SILVA FENNICA
Silva Fennica 43(1) research articles
www.metla.fi/silvafennica · ISSN 0037-5330
The Finnish Society of Forest Science · The Finnish Forest Research Institute

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

Md. Nurul Islam, Mikko Kurttila, Lauri Mehtätalo and Arto Haara

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.

Accurate inventory data are required for ensuring optimal net return on investment from the forest. Erroneous data can lead to the formulation of a non-optimal plan that can cause inoptimality losses. Little is known of the effect of using erroneous stand inventory data in preparing holding-level forest plans. This study reports on an approach for analyzing such inoptimality losses. Furthermore, inoptimality losses caused by measurement errors in the basal area of the dominated tree species were investigated in a case study. Based on the inventory data including routine measurements by 67 measurers, four measurer groups were created with different measurement error profiles for the basal area of the dominated tree species. This was followed by measurement error simulations for each group and by adding these to the accurate control inventory data to create erroneous data of different error profiles. Three different forest plans were then constructed by using erroneous data of each group. The plans were then analyzed and compared with plans based on correct data. The effect of measurement errors on the net present value from the whole planning period, and on the amount of remaining growing stock at the end of planning period, were analyzed and utilized in calculating the inoptimality losses. It was concluded that even errors involving dominated tree species can cause significant changes in the holding-level forest plans.

Keywords dominated tree species, erroneous inventory data, forest plan, inoptimality loss
Addresses *Islam* & *Haara*: University of Joensuu, Faculty of Forest Sciences, FI-80101 Joensuu, Finland; *Kurttila*: Finnish Forest Research Institute, Joensuu Research Unit, FI-80101 Joensuu, Finland; *Mehtätalo*: University of Helsinki, Dept. of Forest Resource Management, FI-00014 University of Helsinki, Finland
E-mail nurul.islam@joensuu.fi; arto.haara@joensuu.fi; mikko.kurttila@metla.fi; lauri.mehtatalo@helsinki.fi
Received 21 April 2008 **Revised** 13 November 2008 **Accepted** 18 December 2008
Available at http://www.metla.fi/silvafennica/full/sf43/sf431071.pdf

1 Introduction

The task of forest planning is to find a combination of treatment schedules for stands that best meets the forest owner's multiple goals expressed at the level of the forest holding. Due to the huge number of different combinations of treatment schedules even in a small forest holding, numerical optimization methods need to be used to define efficient forest plans. The selection of treatment schedules for stands during optimization is based on the outcomes of simulated treatment schedules and on the objectives of the forest owner. Errors in initial inventory data can result in errors in the outcomes and also in the simulation of treatment schedules. These errors can lead to the drawing up of a non-optimal forest plan at the holding-level.

Inventory data can be collected through field sampling either by maintaining statistical principles or by applying subjective sampling procedures. Remote sensing methods, such as laser scanning, are expected to become a commonly used inventory method in the near-future (e.g. Næsset et al. 2004). However, the stand-wise inventory method continues to play an important role in data collection for detailed management planning, especially in non-industrial private forests (Koivuniemi and Korhonen 2006). In Finland, for example, every year about one million hectares of non-industrial private forests (Tapion vuositilastot... 2005) are inventoried by applying this method. This method is a type of sampling method where the sampling points are selected subjectively – according to the measurers' choice. The method is conducted in two steps. In the first step, the stands (or compartments, which are geographically contiguous parcels of forest land whose site type, species composition, and tree age are homogenous) are delineated on the map. In the second step, the measurer subjectively selects some sampling points or sometimes only one sampling point per stand. The measurer takes the measurements for various forest variables, such as basal area, diameter and height of basal area median tree, and mean age for each tree species and canopy layer in the stand. Some of these variables may be taken only visually, without using any instruments. If the number of sampling points is low, the results of the measurements can be adjusted according to the measurers' conception of the stand (e.g. Saari and Kangas 2005). The data obtained are then used for simulating alternative treatment schedules for each stand by applying simulation models that mitigate different management operations and natural processes, such as growth and mortality. Furthermore, stand and holding-level management plans are drawn up by selecting the schedules that best meet the forest owner's objectives.

The accuracy of the stand-wise inventory method is considered to be fairly low (Maltamo et al. 2003, Haara 2005). Since the sampling points are selected subjectively and the measurements are not always taken very accurately, the results may become biased. The reasons for this include that the measurers do not visit every part of the stand, they may purposely underestimate the values of some variables (Haara and Korhonen 2004; Haara 2005; Saari and Kangas 2005) so that the forest owner's actual timber yield would not fall short of the measurers' estimate, their measurement practices may not be appropriate for the stand (Saari and Kangas 2005), and they may systematically select biased locations for their representative plots. The results of several studies have demonstrated that stand basal area in closed mature stands is often underestimated (e.g. Poso 1983, Laasasenaho and Päivinen 1986, Pussinen 1992, Ståhl 1992, Anttila 2002, Haara and Korhonen 2004). Kangas et al. (2004) studied the accuracy of the stand-wise inventory method and found that some variables (e.g. basal area per hectare, diameter and height of basal area median tree, and number of stems per hectare) in older stands were underestimated while some other variables (e.g. basal area and number of stems per hectare) in younger stands were overestimated.

The accuracy of inventory data affects the quality of management decisions (Duvemo and Lämås 2006) and the effects on forest owners may be twofold. Firstly, stand-level and holding-level yield outcomes (e.g. harvestable timber volume) are not predicted correctly, but the treatment recommendations for the stands are correct, i.e. similar to those if correctly predicted outcomes were used. As a result, the 'correct' outcome from the stands will be realized when the treatments are carried out. Secondly, stand-level outcomes are not predicted correctly and also the treatment recommendations for some stands (e.g. timing of thinning) are not correct. In both cases, errors affect the holding-level outcomes of the objective variables, and the owner may end up selecting a non-optimal or even an unfeasible holding-level forest plan. For example, an error in the inventory data can lead to the early final cutting of a stand (Eid 2000). Holopainen and Talvitie (2006) also studied the effect of errors in inventory data on the expected net present value (NPV). They found that the NPV losses were different for different types of inventory methods. The NPV losses that result from not following the optimal treatments of the stands are called inoptimality losses (e.g. Jacobsson 1986). Inoptimality loss is calculated from the difference between the correct NPV and the NPV of the erroneous treatment schedules implemented in the correct data.

In a practical inventory, measurers sometimes concentrate on the dominant tree species and neglect the dominated tree species. The reason for this may be that the measurers think that a small proportion of the other tree species in a stand will not significantly affect the planning calculations. For example, according to Haara and Korhonen (2004), relative measurement errors were at their highest in mixed stands, where the proportion of dominant tree species was below 70% of the stand basal area. However, the effect of totally excluding dominated tree species from measurements can lead to delayed thinning due to underestimation of the basal area. If the basal area of the dominated tree species is merged in the basal area of the dominant tree species, the differences in the timber-prices of different tree species may result in over- or underestimated cutting income estimates. At the holding-level, the effects can be significant due to holding-level goals and constraints. Moreover, the resulting plan may include incorrect timing decisions of forest management operations in some other stands.

The magnitude and effect of errors in inventory data have so far been studied mainly at stand-level (e.g. Poso 1983, Laasasenaho and Päivinen 1986, Pussinen 1992, Ståhl 1992, Eid 2000, 2001, Anttila 2002, Holmström et al. 2003, Eid et al. 2004, Haara and Korhonen 2004, Kangas et al. 2004, Holopainen and Talvitie 2006, Duvemo et al. 2007). At holding-level the effect has been studied in a limited scale. For example, Sprängare (1975) has studied the effect on selecting stands for final harvest and found higher harvest levels than optimal due to erroneous data. Eid (1991, 1993) studied the effects of systematic and random errors in inventory data at the holding level. He found that errors in the site index estimations had the most significant effects on the contents of the plans and particularly for the treatments of young stands, and that the relative deviations in calculated NPVs were smaller than the uncorrelated errors that were generated to the inventory data. The effect of errors on holding-level forest plans, that often include different objectives and constraints related to factors such as remaining growing stock at the end of the planning period and/or even timber-flow from the forest, has, however, not received enough attention.

The objective of this study is to present an approach for estimating the inoptimality losses of using erroneous data in holding-level planning calculations. The approach is illustrated through a case study, where the effects of errors in the basal area of dominated tree species are investigated (In this study, dominated tree species aggregately make up less than 50% of total basal area in a stand). The inventory data used in the case study were collected by applying stand-wise inventory method.

2 Analyzing the Effect of Erroneous Stand-Level Data on Holding-Level Forest Plans

The end result of a forest planning process is usually a holding-level forest plan. It includes the planned treatments and their timings for the individual stands of the forest holding. It also includes holding-level estimates about characteristics of interest, e.g. net present value (NPV) of incomes and remaining growing stock at the end of planning period. A common approach to construct such a forest plan is to apply a three-stage procedure, where 1) alternative feasible treatment schedules are simulated for each stand of the forest holding and characteristics of interest are estimated for each treatment schedule. This is followed by 2) the formulation of the optimization problem describing the forest owner's management objectives. Finally, 3) the holding-level forest plan is constructed by selecting the combination of treatment schedules that best satisfies the objectives formulated at stage (2). In practice, the procedure would include iteration of these stages, i.e. the use of interactive optimization where the final plan is searched through iteratively examining the production possibilities of the forest holding (Pykäläinen 2000).

Errors in inventory data may have three different effects on the simulation of the treatment schedules: 1) the outcome of a forest operation (e.g. net incomes) may be incorrectly estimated, 2) a realistic treatment schedule may be missing among the simulated alternative schedules, and 3) a possibly unrealistic schedule may be simulated. Having the outcome of a schedule being estimated incorrectly is not a problem itself, because the correct outcome will be seen after the harvest operation is conducted. However, errors in the variables utilized in the formulation of the optimization problem may cause a non-optimal schedule to be selected (e.g. Kangas and Kangas 1999). The missing realistic schedule makes the search space of the optimal plan smaller, thus causing inoptimality losses. The unrealistic schedules may include, for example, early thinnings and regeneration cuttings, and this also affects the results of optimization. In this study, however, it is assumed that any early thinning schedules would not be conducted in practice, because the error would be noticed at the operational planning phase and the treatment would be postponed to the future.

Errors in data may also have an impact on the formulation of the optimization problem if the erroneous information on the current state of the holding is used in formulating planning problems. For example, the remaining growing stock volume at the end of planning period may be constrained to be not less than the current, implying that errors in the current growing stock volume will have an effect on determining the value of growing stock constraint.

With a holding-level forest plan, errors in the simulated treatment schedule also affect other stands in the holding through holding-level constraints and objectives. This effect may result in solutions that have lower or higher value for the objective function than possible, and are below or outside the production possibility frontier. As an example, consider a plan where the objective is to maximize NPV of the cutting income subject to a strict lower limit of the remaining growing stock. If the remaining growing stock of one stand is underestimated, this measurement error allows less harvest in other stands. As a result, the true remaining growing stock will be higher than demanded. Thus, the plan would be below the real production possibility frontier, thereby implying ineffective utilization of the production possibilities. On the other hand, having the remaining growing stock of one stand overestimated because of the absence of possible thinning schedules due to underestimation of stand basal area would allow more cuttings in the other stands. This will lead to the growing stock to be lower than the limit which in fact implies that the plan prescribes over-exploitation of the production possibility.

This study proposes an approach to determine holding-level inoptimality losses caused by the errors in the inventory data. The approach includes the following three stages:

A. Determine the realized NPV of a forest plan when measurements include errors, NPV_{err} (NPV for the plan based on erroneous data).
B. Determine the NPV that would be reached if measurements were correct, NPV_{corr} (NPV for the plan based on correct data).
C. Determine the inoptimality loss as difference $NPV_{corr} - NPV_{err}$

Stage A includes the following four steps

A1. Treatment schedules are simulated for each stand using erroneous data
A2. A forest planning problem is specified and a forest plan is obtained as a solution to the problem.
A3. The treatment schedules of the plan are applied to the correct inventory data. The unrealistic schedules are replaced with the most similar realistic schedules. For example, early thinnings are postponed to the earliest realistic period because this would happen also in practice. Timing of the delayed treatments is not changed.
A4. The realized outcomes (i.e., NPV_{err}, remaining growing stock, amount of periodical cuttings) of the forest plan are computed.

Stage B includes three steps. If the aim is to maximize the NPV with no constraint, the steps are as follows:

B1. Treatment schedules are simulated for each stand using correct data.
B2. A plan with the same planning problem as in step A2 is constructed
B3. The outcomes (i.e., NPV_{corr}, remaining growing stock etc.) of the plan are computed.

In stage C, the difference between the NPV obtained in Step A4 is compared with the NPV obtained in step B3.

If the planning problem includes holding-level constraint(s), the situation becomes more complicated. This is because the measurement errors may cause the constraints to be violated, which in turn can lead to even higher NPV than is the case with the optimal plan of stage B2. Thus, the violation of the constraints should also be taken into account in computing NPV_{corr}. In our approach, stages A and C were retained as they were before, but a new procedure to compute the optimal plan in stage B was introduced. The new procedures for steps B1 and B2 are as follows:

B1* Treatment schedules for the correct data are simulated ensuring that the treatment schedules (especially delayed thinning schedules) that were selected for the plan based on erroneous data are also available.
B2* A planning problem is formulated by adjusting the constraints of the problem specified in step A2, given that the obtained plan would become a feasible solution to the problem. For a plan having only the remaining growing stock constraint, this implies using the realized remaining growing stock of step A4 as the value of the remaining growing stock constraint. For a plan with additional even-flow constraints, in addition to the above-mentioned remaining growing stock constraint the amount of periodical cuttings from each sub-period needs to be equal to or greater than the respective periodical amount obtained in step A4.

3 Materials and Methods of the Case Study

3.1 Data

3.1.1 Data for Calculating Measurement Error

The data used in the case study were taken from Haara and Korhonen (2004). The data included both erroneous and correct data from a total of 1158 stands located in Eastern Finland. The erroneous data were collected through the regular stand-wise inventory method done by 67 different measurers. The correct data were collected by means of the checking inventory method which was carried out by measuring all trees within a systematic network of circular sample plots within each stand. The numbers and sizes of the plots within a stand were determined on the basis of the size of the stand, the development class of the trees in the stand, and the tree species composition of the stand. The average size of the plot was 177 m^2 and, on average, 6.2 plots were measured per stand (for more details, see Haara and Korhonen 2004). We ignored the sampling errors of the correct data which was 3.6 percent units for the stand basal area (Haara and Korhonen 2004) and assumed it as error free. Because the error in the basal area of dominated tree species was considered in the case study, 112 stands were excluded from the study due to the absence of dominated tree species in both correct data and in erroneous data. Thus, the data from 1046 stands were used as material for defining the extent of measurement errors in the basal area of dominated tree species. The data used did not include stands where the mean diameter was less than 5 cm.

3.1.2 Forest Planning Datasets

A sample of 118 stands from among the above stands was randomly selected for planning purposes. The selected stands were considered to constitute a forest holding in order to enable analysis at the holding-level. Data for the selected 118 stands were called Dataset #1 (Table 1).

Table 1. Characteristics of Dataset #1 and Dataset #2 based on correct data.

Variable	Dataset #1	Dataset #2
Total area (ha)	247.9	103.2
Current growing stock (m^3)	41742	21882
Current growing stock (m^3/ha)	168.4	212.0
Growth (m^3/ha/yr)	5.9	5.7
Species composition (proportion of growing stock volume)		
Pine (*Pinus sylvestris*):	53%	30%
Spruce (*Picea abies*):	39%	60%
Birch (*Betula* spp):	7%	9%
Other broadleaved trees	1%	1%
Age class distribution (proportion of total area, ha)		
0–19 years	0%	0%
20–39 years	39%	30%
40–59 years	16%	13%
60 and above	45%	57%

To compare the result of Dataset #1, another dataset of 59 stands from a real forest holding within the study area of Haara and Korhonen (2004), was taken and called as Dataset #2. It was based on similar data collection and preprocessing procedures than Dataset #1 except that no sample but all stands in the holding were selected. Dataset #2 was characteristically different from Dataset #1 in that it contained higher proportions of spruce and stands over 60 years of age (Table 1).

3.2 Measurement Errors in the Basal Area of Dominated Tree Species

The measurer-specific bias, relative bias (bias%), and variance of measurement error of the dominated tree species' basal areas were computed by using Equations 1, 2 and 3, respectively. A one-sample t-test (Ary and Jacobs 1976) was applied to test whether the measurers' bias was significantly different from zero.

The bias, bias% and variance were calculated as follows:

$$Bias = \frac{\sum_{i=1}^{n} Y_i}{n} \quad (1)$$

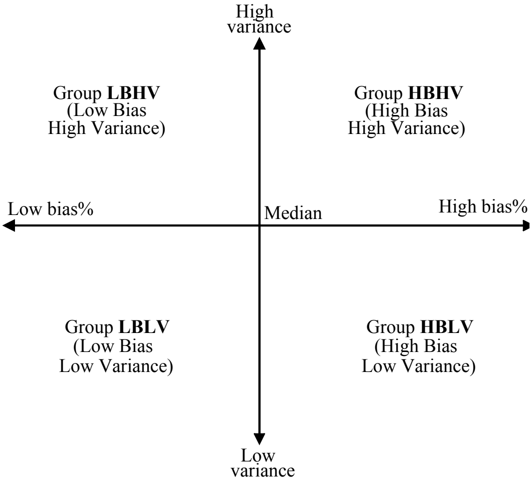

Fig. 1. Categorization of measurers based on the bias% and variance of measurement errors in measuring the basal area of dominated tree species. Bias% and variance crossed at their medians.

$$Relative\ bias\ (\text{bias\%}) = \frac{bias}{mean\ of\ basal\ area\ in\ correct\ data} \times 100 \quad (2)$$

$$Variance(Y) = \frac{\sum (bias - Y_i)^2}{n-1} \quad (3)$$

where n is the number of observations for the measurer, *mean of basal area in correct data* includes only the dominated tree species and Y_i = *Measured value*$_i$ – *Correct value*$_i$.

3.3 Simulation of Erroneous Data

The measurers were categorized into four groups according to the extent of the relative bias and variance in the basal area of dominated tree species. The limit values of the group categories were placed at the medians of bias% and variance (Fig. 1). The groups were named as 'LBLV'(Low Bias and Low Variance), 'LBHV'(Low Bias and High Variance), 'HBLV'(High Bias and Low Variance), and 'HBHV' (High Bias and High Variance) (Fig. 1).

Table 2. Description of the objectives and constraints of the three forest plans.

Plan	Objective function
Plan #1	Maximize NPV
Plan #2	Maximize NPV s.t. Remaining growing stock at the end of planning period > Current growing stock
Plan #3	Maximize NPV s.t. Remaining growing stock at the end of planning period > Current growing stock Even flow of cuttings in each sub-period

To study the effect of measurement error on the quality of a forest plan, four erroneous data were simulated, based on the grouping of measurers. In simulations, the measurement errors were simulated and added to the correct value of the basal area of dominated tree species. All other variables, including the basal area of dominant tree species, were kept as they were in the correct data. The errors were simulated by applying a nearest neighbor method (NN method) (Härdle 1989, Altman 1992, Haara et al. 1997), where the error of the nearest neighbour was added to the correct value of the target stand. In simulation of erroneous data, the neighbours were defined as such stands of the same measurer group which had the same dominant tree species as the target stand had. The nearest neighbour was the one that yielded the smallest sum of absolute differences in the standardized basal area and dominated tree species proportion between the neighbor and target stand. The relative error in the dominated basal area of the nearest neighbour was multiplied by the correct value of dominated basal area to get the absolute error for the target stand. Erroneous data for Dataset #2 were simulated similarly, using neighbors from among the same reference stands as with Dataset #1. Thus, we got data for four groups LBLV, LBHV, HBLV and HBHV for both datasets, along with the data that include the correct values.

3. 4 Planning Calculations

Separate forest plans were formulated based on correct data and erroneous data for the four measurer groups of two datasets. The MONSU forest planning software (Pukkala 2004) was used for the simulation of treatment schedules and optimization calculations for creating alternative forest plans. The planning period was fifteen years, consisting of three sub-periods of five years each.

The first step was to simulate treatment schedules for each stand. In the simulation of treatment schedules, thinning was always simulated when the basal area exceeded the thinning limit based on the stand basal area and dominant height, specified in Finnish silvicultural recommendations (Hyvän metsänhoidon suositukset 2006). Final cutting was simulated when the stand fulfilled the conditions of the regeneration criteria based on diameter and age of the stand set by silvicultural recommendations. In addition, delayed final cuttings were simulated, including the ‘No final cutting’ option, where the final cutting is delayed to happen after the planning period. The regeneration methods and the subsequent treatments were not varied within one stand.

In the next step, three different plans designated as ‘Plan #1’, ‘Plan #2’ and ‘Plan #3’ for each group of both datasets were formulated by using typical linear programming (LP) problem formulation which ensured global optimality of the solution. The objective variable was the NPV of harvest incomes which was maximized subject to different constraints such as remaining growing stock at the end of planning period and even-flow of cuttings (Table 2). In total, 30 plans (6 plans from the correct data of two datasets and 24 plans from the erroneous data of the four groups for the two datasets) were prepared.

NPV was computed as

$$\sum_{j=1}^{n}\sum_{i=1}^{t_j} c_{ij} x_{ij} \tag{4}$$

and the growing stock was computed as

$$\sum_{j=1}^{n}\sum_{i=1}^{t_j} g_{ij} x_{ij} \tag{5}$$

The even flow of cuttings was implemented by constraining the difference between cuttings of the consecutive sub-periods to be zero,

$$\left[\sum_{j=1}^{n}\sum_{i=1}^{t_j} r_{ijp'} x_{ij} - \sum_{j=1}^{n}\sum_{i=1}^{t_j} r_{ijp} x_{ij}\right] = 0 \tag{6}$$

The area of a particular stand *j* was fixed by

$$\sum_{i=1}^{t_j} x_{ij} = A_j,\ j = 1,2,3.........n \tag{7}$$

so that the area under different treatment schedules for stand *j* would be equal to the total area, A_j of the stand.

The symbols in the above equations are:

c_{ij} = Discounted value (NPV) of net incomes of treatment *i* on stand *j*. Timber prices in Finnish market from the year 2005 (Finnish Statistical yearbook 2006) and 3% interest rate were used in NPV calculation.
x_{ij} = Area of stand *j* under treatment *i*
n = Number of stands
t_j = Number of treatments in stand *j*
g_{ij} = Growing stock of stand *j* under treatment *i* at the end of planning period (m^3)
r_{ijp} = Amount of removals (m^3) per hectare from stand *j* by treatment *i* during sub-period *p*
$r_{ijp'}$ = Amount of removals (m^3) per hectare from stand *j* by treatment *i* during sub-period *p*+1
A_j = Total area of stand *j*

3.5 Comparing Forest Plans and Estimating Inoptimality Losses

In the resulting forest plans, the differences between simulated treatment schedules based on correct data and erroneous data were investigated. The timing of thinning of the stand based on erroneous data was compared with that based on correct data. Total areas under thinning schedules based on correct data and erroneous data were also calculated and compared. The occurrence and timing of final cuttings was determined by solving the LP problems of the three different plans in each data for Datasets #1 and #2, respectively. The values for NPV and remaining growing stock (growing stock at the end of planning period) were estimated based on the plans of both correct data and erroneous data. The realized outcomes of the plans based on erroneous data were estimated by applying the obtained plan into the correct data, as described in detail in Chapter 2. And the holding-level inoptimality losses were determined by following the stages described in Chapter 2.

4 Results

4.1 Categorization of Measurers

The grouping of measurers is shown in Table 3. The measurers made both positive and negative biases, but most of them made negative bias. This means that they underestimated the basal areas of dominated tree species. Most of the measurers from Groups HBLV and HBHV made significant ($P < 0.05$) bias.

4.2 The Effects of Measurement Errors on Treatment Schedules

The use of erroneous data clearly affected the extent of thinnings at holding-level during the planning period. Underestimation of the basal area always leads to correct or delayed timing of thinning, whereas overestimation can lead to correct or early timing. This is clearly seen from Table 4, where high bias with the systematic underestimation of the total basal area of dominated tree species' caused lot of delayed thinnings (Groups HBHV and HBLV), whereas low bias allows both early and delayed thinnings (Groups LBHV and LBLV). In some stands, delayed thinning meant that thinning was postponed to occur beyond the 15-year planning period. This is why the thinning areas vary between the groups: the smaller the thinning area is, the more thinnings

Table 3. Groups and grouping of measurers based on the bias% and variance of measurement error in measuring the basal area of dominated tree species.

Groups	Criteria (according to Fig. 1)	Number of measurers (% of measurers)	Group mean (standard deviation in parentheses) Bias %	Variance	Number of measurers who made significant bias ($P < 0.05$)
LBLV	Bias% < 29* Variance < 6*	14 (21%)	12.52 (10.92)	3.53 (1.84)	4
LBHV	Bias % < 29 Variance > 6	18 (27%)	11.83 (10.90)	11.62 (5.04)	1
HBLV	Bias % > 29 Variance < 6	18 (27%)	44.47 (17.41)	3.55 (1.72)	13
HBHV	Bias % > 29 Variance > 6	17 (25%)	49.45 (16.29)	10.75 (5.39)	11

* Bias% 29 and variance 6 were medians

are delayed beyond the planning period.

In Plan #1, the final cutting areas were same for all groups of Dataset #2, whereas in Dataset #1 the final cutting areas were slightly smaller for HBLV and HBHV (Table 5 and 6). In Plans #2 and #3, the additional constraints made the final cutting areas based on erroneous data differ from those based on correct data. In Plans #2 and #3 of Dataset #1, the final cutting area was higher for all measurer groups than was the case for correct data, whereas in the same plans (Plans #2 and #3) of Dataset #2 the final cutting area was lower in all measurer groups than was the case for correct data; the exception was Group HBLV.

The harvest removals and the resulting NPVs and remaining growing stocks were found to be different between correct data and the erroneous data. In Plans #2 and #3 of Dataset #1, the sawlog removals and estimated NPVs were higher for the erroneous data than for correct data (Table 5). Correspondingly, the remaining growing stocks and harvested pulpwood for the measurer groups were lower than for the correct data. The delayed thinnings led to lower thinning removals, which in turn allowed higher final cut removals and higher estimated NPVs with lower amount of the remaining growing stocks.

A reverse result was observed in Dataset #2 for Plans #2 and #3. Inefficient plans with lower final cut removals, lower NPV of incomes, and higher remaining growing stock than in the data based on correct data were observed for all other measurer groups except for group HBLV (Table 6). Although the thinning area was smaller than in the correct data, the thinning removals did not differ so much. As a result, the amount of final cuts could not be increased much from the level of correct data.

Table 4. Changes in timings of thinnings based on erroneous data compared to the correct data for schedules of all plans. The thinning area during the planning period in the plan based on correct data was 58.3 ha for Dataset #1 and 13.5 ha for Dataset #2.

Data source	Thinning area (ha)	Number of stands* Delayed thinning	Early thinning	No change in timing
Dataset #1				
LBLV	49.0	6	3	25
LBHV	53.8	6	8	23
HBLV	46.0	13	0	21
HBHV	42.2	16	1	18
Dataset #2				
LBLV	13.3	3	4	3
LBHV	11.6	7	3	1
HBLV	5.1	8	0	1
HBHV	6.9	8	0	1

*Number of thinning stands based on correct data is 31 in Dataset #1 and 9 in Dataset #2. Total numbers of stands are 118 and 59 in Dataset #1and Dataset #2 respectively

The Plans #2 and #3 of group HBLV in Dataset #2 seems impossible at first glance, as both the NPV and remaining growing stock were higher than those for the correct data. This can be explained by bearing in mind that thinnings were always forced when the thinning limit is reached, which means that the timings of thin-

Table 5. Final cutting area, harvest volumes, remaining growing stocks (at the end of the 15-year planning period) and NPV under the different planning problems for both correct and erroneous data of Dataset #1.

Data source	Final cutting area (ha)	Harvest volume (m^3)			Remaining growing stock (m^3)	NPV (€)
		Sawlog	Pulpwood	Total		
Plan #1						
Correct	116.5	23509	11441	34950	22678	797311
LBLV	116.5	23294	11095	34389	23232	792500
LBHV	116.5	23320	11023	34343	23276	790796
HBLV	115.4	23506	10986	34492	23511	790261
HBHV	114.2	23139	10922	34061	23776	785859
Plan #2						
Correct	42.6	11940	6161	18101	41700	451489
LBLV	46.6	12764	5715	18479	41187	476634
LBHV	46.5	12868	5838	18706	40882	479225
HBLV	48.1	13314	5765	19079	40649	492230
HBHV	48.7	13022	5698	18720	41045	484035
Plan #3						
Correct	42.5	12925	6113	19038	41700	441149
LBLV	46.8	13760	5724	19484	41111	469212
LBHV	46.5	13391	5836	19227	41262	461115
HBLV	48.3	14113	5734	19847	40652	484629
HBHV	48.5	13882	5672	19554	41054	475197

Table 6. Final cutting area, harvest volumes, remaining growing stocks (at the end of the 15-year planning period) and NPV under the different planning problems for both correct and erroneous data of Dataset #2.

Data source	Final cutting area (ha)	Harvest volume (m^3)			Remaining growing stock (m^3)	NPV (€)
		Sawlog	Pulpwood	Total		
Plan #1						
Correct	65.7	15935	3518	19453	7498	559398
LBLV	65.7	16083	3401	19484	7571	555903
LBHV	65.7	15881	3396	19277	7680	555538
HBLV	65.7	15995	3166	19161	8019	553277
HBHV	65.7	16149	3299	19488	7800	548788
Plan #2						
Correct	18.2	5788	1482	7269	21880	223311
LBLV	17.1	5542	1320	6862	22415	212375
LBHV	17.9	5666	1356	7022	22217	218096
HBLV	20.7	6077	1187	7264	21984	230773
HBHV	18.0	5606	1247	6853	22470	215888
Plan #3						
Correct	18.3	6228	1422	7650	21880	212827
LBLV	17.1	5960	1267	7227	22441	201145
LBHV	17.9	6130	1331	7461	22190	211101
HBLV	19.4	6453	1162	7617	22028	219526
HBHV	17.2	5819	1245	7064	22643	201132

nings in the alternative management schedules based on erroneous data differed from those for the correct data. Thus, the measurement errors could even increase the overall utility due to an additional realistic simulated management schedule (delayed thinning). Had delayed thinnings been simulated into the correct data, this would not have happened. The smaller amount of removals resulting from a smaller thinning area (Table 4) made it possible to carry out more final cuttings, which resulted in a higher value of removals for approximately a similar volume (for Plan #2, 7264 m^3 for HBLV and 7269 m^3 for correct data) as a result of a higher sawlog proportion (Table 6).

In Plan #1 of both datasets, the cutting removals and the estimated NPVs were lower for the measurer groups than those for the correct data. The remaining growing stocks were higher for the measurer groups than for the correct data (Table 5 and 6). It was observed that the thinning extent had no effect on final cutting in Plan #1 because of the non-existence of constraints in the objective function.

4.3 Inoptimality Losses in Holding-Level Plans

We computed NPV$_{corr}$ under the Plan #2 and #3 according to the stages described in Chapter 2. As an example, for the plan based on the data of the group LBLV of Dataset #1 under Plan #2, the NPV$_{corr}$ was computed by maximizing NPV from the treatment schedules based on correct data simulated according to step (B1*) subject to the remaining growing stock constraint of 41 187 m^3 which was equal to the remaining growing stock in the plan based on erroneous data (Table 5, data source LBLV, Plan #2). When computing NPV$_{corr}$ for the plan of group LBLV of Dataset #1 under Plan #3, the remaining growing stock constraint was kept equal to 41 111 m^3 (Table 5, data source LBLV, Plan #3) and the constraints for the amounts of periodical cuttings were equal to or greater than 6665 m^3, 6439 m^3 and 6379 m^3 for period #1, #2 and #3 respectively (Table 7, data source LBLV, Dataset #1). For Plan #2 of group LBLV in Dataset #1, the inoptimality loss (4843 €) was computed from the difference between the NPV$_{corr}$ (481 477 €) and the NPV$_{err}$ (476 634 €) (Table 8, data source LBLV, Plan #2). The computed inoptimality losses for all erroneous data are presented in Table 8 and 9.

Table 7. Amounts of periodical cuttings under Plan #3 based on erroneous data. These amounts were used as periodical cutting constraints for computing NPV$_{corr}$ under Plan #3.

Data source	Amounts of periodical cuttings (m^3) Period #1	Period #2	Period #3	Total (m^3)
Dataset #1				
LBLV	6665	6439	6379	19483
LBHV	8190	4587	6449	19226
HBLV	7189	6555	6104	19848
HBHV	6801	6353	6398	19552
Dataset #2				
LBLV	2242	2377	2609	7228
LBHV	2497	2506	2458	7461
HBLV	2527	2573	2517	7617
HBHV	2403	2169	2493	7065

Inoptimality losses due to the use of erroneous data were different under different planning problems for different measurer groups. For Plan #1, the inoptimality losses increased as the measurement error of the groups increased similarly in both datasets (Tables 8 and 9). The effects of adding the remaining growing stock and even flow constraints to the objective function, i.e. the difference between Plan #1, #2 and #3, were as expected: the NPVs decreased. However, the effects on the holding-level inoptimality losses varied depending on the measurer groups. With at least the same remaining growing stock and periodical cutting amounts as were derived from the schedules of the erroneous data in Plan #2 and Plan #3, the new problem formulation described in stage B2* could improve the NPVs from 1329 € (0.3%) to 6953 € (1.5%) in Dataset #1, and from 885 € (0.4%) to 8103 € (3.7%) in Dataset #2. In both datasets, the lowest inoptimality losses were found from the most accurate group LBLV. Adding the even flow constraint decreased the inoptimality loss in all groups, except the group HBLV of Dataset #2.

Table 8. Inoptimality losses for different measurer groups under different planning problems for Dataset #1. NPV_{corr} for Plan #2 and #3 were computed by following the procedure described in Step B1* and B2* in Chapter 2 and illustrated in Chapter 4.3.

Data source	NPV_{corr} (€)*	NPV_{err} (from Table 5) (€)	Inoptimality loss ($NPV_{corr} - NPV_{err}$) (€)
Plan #1			
LBLV	797311	792500	4811
LBHV	797311	790796	6515
HBLV	797311	790261	7050
HBHV	797311	785859	11452
Plan #2			
LBLV	481477	476634	4843
LBHV	486178	479225	6953
HBLV	495465	492230	3235
HBHV	490703	484035	6668
Plan #3			
LBLV	470541	469212	1329
LBHV	467732	461115	6617
HBLV	491378	484629	6749
HBHV	480247	475197	5050

* NPV_{corr} were same in different data sources for Plan #1 because no constraints were used

Table 9. Inoptimality losses for different measurer groups under different planning problems for Dataset #2. NPV_{corr} for Plan #2 and #3 were computed by following the procedure described in Step B1* and B2* in Chapter 2 and illustrated in Chapter 4.3.

Data source	NPV_{corr} (€)*	NPV_{err} (from Table 6) (€)	Inoptimality loss ($NPV_{corr} - NPV_{err}$) (€)
Plan #1			
LBLV	559398	555903	3495
LBHV	559398	555538	3860
HBLV	559398	553277	6121
HBHV	559398	548788	10610
Plan #2			
LBLV	216422	212375	4047
LBHV	226199	218096	8103
HBLV	237275	230773	6502
HBHV	221143	215888	5255
Plan #3			
LBLV	202030	201145	885
LBHV	216254	211101	5153
HBLV	224707	219526	5181
HBHV	204162	201132	3030

* NPV_{corr} were same in different data sources for Plan #1 because no constraints were used

5 Discussion and Conclusions

Reliability of inventory data is generally examined by studying its extent of errors in terms of so called root mean square error (RMSE). The problem of RMSE is that it does not provide any idea on the amount of monetary loss the forest owner may suffer from erroneous data. From the viewpoint of the forest owner, the real meaning of RMSE can also be rather difficult to understand. The estimated amount of inoptimality loss provides probably a more understandable signal to the forest owner; Can s/he afford the expected monetary loss or is it better to look for more accurate inventory data. The proposed approach can be a useful tool for analyzing the effects of error in inventory data on the holding-level planning process. It provided the guidelines for analyzing the effects. It also pointed out the principles of how errors in inventory data do affect the planning procedures. More importantly, this approach has turned the spotlight on the aspect of holding-level inoptimality loss estimation. The approach was illustrated by a case study in order to show the procedure of its implementation.

The holding-level effects of using erroneous inventory data in forest planning calculations were examined by using two datasets. Measurement errors confined only to the basal area of the dominated tree species were considered. As expected, the basal area of dominated tree species was generally underestimated. The variation in the relative bias and variance of the measurement error for the dominated tree species' basal area was high among the measurers, although absolute errors can be small in stands where the proportion of dominated tree species is small. The measurement errors led to rather modest inoptimality losses, at maximum 3.7 per cent, the loss being generally the higher the more errors were made in the measurements. The effects of measurement errors on the forest plan were, however, different for different measurer groups and for the two datasets used. This unpredictable nature of the effects of errors emphasizes the importance of actually considering the effects of measurement errors on the end result – the contents of the forest plan. This indicates research and development needs for the further development of forest planning practices in this respect.

It can be concluded that the higher the bias in the measured basal area of dominated tree species, the more the contents of holding-level forest plans differ when compared to a plan created based on correct data. Firstly, this difference is caused by the different timings of the thinnings if these are forced to take place when the thinning limit (with respect to basal area and dominant height) is exceeded. In the thinning stands covered by the two datasets, delayed thinnings were more common than early thinnings. Missing thinning treatments (that were actually delayed after the planning period) increased the holding's remaining growing stock. As a result, the remaining growing stock constraint allowed additional clear cutting to take place. As a further result, the forest owner got higher NPVs, but the remaining growing stocks were lower than the initial growing stocks. This was the principal result of Dataset #1 and it is in line with the idea presented in Chapter 2. However, in some stands, overestimates of the basal areas of the dominated tree species also had the result that thinnings were made before thinnings based on the correct data. If thinnings are not forced to take place when the thinning limit is exceeded, the effects of errors depend on the objective function used.

For Dataset #2, the results followed the other argumentation of Chapter 2. The thinnings were delayed in almost all stands where the thinning limit was reached based on correct data (Table 4). When compared to Dataset #1, the results for Dataset #2 were the opposite, except for Group HBLV, i.e. the resulting NPV was lower and the remaining growing stock higher in Plans #2 and #3 when the erroneous data were used than was the case when the correct data were used for these plans. The combined effect of the holding's forest structure (high initial growing stock), the constraints used, and the delayed thinnings together caused this result. In Group HBLV of Dataset #2, the results in Plan #2 and #3 were, however, quite surprising as both the remaining growing stocks and NPVs were higher than those for the correct data. The reason for this was the very low thinning area, i.e. some stands with very high growing stocks were not thinned. As a result, the final cutting area, sawlog removals, and NPVs could be increased without violating the remaining growing stock and even flow constraints. This would not have happened if the alternative treatment schedules had also included delayed thinnings.

The results of the case study showed that the holding-level effects of measurement errors are not easy to anticipate. In two datasets, the effects were opposite. One should also note that errors only in measuring the basal area of dominated tree species were examined. All the other variables were kept at their correct values. Even this fairly minor error was enough to considerably change the contents of the holding-level forest plans, particularly in the case of groups with large systematic measurement errors. Had other errors been accounted for, the changes in holding-level forest plans could have been even bigger and more difficult to anticipate, unless the effects eliminate each other, which can happen, for example, when measuring dominated tree species. In this study, the correlation between the errors in the basal areas of the dominant and dominated tree species was negative and significant at 0.01 level. However, in 41.4% of all the stands (1046) the errors in measuring the basal areas of the dominant and dominated tree species were in the same direction, either underestimation or overestimation. And in 68.3% of the underestimated stands both dominant and dominated tree species were underestimated. Thus, there were quite a large number of stands where the effects did not compensate each other.

At stand level, the results of the present study are fairly trivial: underestimation of basal area causes delayed thinning which ultimately leads to the losses in future timber value, especially through diameter growth (e.g. Mäkinen and Isomäki 2004a, 2004b, Plauborg 2004, Karlsson 2006), whereas overestimation causes early thinning proposals. In practice, however, early thinnings can be canceled in the operational planning phase. At holding-level, the effects are not that easily predicted, and they depend on the objective function used as well as on the current state of the forests, and on the production possibilities of the holding.

In this study, inoptimality losses were estimated by using LP formulation. This approach summarizes loss into a single figure. The loss expresses how much improvement the forest owner could get by rescheduling the harvests, given that the constraints would remain at the level they are in the current plan. However, it does not take into account the loss caused by the constraints not being actually met or being exceeded in the current plan. The use of even-flow constraints, for example, caused some problems in calculating the inoptimality losses. As an alternative to the approach described in Chapter 2, the calculation of inoptimality loss from Plan #3 could have been done by using similar even-flow constraints as were used in optimization procedure. However, the approach we used ensured that the current plan is included among the possible plans.

As an additional alternative to the selected approach for estimating the inoptimality loss, the use of the multi-attribute utility function as an objective function of optimization calculations would have been possible. Had the utility function been used, the inoptimality losses caused by applying non-optimal treatment schedules could have been calculated simply as utility loss. This loss could have been then transformed into monetary values by applying the approach presented by Kurttila et al. (2005), where the opportunity cost of voluntary biodiversity protection was estimated at forest holding-level. However, this alternative approach demands that the parameters of the utility function can be accurately estimated, because they have a major impact on the magnitude of the inoptimality losses.

Only the effects during the 15-year planning period were considered in the case study, and the examination did not take into account the development of the stands during the remainder of the rotation nor the coming rotations. In addition to direct inoptimality losses from only the planning period that were measured in the case study, delayed thinnings and other non-optimal treatments of stands also cause subsequent losses to the forest owner because stand development is delayed, at least when viewed in the light of the applied treatment recommendations.

References

Altman, N.S. 1992. The introduction to kernel and nearest neighbour nonparametric regression. Journal of the American Statistical Association 46: 175–185.

Anttila, P. 2002. Updating stand level inventory data applying growth models and visual interpretation of aerial photographs. Silva Fennica 36(2): 549–560.

Ary, D. & Jacobs, L.C. 1976. Introduction to statistics: purposes and procedures. Holt, Rinehart and Winston. New York, ... 461 p.

Duvemo, K. & Lämås, T. 2006. The influence of forest data quality on planning processes in forestry. Scandinavian Journal of Forest Research 21: 327–339.

— , 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. 1991. Consequenses of erroneous data basis for planning and management of forest holdings. Ås: Skogforsk. Rapp. Skogsforsk 9/91. (In Norwegian with English summary).

— 1993. Random errors and strategic planning in forestry. Ås: Norwegian Forest Research Institute. Medd. Skogforsk 46(7).

— 2000. Use of uncertain inventory data in forestry scenario models and consequential incorrect harvest decisions. Silva Fennica 34(2): 89–100.

— 2001. Models for prediction of basal area mean diameter and number of trees for forest stands in south-eastern Norway. Scandinavian Journal of Forest Research 16: 467–479.

— , 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.

Finnish Statistical yearbook of Forestry. 2006. Finnish Forest Research Institute. SVT Agriculture, Forestry and Fishery. 438 p.

Haara, A. 2005. The uncertainty of forest management planning data in Finnish non-industrial private forestry. Doctoral Thesis. Dissertationes Forestales 8. University of Joensuu 34 p + 5 appendices.

— & Korhonen, K.T. 2004. Kuvioittaisen arvioinnin luotettavuus. Metsätieteen aikakauskirja 4/2004:

489–508. (In Finnish).

— , Maltamo, M. & Tokola, T. 1997. The k-nearest-neighbour method for estimating basal-area diameter distribution. Scandinavian Journal of Forest Research 12: 200–208.

Härdle, W. 1989. Applied nonparametric regression. Cambridge University Press, Cambridge. 323 p.

Holmström, H., Kallur, H. & Ståhl, G. 2003. Cost-plus-loss analyses of forest inventory strategies based on kNN-assigned reference sample plot. Silva Fennica 37(3): 381–398.

Holopainen, M. & Talvitie, M. 2006. Effect of data accuracy on timing of stand harvest and expected net present value. Silva Fennica 40(3): 531–543.

Hyvän metsänhoidon suositukset. 2006. Metsätalouden kehittämiskeskus Tapio. 100 p.

Jacobsson, J. 1986. Optimization and data requirements – a forest management planning problem. Swedish University of Agricultural Sciences, Section of Forest Mensuration and Management, Umeå, Sweden. 143 p.

Kangas, A. & Kangas, J. 1999. Optimization bias in forest management planning solutions due to errors in errors in forest variables. Silva Fennica 33(4): 303–315.

— , Heikkinen, E. & Maltamo, M. 2004. Accuracy of partially visually assessed stand characteristics: a case study of Finnish forest inventory by compartments. Canadian Journal of Forest Research 34: 916–930.

Karlsson, K. 2006. Impact of the thinning regimes on the mean diameter of the largest stems by diameter at breast height in even-aged Picea abies stands. Scandinavian Journal for Forest Research 21: 20–31.

Koivuniemi, J. & Korhonen, K.T. 2006. Inventory by compartments. In: Kangas, A. & Maltamo, M. (eds.). Forest inventory: methodology and applications. Managing Forest Ecosystem 10: 271–278.

Kurttila, M., Pykäläinen, J. & Leskinen, P. 2005. Defining the forest landowner's utility–loss compensative subsidy level for a biodiversity object. European Journal of Forest Research 125(1): 67–78.

Laasasenaho, J. & Päivinen, R. 1986. On the checking of inventory by compartments. Folia Forestalia 664. 19 p. (In Finnish with English summary).

Maltamo, M., Malinen, J., Kangas, A., Härkönen, S. & Pasanen, A.-M. 2003. Most similar neighbour-based stand variable estimation for use in inventory by compartments in Finland. Forestry 76(4): 449–463.

Mäkinen, H. & Isomäki, A. 2004a. Thinning intensity and growth of Norway spruce stands in Finland. Forestry 77(4): 349–364.

— & Isomäki, A. 2004b. Thinning intensity and growth of Scots pine stands in Finland. Forest Ecology and Management 201(2–3): 311–325.

Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, A. & Söderman, U. 2004. Laser scanning of forest resources: The Nordic experiences. Scandinavian Journal of Forest Research 19: 482–499.

Plauborg, K.U. 2004. Analysis of radial growth responses to changes in stand density for four tree species. Forest Ecology and Management 188(1–3): 65–75.

Poso, S. 1983. Basic features of forest inventory by compartments. Silva Fennica 17: 313–343. (In Finnish with English summary).

Pukkala, T. 2004. Monsu Metsäsuunnitteluohjelma Versio 4. Ohjelmiston toiminta ja käyttö. Mimeograph at the University of Joensuu, Finland. 75 p. (In Finnish).

Pussinen, A. 1992. Ilmakuvat ja Landsat TM – satelliittikuva välialueiden kuvioittaisessa arvioinnissa. M. Sc. Thesis. University of Joensuu, Faculty of Forestry. 48p. (In Finnish).

Pykäläinen, J. 2000. Interactive use of multi-criteria decision analysis in forest planning. Academic dissertation. University of Joensuu, Faculty of Forestry. 37 p.

Saari, A. & Kangas, A. 2005. Kuvioittaisen arvioinnin harhan muodostuminen. Metsätieteen aikakauskirja 1/2005: 5–18. (In Finnish).

Sprängare, B. 1975. A method for analyzing the sensitivity of the long-range planning which depends on errors in stand data. Stockholm Royal College of Forestry. Research notes 87. (In Swedish with English summary).

Ståhl, G. 1992. A study on the quality of compartment-wise forest data acquired by subjective inventory methods. Swedish University of Agricultural Sciences, Dept. of forest mensuration and management. Report 24. 128 p. (In Swedish with English summary).

Tapion vuositilastot 2005. Tapios årsstatistik 2005. Metsätalouden kehittämiskeskus Tapio. 56 p. (In Finnish and Swedish).

Total of 38 references