

Influence of Growth Prediction Errors on the Expected Losses from Forest Decisions

Ilona Pietilä, Annika Kangas, Antti Mäkinen and Lauri Mehtätalo

Pietilä, I., Kangas, A., Mäkinen, A. & Mehtätalo, L. 2010. Influence of growth prediction errors on the expected losses from forest decisions. *Silva Fennica* 44(5): 829–843.

In forest planning, forest inventory information is used for predicting future development of forests under different treatments. Model predictions always include some errors, which can lead to sub-optimal decisions and economic loss. The influence of growth prediction errors on the reliability of projected forest variables and on the treatment propositions have previously been examined in a few studies, but economic losses due to growth prediction errors is an almost unexplored subject. The aim of this study was to examine how the growth prediction errors affected the expected losses caused by incorrect harvest decisions, when the inventory interval increased. The growth models applied in the analysis were stand-level growth models for basal area and dominant height. The focus was entirely on the effects of growth prediction errors, other sources of uncertainty being ignored. The results show that inoptimality losses increased with the inventory interval. Average relative inoptimality loss was 3.3% when the inventory interval was 5 years and 11.6% when it was 60 years. Average absolute inoptimality loss was 230 euro ha⁻¹ when the inventory interval was 5 years and 860 euro ha⁻¹ when it was 60 years. The average inoptimality losses varied between development classes, site classes and main tree species.

Keywords growth prediction, uncertainty, forest information, updating, inoptimality loss

Addresses *Pietilä, Kangas and Mäkinen*, University of Helsinki, Department of Forest Sciences, P.O. Box 27, FI-00014 University of Helsinki, Finland; *Mehtätalo*, University of Eastern Finland, School of Forest Sciences, P.O. Box 111, FI-80101 Joensuu, Finland

E-mail annika.kangas@helsinki.fi

Received 2 June 2010 **Revised** 23 November 2010 **Accepted** 1 December 2010

Available at <http://www.metla.fi/silvafennica/full/sf44/sf445829.pdf>

1 Introduction

The purpose of forest planning is to help decision makers to choose appropriate forest management actions for a given planning period. The recommendations are based on information collected in the forest inventory. Nowadays in Finland, a forest inventory is commonly carried out at 10–15-year interval, and the information is updated between inventories using growth models. The predicted growth under different treatment schedules forms the basis for selecting the optimal future treatments for each stand. In Finland, growth models used in forest planning are usually tree-level models (e.g., Hynynen et al. 2002) but stand-level models (e.g., Gustavsen 1977, Saramäki 1977, Nyssönen and Mielikäinen 1978, Vuokila and Väliaho 1980, Oikarinen 1983) can also be used. Commonly used Finnish simulators for forest planning purposes are MELA (Siitonen 1996), MOTTI (Hynynen et al. 2005), SIMO (Rasinmäki et al. 2009) and MONSU (Pukkala 2004). MELA, MOTTI and MONSU are based on tree-level models, SIMO includes both tree- and stand-level growth models.

Future development of forests predicted by growth models always includes errors, the amount of which depends on (i) random variation in growth, (ii) inaccurate description of the present state of the forest and (iii) the parameters of growth models and their estimation errors (Ståhl and Holm 1994). In addition, there may be uncertainty concerning the model misspecification, i.e. the number of relevant predictors and the form of their effect, both of which may introduce systematic error or bias (e.g. Kangas 1999). The errors from the parameters of growth models or errors due to model misspecification can be reduced by making an improved model. Likewise, the errors due to inventory can be reduced by making a new, more accurate, inventory. Uncertainty caused by natural randomness, on the other hand, cannot be decreased with additional information. When growth is predicted for several consecutive periods, growth model variables include errors, which result from the prediction errors of previous periods (Kangas 1997). This error will accumulate in simulation and the longer the prediction period is the greater the variance of a given forest variable is.

The effects of growth prediction errors on the reliability of the projected forest variables and the treatment propositions in the planning system have been discussed in a few studies. Ojansuu et al. (2002), Hyvönen and Korhonen (2003) and Haara and Korhonen (2004) have studied prediction periods of less than 15 years. In all of these studies, the correct treatment propositions were defined by applying the silvicultural guidelines (Hyvän metsänhoidon... 2006) to the known data on initial state and realized growth. These were then compared with recommendations generated by the MELA software. Information updated using growth models produced only a few or no incorrect treatment propositions according to all of these studies. The economic losses due to them were not analyzed, however.

According to Haara and Leskinen (2009), uncertainty due to MELA growth models (Hynynen et al. 2002) commonly used in Finland leads to relative RMSE-values of 5–10% when prediction period was five years and relative RMSE-values of 5–20% when the prediction period was ten years at stands where no treatments were done. The studied variables were basal area median diameter, height of basal area median tree, basal area and the stand volume. Corresponding values at stands where the effect of treatments was included were 5–20% and 5–40%. This is the best available information concerning the accuracy of growth models used in Finland. Mäkinen et al. (2008) had a longer period (20 years), but they only had a small data set from central Finland, and the stands were selected only from among very dense stands.

Mostly the different error sources are considered one at a time. Holopainen et al. (2010), however, considered the effects of forest inventory errors, stochastic variation in timber assortment prices and errors in growth predictions in predicting net present values (NPV) of individual stands. When these factors were considered one at a time, growth prediction errors had the greatest impact on relative variance of the NPVs. When all of the factors were included, interaction between factors led to a greater variance of NPVs. The variance was smaller, however, than the sum of variances of all of the factors, indicating that the different errors may somewhat cancel each other out.

A lot of effort has lately been put to decreasing forest inventory costs. New methods have been developed for collecting data using remote sensing and for updating the information continuously using growth models. The aim has typically been to achieve at least the same level of accuracy as with the traditional method. The various implications of information accuracy have been researched only recently. Improving the accuracy of the information is important only if the improvement has a clearly positive influence on the decisions made using the information. Erroneous information can lead to incorrect decisions which cause economic loss (Burkhart et al. 1978, Hamilton 1978), a so-called inoptimality loss. Such a loss can be calculated as the difference between net present values produced by error-free information and those produced by erroneous information (as in Eid 2000). Calculating the inoptimality loss requires that the error-free, optimal result is known.

The research on the effects of inoptimality losses has so far focused on losses due to forest inventory errors (e.g., Ståhl et al. 1994, Eid 2000, Holmström et al. 2003, Eid et al. 2004, Holopainen and Talvitie 2006, Juntunen 2006, Duvemo et al. 2007, Borders et al. 2008, Islam et al. 2009, Mäkinen et al. 2010). In one study, the effect of fluctuations of prices has been included (Ståhl 1994). Effects of growth prediction errors on inoptimality losses have been investigated much less, even though they have an effect on how long the inventory results can be used in decision making. So far it has been assumed that the information, once collected, is equally usable throughout the whole planning period, while in reality the quality of information deteriorates over time.

If the economic losses caused by lengthening the inventory interval are known, the decision maker can evaluate when new information should be collected. In some time-point, collecting new information will reduce the losses more than the new inventory introduces costs, and this time point defines the optimal life-span of data. In order to make this kind of choice, the decision maker must also know the quality and cost of a new inventory. In general, the more uncertain the decision is, the more useful new information should be (Eid 2000). For instance, sometimes

the choice between thinning a stand and final harvesting is very uncertain (i.e. the two options may seem equally profitable) and the decision made may affect the income of forest owner considerably. If the stand meets the regeneration criteria and will be harvested regardless of errors in growth modelling, however, growth prediction errors are negligible.

In this study, we evaluated the influence of growth prediction errors on the expected losses from forest decisions at various inventory intervals. It was assumed that the longer the inventory interval is, the lower the quality of information becomes because of errors in growth prediction. These losses can further be used in defining the life-span of the information.

2 Material and Methods

2.1 Material

The input data for the forest planning simulations consisted of 99 stands in central Finland. The various stand attributes were estimated using a traditional stand-wise field inventory, but for the purposes of this study they were treated as if they were true values. The average stand size was 2 hectares and thus the total area of the stands was approximately 200 hectares. Different development classes, site classes and main tree species were well represented in the material. Average, maximum and minimum values of aggregate stand attributes are shown in Table 1a, the site quality distribution in Table 1b and the species distribution in Table 1c.

2.2 Methods

The simulation period in this study was 60 years and the effects of the growth prediction errors were determined for inventory intervals of 5, 10, 15, 20, 30 and 60 years. In order to be able to capture the effects of growth prediction errors, we assumed that inventory errors did not occur. Thus, new error-free information on stand attributes (in this case more specifically basal area G and dominant height H_{dom}) was obtained in each inventory

Table 1a. Average, maximum and minimum values of aggregate stand attributes in forest planning input data.

Stand variable	Average	Max	Min
Basal area (m ² ha ⁻¹)	13.7	35.3	0
Basal area median diameter (cm)	13.2	34.4	0
Age (years)	38.8	180	1
Dominant height (m)	13	27.4	0.3
Volume (m ³ ha ⁻¹)	109.8	375.3	0

Table 1b. Distribution of sites in forest planning input data.

Site class (mineral soil or corresponding peatland)	Proportion of stands (%)
Grove	1
Grove-like heath	14
Fresh heath	46
Dryish heath	26
Dry heath	12

Table 1c. Distribution of species in forest planning input data.

Main tree species	Proportion of stands (%)
Scots pine	57
Norway spruce	22
White birch	10
Pubescent birch	10

done at the given intervals. We assumed that the basic growth models were able to predict the true development of the stands and the models themselves contained no errors. Erroneous developments were produced by adding random growth prediction errors generated by error models developed for this study to the growth prediction. The treatments were assumed to occur at the end of each period. Error-free and erroneous developments, with thinnings and final harvests, were simulated and optimized for each stand. We calculated economic losses caused by growth prediction errors using the results of simulation and optimization.

Phase I: Simulation and Optimization of Error-Free Developments

The stand-level growth models used in the forest planning simulations were basal area and dominant height growth models for Scots pine and Norway spruce (Vuokila and Väliäho 1980), basal area growth models for white and pubescent birch (Mielikäinen 1985), a dominant height growth model for white birch (Oikarinen 1983) and dominant height growth model for pubescent birch (Saramäki 1977). Different models were used for seedling stands, i.e. for stands where the mean height is less than 1.3 m. For pine, a growth model by Huuskonen and Miina (2007) was used. For other tree species, no growth models for seedling stands exist. In these stands, it is just assumed that the seedling stand will reach the 1.3 m mean height at a given age, and at that point the stand has certain “target values” for the needed variables (i.e. dominant height, basal area, mean diameter, volume), depending on the tree species and site quality. The target values are based on the published growth and yield tables. After that growth models for non-seedling stands can be used. Thus, ingrowth is predicted for Norway spruce and birch seedling stands rather than growth.

The growth prediction of basal area and dominant height were calculated separately for every tree species (each stratum) in each stand.

Simulation and optimization were done with the SIMO software (see Kangas and Rasinmäki 2008, Rasinmäki et al. 2009). The aim in the simulation and optimization computations was to determine the optimal harvest schedule for each stand. In order to do this, we first simulated a number of alternative harvest schedules for each stand by predicting the future development using the growth models, and simulating thinnings and final harvests with alternative timings. The simulated thinnings follow the published guidelines (Hyvän metsänhoidon... 2006). In this study, the length of each simulation period was 5 years. Harvests were done at the end of a five-year period. For a description of the simulator, see Fig. 1.

After simulating a number of alternative harvest schedules and future developments for each stand, the optimal harvest schedules were selected from the set of alternative harvest schedules using an optimization algorithm. The optimization task

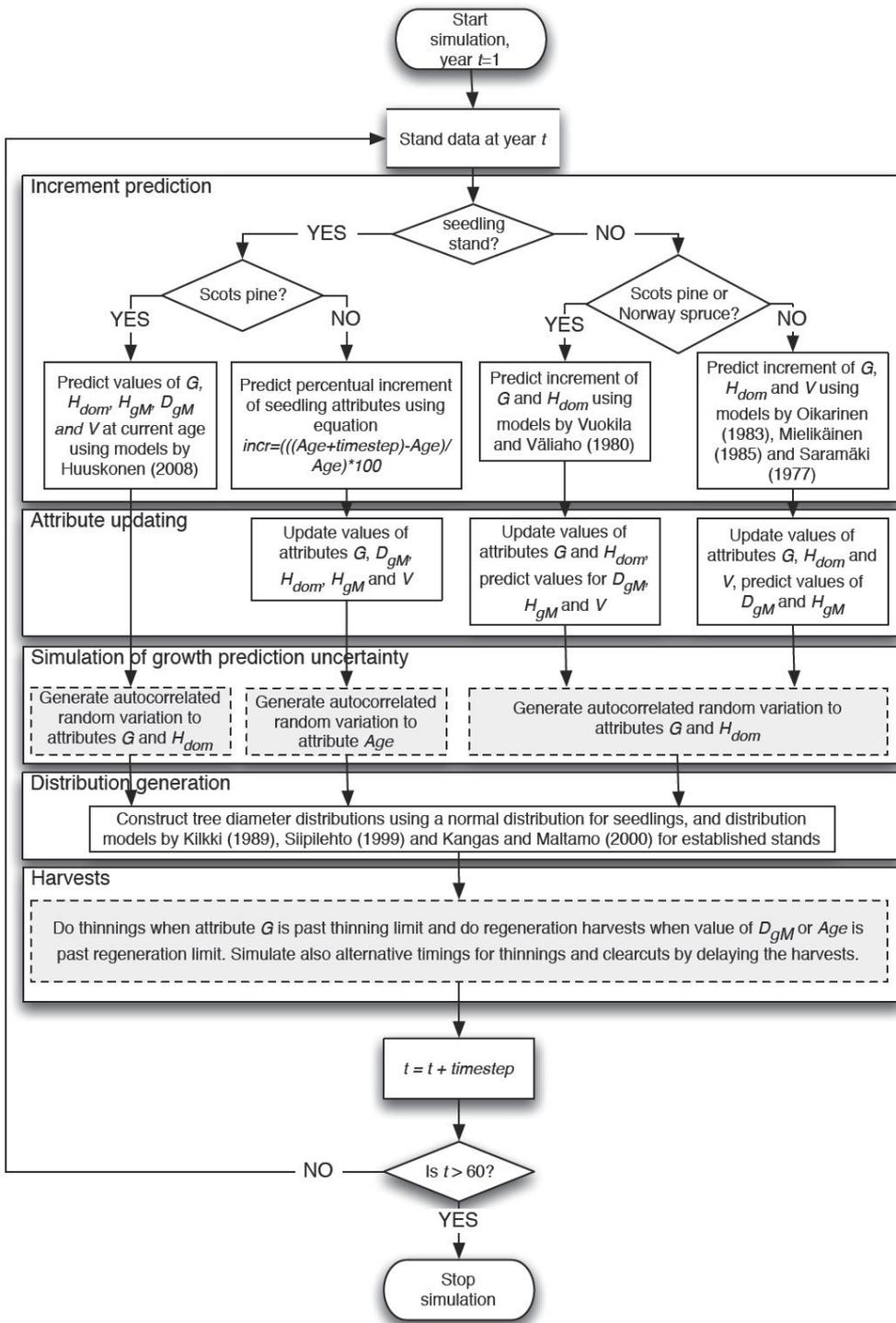


Fig. 1. A flow chart of the simulation logic.

was to maximize the net present value of the stand, which was calculated as

$$\begin{aligned}
 NPV &= NPV_{[0,60]} + \frac{NPV_{60}}{(1+r)^{60}} + \frac{NPV_{60land}}{(1+r)^{60}} \\
 &= \sum_{t=0}^{12} \frac{N_{5t}}{(1+r)^{5t}} + \frac{NPV_{60}}{(1+r)^{60}} + \frac{NPV_{60land}}{(1+r)^{60}}
 \end{aligned}
 \tag{1}$$

where NPV is the net present value of the stand, $NPV_{[0,60]}$ is the net present value at the simulation period (60 years), NPV_{60} and NPV_{60land} are the net present value of the stand and the land at year 60 calculated with the models presented by Pukkala (2005), r is the interest rate used and N_{5t} the net income at time $5t$ (t being the ordinal number of the simulation period). Pukkala's (2005) models predict the expected NPV based on, for instance, the quality of the site (land NPV) and the forest characteristics such as species, age, volume (stand NPV). The interest rate used for discounting future net incomes was 3%. The net incomes consisted of the income from thinnings and final harvests and cost of regeneration after final harvest.

The optimization algorithm was a simple heuristic search algorithm called HERO (see Pukkala and Kangas 1993). The solution space was limited: harvests were done at intervals of five years and alternative timings of harvests were limited by the number of branches in the simulation. Because no constraints were included in the optimization task and whole allowable solution space was gone through, the method always identified the optimal solution from the set of simulated alternatives.

Phase II: Simulation and Optimization of Erroneous Developments

The erroneous growth prediction was generated by the error model, the erroneous growth being the sum of the assumed true growth and the random error component

$$\hat{I}_t = I_t + \varepsilon_t \tag{2}$$

where \hat{I}_t is predicted erroneous growth, I_t is true growth and ε_t is the growth prediction error.

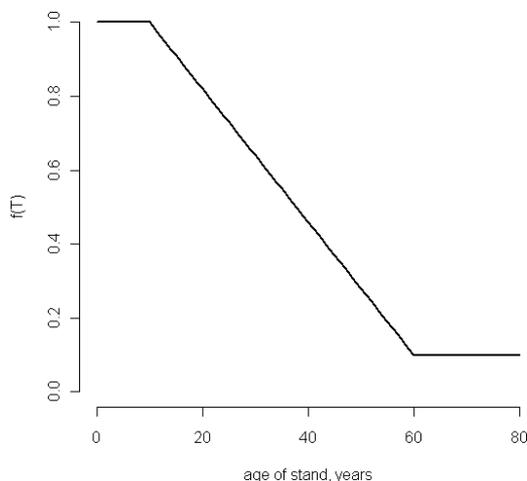


Fig. 2. Function $f(T)$, which limited inter-stand error u .

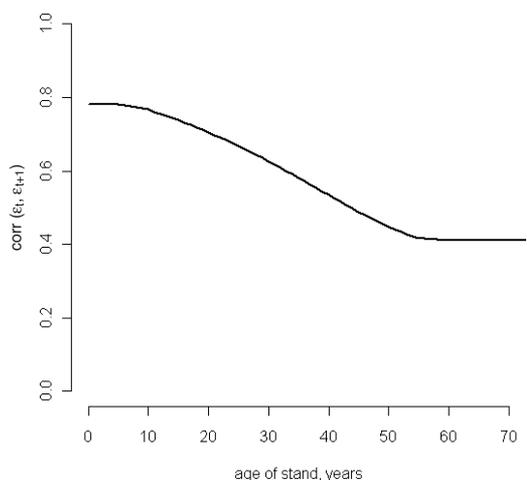


Fig. 3. Correlation between consecutive period growth prediction error of basal area and dominant height as a function of stand age.

The growth prediction error was assumed to consist of stand effect (inter-stand variation) and period effect (intra-stand, or between growth period variation). Thus the growth prediction error in each stand was described as

$$\varepsilon_t = u + e_t \tag{3}$$

where ε_t is total growth prediction error, u is stand effect and e_t is period effect within the stand.

Table 2. Relative RMSE values and biases of basal area (G) and height (H) after a 5-year growing period (Haara and Leskinen 2009) and the variances of growth prediction errors based on these values.

	G (m ² ha ⁻¹)	H _{dom} (m)
RMSE	0.076	0.086
Bias	-0.008	0.007
var(ε)	0.00571	0.00735
var(u)	0.0036	0.00467
var(e)	0.00208	0.00268

Intra-stand error, e_t , was assumed to follow the autoregressive process AR(1), where error e at period $t + 1$ depends on error e at period t so that

$$e_{t+1} = Ae_t + b_{t+1} \tag{4}$$

where A is the correlation coefficient of growth prediction error between consecutive periods and b_{t+1} is the random error. For the random error

$$\begin{aligned} \text{corr}(\varepsilon_t, \varepsilon_{t+1}) &= \frac{\text{cov}(\varepsilon_t, \varepsilon_{t+1})}{sd(\varepsilon_t)sd(\varepsilon_{t+1})} \\ &= \frac{\text{cov}(u_t + e_t, u_{t+1} + e_{t+1})}{sd(\varepsilon_t)sd(\varepsilon_{t+1})} \\ &= \frac{\text{cov}(f_t(T)u, f_{t+1}(T)u) + \text{cov}(e_t, e_{t+1})}{sd(\varepsilon_t)sd(\varepsilon_{t+1})} \\ &= \frac{f_t(T)f_{t+1}(T)\text{var}(u) + A\text{var}(e)}{\sqrt{(f_t(T))^2\text{var}(u) + \text{var}(e)}\sqrt{(f_{t+1}(T))^2\text{var}(u) + \text{var}(e)}} \end{aligned} \tag{6}$$

The variance of growth prediction error (ε_t) was based on the results of Haara and Leskinen (2009, Table 2). The total variance of ε_t was divided into the variance of intra-stand variance ($\text{var}(e_t)$) and inter-stand variance ($\text{var}(u)$), applying the results from Kangas (1999) for volume and assuming the proportions of 0.365 and 0.635 for the intra-stand and the inter-stand variances respectively (Table 2). The value of the correlation coefficient A was set at 0.4. In the study of Kangas (1999), if the variation of volume was divided to an experiment effect and AR(1)-distributed error for intra-stand correlation, the correlation coefficient was 0.37, which supports this choice. As both the inter-stand effect and the intra-stand effect were correlated in time, the resulting total autocorrelation was 0.41–0.78, depending on age of stand (Fig. 3).

component b_t , a new normally distributed value was generated on each time step.

The inter-stand error u was also assumed to follow normal distribution. In addition, it was assumed that the value of u was constant over the whole rotation period and the same for all tree species. Thus, the value of u was generated once at the beginning of the simulation for each stand. When the stand was regenerated, a new normally distributed value of u was generated. The error was further limited by an age-dependent factor $f(T)$, having a value of 1 in young stands and 0.1 in old stands (Fig. 2), resulting in a specific value u_t for each planning period as

$$u_t = f_t(T)u \tag{5}$$

Thus, although u is constant, when the stand becomes older, $f_t(T)$ becomes smaller, so that $\text{cov}(u_t, u_{t+1})$ diminishes and the correlation of the growth prediction error between prediction periods changes as a function of age of stand (Eq. 6).

Again, in Kangas (1999), assuming a model with intra-stand AR(1) autocorrelation but no experiment effect gave as a total autocorrelation 0.72. Thus, the chosen parameters seem reasonable enough based on earlier studies available.

Given these assumptions, the development of the basal area of the stand with growth prediction error was described as

$$\begin{aligned} G_{t+1} &= G_t \left(\left(1 + \frac{P_G}{100} \right) + \varepsilon_{t+1} \right) \\ &= G_t \left(\left(1 + \frac{P_G}{100} \right) + (f_{t+1}(T)u + e_{t+1}) \right) \end{aligned} \tag{7}$$

where G is basal area, P_G is the growth percentage predicted by the growth model and ε_{t+1} is the growth prediction error at period $t + 1$.

Because growth modelling in seedling stands was different for different tree species, modelling of the growth prediction errors was also different for pine seedling stands and the other seedling stands. The growth prediction error for pine seedling stands was calculated in a similar way to non-seedling stands. For other tree species, the growth prediction error was simulated by generating a random error into the age at which the stand reaches the target values, i.e. the ingrowth age. Thus, it is assumed that the stand either reaches the target values at a lower or at a higher age than the growth and yield tables predict (ingrowth is slower or faster than assumed), but the target values as such were assumed deterministic. The random error in stand age was also divided into intra-stand and inter-stand components. The ingrowth age error was generated from a normal distribution with zero mean and variance estimated from Kalliovirta (unpublished) and divided into intra-stand and inter-stand variance applying the results from Kangas (1999), similarly to modelling growth prediction errors for non-seedling stands (Table 3).

Inter-stand growth prediction error (u) was assumed to correlate between forest variables. Correlation was assumed between basal area and dominant height (non-seedling stands and seedling pine stands) and between seedling spruce and birch stratum ages. The assumed correlations were positive as we presumed that the same environmental factors (such as temperature and humidity) have an influence on the growth prediction errors of both variables. However, no data exist to support this choice.

We also assumed that spruce and birch seedling stand ingrowth ages were correlated with non-

Table 3. Relative variance of growth prediction error of age of spruce and birch seedling stands according to Kalliovirta (unpublished) and its subdivision into inter-stand and intra-stand variances $var(u)$ and $var(e)$ respectively.

	T_{spruce}	T_{birch}
$var(\varepsilon)$	0.01254	0.01464
$var(u)$	0.00797	0.0093
$var(e)$	0.00457	0.00534

seedling stand basal area and dominant height. In spruce and birch stands, the growth prediction error was generated to ingrowth age in seedling stands and to basal area and dominant height in non-seedling stands. The correlation between seedling stand ages and basal area or dominant height of non-seedling stands is assumed to be negative because the underestimation of seedling ingrowth age means that the stand reaches the development stage required by the non-seedling growth models earlier than in reality, i.e. it grows faster than in reality. Degree of correlation was defined subjectively, as no data are available. In pine stands, the correlation of basal area or dominant height between seedling and non-seedling stands was 1 because the growth prediction error was added to the same variables the whole rotation period and thus the same value of u could always be used.

Covariances were calculated with variances of the inter-stand growth prediction error u (see Tables 2 and 3) and the correlation matrix

$$corr(u(G), u(H), u(T_{\text{spruce}}), u(T_{\text{birch}}))' = \begin{bmatrix} 1 & 0.5 & -0.8 & -0.8 \\ 0.5 & 1 & -0.8 & -0.8 \\ -0.8 & -0.8 & 1 & 0.8 \\ -0.8 & -0.8 & 0.8 & 1 \end{bmatrix} \quad (8)$$

Erroneous developments were simulated similarly to error-free developments, by adding random errors generated by the error models (Fig. 1). We adopted a Monte Carlo simulation approach and multiple erroneous developments were simulated for each stand and inventory interval combination.

We used 100 simulation iterations, but growth prediction errors led in some stands to such illogical developments that not all iterations could be finished by the SIMO simulator. The total number of such iterations was 10 out of 54 900, i.e. 0.0168% of the simulated developments.

Aborted simulation iterations were rejected from further calculations.

When the erroneous developments were simulated, we assumed that new error-free forest inventory information on the forest attributes was derived at particular intervals. In practice, this meant that the erroneous values were substituted by the error-free values when forest inventory was simulated. Thus, the simulation of growth prediction errors started again from zero after an inventory. Still, when erroneous development led incorrectly to harvest, the impact of this incorrect decision was not cancelled out when forest inventory was simulated. This meant that the erroneous values were substituted by the error-free values in relation to the group of trees remaining after the incorrect decision.

Erroneous developments were optimized similarly to error-free developments.

Phase III: Simulation of Error-Free Developments Forced to Erroneous Harvest Schedules

The optimized harvest schedules based on the reference simulations (without simulated errors) were assumed to represent the real optimal harvest decisions. The optimized harvest schedules that were based on the erroneous simulations (with simulated errors) were assumed to lead to sub-optimal decisions. Thus, for every stand there were as many alternative sub-optimal harvest schedules as there were simulation iterations. To assess the losses due to the growth prediction errors and the resulting incorrect harvest decisions, error-free developments were simulated for each stand with the optimized harvest schedules from the erroneous simulations. This meant that the correct development of the stand was simulated again with the harvests and timings of harvests of each erroneous development.

The inoptimality loss was determined as the difference between the net present value of the error-free harvest schedule and that of the erroneous harvest schedule simulated without the growth prediction errors. The relative inoptimality loss was calculated relative to the net present value of the stand with the formula

$$L_{i\%} = \frac{\text{NPV} - \text{NPV}_i}{\text{NPV}} \times 100\% \quad (9)$$

where $L_{i\%}$ is the relative inoptimality loss in iteration i , NPV is the net present value of error-free development and NPV_i is the net present value in iteration i .

The number of inoptimality loss values for any stand and inventory interval combination was equal to the number of successful simulation iterations.

3 Results

Average absolute and relative RMSE values of basal area and dominant height produced by the used error model were calculated for the 5-year and 10-year periods (Table 4). They are compared with the values from the study by Haara and Leskinen (2009). The relative RMSEs in dominant height were slightly higher than the ones observed by Haara and Leskinen (2009) and those for basal area slightly smaller. Overall, the values used in this study were close enough to the observed error levels for the purpose of the study: the models used in this study represent well enough the accuracy available with the models currently in use.

Average inoptimality losses at different inventory intervals are shown in Table 5 and in Fig. 4. Both absolute and relative inoptimality losses increased when the inventory interval became longer. Absolute losses were approximately

Table 4. Absolute and relative RMSE values of basal area and height when the simulation period is 5 years and 10 years. Upper values are results of this study and values in parentheses results from Haara and Leskinen (2009).

Length of simulation period	RMSE (abs)		RMSE (%)	
	G (m ² ha ⁻¹)	H _{dom} (m)/ H _{gM} (m)	G (m ² ha ⁻¹)	H _{dom} (m)/ H _{gM} (m)
5 years	1.3	1.6	6.5	10
	(1.4)	(1.1)	(7.6)	(8.6)
10 years	2.2	2.3	9.1	12.3
	(2.4)	(1.4)	(11.6)	(10.4)

Table 5. Average, highest and lowest absolute (euro ha⁻¹) and relative (%) inoptimality losses and deviations of losses at different inventory intervals. Results are calculated from stand wise average values.

Length of inventory interval (years)	Average	Inoptimality loss (euro ha ⁻¹)			Average	Inoptimality loss (%)		
		Max value	Min value	Standard deviation		Max value	Min value	Standard deviation
5	229.6	1287.7	0	205.1	3.3	19.9	0	3.5
10	341.3	3581.2	0	481.9	4.5	28.5	0	5.4
15	392.8	3316.3	0	493.8	5.7	40.9	0	7.4
20	685	3931.3	0	659.6	9.1	35.2	0	6.9
30	768.3	4132.8	0	717.1	10.4	40.9	0	8.4
60	859.6	3358.2	0	712.5	11.6	36.3	0	8.3

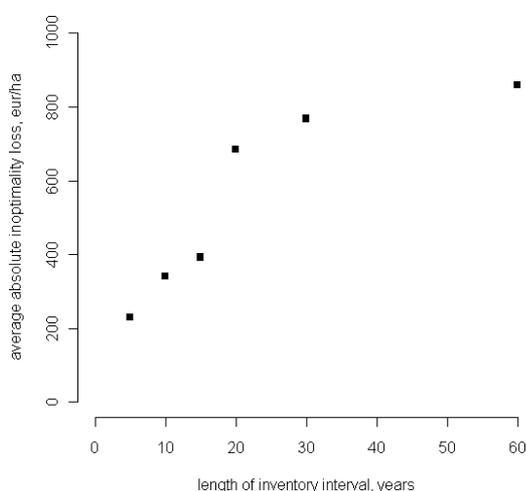


Fig. 4. Average absolute inoptimality losses at different inventory intervals.

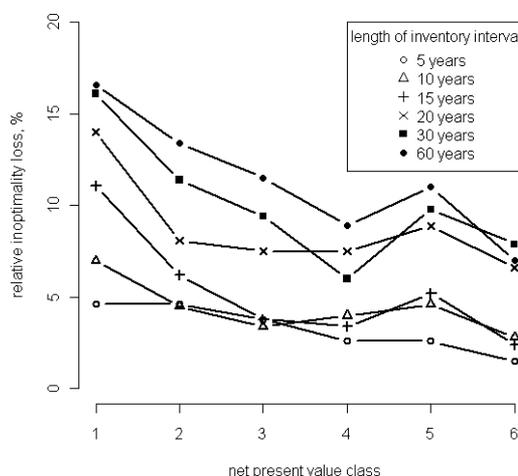


Fig. 5. Average relative inoptimality losses at different inventory intervals by productive value class. 1=0–4000 euro ha⁻¹, 2=4000–6000 euro ha⁻¹, 3=6000–8000 euro ha⁻¹, 4=8000–10000 euro ha⁻¹, 5=10000–12000 euro ha⁻¹, 6=over 12000 euro ha⁻¹.

230–860 euro ha⁻¹ depending on the length of the inventory interval. Relative losses were 3.3–11.6%. Inoptimality losses grew most when the inventory interval was extended from 15 to 20 years. Extending the interval from 20 to 60 years increased losses by only 2.5 percentage units.

Average relative inoptimality losses were lower in stands with a high NPV than in stands with a low NPV (Fig. 5), but in absolute terms the losses were smallest in sites with low NPV. Average relative inoptimality losses differed between develop-

ment and site classes. Losses were mainly bigger in young stands than in mature stands (Fig. 6), but again the absolute losses were quite small in young stands, i.e. from 100 to 400 euro ha⁻¹. Average relative losses were larger in barren stands than in rich stands (Fig. 7), but likewise the absolute losses were smallest in these sites. Inoptimality losses were also different between main tree species but no apparent trend can be seen (Fig. 8).

Fig. 9 illustrates how inventory costs and losses

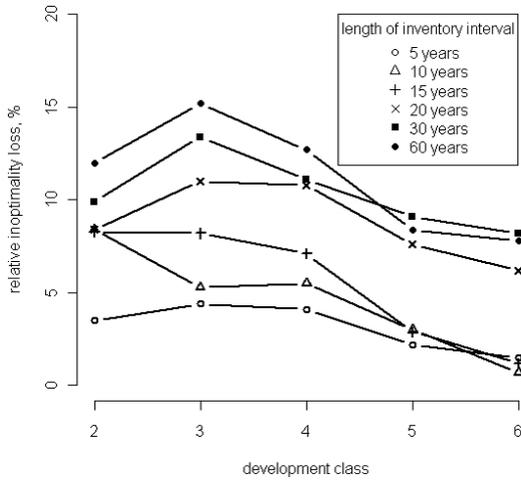


Fig. 6. Average relative inoptimality losses at different inventory intervals by development class. 2=young seedling stand, 3=advanced seedling stand, 4=young thinning stand, 5=advanced thinning stand, 6=mature stand. Four seedling stands, which also contained upper storey (initial development class, e.g., seed tree or shelter tree stand), were left out because only one or two stands belong to all of these types.

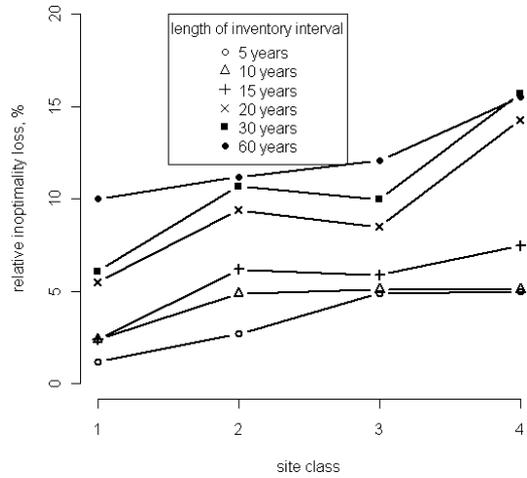


Fig. 7. Average relative inoptimality losses at different inventory intervals by site class. 1=grove-like heath and corresponding peatland, 2=fresh heath and corresponding peatland, 3=dryish heath and corresponding peatland, 4=dry heath and corresponding peatland. One stand whose site type was grove was left out.

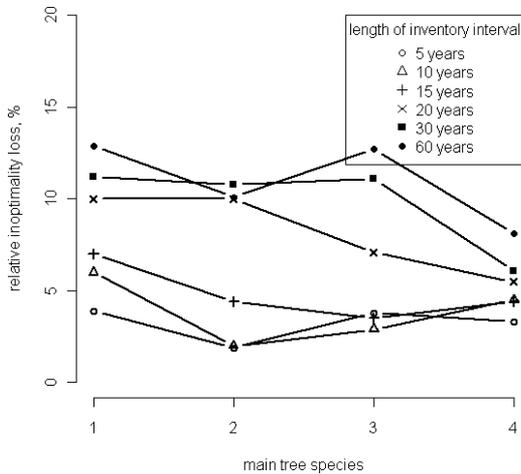


Fig. 8. Average relative inoptimality losses at different inventory intervals by main tree species. 1=Scots pine, 2=Norway spruce, 3=white birch, 4=pubescent birch. Two stands whose main tree species was not Scots pine, Norway spruce or birch were left out.

can affect the optimal inventory interval, i.e. how the costs define the life-span of data. It is assumed that inventory errors and growth prediction errors are independent of each other. It means that the total loss is calculated as the sum of losses due to these two sources, i.e. that inventory errors do not cancel out any of the prediction errors or vice versa. Moreover, we assume that in each inventory the inventory costs and expected losses due to inventory errors are similar, and applied only once when a 60-year interval is assumed and 12 times when a 5-year interval is assumed. Thus, the total costs and losses are the sum of inventory costs, inoptimality losses caused by inventory errors and inoptimality losses of growth prediction errors, all discounted. According to the figure, a five year inventory interval is optimal when inventory costs and losses are less than 40 euro ha⁻¹. If inventory costs and losses are more than 40 euro ha⁻¹, a 15 year inventory interval is optimal until inventory costs and losses are more than 350 euro ha⁻¹ when a 60 year inventory interval is optimal.

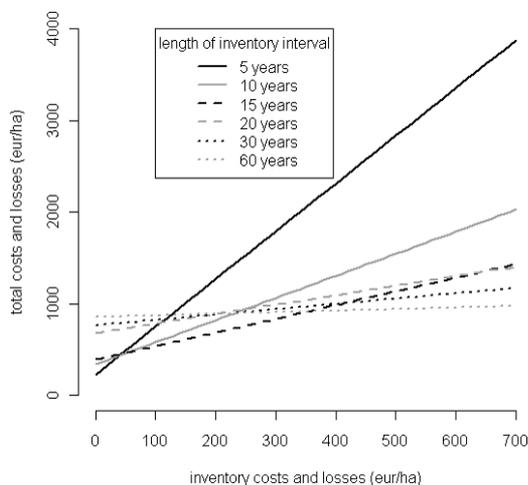


Fig. 9. Effects of inventory costs and losses on optimal inventory interval. The costs and losses of inventory are in the x-axis, the total costs and losses due to inventory and growth prediction errors in the y-axis.

4 Discussion

Inoptimality losses caused by some inventory methods have been on average less than 700 euro ha⁻¹ according to earlier studies (e.g., Holmström et al. 2003, Juntunen 2006, Duvemo et al. 2007). In this study, the growth prediction errors led to losses of 700 euro ha⁻¹ in an inventory period of 30 years. According to Mäkinen et al. (2010), inventory errors caused inoptimality losses of less than 6.4% on average. In this study, inventory intervals whose length was 20 years or more led to higher losses. Based on this kind of comparison only, the maximum information updating period is 15 years; thus, a 20-year inventory interval led to higher relative losses than a new inventory.

However, many other elements should also be taken into account when the profitable inventory interval is adjusted. Because carrying inventory out is more expensive than updating the information using models, inventory costs have a significant effect on the profitable length of the inventory interval. When the losses from the inventory were assumed to be independent of the losses from growth prediction errors, the optimal

inventory interval is 15 years when the inventory costs and losses are more than 40 but less than 350 euros per hectare. This assumption of independency probably overestimates the total losses, as there is evidence that different error sources actually compensate each other to some extent (Eid 2000, Holopainen et al. 2010). However, if the overestimation does not depend on the inventory interval, it does not affect to the relations between the different intervals, i.e. the optimal life-span.

For comparison, the observed costs of the traditional inventory in Finland have varied from 7.9 euro ha⁻¹ (Utterter et al. 2002) to 9.4 euro ha⁻¹ (Utterter et al. 2006), meaning that the actual costs of inventory are far less than the expected losses. Forestry Centres have, however, decided to stop the traditional forest field inventory, and laser scanning based inventory was started in summer 2010. The costs of this new method at stand level are not yet known, but the laser-scanning based inventory in Eid et al. (2004) cost about 11 euro ha⁻¹. The accuracy of the new method is expected to be much better than the previous one, but real large-scale accuracy analysis is still lacking.

In this study, when the inventory interval was 20 years or more, lengthening the interval had only a slight effect on the amount of inoptimality losses. One explanation of this is that the errors in the near future have the most significant effect on the NPV and then on the inoptimality losses. Moreover, the compilation of the initial data set may have an effect on how much extension of the inventory interval affects losses.

According to this study, growth prediction errors have different effects for instance, in different site classes and development classes. However, this finding is based on the assumption that the prediction errors do not depend on site quality. In the future, the error models should probably depend on site and species. For instance, when Haara and Leskinen (2009) modelled the error variance of the basal area, the variance was higher in richer soils since site quality was a significant predictor.

Because the effects of inventory errors are also different in different stands (see also Eid 2000, Mäkinen et al. 2010), relations between errors probably vary between stands. If the quality of the initial information is poor, it is possible that

updating will not make it worse (Kangas 2010). Inventory errors and growth prediction errors may also partly cancel each other out (Holopainen et al. 2010) and relative RMSEs of forest variables may become even smaller when information is updated using growth models (Välimäki 2006). On the other hand, quality of inventory information has an influence on the quality of independent variables in the models (Eid 2003) and on what variables should be selected when models are constructed. Quality of inventory information thus also affects losses caused by growth prediction errors.

According to Mäkinen et al. (2010) relative inoptimality losses caused by inventory errors were low in young stands and high in mature stands. In this study, since the influence of growth prediction errors was the reverse, it may be profitable to carry out an inventory in young stands often and to predict the development of mature stands with growth models.

The growth prediction errors in this study were produced with a model based on real observed errors. However, there are many assumptions that could not be verified with the available data. The autocorrelation between periods is a very important parameter, but very scant information was available. The division of error to stand-effect and period-effect was also based on real data (Kangas 1999), but the data set used was very small. The defined intra-stand autocorrelation parameter can also be supported with the same study. The effect of stand age on correlation was based on an assumed model. Limit term (function $f(T)$) was used in order to avoid errors growing unrealistically large. It was assumed that with the term $f(T)$, limited errors describe real situations better than unlimited errors. All the assumptions related to error model incur uncertainty in the results. However, error model formulated led to quite realistic error levels compared to previous studies (see Table 4). In future, the sensitivity of the results to the various assumptions needs to be examined, and better information on these parameters obtained.

All the error distributions were assumed to be normal. Normal distribution assumption was used as no data are available to analyze the true distribution. This may not be the best option in all cases, but fortunately Mäkinen et al. (2010) noted

that the shape of distribution is not crucial. Also, we had to make assumptions on the correlation of the errors of different variables without any real data to support the choices.

This study was carried out using a stand-level simulator. This approach was selected because returning the tree-level variables to the error-free development path (depending on the correct and incorrect decisions carried out in the period) after each inventory would be computationally problematic. Moreover, it was assumed that the model describes the true growth, and erroneous growth paths vary around it. In reality the erroneous paths are described by the model, which is an expected value of possible true paths. However, since the latter procedure would be more complicated to compute, the simpler approach was adopted.

This study focused on the growth prediction errors and their economic effects. For purposes of forest planning decision making, it will be important to know the total cost of information used. Total cost consists of costs and losses of the inventory and of updating information as well as losses caused by uncertainty of such things as future timber prices and silvicultural treatment costs. If an estimate of the total cost of various decision and activity alternatives existed, it would be possible for the decision maker to decide when and by what methods the new information should be collected, how much the decision maker would pay for information and how great a loss decisions could lead to. In future, the various sources of uncertainty and their relations to each other should be taken into account.

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.
- Burkhart, H.E., Stuck, R.D., Leuschner, W.A. & Reynolds, M.R. 1978. Allocating inventory resources for multiple-use planning. *Canadian Journal of Forest Research* 8: 100–110.
- Duvemo, K., Barth, A. & Wallerman, J. 2007. 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(2): 89–100.
- 2003. Model testing by means of cost-plus-loss analyses. In: Amaro, A., Reed, D. & Soares, P. (eds.) *Modelling Forest Systems*. CABI Publishing, p. 295–305.
- , Gobakken, T. & Næsset, E. 2004. Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analyses. *Scandinavian Journal of Forest Research* 19: 512–523.
- Gustavsen, H.G. 1977. Valtakunnalliset kuutiokasvuyhtälöt. *Folia Forestalia* 331. (In Finnish).
- Haara, A. & Korhonen, K.T. 2004. Toimenpide-ehdotusten simulointi laskennallisesti ajantasaistetusta kuvioaineistosta. *Metsätieteen aikakauskirja* 2/2004: 157–173. (In Finnish).
- & Leskinen, P. 2009. The assessment of the uncertainty of updated stand-level inventory data. *Silva Fennica* 43(1): 87–112.
- Hamilton, D.A. 1978. Specifying precision in natural resource inventories. In: *Integrated inventories of renewable resources: proceeding of the workshop*. USDA Forest Service, General Technical Report RM-55: 276–281.
- Holmström, H., Kallur, H. & Ståhl, G. 2003. Cost-plus-loss analyses of forest inventory strategies based on *k*NN-assigned reference sample plot data. *Silva Fennica* 37(3): 381–398.
- Holopainen, M. & Talvitie, M. 2006. Effect of data acquisition accuracy on timing of stand harvests and expected net present value. *Silva Fennica* 40(3): 531–543.
- , Mäkinen, A., Rasinmäki, J., Hyytiäinen, K., Bayazidi, S., & Pietilä, I. 2010. Comparison of various sources of uncertainty in stand-level net present value estimates. *Forest Policy and Economics* 12: 377–386.
- Huuskonen, S. 2008. Nuorten männiköiden kehitys – taimikonhoito ja ensiharvennus. *Dissertationes Forestales* 62. (In Finnish).
- & Miina, J. 2007. Stand-level growth models for young Scots pine stands in Finland. *Forest Ecology and Management* 241: 49–61.
- 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.
- , Ahtikoski, A., Siitonen, J., Sievänen, R. & Liski, J. 2005. Applying the MOTTI simulator to analyse the effects of alternative management schedules on timber and non-timber production. *Forest Ecology and Management* 207: 5–18.
- Hyvän metsänhoidon suositukset. 2006. Metsätalouden kehittämiskeskus Tapio. Metsäkustannus Oy. (In Finnish).
- Hyvönen, P. & Korhonen, K.T. 2003. Metsävaratiedon jatkuva ajantasaistus yksityismetsissä. *Metsätieteen aikakauskirja* 2/2003: 83–96. (In Finnish).
- Islam, M.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(1): 71–85.
- Juntunen, R. 2006. Puustotiedon laadun vaikutus metsänkäsittelyn optimoinnin tuloksiin – UPM Metsän laserkeilausaineiston ja kuviotiedon vertailu. Pro gradu -työ. Helsingin yliopisto, metsäekonomian laitos. (In Finnish).
- Kalliovirta, J. (unpublished). Age models for Scots pine, Norway spruce and pine.
- Kangas, A. 1997. On the prediction bias and variance in long-term growth projections. *Forest Ecology and Management* 96: 207–216.
- 1999. Methods for assessing uncertainty of growth and yield predictions. *Canadian Journal of Forest Research* 29: 1357–1364.
- 2010. Value of forest information. *European Journal of Forest Research* 129: 863–874.
- & Maltamo, M. 2000. Percentile based basal area diameter distribution models for Scots pine, Norway spruce and birch species. *Silva Fennica* 34(4): 371–380.
- & Rasinmäki, J. (eds.) 2008. SIMO – adaptable simulation and optimization for forest management planning. University of Helsinki, Department of Forest Resource Management, Publications 41.
- Kilkki, P. 1989. Use of the Weibull function in estimating the basal area dbh-distribution. *Silva Fennica* 23(4): 311–318.
- Mäkinen, A., Kangas, A., Välimäki, E., Kalliovirta, J., & Rasinmäki, J. 2008. Comparison of tree-wise and stand-wise forest simulators with the aid of quantile regression. *Forest Ecology & Management* 255: 2709–2717.
- , Kangas, A. & Mehtätalo, L. 2010. Correlations,

- distributions and trends of forest inventory errors and their effects on forest planning. *Canadian Journal of Forest Research* 40: 1386–1396.
- Mielikäinen, K. 1985. Koivusekoituksen vaikutus kuusikon rakenteeseen ja kehitykseen. *Communications Instituti Forestalis Fenniae* 133. (In Finnish).
- Nyysönen, A. & Mielikäinen, K. 1978. Metsikön kasvun arviointi. *Acta Forestalia Fennica* 163. (In Finnish).
- Oikarinen, M. 1983. Etelä-Suomen viljeltyjen rauduskoivikoiden kasvatusmallit. *Communications Instituti Forestalis Fenniae* 113. (In Finnish).
- Ojansuu, R., Halinen, M. & Härkönen, K. 2002. Metsätalouden suunnittelujärjestelmän virhelähteet männyn ensiharvennuskypsytyden määrittämisessä. *Metsätieteen aikakauskirja* 3/2002: 441–457. (In Finnish).
- Pukkala T. 2004. Monikäytön suunnitteluohjelma Monsu. Ohjelmiston toiminta ja käyttö. University of Joensuu, Faculty of Forestry. (In Finnish).
- 2005. Metsikön tuottoarvon ennustemallit kivennäismaan männiköille, kuusikoille ja rauduskoivikoille. *Metsätieteen aikakauskirja* 3/2005: 311–322. (In Finnish).
- & Kangas, J. 1993. A heuristic optimization method for forest planning and decision making. *Scandinavian Journal of Forest Research* 8: 560–570.
- Rasinmäki, J., Mäkinen, A. & Kalliovirta, J. 2009. SIMO: an adaptable simulation framework for multiscale forest resource data. *Computers and Electronics in Agriculture* 66: 76–84.
- Saramäki, J. 1977. Ojitettujen turvemaiden hieskoivikoiden kehitys Kainuussa ja Pohjanmaalla. *Metsäntutkimuslaitoksen julkaisuja* 91(2). (In Finnish).
- Siipilehto, J. 1999. Improving the accuracy of predicted basal-area diameter distribution in advanced stands by determining stem number. *Silva Fennica* 33(4): 281–301.
- Siitonen, M. 1996. Introduction to the MELA system and its use. In: Päivinen, R., Roihuvuo, L., Siitonen, M. (eds.). *Large-scale forestry scenario models: experiences and requirements. Proceedings of the International Seminar and Summer School, 15–22 June 1995, Joensuu, Finland. EFI Proceedings* 5. p. 171–174.
- Ståhl, G. 1994. Optimal forest inventory intensities under deterministic and stochastic stumpage value assumptions. *Scandinavian Journal of Forest Research* 9: 405–412.
- & Holm, S. 1994. The combined effect of inventory errors and growth prediction errors on estimations of future forestry states. In: Ståhl, G. *Optimizing the utility of forest inventory activities. Swedish University of Agricultural Sciences, Department of Biometry and Forest Management, Umeå, Report* 27.
- , Carlson, D. & Bondesson, L. 1994. A method to determine optimal stand data acquisition policies. *Forest Science* 40: 630–649.
- Utterä, J., Hiltunen, J., Rissanen, P., Anttila, P. & Hyvönen, P. 2002. Uudet kuvioittaisen arvioinnin menetelmät – arvio soveltuvuudesta yksityismaiden metsäsuunnitteluun. *Metsätieteen aikakauskirja* 3/2002: 523–531.
- , Anttila, P., Suvanto, A. & Maltamo, M. 2006. Yksityismetsien metsävaratiedon keruuseen soveltuvilla kaukokartoitusmenetelmillä estimoitujen puustotunnusten luotettavuus. *Metsätieteen aikakauskirja* 4/2006: 507–519.
- Välimäki, E. 2006. Kasvumallien toiminnan validointi ylitiheissä metsiköissä. Pro gradu -työ. Helsingin yliopisto, metsävarojen käytön laitos. (In Finnish).
- Vuokila, Y. & Väliaho, H. 1980. Viljeltyjen havumetsiköiden kasvatusmallit. *Metsäntutkimuslaitoksen julkaisuja* 99(2). (In Finnish).

Total of 48 references