Sensitivity of Harvest Decisions to Errors in Stand Characteristics

Annika Kangas, Lauri Mehtätalo, Antti Mäkinen and Kalle Vanhatalo

In forest planning, the decision maker chooses for each stand a treatment schedule for a predefined planning period. The choice is based either on optimization calculations or on silvicultural guidelines. Schedules for individual stands are obtained using a growth simulator, where measured stand characteristics such as the basal area, mean diameter, site class and mean height are used as input variables. These characteristics include errors, however, which may lead to incorrect decisions. In this study, the aim is to study the sensitivity of harvest decisions to errors in a dataset of 157 stands. Correct schedules according to silvicultural guidelines were first determined using error-free data. Different amounts of errors were then generated to the stand-specific characteristics, and the treatment schedule was selected again using the erroneous data. The decision was defined as correct, if the type of harvest in these two schedules were similar, and if the timings deviated at maximum ±2 for thinning and ±3 years for clear-cut. The dependency of probability of correct decisions on stand characteristics and the degree of errors was then modelled. The proposed model can be used to determine the required level of measurement accuracy for each characteristics in different kinds of stands, with a given accuracy requirement for the timing of treatments. This information can further be utilized in selecting the most appropriate inventory method.

1   Introduction

In forest planning, the decision maker chooses for each stand a treatment schedule for a predefined planning period. Typically the choice is based on maximizing the net present value (NPV) of the stand. The optimal schedule is selected from alternative harvest schedules simulated for each stand using a growth simulator. Another option is to simulate the treatment schedule according to silvicultural guidelines. These guidelines are often based on optimization calculations for "typical" forest stands, but are not necessarily optimal for a given stand.

Observed values for stand characteristics such as the basal area, mean diameter, site type and mean height are used as input variables in simulation. While generally better decisions can be obtained with more accurate data, accuracy alone does not determine the usefulness of the data in decision making (e.g. Kangas 2010). The usefulness depends on whether the errors result in inoptimal harvest decisions. When the adopted schedule differs from the optimal, it incurs inoptimality losses. The loss is defined as the difference between the outcome (typically NPV) of the optimal schedule and the selected schedule.

In cost-plus-loss analysis, the expected inoptimality losses are added to the total costs of the forest inventory (Hamilton 1978, Burkhart et al. 1978, Ståhl 1994, Ståhl et al. 1994). The decision maker is assumed to maximize the net present value (NPV) of the forest area from decisions concerning the timing and intensity of different silvicultural treatments in each stand, and the losses are also defined in terms of NPV (e.g. Eid 2000, Holmström et al. 2003, Eid et al. 2004, Duvemo et al. 2007, Borders et al. 2008, Islam et al. 2009). The losses, in turn, can be interpreted as the expected value of perfect information, EVPI (Lawrence 1999, Kangas 2010). According to Eid (2000), the losses were highest in stands close to their commercial maturity, but in very old stands the losses were negligible as the optimal decision was always to do a clear-cut immediately. Similar observations were made by Holmström et al. (2003). Consequently, the average losses depend on the age distribution of the area in consideration (Holmström et al. 2003).

In practical forest management, the selection of treatments is often based on silvicultural guidelines. In this case, inoptimality losses cannot be calculated, but the correctness of the scheduling decision can still be defined. A question of interest is, how large errors can be tolerated, i.e. which is the magnitude of error that leads to a different decision than error-free data. This depends on whether the decision to be made is about final harvest or thinning, as the stand characteristics that trigger these decisions are different. It is obvious that also the direction of the errors matters: in some cases overestimates may be more detrimental than underestimates and vice versa. Obviously, optimal or correct decision in each and every case requires perfect information. Thus, assessing the accuracy with which the probability of making an optimal or correct decision is within a given level, say 95%, is a more useful approach.

In this study, the aim is to study the sensitivity of the harvest decisions to errors in a dataset of 157 stands. First, we defined the correct treatment schedule for each stand for a 10-year period based on the silvicultural recommendations of UPM-Kymmene using error-free data. Next, we generated errors to the input variables and defined the treatment schedule based on this erroneous data. The probability of obtaining a correct schedule decision within different error level, stand characteristics and treatment classes was calculated. The main aim is to define the maximum error in the input variables that produces an acceptable probability of correct decisions in a given stand. Ultimately, this information can be used to select an appropriate inventory method for a given forest area.

The probability of correct decisions was assumed to depend not only on the error of the input variables but also on the stand characteristics. Thus, the dependence of the probability on stand characteristics and the magnitude of errors was then studied graphically and modelled using logistic regression and classification and regression trees (CART). Two approaches were tested, as these approaches have different uses.

2 Material and Methods

2.1 The Problem Formulation

First, we defined the correct treatment schedule – the correct decision – for each of the stands based on the silvicultural recommendations using error-free data. Then we generated errors to the data, and defined the treatment schedule for each stand with each error combination. The next step was to define if the treatment schedule decisions obtained with the erroneous data were correct. The decision was defined as correct, if the type of harvest in these two schedules were similar (first thinning, later thinning, clear-cut), and if the timings deviated at maximum ±2 years for thinning and ±3 years for clear-cut. Finally, we analysed the probability of obtaining a correct decision within different error, stand characteristics and treatment classes, and modelled the probability as a function of them.

The forest planning computations, ie. growth simulation and harvest scheduling, in this study were done using SIMO (SIMulation and Optimization) software (Rasinmäki et al. 2009). The guidelines are modified for UPM-Kymmene from the Recommendations of Tapio (2006). A stand is defined as mature (i.e. ready for clear cut) when either the diameter or age is above a specified limit value depending on the site type. The criteria for the need of first thinning are site type, dominant height and number of stems. In later thinnings, the need for thinning is defined using site type, dominant height and basal area, according to a thinning model (Fig. 1). In thinnings, the limiting number of stems or basal area thus depends on both dominant height and site type. The main difference between the two thinnings in practise is that the first thinning is carried out for silvicultural reasons, but the later thinnings also for incomes.

The whole simulation period used was 30 years, but the correctness of the decisions was analysed only for the first ten years. The rest of the simulation period was used to define the difference in timing, if the next treatment in either the correct or erroneous schedule occurred after the first 10 years. In the first 10 years, a one-year step was utilised in the simulation, and in later years the step was longer.
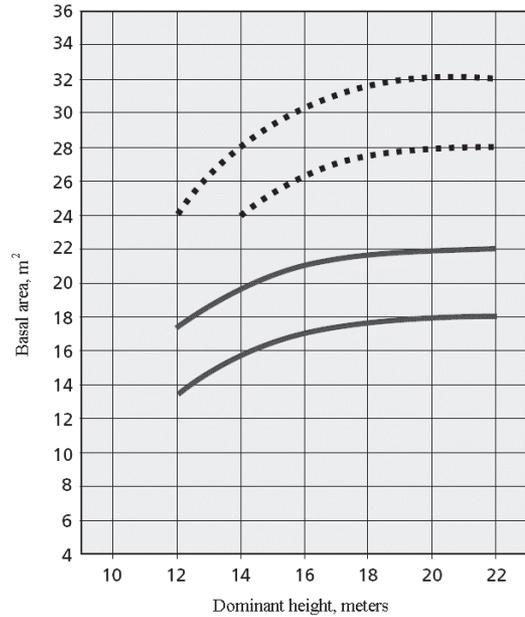
Fig. 1. The Finnish thinning model for medium fertility (MT) Scots pine and Norway spruce stands in Southern Finland. (Recommendations of Tapio, 2006). A thinning is scheduled when the basal area (y-axis) at given dominant height (x-axis) is above the thinning recommendation (the lower dotted line), and urgently recommended when basal area is above the upper dotted line. In thinning, the basal area is suggested to be reduced to between the two solid lines of the model.

The simulator includes growth models for all Finland's main tree species and forest types (Hynynen et al. 2002). The simulation was carried out with a tree-level growth simulator. The dependent variables in these models were growth in height and basal area, while the independent variables included a substantial number of variables describing the characteristics of a single tree, the stand and the geographical area. However, the only input variables in the system were basal area G, mean diameter D, mean height H and site type ST, all other variables were calculated based on these. For instance, as the tree-level models operate on individual trees, diameter and height of all trees needs to be available. In this study, the diameters of the trees were predicted using a diameter distribution model and the heights with

a H/D curve included in the system. In addition, the simulator includes models for different forest treatments.

2.2 Material

The data in this study is from the stand register data of UPM-Kymmene. 157 spruce stands from three different development classes (young, middle-aged and mature) were selected with stratified sampling based on basal area. These data are assumed error-free in this study. The data included basal area (G), basal area median diameter (D, hereafter called mean diameter), height of basal area median tree (H, mean height) and site type (ST), commonly measured in the Finnish compartmentwise inventory system.

For each of the selected stands, errors from –30% to +30% with 1% steps (i.e. 61 different error levels) were simulated for either G, D or H, assuming the other variables were assumed error-free. Optimally, the dataset used would include also all possible combinations of these errors. This would, however, mean 61·61·61 error combinations for each stand. Therefore, for each two-variable combination, we used the errors generated above for the first variable, and added random variation for the second and third variables. For continuous variables a normal distribution was used with either 10%, 20% or 30% standard deviation (61·3 combinations). In site type (ST), the simulated errors were ±2 classes (the number of site type classes is 6, and most of the stands belong to two most common classes, 3 and 4) (61 combinations). The error combinations considered were

  G,
  D,
  H,
  G+H,
  D+H,
  D+ST,
  G+ST,
  D+H+ST,
  G+H+ST,
  D+G+H,

resulting altogether 182 438 realizations. It means that we simulated 1032 different error levels for

Table 1 The main characteristics of the data.

	G, m^2/ha	D, cm	Site type
Min	5	6	1
Median	22	19	3
Max	39	36	6
SD	5.15	5.69	

each stand. The error distribution was not based on any pre-defined forest inventory method, nor did it follow any pre-defined statistical distribution. The purpose is to get observations from all possible combinations or errors and stand characteristics, in order to be able to define the maximum error that still produces an acceptable accuracy of the decisions with given stand characteristics.

No errors were generated to number of stems nor to dominant height, which are also important in defining the correct treatment, as these were calculated within the simulator used. However, the calculation of dominant height depends on both mean diameter and mean height, meaning that error in either has an effect on the estimate of dominant height. Likewise, the number of stems estimate is calculated based on basal area and mean diameter, which means that error in either causes errors in number of stems estimate (see e.g. Mäkinen et al. 2010, Holopainen et al. 2010).

2.3 The Probability of Making a Correct Decision

The dataset was first classified according to the error level in each input variable (G, D and H). The errors in G and D were classified into intervals [–30%,–20%[, [–20%,–10%[, [–10%,0%[, [0%,10%[, [10%,20%[, [20%,30%]. The errors in H were also classified into intervals [–30%,–5%[, [–5%,5%[, [5%,30%[. The probability of making a correct decision was calculated in each class. However, as the probability of correct decision depends on the stand characteristics besides the error level, we decided to model the probability. The probability of correct decision was modelled

with i) logistic regression using function glm in the R statistical package, and ii) with a classification and regression tree (CART) using function rpart (R development core team 2009).

A binomial model with logit link was fitted (with ML) to the data. Let random variable OK_i determine whether schedule i was ok ($OK_i = 1$) or not ($OK_i = 0$). Assume that the binary indicator variable follows the Bernoulli distribution with probability of success π:

$$OK_i \sim Bernoulli(\pi) \tag{1}$$

Furthermore, assume that the probability of success can be written using the logit link as

$$\ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p \tag{2}$$

The model can later be used to calculate the error level that will give the given accuracy with given stand characteristics, i.e. solving the error from

$$1 - \alpha = \hat{\pi} = \frac{\exp\left(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_p x_p\right)}{1 + \exp\left(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_p x_p\right)} \tag{3}$$

where α is the tolerated probability for incorrect decision. The goodness-of-fit is analyzed with the ratio of residual deviance to the null deviance. In this analysis, the observations were assumed to be independent, even though there are very many observations from the same stands. The reasons for this decision are given in discussion.

For modelling purposes, the errors in stand characteristics were divided into negative and positive errors (i.e., negative error is 0 when error is positive and vice versa). To take into account the nonlinear relationship between logit of the response and the predictors, the positive and negative errors entered into the model in a polynomial form with up to 3 degrees. Interactions of stand characteristics and terms of the polynomials were included to allow different relationship for different values of stand characteristics. The different responses with different mean diameter and basal area classes were accounted by forming dummy variables for small (<15), medium (>20), large (>25) and very large (>30) diameters (SD, MD, LD and XLD, respectively), and for small (<15),

large (>20) and very large (>25) basal areas (SG, MG, and LG, respectively). This made it possible to depict the differences in the responses in these different areas. The dummy variables were also used in interaction terms, so that each independent variable could have a different coefficient in the different areas. The cutting points for the dummy variables (15, 20, 25 and 30) were selected by testing different values and selecting the ones with biggest improvement in information criterion AIC.

In the CART approach, we modelled the same dependent variable OK as in logistic case. In CART, the approach is to divide the dataset hierarchically to homogeneous groups (terminal nodes or leaves). In applications the whole group will have the same prediction, in this case the probability of correct decision. In a regression tree, the division is based on the sum of squares of the dependent variable (also called deviance in rpart), and the division is carried out in order to maximize the difference of the original sum of squares and the sum of the sum of squares in the two branches, i.e. max(SSparent – (SSleft + SSright)) (see e.g. Breiman et al.1984). The size of the tree is controlled through a complexity parameter cp. In this case, the model was estimated using first cp=0.00005 and then pruned. The cp parameter was set to a very low level at first, in order to find the cp that gave the minimum variance in the studied case. However, very low values of cp produce highly complicated regression trees, and the tree was then pruned to value cp=0.005, giving a tree with both reasonable size and fit.

3 Results

## 3.1 The Probabilities within Classes

The probability of correct decisions with each level of errors in basal area (eG) and mean diameter (eD), assuming the absolute error in mean height (eH) lower than 5% is given in Table 2. If the required accuracy for the decisions were 95%, the errors in the input variables G and D should always be in the interval [0%,10%[. The interval [–10%,0[was already much weaker (probability of correct decisions 84%). However, the differ-

Table 2. The probability of making a correct decision in different error classes for G and D. The error of H is assumed to be [–5%–5%].

G\D	–25	–15	–5	5	15	25
–25	0.55	0.54	0.40	0.39	0.31	0.23
–15	0.75	0.74	0.62	0.47	0.44	0.37
–5	0.66	0.78	0.84	0.90	0.59	0.50
5	0.54	0.79	0.93	0.96	0.81	0.62
15	0.31	0.41	0.48	0.54	0.67	0.63
25	0.26	0.31	0.38	0.33	0.43	0.44

ence between these two classes is mainly due to the observations with zero error with respect to one or two characteristics.

Besides the errors, the probability of errors also depends on the stand characteristics, basal area (G) and the basal area median diameter (D) and height (H), and the type of the true next treatment (final felling, first thinning or later thinning). To illustrate the effects of these characteristics, the proportion of correct schedules were plotted on the realized relative errors in basal area, mean
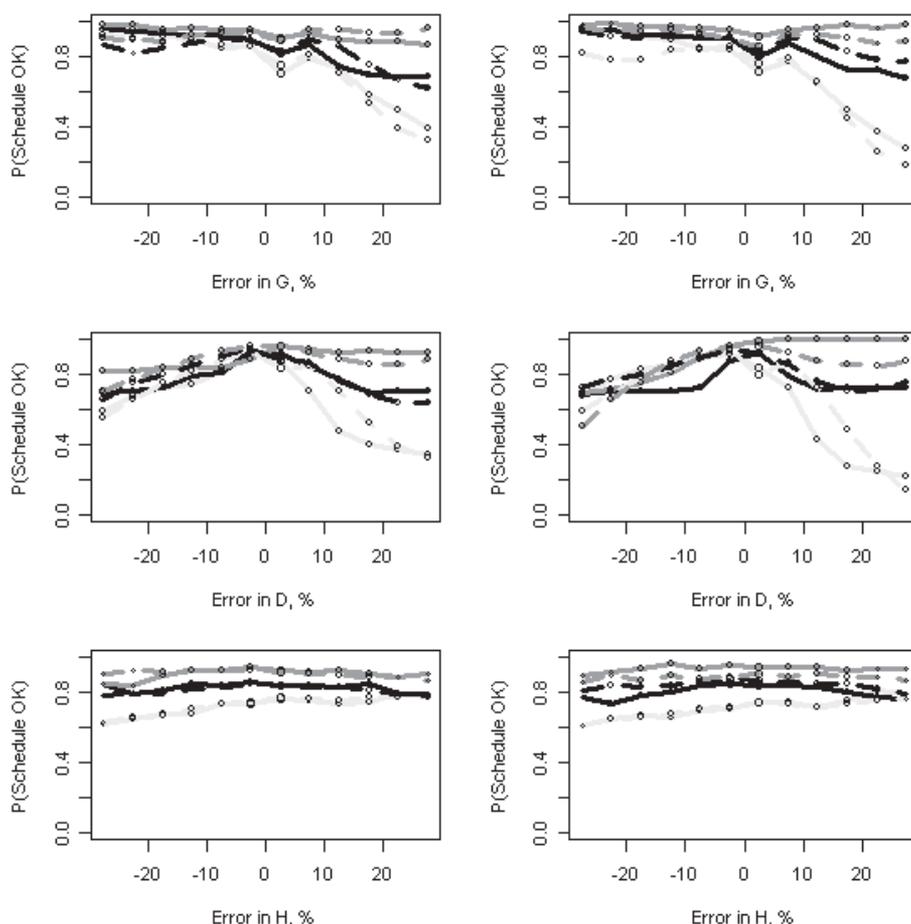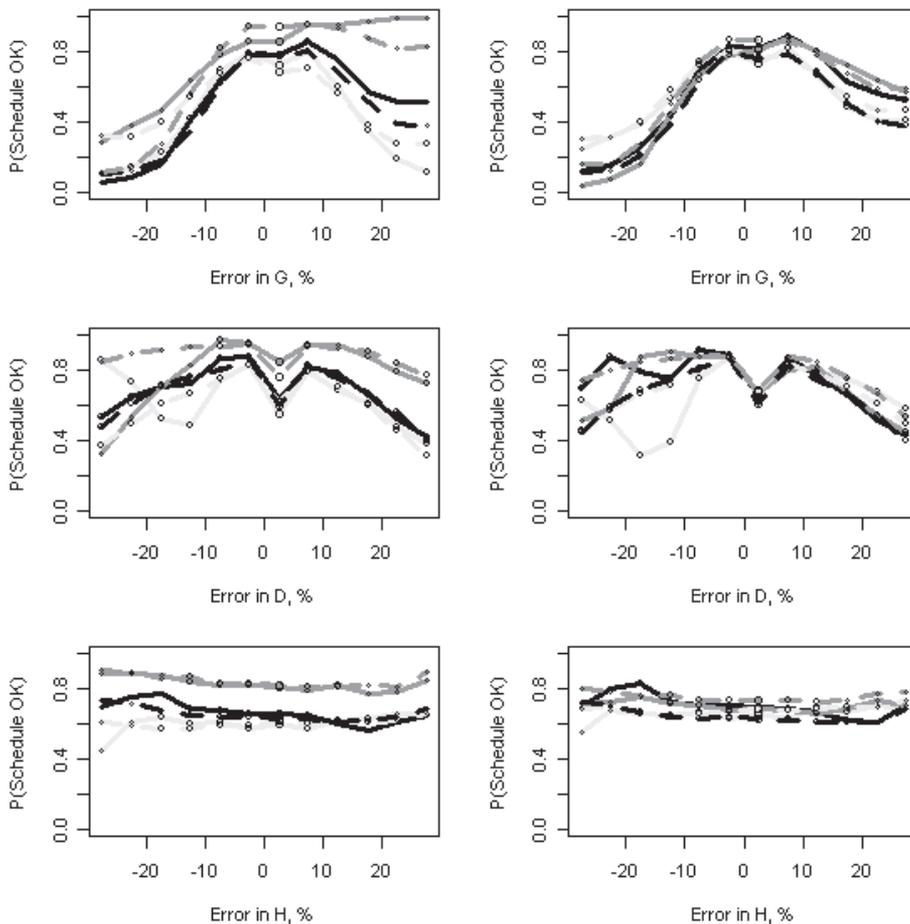

Fig. 2. The probability of correctly defined *final felling* as a function of the relative errors in basal area (up), mean diameter (middle), and height (low) in different classes of true basal area (left, light grey: G<22, black:22<G<26, grey:G>26) and mean diameter (right, light grey: D<25, black: 25<D<29, grey:29>D). Solid lines shows the proportion of correct schedules in each class, and the dashed line the mean of predictions in the class.

diameter and mean height in different classes of basal area and mean diameter. Fig. 2 shows these relationships in stands where the next correct treatment was final felling, Fig. 3 in stands where the next correct treatment is first thinning, and finally, Fig. 4 in stands where the next correct treatment is later thinning. In this classification, the next correct treatment may happen at any point in time within the simulation period of 30 years, not necessarily within the first ten years.

The plots show the relationship in classes of basal area and mean diameter. These plots are somewhat similar, because there is a rather high, linear correlation between basal area and mean diameter in the data (r=0.52).

The solid lines in upper plots of Fig. 2 show that overestimation of basal area decreases the probability of having the final felling correctly scheduled, especially in stands with low mean diameter and basal area. In contrast, underestimation of basal area seems to increase the probability of correctly scheduled final felling. This may be

Fig. 3. The probability of correctly defined *first thinning* as a function of the relative errors in basal area (up), mean diameter (middle), and height (low) in different classes of true basal area (left, light grey: G<18, black:18<G<22, grey:G>22) and mean diameter (right, light grey: D<11, black: 11<D<13, grey:13<D). Solid lines shows the proportion of correct schedules in each class, and the dashed line the mean of predictions in the class.
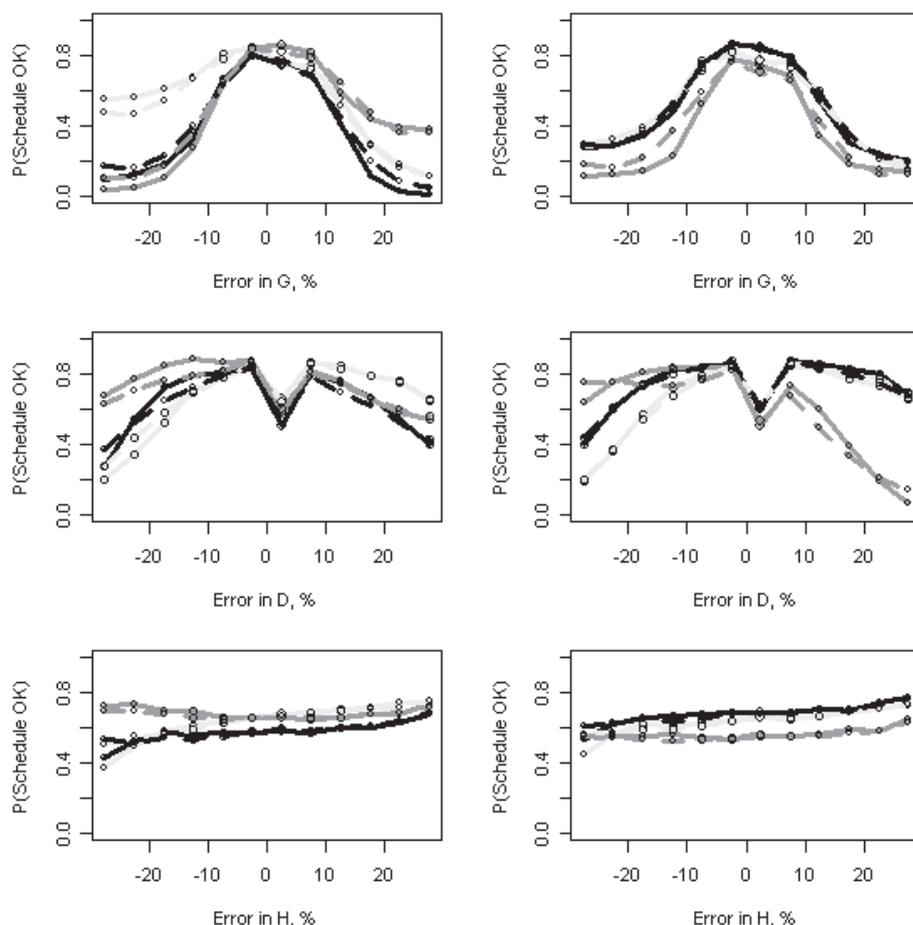
Fig. 4. The probability of correctly defined *later thinning* as a function of the relative errors in basal area (up),
mean diameter (middle), and height (low) in different classes of true basal area (left, light grey: G<19,
black:19<G<23, grey:G>23) and mean diameter (right, light grey: D<17, black: 17<D<20, grey:20>D).
Solid lines shows the proportion of correct schedules in each class, and the dashed line the mean of predic-
tions in the class.

explained by the fact that overestimation of basal
area may lead to scheduling a thinning instead of
the final felling.

The solid lines in middle plots of Fig. 2 show
the effect of errors in mean diameter on the sched-
uling of final felling. In stands with high mean
diameter, underestimation of diameter leads to
delayed final fellings. On the other hand, over-
estimation does not lead to incorrect decision,
because the initial diameter is already above the
final felling limit. In stands with medium mean
diameter (25–29 cm), both over- and underestima-

tion decrease the probability of correct schedule:
overestimation leads to early and underestimation
to late clear-cuts. In the lowest diameter class
(21–25 cm) the correct timing for final felling is
usually after the 10 year simulation period. Thus,
underestimation or slight overestimation of diam-
eter does not necessarily change the schedule for
the next 10 years. However, large overestimation
of diameter may make it happen already within
the next ten years.

Underestimation of stand density causes
delayed thinnings in all classes of basal area

and mean diameter (solid lines in upper plots of Figs. 3 and 4). The effect of overestimation of basal area varies according to initial basal area and treatment. In dense stands overestimation does not decrease the probability of a correctly scheduled first thinning (solid lines in upper left Fig. 3), but in later thinnings it does (solid lines in upper left Fig. 4). The reason may be that in case of later thinnings also the highest class of basal area (23–40) includes quite many stands that are not to be thinned immediately (thinning limit is usually above 30 m²/ha, see Fig. 1).

Both positive and negative errors in mean diameter decreased the probability of correctly scheduled first and later thinnings. The greatest decrease in the probability was observed when the mean diameter was overestimated in later thinning stands with a high mean diameter. These stands were then incorrectly scheduled to a clearcut. The solid lines in lowest plots of Figs. 3 and 4 show that errors in mean height do have only a slight effect.

A striking feature in the middle plots of Figs. 3 and 4 is a clear decrease in the probability of correctly defined thinning when there is no error in the mean diameter. However, this is an artefact, which results from the way in which the data were generated. Almost all realizations where error in mean diameter is 0 have nonzero errors in other variables, whereas high proportion of realizations with nonzero error in mean diameter has zero errors in other characteristics. This can also be seen from Table 2, where all the errors were controlled at the same time.

3.2 Modeling the Errors

3.2.1 The Logistic Model

According to the Figs. 2–4, the dependency between the probability of correct schedule, π, the stand characteristics and the errors in them seems to be nonlinear (this was confirmed by plotting the graphs of Figs. 2–4 with y-axis in the logit scale). Furthermore, there seems to be several important interactions. All three different errors (eG, eD, and eH) seem to have different effect on the probability of correct decision with different values of basal area, mean diameter, and

the next operation. The effects are usually different for positive and negative errors.

The estimated model is shown in Table 3. The dashed lines in Figs. 2–4 show the mean of fitted values in the classes of errors and stand variables. The final model estimated with logistic regression is very complicated, with almost 100 independent variables that are all modifications of the basic variables D, G, H, ST, the future operation and the errors. All these independent variables are statistically significant, however, except the coefficients for site classes 5 and 6, but these were included in the model to ensure model logicality. With this model, the predicted probabilities in the different classes were fairly well in accordance with the observed probabilities (see Figs. 2–4). The deviance could be reduced from the original 227498 to 132398, i.e. to 58.2% of the original (meaning that the explained variation was 41.8%). This model was able to correctly classify the future decisions (as correct or incorrect) in 84.05% of cases.

Calculating analytically the acceptable error levels for each type of stand with this model (formula 3) is somewhat complicated.

3.2.2 The CART Model

The CART model was first fitted with cp = 0.00005, and then pruned to a cp providing the minimum error. This cp was 0.00005841 (Fig. 5). This model had 1203 splits and over 130 million different nodes. With this cp, the model was very flexible, and could follow the observed probabilities almost perfectly. The proportion of correctly classified decisions was 92.14%. The original residual error could in this case be reduced to 27.7% % of the original variation, i.e. the explained variation was 72.3%. However, interpretation of the model is practically impossible because it is so complicated. This model is way too complex for practical use. Pruning the tree to the same proportion of explained variation than the glm model, would give cp = 0.0027 and 34 splits, which would classify correctly 83% of the decisions. However, the model was further pruned to cp = 0.005 and 15 splits, after which the improvements with new splits were already small. Then, the total number of nodes was 263, of which the number of terminal nodes (i.e. different predictions) was 16.

Table 3. The fitted model. G and D denote the basal area and mean diameter. Variables beginning with "e" denote the error in the corresponding variable and variables ending with "neg" or "pos" denote negative or positive errors. Other variables are dummy variables for site types 2–6 (1 is the default), dummy variables for the mean diameter classes (SD, MD, LD, XLD) and basal area classes (SG, MG and LG) and dummy variables for the true next treatment.

|	Estimate	Std. Error	Pr(>|z|)		Estimate	Std. Error	Pr(>|z|)		Estimate	Std. Error	Pr(>|z|)
Intercept	-9.0460	0.4956	0.0000	eHpos	0.0168	0.0064	0.0086	D:eGneg	-0.0285	0.0012	0.0000
G	-1.0630	0.0222	0.0000	eHneg	0.2138	0.0062	0.0000	G:eDneg	-0.0226	0.0009	0.0000
D	1.0050	0.0228	0.0000	XLD:eDneg	-0.8357	0.1732	0.0000	G:eHneg	-0.0068	0.0003	0.0000
eST	-0.4307	0.0303	0.0000	MD:eDpos	-0.1546	0.0085	0.0000	G:eHpos	-0.0010	0.0003	0.0001
ST2	-0.5278	0.0601	0.0000	MD:eDneg	0.0877	0.0081	0.0000	G:eGneg	0.0230	0.0011	0.0000
ST3	-0.3808	0.0603	0.0000	SD:eDpos	0.0658	0.0098	0.0000	GD:eDpos	-0.0279	0.0065	0.0000
ST4	-0.5247	0.0608	0.0000	SD:eHpos	-0.0168	0.0031	0.0000	GD:eDneg	0.4743	0.0143	0.0000
ST5	-0.1553	0.0862	0.0715	SD:eHneg	-0.0532	0.0031	0.0000	GD:eGpos	0.0181	0.0062	0.0036
ST6	-0.3301	0.2091	0.1143	D:XLD	2.1300	0.1352	0.0000	GD:eGneg	-0.2421	0.0177	0.0000
GD	15.8400	0.4094	0.0000	D:LD	-0.5582	0.0243	0.0000	LG:eGpos^2	0.0025	0.0004	0.0000
XLD	-63.0700	4.2290	0.0000	D:MD	-0.5263	0.0150	0.0000	LG:eGneg^2	0.0007	0.0002	0.0019
LD	14.3400	0.6564	0.0000	D:SD	02.6567	0.0187	0.0000	LD:eDpos^2	0.0056	0.0002	0.0000
MD	10.7700	0.3257	0.0000	LG:eGpos	0.0230	0.0094	0.0147	MG:eGneg^2	0.0029	0.0003	0.0000
SD	-10.4500	0.2666	0.0000	MG:eGneg	0.1776	0.0075	0.0000	MD:eDpos^2	0.0023	0.0003	0.0000
LG	-5.4920	0.3202	0.0000	SG:eGneg	-0.2018	0.0192	0.0000	MD:eDneg^2	0.0042	0.0003	0.0000
MG	-8.5890	0.2756	0.0000	G:LG	0.2019	0.0123	0.0000	SG:eGneg^2	-0.0031	0.0008	0.0001
SG	15.1100	0.4150	0.0000	G:MG	0.4603	0.0131	0.0000	SD:eDpos^2	-0.0010	0.0004	0.0053
eGpos	-0.2218	0.0112	0.0000	G:SG	-1.0870	0.0297	0.0000	SD:eDneg^2	0.0021	0.0002	0.0000
eGpos^2	-0.0107	0.0008	0.0000	XLD:eGpos	-2.0850	0.2482	0.0000	LG:eDneg^2	0.0036	0.0003	0.0000
eGpos^3	0.0003	0.0000	0.0000	LD:eGpos	0.1859	0.0112	0.0000	LD:eGpos^2	-0.0023	0.0004	0.0000
eGneg	0.4862	0.0262	0.0000	LD:eGneg	-0.2162	0.0135	0.0000	LD:eGneg^2	-0.0032	0.0006	0.0000
eGneg^2	-0.0069	0.0008	0.0000	MD:eGpos	-0.0097	0.0033	0.0032	eDpos:eGpos	0.0094	0.0002	0.0000
eGneg^3	-0.0003	0.0000	0.0000	SD:eGpos	0.0864	0.0035	0.0000	eDneg:eGpos	-0.0045	0.0002	0.0000
eDpos	-0.0746	0.0184	0.0000	SD:eGneg	-0.0138	0.0048	0.0042	eDpos:eGneg	-0.0039	0.0002	0.0000
eDpos^2	0.0031	0.0008	0.0002	eHpos:SG	0.0264	0.0062	0.0000	eDneg:eGneg	0.0116	0.0002	0.0000
eDpos^3	-0.0001	0.0008	0.0000	eDpos:LG	0.0558	0.0036	0.0000	eDneg:eHpos	-0.0014	0.0001	0.0000
eDneg	-0.0481	0.0221	0.0296	eDneg:LG	0.0644	0.0093	0.0000	eDpos:eHneg	-0.0037	0.0001	0.0000
eDneg^2	0.0046	0.0008	0.0000	D:eDneg	0.0130	0.0010	0.0000	eDpos:eHpos	-0.0019	0.0001	0.0000
eDneg^3	0.0002	0.0000	0.0000	D:eDpos	-0.0047	0.0006	0.0000	G: first_thinning	-0.2066	0.0159	0.0000
G: thinning	-0.1252	0.0080	0.0000	eDpos: first_thinning	-0.0374	0.0061	0.0000	eDneg: thinning	-0.0843	0.0038	0.0000
D: first_thinning	0.1474	0.0205	0.0000	eDpos: thinning	0.0487	0.0038	0.0000	D:XLD: eGneg	0.0689	0.0077	0.0000
D: thinning	0.0324	0.0083	0.0000	eDneg: first_thinming	-0.2240	0.0060	0.0000	D:XLD: eDneg	0.0330	0.0055	0.0000
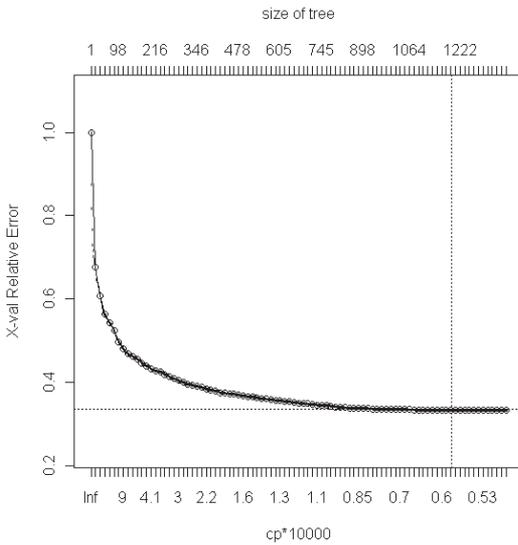
Fig. 5. The relative error in CART, as a function of complexity parameter cp (times 10000) and the tree size (number of splits). The horizontal dashed line describes the minimum error level and the vertical dashed line the cp parameter producing the minimum relative error.

The relative error of the model was 65% of the initial variation. This model is already pretty simple to interpret (Fig. 6), and the predictions are still fairly good (Figs. 7–9). The proportion of correctly classified decisions with this model was 80.05%.

4 Discussion

In this study, we analyzed the errors in the forest data from the decision making point of view. The aim was to define what is the accuracy required in stand characteristics, in order to achieve an acceptable accuracy in the decisions of harvest scheduling. The analysis carried out also helps in defining in which kinds of stands the good quality data is most important. The analysis was carried out by analysing the probabilities in different error classes, by visual inspection in stands with different next operations and by making a logistic or a CART model for the probability of correct decisions.
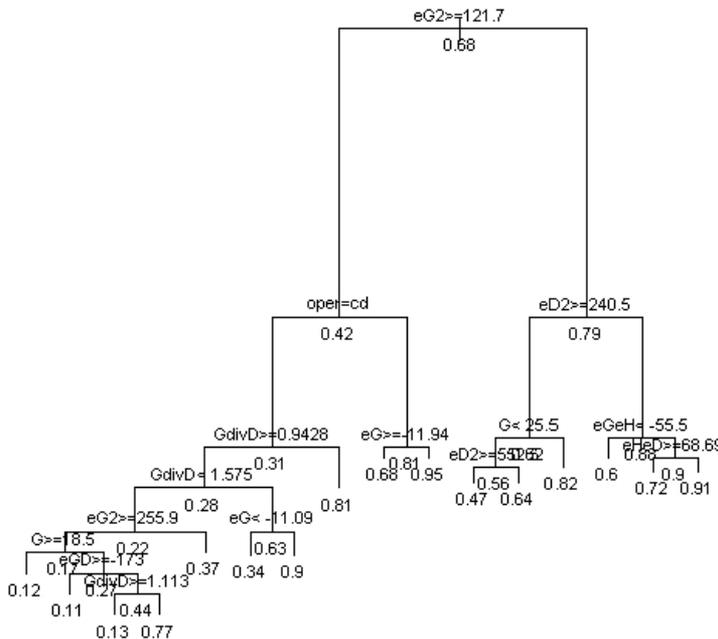

Fig. 6. The CART model with cp=0.005. The definitions of the variables are the same as in Table 2. For example, G is basal area, eG is error in basal area, and eG2 its second power. GdivD means basal area divided by mean diameter. Oper=b means clear-cut, c means first thinning and d thinning.

703
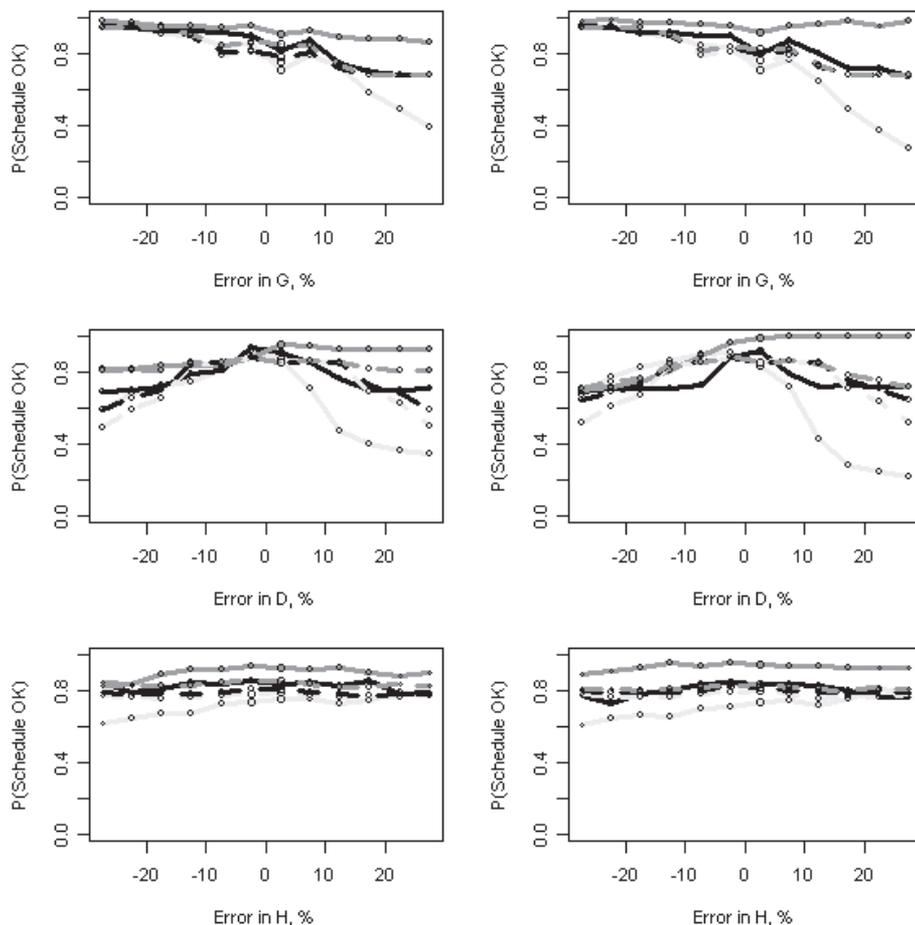
Fig. 7. The probability of correctly defined *final felling* as a function of the relative errors in basal area (up), mean diameter (middle), and height (low) in different classes of true basal area (left, light grey: G<22, black:22<G<26, grey: G>26) and mean diameter (right, light grey: D<25, black: 25<D<29, grey: D>29). Solid lines shows the proportion of correct schedules in each class, and the dashed line the mean of predictions in the class.
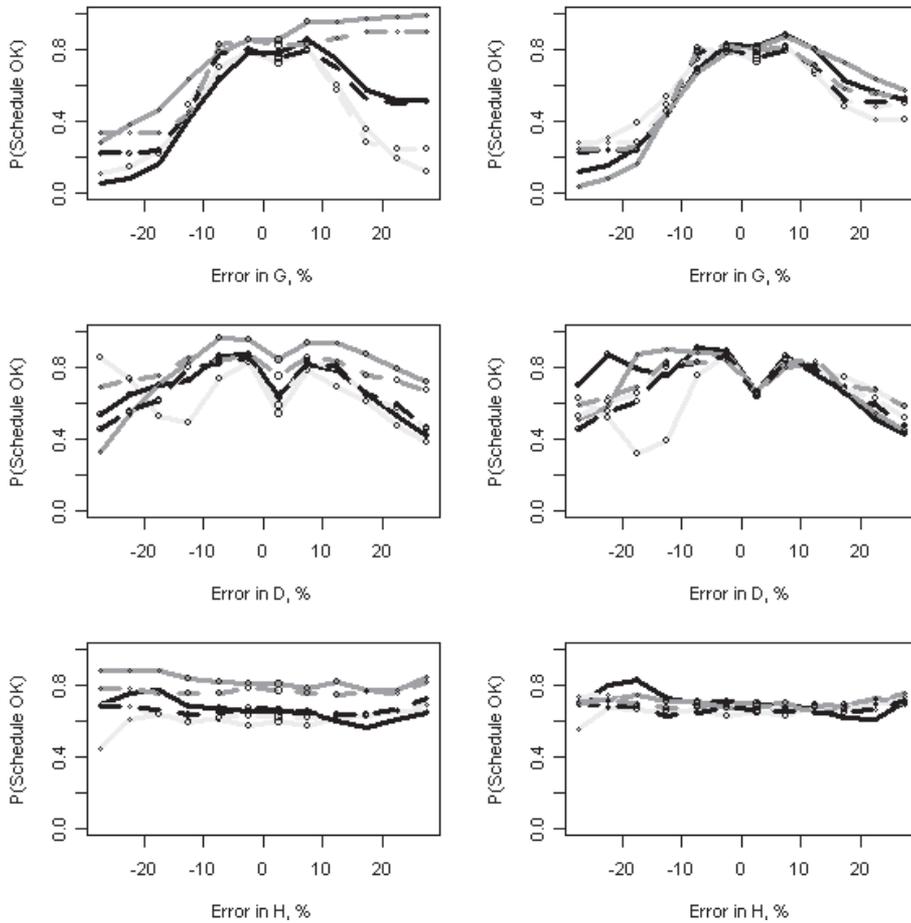
Fig. 8. The probability of correctly defined *first thinning* as a function of the relative errors in basal area (up), mean diameter (middle), and height (low) in different classes of true basal area (left, light grey: G<18, black:18<G<22, grey:G>22) and mean diameter (right, light grey: D<11, black: 11<D<13, grey:D>13). Solid lines shows the proportion of correct schedules in each class, and the dashed line the mean of predictions in the class.
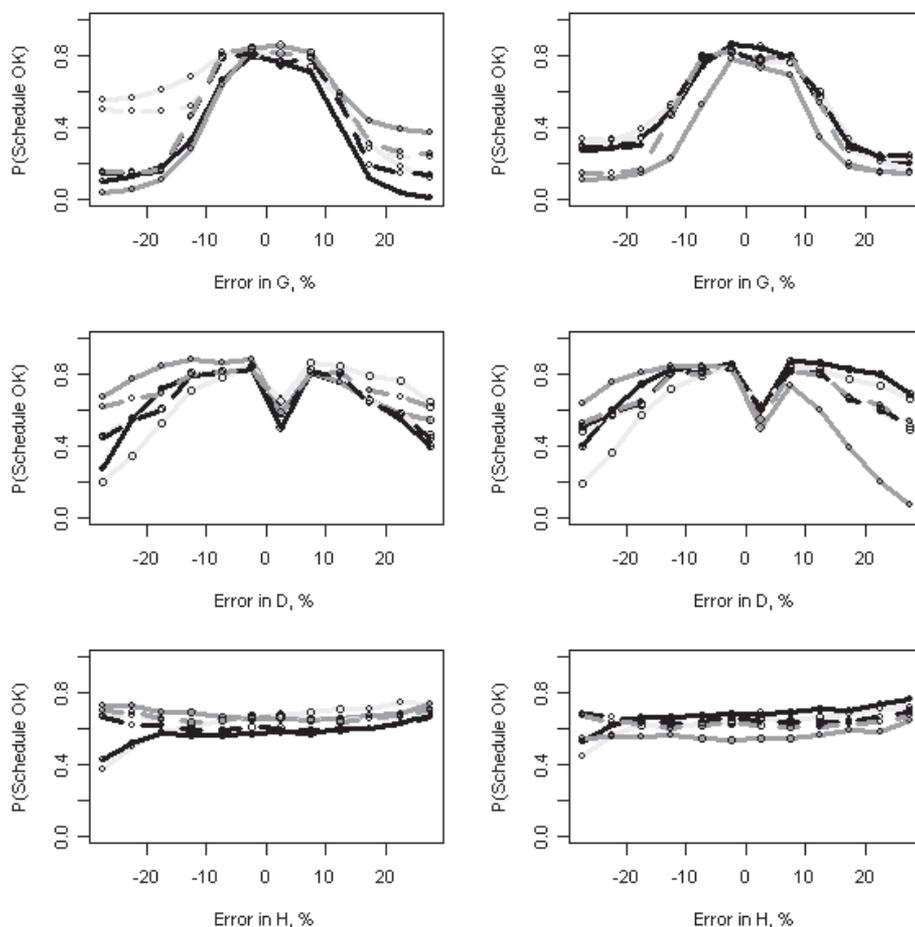
Fig. 9. The probability of correctly defined *later thinning* as a function of the relative errors in basal area (up), mean diameter (middle), and height (low) in different classes of true basal area (left, light grey: G < 19, black:19 < G < 23, grey:G > 23) and mean diameter (right, light grey: D < 17, black: 17 < D < 20, grey:D > 20). Solid lines shows the proportion of correct schedules in each class, and the dashed line the mean of predictions in the class.

In this study, we assumed that the treatment recommended in the silvicultural guidelines was the optimal one. This was done in order to be able to simulate just one treatment schedule (i.e. the one recommended in the silvicultural guidelines) for each stand and each error scenario. The reasoning is that the guidelines have been formed to produce the optimal treatment for an average stand (e.g. Hyytiäinen et al. 2004). Another justification to our approach is that the applied guidelines are the main tools of practical harvest scheduling in Finland. Here, as the selected stands are not so important as such, but tolerable average error levels in different kinds of stands are of main interest, this can be justified. In principle it is, however, possible that some "incorrect" schedules were actually closer to the optimal than the "correct" schedules.

In the SIMO simulator (like in other computer simulators) the rules for the treatment schedules are exact: if the stand characteristic considered exceeds a given exact limit, the treatment is scheduled for the year in question, but not otherwise. This makes the system very sensitive to the errors. Therefore, we utilised the tolerance limits of ±2 or ±3 years. It also means that the tolerable error limit for each input variable could be inferred directly from the given limits for the treatments, if the treatments were to be carried out immediately. For instance, if the diameter limit for clear-cut were 28 cm, then for 32 cm stand the largest tolerable underestimate would be 4 cm, and overestimation had no effect. Since the treatments can happen at any time during the 10-year period, but the decisions are based on the data measured in the beginning of the period, the situation is not quite as simple, as the effect of growth needs to be considered. Further complication comes from the tolerance limit given to the timing and from the interactions of errors in several variables. Therefore, we decided to model the probability.

Islam et al (2010) used the realized errors to predict the monetary inoptimality losses due to inventory errors. Their data did not show any reason to treat positive and negative errors separately, whereas our data showed. One possible reason for this difference may be that their schedules were based on optimization using a stand level growth simulator, whereas we applied the

silvicultural guidelines, i.e. rule-based reasoning. Another possible explanation is that our data set was much larger, and included a large number of both under- and overestimates for each stand and variable: the larger the data, and the wider the ranges of independent variables, the easier it is to observe the trends and obtain statistically significant differences.

The results (Table 2, Figs. 2–4) show that precise estimates for basal area are very important for thinning decisions. In practise, the errors should be within ±10% in order to obtain correct decisions with high probability (>80%). This degree of accuracy is not often obtained with the current inventory methods: in the traditional compartmentwise inventory with a RMSE of 25%, within these limits are about 39% of the basal area assessments (Kangas and Lappi 2011). With the currently adopted laser scanning methodology the proportion is about 54% (Kangas and Lappi 2011). The errors in mean diameter and height do not seem quite as detrimental, but obtaining estimates within ±20% would give high probabilities for good decisions in most cases. This accuracy is obtained with both above mentioned methods at least in 70% of cases (Kangas and Lappi 2011). Consequently, using either of these methods the overall probability of correct decisions in any given stand is much lower than 80%.

This result is confirmed in the CART analysis. The first division in CART model is based on the second power of error in basal area (eG2), and if this error is low enough (<121.7, i.e. $|eG| < 11.03\%$), the probability of making a correct decision is 0.79, and in the other branch 0.42. If the error in basal area is higher than 11%, then it is still possible to obtain good accuracy of decisions (0.81) if the next operation is clear-cut (b) instead of first thinning (c) or thinning (d) (0.31). In the thinning stands, if basal area G is small compared to mean diameter D (GdivD < 0.94), it is still possible to make accurate decisions (probability 0.81). If the error of basal area is small, and the absolute error of mean diameter is also less than 15.5%, the probability of correct decisions is as high as 0.88. It is notable that in one of the leaves only the probability is as high as 0.95.

Our results apply at the stand level, and can be used for selecting the accuracy needed in measuring a given stand. The overall accuracy of the

decisions in a forest area, would depend on the age distribution of the area, for instance, besides the measurement accuracy. The forest-area level analysis is out of the scope of this study, but the models estimated could be used to simulate the accuracy of decisions in a forest area where the inventory is carried out with a given method, with a given error distribution. Such a tool could be used for defining if a planned inventory method is accurate enough for decision making.

In this study, we modelled the probability both with a parametric logistic regression and with a non-parametric regression tree CART. Two methods were used, as both of them have their pros and cons. With logistic model, it would be possible to analytically solve the error level that would produce a given probability of accuracy for the decisions in a given stand. As the error is included as a signed error and in different powers, it means that the equation has to be separately solved for positive and negative errors and the correct sign of the possible solutions needs to be chosen accordingly. The produced model is overwhelmingly complicated, however, and very difficult to interpret. Interpreting the results of the model would require a computer tool for predicting the probability in different situations. CART model, on the other hand, is very simple to interpret visually from the tree plot. The problem is that the model does not allow for calculations of probability for any other combinations of stand characteristics and errors than the ones predicted. It is not, therefore, possible to determine a combination where the accuracy of decisions would be, say 97%. The CART results would thus be useful for forest practitioners, who could very efficiently see how reliable data they would need for making the decisions, and the logistic model for researchers, who could use it to simulate area-level inventories and analyse the accuracy of the overall results.

A notable difference between the CART model and logistic regression model was that in the CART model positive and negative errors needed not to be treated separately, dummy variables were not useful and also the interactions were less useful than in the logistic model. It means that with CART, the modeller does not need to worry about the model shape, at least as much as in parametric regression. It also means that it is much

easier to get reasonable results in a complicated modelling task like this. For instance, it is probable that the parametric model could be improved to the same level as the variance minimizing CART model, with R^2 72%, if new interaction terms were discovered. This would, however, mean a lot of work. Therefore, CART models are also used in datamining applications.

In principle, the regression models should be estimated using methods that take into account the correlation of observations that are simulated for the same stands. However, as the OLS estimates are still unbiased (although the p-values may be overestimated), this further complication was not deemed necessary here. The reasoning is that the stand-level variation would describe such between-stand variation that cannot be explained with the used predictors. In this case, the possibility of such variation was deemed negligible, as all correct treatments were generated using the same, exact rules, and the same growth models, that were assumed correct. Thus, all the variation in the data set is due to simulated errors, and none due to random variation, all the predictors that have significant values are truly useful in the model.

References

Borders, B.E., Harrison, W.M., Clutter, M.L., Shiver, B.D. & Souter, R.A. 2008. The value of timber inventory information for management planning. Canadian Journal of Forest Research 38: 2287–2294.

Breiman, L., Friedman, J.H., Olshen, R.D. & Stone, C.J. 1984. Classification and regression trees. Wadsworth & Brooks, Belmont, CA.

Burkhart, H.E., Stuck, R,D., Leuschner, W.A. & Reynolds, M.A. 1978. Allocating inventory resources for multiple-use planning. Canadian Journal of Forest Research 8: 100–110.

Duvemo, K. 2009. The influence of data uncertainty on planning and decision processes in forest management. Doctoral Thesis. Faculty of Forest Sciences, Department of Forest Resource Management, Umeå. Acta Universitatis Agriculturae Sueciae. 60 p. + appendices.

— , Barth, A. & Wallerman, J. 2007. Evaluating

sample plot imputation techniques as evaluating sample plot imputation techniques as input in forest management planning. Canadian Journal of Forest Research 37: 2069–2079.

Eid, T. 2000. Use of uncertain inventory data in forestry scenario models and consequential incorrect harvest decisions. Silva Fennica 34: 89–100.

— , Gobakken, T. & Næsset E. 2004. Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analysis. Scandinavian Journal of Forest Research 19: 512–523.

Hamilton, D.A. 1978. Specifying precision in natural resource inventories. In: Integrated inventories of renewable resources: proceedings of the workshop. USDA Forest Service, General Technical Report RM-55: 276–281.

Holopainen, M., Mäkinen, A., Rasinmäki, J., Hyyppä, J., Hyyppä, H., Kaartinen, H., Viitala, R., Vastaranta, M. & Kangas, A. 2010. Effect of tree-level airborne laser-scanning measurement accuracy on the timing and expected value of harvest decisions. European Journal of Forest Research 129: 899–907.

Hynynen, J., Ojansuu, R., Hökkä, H., Siipilehto, J., Salminen, H. & Haapala, P. 2002. Models for predicting stand development in MELA system. Finnish Forest Research Institute Research Papers 835.

Hyytiäinen, K., Hari, P., Kokkila, T., Mäkelä, A., Tahvonen, O. & Taipale, J. 2004. Connecting a process-based forest growth model to stand-level economic optimization. Canadian Journal of Forest Research 34: 2060–2073.

Islam, N., Kurttila, M., Mehtätalo, L. & Haara, A. 2009. Analyzing the effects of inventory errors on holding-level forest plans: the case of measurement error in the basal area of the dominated tree species. Silva Fennica 43: 71–85.

Islam, Md.N., Kurttila, M., Mehtätalo, L. & Pukkala, T. 2010. Inoptimality losses in forest management decisions caused by errors in an inventory based on airborne laser scanning and aerial photographs. Canadian Journal of Forest Research 40: 2427–2438.

Kangas, A. 2010. Value of forest information. European Journal of Forest Research 129: 863–874.

— & Lappi J. On comparing agreement of forest variable estimates obtained with different methods. Submitted manuscript.

— & Rasinmäki, J. 2008. (eds.). SIMO Adaptable simulation and optimization for forest management planning. Department of Forest Resource Management Publications 41. 40 p.

Lawrence, D.B. 1999. The economic value of information. Springer. 393 p.

Mäkinen, A., Holopainen, M., Kangas, A. & Rasinmäki, J. 2010. Propagating the errors of initial forest variables through stand- and tree-level simulators. European Journal of Forest Research 129: 887–897.

Rasinmäki, J., Kalliovirta, J. & Mäkinen, A. 2009. An adaptable simulation framework for multiscale forest resource data. Computers and Electronics in Agriculture 66: 76–84.

R Development Core Team. 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: http://www.R-project.org.

[Recommendations of Tapio for good silviculture]. 2006. Metsäkustannus. 100 p. (In Finnish).

Ståhl, G. 1994. Optimizing the utility of forest inventory activities. Swedish University of Agricultural Sciences, Department of Biometry and Forest Management. Report 27.

— , Carlson, D. & Bondesson, L. 1994. A method to determine optimal stand data acquisition policies. Forest Science 40:630–649.

Total of 23 references