one attribute in relation to the rest of the stand may be irrelevant.

Finally, a crucial problem concerning the model input is its availability. The majority of thinning models require the total number or proportion of removed trees to be provided as input. However this information is usually not known in advance. In fact, if a thinning model is applied to a new stand we would expect the number of trees to be removed to be part of the model's output rather than its input.

1.1.2 Algorithm

The algorithm employed in a thinning model is responsible for combining the available data and selecting trees for removal based on this combination of the provided data. The algorithm in Nagel (1996) for example is a recursive procedure that randomly selects a tree in a certain dbh class until the specified number of trees in this class has been removed. The algorithm used by Kahn (1995) is more elaborate. It consists of an assignment of thinning urgencies to each tree, then all trees are ordered according to this value and the tree with the highest thinning urgency is removed. This procedure is repeated until the specified number of trees has been removed.

A major criticism concerning many existing algorithms is that they do not seem to reflect the decision procedures likely to be employed by a forester. The strongest argument for that is the fact that the information used by many models – for example ordering the thinning urgencies of all trees in the stand (as in Kahn 1995) – is simply not available to a forester. In practice he/she is restricted to using the information of those parts of the stand in which he/she is standing. Furthermore, for all similar ordering approaches the problem of misrelated information mentioned in the previous section occurs. Comparing the thinning urgency, probability or priority of a tree with those of all other trees in a stand implies that the removal of a tree in one part of the stand influences another decision in a distant part of the stand. It is however more likely that those decisions are independent of each other and that the removal of trees that are a large distance apart have nothing to do with each other. By considering the spatial constraints of an actual thinning (Daume et al. 1998), we may view a thinning as consisting of multiple independent small-scale decisions. If we incorporate this view in our model we are more likely to come up with an algorithm that reflects a forester's decision procedures.

In summary 1) currently employed algorithmic procedures do not seem to reflect the decision procedures of a forester and 2) current thinning algorithms are based on assumptions that are not plausible with respect to a forester's decision procedure.

1.1.3 Performance Evaluation

All distance dependent single tree models are able to identify a set of trees for removal. Together with the information about the attributes of these trees the output of such a model can be used to evaluate a model's performance. We assume that a model whose output matches a given observation to a satisfactory degree will give reliable thinning predictions in similar stands. It is common practice to evaluate the performance of thinning models by measuring the agreement between its output and an observation on the basis of stand parameters. Examples for such parameters are dbh distributions, harvested timber volume, stand density measures, etc.

Judging the model's performance according to a certain parameter however makes the implicit assumption that a particular parameter captures the purpose of the whole thinning or at least captures the most important effect of the thinning. Thus, using the dbh distribution suggests that it was the purpose of the thinning to remove a collection of trees with certain diameters. In many cases it will be obvious that the purpose of a thinning was different and in many cases it will be difficult to quantitatively describe the purpose of a thinning at all. Yet another problem is that of misrelated information similar to that mentioned in Section 1.1.1. If it is assumed that a thinning consists of a number of independent small-scale decisions then describing the thinning effect as an average over the whole stand does not seem suitable.

Whatever single parameter is chosen for quantifying the agreement between model and observation it is doubtful if it captures all the important aspects of a thinning. In general more information is required and such information is difficult to obtain. Neutral and more restrictive measures are required to evaluate the performance of a thinning model and guide its development.

1.2 Aim of This Study

The discussion of problems associated with existing thinning model approaches served as an orientation in the development of our alternative heuristic model. In this study we were interested in distance-dependent, single-tree models. With respect to an empirical observation we expect our model to give an explanation as to why a particular tree has been removed during a thinning. We aim to provide an approach which increases a model's accuracy with respect to its explanation of an actual thinning and the reliability of its thinning predictions when applied to stands similar to the one it was developed for.

2 A Heuristic Thinning Model

Existing thinning models seem to neglect the fact that a thinning is a process involving a forester. Therefore, it seems more natural that the first step in developing a thinning model should be to identify the information and procedures likely to be applied by a forester carrying out a thinning. The main emphasis should not be on an explanation of the thinning result quantified by, for example, dbh distribution or similar parameters, but on how a forester makes removal decisions. Provided that a model has access to the same information as a forester and employs the same decision procedures an identical result would be achieved naturally. Our modelling task has therefore changed from explaining a thinning output to explaining the reasoning of a forest expert. We will describe the implications of this modelling approach in the following sections.

2.1 Input

Our alternative modelling restricts attributes to those accessible to the forester, such that they could be easily measured, estimated or calculated. This includes attributes like tree position, dbh, species or distance between trees. Furthermore, the model includes attributes like vitality or quality of a tree that might be of importance for removal. We exclude competition indices or parameters like relative dbh. The former requires complex calculations that are unlikely to be performed by a forester and the latter contradicts our view that a thinning consists of multiple, small-scale decisions for which the average properties of the whole stand are irrelevant.

We already mentioned that one reason for inaccuracy of thinning models is static use of available data. To overcome this problem we introduce another form of input: heuristic rules. These heuristics describe how and which data might be used by a forester to make a thinning decision. This knowledge of how to combine data is essential for a thinning decision and is actually our main focus in explaining how a forester makes his/her decisions. In the thinning models mentioned above, representation of this knowledge is implicit, because it is hidden in the algorithm. The sorting algorithm of Kahn (1995), for example, tells us something about the importance of certain attribute combinations. The problem is that more than one way of combining data might exist and that it is more instructive to represent this knowledge explicitly. Heuristic rules also enable easier checking of model plausibility. Rather than committing ourselves to a certain way of solving the thinning problem, we will describe the kind of problems or concepts that might be encountered in a thinning (in the form of heuristic rules) and leave the choice of which heuristics apply in a certain situation to an appropriate search algorithm, which will be described in Section 2.2.

It is in the nature of heuristics that they lead to valid conclusions in many cases but not in all. Heuristics are 'rules of thumb' that allow us to make decisions in situations where either our knowledge or the available data is incomplete (Giarratano and Riley 1998). This is particularly useful in the case of thinning decisions. A forester will never have a complete rule set covering every possible situation that might be encountered in a forest. Two groups of trees can be judged similar, but are never the same. Heuristics can help focus on those properties of trees that make them similar to previously encountered situations and that are most important for a thinning decision.

In the following, we present some rules that can be considered as likely approximations of the knowledge employed by a forester. We describe three types of heuristics that are required for a complete thinning model reflecting steps in the decision-making process of a forester. In our implemented system these are written in the formal notation of the logic programming language Prolog, but to make them easier to read, we present them here in English, following a similar structure to our Prolog definitions. We say more about Prolog in Section 2.2.

As mentioned earlier we view a thinning as a number of small-scale decisions. At any point in the stand the forester's decisions are constrained by the information close at hand. Thus, the first step in the decision-making process and also the first type of rule required is the identification of a stand section to focus on. We will illustrate this type of heuristic and all following with an example of a so-called selective thinning, where trees are removed to support pre-selected elite trees in a stand. The following heuristic represents a simple example for the first type of rules which we will call focus rules. It would be applied repeatedly throughout a thinning until the whole stand has been thinned.

> *"Pick the next elite tree and make it the centre of your focus. Include all neighbours of the selected elite tree in your focus that are within distance of three times the crown radius of the elite tree."*

This heuristic is not only easy to understand and apply but is also plausible. As the overall aim of this particular thinning is the support of elite trees, it makes sense to make an elite tree the centre of your focus and consider its surrounding area. Distance to the elite tree is a plausible way of distinguishing between trees that could have an effect on the elite tree's growth and those that have none at all. Finally, it seems plausible that the radius to be considered will vary with the space the elite tree occupies itself.

The second step in our decision process identifies potential candidates for removal. This requires rules that tell us whether a tree in the current focus could be removed or not. The following rule is an example for this second type of rule which we will call select rules.

> *"If a tree is a competitor of the elite tree under focus and there are no reasons for objecting to its removal than this tree is a candidate for removal."*

This obviously requires further rules that define a competitor or state reasons not to remove a tree. A competitor that is itself an elite tree or belongs to a scarce species might for example be excluded from removal. A simple competitor rule might look like this:

> *"If a tree overlaps 1/4 of the crown radius of an elite tree it is considered to be a competitor of that elite tree."*

In a more general form, this simple rule states that a tree is a competitor of an elite tree if its crown overlaps the crown of the elite tree to a certain extent. Thus, definition of this rule requires quantification of the extent to which crowns overlap. More rules will be required to specify these conditions and complete the set of select rules.

In the final step we have to commit to certain trees for removal. Given a complete set of focus and select rules we are able to identify potential candidates for removal. Committing to one of these candidates or a combination of them requires another type of heuristic that we will call goal check rules. These heuristics will guide us in how to best achieve the effect we have in mind for our thinning. Again the following heuristic serves as a simple example for that type of rule.

> *"Among the identified candidates for removal remove the tree that has the largest crown overlap with the elite tree under focus."*

This heuristic restricts itself to one choice only. It could however be easily extended. For example, by considering the number of trees competing with an elite tree. The more candidates for removal that can be identified the more that will be removed. We could also imagine using this extension on its own, which would specify a rule that leaves room for alternative choices.

A forester might not agree that, for example, a certain attribute should be considered in the way it is done in our rules (for example the extent of crown overlap). But although these rules might be disputable they are plausible and explicit. A forest expert should be able to understand how a certain decision was reached using these rules and how it could be repeated. Furthermore, the forester would be able to explain which rule or part of a rule is disagreeable and perhaps how heuristics should be changed or extended to reflect his/her own view of the problem. This is not the case for the approaches mentioned in the introduction where it might be possible to state disagreement with the result of a decision, but not to point to the parts of the model that would require change in order to reach an alternative decision.

The explicit representation of thinning knowledge by means of these heuristics and the possibility to identify parts that would require adaptation to suit the decision of a certain forester points to a useful distinction we can make for these rules: generic and subjective rules and rule components. Generic rules are those most foresters will agree with, e.g. that the crowns of two trees must overlap to make them competitors. Subjective components vary according to the individual judgement of a forester (e.g. the extent to which crowns must overlap to make them competitors). Separating those generic and subjective parts gives us a generic and reusable set of heuristics on one side and knowledge about how and where to adapt rules to explain individual thinnings on the other.

2.2 Algorithm

As mentioned earlier the main task of our model has changed from explaining a thinning result to explaining the reasoning of a forester in order to achieve a certain result. In contrast to the algorithms that characterise the approaches briefly mentioned in Section 1.1 our algorithm reflects this deductive process*. On the basis of the described rules and the available stand data we draw conclusions concerning the removal of certain trees. In order to illustrate that we will take the competitor heuristic from the previous section and represent it in a more formal way:

If tree **X** is an elite tree
and tree **Y**'s crown overlaps tree **X**'s crown
and the crown overlap between tree **Y** and tree **X** is greater than 1/4
then tree **Y** is a competitor of tree **X**.

* The algorithm explained in this section is basically a description of a procedure involving depth-first search and unification (Luger and Stubblefield 1997 for details) as for example employed by the Prolog programming language.

This representation highlights the conditional parts and the conclusion of the rule and makes use of variables. For illustration purposes we look at this rule in isolation. It is assumed that our focus rule has been applied and produced the trees listed below as those currently in focus. Thus, having identified tree 1 as elite tree we are now trying to find out which of its neighbours 2, 3 and 4 are its competitors. The following are the facts we assume to be available in this example:

Tree **1** is an elite tree.

Tree **2**'s crown overlaps tree **1**'s crown.
Tree **3**'s crown overlaps tree **1**'s crown.
Tree **4**'s crown overlaps tree **1**'s crown.

Crown overlap between tree **2** and tree **1** is 1/5.
Crown overlap between tree **3** and tree **1** is 1/3.
Crown overlap between tree **4** and tree **1** is 1/2.

There are mainly two types of questions we want our algorithm to solve. Given a thinning model we want to apply it to new stands returning the trees that should be removed according to our model. The equivalent question we would ask in our scenario would therefore be: *Which tree Y is a competitor of tree 1*? The second type of question will be important during the development and refinement of the thinning model. We want to make sure that our model explains a given observation. Assuming that we know that in the actual thinning tree 4 was considered as a competitor of the elite tree 1, we would expect our model to come up with a positive result when queried: *Is tree 4 a competitor of tree 1*?

In order to solve the first query our algorithm has to search for a rule that matches our query and will find the rule we provided above. The next step in solving the question is to prove the first condition respectively to find a fact that matches this condition – that tree 1 is an elite tree. Our algorithm will succeed as it will come across the appropriate fact in our database. Variable X in our rule is now bound to 1.

Having succeeded in proving the first condition the next step for our algorithm is to find a fact which matches the second condition of our rule. The second fact in our database can be matched against the second condition resulting in the variable Y to be bound to 2.

At this point all variables are instantiated and the last condition to be proved reads now: *'the crown overlap between tree 2 and tree 1 is greater than 1/4'*. Although our algorithm will be able to find three matching facts of this form, none of the facts in our database will satisfy the condition. In trying to establish a successful proof the algorithm will now backtrack to the point where an alternative choice would have been possible, thus another matching fact could have been used. In our simple example this means that instead of matching the second fact against the second condition, the third fact is now used resulting in Y to be bound to 3 instead of 2. Proceeding in the same manner as before the algorithm will successfully conclude the proof and eventually return that "*Tree 3 is a competitor of tree 1*" as the result of our query. If forced to search for further solutions the algorithm will backtrack again on its previous search and eventually return "*Tree 4 is a competitor of tree 1*" as the second possible answer to our query.

The second type of question "*Is tree 4 a competitor of tree 1*?" expected to be used during model construction and refinement is solved by the same algorithmic search process of pattern-matching and backtracking.

It is important to note that the process of pattern-matching and backtracking works over both facts and rules. Thus, if the first matching rule our algorithm comes across does not result in a successful proof it will use another rule that matches. Having model construction and refinement in mind the algorithm will therefore also be able to identify the appropriate rules that allow to explain an observation. In case that our database would not contain an appropriate rule to explain a given observation our deductive approach allows at least to trace those conditions that could not be satisfied and would therefore need to be changed.

The presented examples illustrate that the algorithm is reduced to a search procedure that scans the database for matching facts and rules that allow to solve a given query. The algorithm used does not rely on any thinning specific information and can be applied to any set of heuristics and facts presented as an input, a typical characteristic of knowledge based systems and one of their main advantages.

The described heuristics and algorithm were implemented in this project using Prolog – a programming language based on formal logic. The formalisation of the heuristics proposed in Section 2.1 in Prolog is a straightforward process without loosing significantly the expressiveness of the initial heuristics and the algorithm described in this section is built into Prolog. For a more detailed description of the actual implementation, the complete set of heuristics used and the Prolog programming language we refer to Daume (1998) and Sterling and Shapiro (1994) respectively.

2.3 Performance Evaluation

The problems related to the evaluation of thinning models were discussed in section 1.1.3. It was mentioned that usually stand or thinning parameters measuring the thinning's effect on the stand are used to quantify the agreement between the model and an actual thinning and evaluate the performance of the model. In order to avoid the problems related to this approach we will use the strongest possible measure: a tree wise comparison of the models and the foresters decisions as for example used in Daume (1995). We simply assume that the higher the tree wise agreement between two thinnings (the more trees the forester and the model agree upon) the more reliable our model's predictions will be when applied to similar stands. If our model selects exactly the same trees for removal as a forester in an actual thinning, its accuracy is perfect and its reliability should be high. We also expect our model to select the same number of trees as the forester in an actual thinning. When discussing the accuracy of our model we will refer to these two measures: the proportion of trees selected in the actual thinning that were also selected by the model and the total number of trees selected for removal according to the model.

Besides the set of trees selected for removal by the model we can name another form of output for our model: the explanation for the removal of a certain tree. Tracing the heuristics and data used to conclude on the removal of a tree this information could be returned together with the

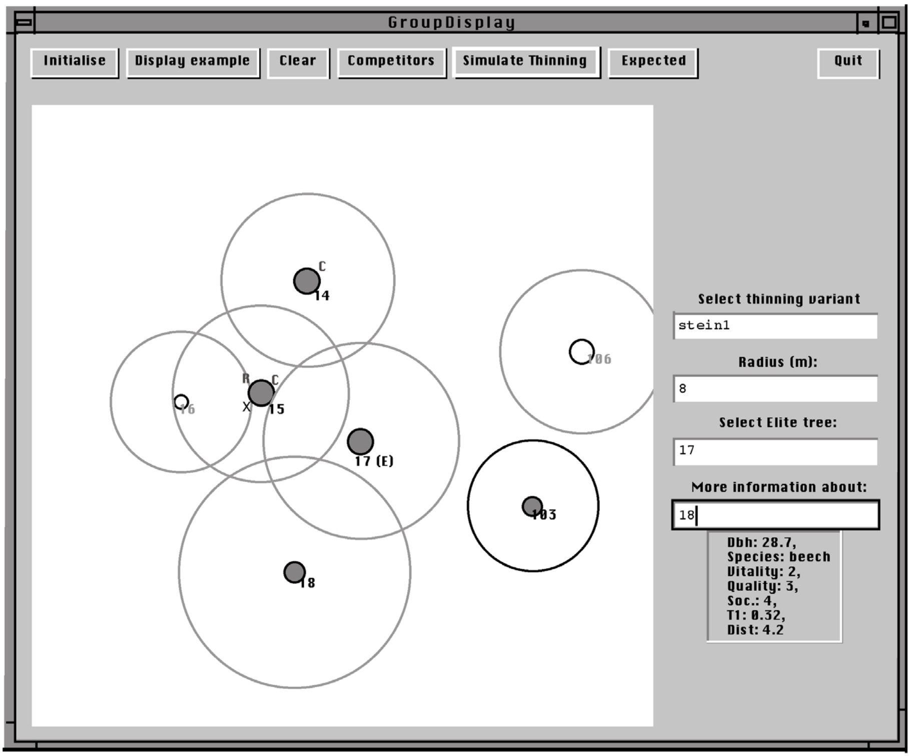

Fig. 2. A simple interface displaying tree groups and thinning results.

tree to be removed. A simplified way of visualising certain steps of the reasoning process is presented in Section 3.

3 Visualisation of the Decision Process

For the refinement of the model as well as the understanding of the decision-making process it is helpful to visualise the situation the model is dealing with and the results of thinning decisions both by the model and a forester. It is possible to provide an extensive explanation of the models decision by tracing all facts and rules employed to reach it. We will however use a much simpler approach following the main steps of the process outlined in Section 2.1: focusing on a group and an elite tree, identifying competitors and selecting trees for removal. For this purpose a graphical user interface (GUI) was developed that can display the above information as well as the outcome of the thinning that our model should capture. The GUI was interfaced with the rule-based thinning model implemented in Prolog. Fig. 2 shows an example session with the interface*.

The display shows an idealised map of a tree group. The different grey shades indicate certain species, the filled circles represent the trees stems**, and the concentric circles around them their crowns***. All trees carry a numeric label and the E attached to tree 17 identifies this tree as an elite tree. The displayed data is part of the basic information that has to be available in order to reason about the removal of trees.

* The graphical user interface was implemented in Tcl/Tk which can easily be interfaced with a Prolog program.

** More precisely it represents the trees' cross-sectional area at breast height (expanded by a certain factor for illustration purposes).

*** As crown measurements for this example stand were not available, the crown radius had to be calculated on the basis of the trees' dbh using a regression model by Nagel (1996).

As mentioned above the interface can support the development and refinement of a thinning model. In order to do so it must first of all display the result we want to achieve, in our case the trees removed in an actual thinning. In our example the tree labelled 15 is marked with an X which indicates that it was removed in the actual thinning that should be modelled. Even with limited knowledge about the nature of this thinning the displayed map might already hint that the crown overlap between the elite tree and its neighbours could have an influence on the thinning decision and that a corresponding rule should be part of our heuristics.

The thinning heuristics are applied in response to queries given by the user to the thinning model implemented in Prolog. The button *Competitors* invokes a query to the Prolog rule-base returning all trees that are considered as competitors of the elite tree according to the model marking them with a C in the display. *Simulate Thinning* forces the model to decide on tree(s) to be removed which are marked with an R in Fig. 2.

Displaying this information visualises the main steps in the model's reasoning process. In the case illustrated above the model's decision matches the observed decision of the actual thinning suggesting that no further refinement of the thinning model is necessary in order to explain this case. If our model had not reached a matching decision the displayed information could have been used to guide the refinement of the model. Assuming for example that, besides trees 14 and 15, tree 18 would also have been considered as a competitor and would have been removed by our model instead of tree 15 we might have concluded that more restrictive competitor rules are required excluding tree 18 as a competitor. Using this interface during model development, with the forester responsible for the thinning to be modelled present, can speed up model construction considerably.

4 Testing the Rule-based Model

Our model was put to test using the data of a sample thinning in a 110 year old mixed beech-spruce forest with 742 trees covering an area of about 1.15 hectares. The stand information available included the position, species, dbh and qualitative attributes vitality, quality and social class of the trees. Furthermore, 93 trees were preselected as elite trees. Vitality and quality of a tree were each estimated according to a system applied locally. A number of tree characteristics (like damages, crown size, etc.) are condensed into three classes which allow to characterize the quality and vitality as low, medium and high. The social class has five possible values ranging from very low to very high and describes the social standing of a tree compared to other stand members.

The purpose of the thinning carried out in this stand was the removal of trees in order to improve growth conditions for the elite trees. The thinning was carried out by a local forester and resulted in the removal of 90 trees. Some of the rules described previously were defined after consultation with the forester others according to the authors own knowledge about thinnings. Throughout the process of model construction and refinement the subjective parts of these rules were gradually adapted in order to improve the agreement between the models predictions and the actual decisions of the forester. This process finally resulted in a refined model that suggested the removal of 92 trees in our example stand including 52 % of those trees originally selected by the forester.

The purpose of our model is of course to predict the outcome of a thinning in stands similar to the one on which it was modelled. We therefore tested our completed model and applied it unchanged to another mixed beech-spruce forest. This new stand consisted of 930 trees with an average age of 80 years covering an area of about 1 hectares. For this stand 94 elite trees had been preselected. The stand had been thinned by the same forester as the previous stand resulting in the removal of 103 trees. When applying our model to this new stand it suggested the removal of 100 trees, 54 % of which were also selected by the forester.

5 Discussion

A major problem we face discussing the performance of our model compared with alternative approaches is that they usually apply other measures to quantify the agreement between a model's predictions and an observation. We described in Section 1.1.3 that we view this as a major weakness of these models. Therefore, we restricted ourselves to a tree wise comparison of the actual thinning outcome and the model's predictions to evaluate the accuracy of our model. We claimed that this is the strongest possible and the most neutral measure available. Setting the quantitative comparison with other models aside we can however ask if a model that explains 52 % of an observation is satisfactory.

In general a model is always only an approximation of the real process that it should explain and our model is no exception. Although we work at the level of individual trees we are obliged to be selective and parsimonious in our choice of the attributes to include in our model. This begins with the available data. Using attributes like the vitality or quality of a tree results in a much more detailed description of a tree that many other models are lacking. These attributes do however represent a summary of details that are available to the forester carrying out a thinning, but subsequently not to the model. A forester will be able to spot the nesting place of a rare bird in the crown of a tree and therefore refrain from selecting a tree that otherwise would have been removed. Our model however has no access to this information and may decide to remove the tree. This is just one example of information that might influence a foresters decision, but usually is not available to the model because the stand data that can be sampled is constrained by time and costs.

The same point has to be made for the heuristics described, since they are also only approximations of the knowledge employed by a forester. They are approximate in the sense that we cannot really verify that our heuristics are the correct representation of a forester's knowledge. We can at best argue that they are the most likely representation. They are also approximate in the sense that we can name heuristics that should be included but are missing. As an example we can again refer to the previously mentioned example of the bird's nest. A suitable heuristic could be included in our rule set, but is omitted due to the lack of appropriate information.

But how could the accuracy of our model be improved and how big an improvement could we expect? The suggested heuristics were quite simple. For example we made no distinction between different species, but it is likely that a forester might make a distinction when encountering two competing beeches or a competing beech and spruce and subsequently apply different heuristics for these cases. Providing appropriate heuristics offers in our opinion the best chance for increasing the models accuracy, but has the drawback of making the model and its refinement more complicated. In a stand with two species we would only have to cover for three combinations of species. However, for stands with more species the trade off between the models complexity and the gained accuracy is likely to become disadvantegous. Furthermore, our heuristics would have to cover species that did not appear in the stand but might occur in a new stand the model should be applied to.

In addition the models accuracy concerning a certain observation is only part of its performance. The second part is its reliability when applied to new situations. We assume that high accuracy will give high reliability, but this is not necessarily the case. Overfitting the model with respect to the example thinning can in fact decrease its reliability concerning predictions for thinnings in similar stands. Testing the reliability of our model was the purpose of testing it in a stand similar to the one it was adapted to which was also thinned by the same forester (see Section 4). The results of this test were very encouraging as the agreement between model and forester almost equalled those for the example for which our model was built.

Given the heuristic nature of our model it is unlikely that we will ever get a perfect agreement between model and observation. Tests show that for a repetition of the same thinning in the same stand by the same forester on average only 56 % of the trees selected in one thinning are also selected in another (Kahle 1995). Assuming that the forester is consistent in the heuristics he/she applies these results show that an incomplete agreement between two thinnings is not necessarily proof for poor performance but simply suggests that alternative ways of doing the same thinning exist. Discussing the accuracy of our model in the light of these results an agreement on 52 % of the selected trees seems to support our view that this is indeed an encouraging result. Trees selected by our model but not selected by our forester might in fact be considered as valid alternatives by the forester. This however remains speculation as we could not consult the forester responsible for the thinning for our tests.

The results of Kahle (1995) also suggest that a thinning is not the precise science assumed by many thinning models. It seems to be in the nature of a thinning that its result is not completely deterministic. This is reflected in our modelling system which represents heuristics and which is non-deterministic in the sense that it allows for alternative solutions to be found on backtracking.

6 Conclusions

Although an extensive quantitative comparison with other thinning models does not seem appropriate, due to different approaches in the evaluation of model performance, the test results are encouraging. Restricting the model to data that is likely to be used by a forester during a thinning produces satisfactory results. The thinning heuristics, though simple, seem to capture the most important aspects of an individual thinning as its performance for the second stand suggests. In our discussion of the model's performance we made a distinction between a model's accuracy and its reliability. The reliability of a thinning model with respect to thinning predictions in stands similar to the one it was built for has in fact an enormous influence on how useful a model can be in forest practice (i.e. forest planning).

One practical benefit of the heuristic approach concerns the communication of thinning heuristics between forest experts. We tried to provide an explicit representation and formalisation of rules that a forester might apply in order to conclude on the removal of individual trees in a forest. The presented heuristics were the result of this effort. Although we focused on the application of rules in small parts of a stand they can be applied to a forest as a whole and provide not only an explanation for an actual thinning but, as we have seen, also a fairly reliable prediction of a thinning in a new though similar stand. Usage of our heuristics however does not have to stop there. These rules could for example be used to pass on thinning knowledge from experienced foresters to novices. In a training scenario a novice forester would be presented with tree groups and asked to apply given heuristics under the guidance of a senior forester. The current stage of our work (now at prototype stage) aims to provide an automated solution to this training scenario where a forester will be equipped with a mobile computing device with a heuristic thinning model built into it. The device allows the forester to enter information about tree groups he wants thinning guidance on. In turn the thinning model suggests the removal of certain trees and on request will explain why. This is possible because the thinning heuristics tried to capture the reasoning of a forester in the first place and one of the outputs our models can provide is a trace of this reasoning chain. After the forester has solved a number of training examples under the guidance of and in dialog with the mobile thinning model he/she should have learned the thinning heuristics and complete the thinning of a stand on his/her own.

Finally, in Section 1 it was mentioned that thinning models provide a valuable source of information with respect to the selection of silvicultural strategies. In practice, different thinning models might be applied to data of a particular stand to be thinned in order to find the most appropriate thinning strategy. If the contents of the model that represents the most suitable strategy are not explicit enough it will be impossible or at least difficult to communicate the characteristics of a thinning to a forester such that he/she can replicate it as intended. Thus, a gap would remain between the planning and the realisation of a thinning leaving the reliability of the planned action and the value of the planning effort itself in doubt.The training scenario explained above might help to narrow this gap.

Acknowledgements

This work is supported by the European Commission with a Marie Curie Research Training Grant (FAIR-BM-973975). We would also like to thank Klaus von Gadow and Matthias Albert for their helpful comments.

References

Albert, M. 1999. Analyse der eingriffsbedingten Strukturveränderung und Durchforstungsmodellierung in Mischbeständen. Diss. Forstliche Fakultät Göttingen, Universität Göttingen.

Biging, G. & Dobbertin, M. 1992. A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. Forest Science 38(3): 695–720.

Daume, S. 1995. Durchforstungssimulation in einem Buchen-Edellaubholz Mischbestand. Diplomarbeit, Forstliche Fakultät, Göttingen.

— 1998. A knowledge-based system to model thinnings in Central European forests. M.Sc. thesis, Dept. of Artificial Intelligence, University of Edinburgh.

— , Füldner, K. & Gadow, K. von. 1998. Zur Modellierung personenspezifischer Durchforstungen in ungleichaltrigen Mischbeständen. Allgemeine Forst- und Jagdzeitung 169: 21–26.

Gadow, K. von. & Hui, G. 1998. Modelling thinnings in even-aged plantation forests. Plant Research and Development 47/48: 98–107.

Giarratano, J.C. & Riley, G.D. 1998. Expert systems. PWS.

Hegyi, J. 1974. A simulation model for managing jack-pine stands. In: Fries, J. (ed.), Growth models for tree and stand simulation. Royal Coll. of For., Stockholm. p. 74–76.

Kahle, M. 1995. Die Analyse von Bestandeseingriffen in einem Buchen-Edellaubholz Mischbestand. Diplomarbeit, Forstliche Fakultät Göttingen.

Kahn, M. 1995. Die Fuzzy Logik basierte Modellierung von Durchforstungseingriffen. Allgemeine Forst- und Jagdzeitung 166(9–10): 169–176.

Lemm, G. 1991. Ein dynamisches Forstbetriebssimulationsmodell. Diss. Professur Forsteinrichtung und Waldwachstum, ETH Zürich.

Luger, W. & Stubblefield, A. 1997. Artificial intelligence. Addison Wesley.

Murray, D.M. & Gadow, K. von. 1991. Relationship between the diameter distributions before and after thinning. Forest Science 37(2)): 552–558.

Nagel, J. 1996. Anwendungsprogramm zur Bestandesbewertung und zur Prognose der Bestandesentwicklung. Forst und Holz 51: 76–78.

Sterling, L. & Shapiro, E. 1994. The art of Prolog. MIT Press.

Vanclay, J.K. 1989. Modelling selection harvesting in tropical rain forests. Journal of Tropical Forest Science 1(3): 280–294.

Total of 16 references