Author(s): Kangas, Annika; Hartikainen, Markus; Miettinen, Kaisa

Title: Simultaneous optimization of harvest schedule and data quality

Year: 2015

Version:

1   **Simultaneous optimization of harvest schedule and data quality**

2

3   Kangas, Annika[1], Hartikainen Markus[2] and Miettinen Kaisa[2]

4   1) Department of Forest Sciences, P.O. BOX 27, FI-00014 University of Helsinki, Finland,

5   Email: Annika.kangas@helsinki.fi, Phone +358-50-3750870

6   2) University of Jyväskyla, Department of Mathematical Information Technology, P.O.

7   BOX 35 (Agora), FI-40014 University of Jyväskylä, Finland

8

1   **Abstract**

2

3   In many recent studies, the value of forest inventory information (VOI) in the harvest scheduling

4   has been examined. In a previous paper, we demonstrated that selecting only a small part of stands

5   for additional measurements simultaneously with making the harvest decisions may be highly

6   profitable. In that study, the quality of additional measurements was not a decision variable, and the

7   only options were between making no measurements and measuring perfect information. In this

8   study, we introduce the data quality into the decision problem, i.e., the decision maker can select

9   between making imperfect or perfect measurements. The imperfect information is obtained with a

10   specific scenario tree formulation. Our decision problem includes three types of decisions: harvest

11   decisions, measurement decisions, and decisions about measurement quality. In addition, we decide

12   about the timing of the harvests and measurements. These decisions are evaluated based on two

13   objectives, discounted aggregate incomes for the planning periods and the end value of the forest at

14   the end of the planning horizon. Solving the biobjective optimization problem formed shows that

15   imperfect information is mostly sufficient for the harvest timing decisions during the planning

16   horizon but perfect information was required in order to meet the end inventory requirement. The

17   relative importance of the two objectives affects the measurements indirectly, through increasing or

18   decreasing the number of decisions that are identical in all scenarios, i.e., certain decisions.

19

1. Introduction

Any non-trivial decision typically includes uncertainty concerning the prevailing state of nature (Hirshleifer and Riley 1979). The decision maker can then either make an optimal choice between different alternatives with the current information or reduce the uncertainty by collecting more information. Value of information (VOI) in decision making (ex ante) can hereby be defined as the difference between the expected value of this choice with and without the information (e.g. Hirshleifer and Riley 1979, Lawrence 1999, Birchler and Bütler 2007 and Kangas 2010).

In harvest scheduling, the VOI in forest inventory can be calculated to support the decision for measuring the area in question as a whole. The measurements are profitable, if VOI is higher than the measurement costs. It is also possible to formulate the forest planning problem in such a way that the data acquisition decisions are optimized at the same time as the harvest decisions. In this case, it is possible to select a sub-set of stands or some sub-areas as stands to be measured. For instance, Duvemo et al. (2014) formulated a problem where the stands that were eligible to be harvested during the first years of planning horizon in a tactical plan were selected for a pre-harvest inventory, for which more accurate data was acquired before the final operational decisions. The selection of the stands for a pre-harvest inventory was not optimized, though. Kangas et al. (2014) included the measurement decisions into the harvest scheduling optimization problem, and then it was possible not only to select some stands to be measured, but also to define when such data would be needed.

If it was assumed that the uncertainty can be resolved, i.e., perfect information of the forests will become available at some time point, two-stage optimization (also programming with recourse) could be utilized. In the first stage, decisions are made under uncertainty, but the second stage decisions are assumed to be made under certainty. In a multi-stage problem, new information

1   becomes available after each stage and the uncertainty involved diminishes in time. If the decision

2   of forest inventory in a given stand is optimized at the same time as the harvests, the harvest

3   decision gives the optimal timing of the harvest and the inventory decision indicates whether a

4   stand is measured or not. This decision depends on the costs of the inventory per hectare. This

5   means that if the stand is measured, optimization with perfect information is carried out for that

6   stand, and if it is not measured, optimization is carried out under uncertainty. In this case, moving to

7   a second stage is an active decision, rather than happening in time as in common "wait-and-see"

8   problem formulations (see e.g. Birge and Louveaux 2011).

9

10   A generalization of the described problem is that besides the harvest and inventory decisions we

11   also decide about the quality of the inventory. Then, for each of the stands we can choose whether it

12   requires additional information, and also whether imperfect information is sufficient, or whether

13   perfect information is needed. That may depend on the decisions to be made, the relative

14   importance of the different objectives, the quality ratio of imperfect information and perfect

15   information, and the costs of the different measurements.

16

17   In a previous paper (Kangas et al. 2014), we demonstrated that selecting only a small part of stands

18   for additional measurements simultaneously with making the harvest decisions may be highly

19   profitable. In that study, the quality of additional measurements was not a decision variable but,

20   instead, the only options were either to make no measurements and to measure perfect information.

21   In this study, we extend the consideration and introduce the data quality in the decision problem,

22   i.e., the decision maker can select between making imperfect or perfect measurements. We study

23   this by formulating an optimization problem where we optimize simultaneously both the decision of

24   whether to harvest the stand or not, whether to measure the stand or not, and the optimal quality of

25   the information acquired. In addition, we study the optimal timings of the harvests and

1   measurements. We formulate this as a multiobjective optimization problem (see e.g. Miettinen

2   1999) and solve it using the ε-constraint method (Haimes et al. 1971, Miettinen 1999), using as the

3   objectives both the end value of the forests and the aggregated discounted incomes from the

4   planning periods within the planning horizon. We study the effects of measurement costs, interest

5   rate and values of end-inventory, which is measured as a relative maximum decline allowed in end-

6   inventory value during the planning horizon.

7

8   **2. Material**

9

10 The study site consists of 29 stands, i.e., $J = 29$ and 67.29 hectares located in the municipality of

11 Juuka in Eastern Finland (Table 1). This estate is a part of a larger estate owned by UPM-

12 Kymmene. The stands were measured using a compartment-wise inventory (for details, see

13 Mustonen et al. 2008).

14

15 The initial data was assumed to include inventory errors. The prior distributions of the basal area

16 and dominant height in the beginning of the planning horizon were formulated using Monte Carlo

17 simulation. Correlated multinormal errors were introduced to the initial variables using the

18 rmvnorm package of R (R Development Core Team 2011). The relative standard error of dominant

19 height and stand age was assumed to be 10%, and that of basal area 27.5%. Correlation was

20 assumed to be 0.1. The means of errors were assumed to be zero. This level corresponds to the level

21 of five-year old data obtained with a traditional stand-level forest inventory (Haara & Leskinen

22 2009).

23

24 The errors were simulated so that the prior variance of each variable (dominant height (H_{dom}), stand

25 age (T) and basal area (G)) was divided into two components: $\sigma^2_{prior} = \sigma^2_{imperfect} + \sigma^2_{residual}$. First, 10

1   scenarios of an error component $\varepsilon_{imperfect}$ were generated using random number generation. Then, for

2   each of these scenarios 10 additional error components $\varepsilon_{residual}$ were generated, giving a total of $I =$

3   100 scenarios of stand characteristics (see Figure 1 for a case with 3 replications). Then, in the

4   optimization calculations the mean basal area (or mean height or age) from each set of 10 scenarios

5   served as one estimate of the imperfectly measured basal area $\hat{G}_{imperfect}$, i.e. altogether these sets

6   formed 10 different values of imperfectly measured basal area. In this study, the imperfect

7   information was assumed to resolve 70% of the prior variation, which means that the variance of

8   the imperfect information was 70% and the residual uncertainty 30% of the prior variance for each

9   of the variables. That would reflect that the imperfect information for basal area for instance, is in

10   the level of modern ALS forest inventory (Packalén & Maltamo 2007).

11

12   When the initial data was generated, a simple stand-level simulator was used to predict the

13   development of the stands for $K = 9$ five-year periods, for each of the I scenarios. The simulator

14   consisted of a model for the relative growth of basal area as a function of stand age, mean height,

15   site index (dominant height at the age of 100 years, H_{100}) and basal area, and of a growth model for

16   the relative growth of dominant height (H_{dom}) as a function of stand age and dominant height for

17   spruce and pine (Vuokila and Väliaho 1980). The latter model was also used to predict the site

18   index.

19

20   Stand volume (V) was calculated with a model $V=G \cdot H_{gm} \cdot f$, where f is the form factor. The form

21   factor was predicted using models by Gustavsen and Fagerström (1983) as a function of mean

22   height and site type. The relation between mean height and dominant height was assumed to be

23   constant in the data, and a model of the form

24

25   $\hat{H}_{gm} = 0.5784 + 0.7807 H_{dom}$

1

2 for the mean height was estimated from the dataset in the beginning of the simulation.

3

4 The alternative decisions available for the decision maker were the different timing options for

5 clearcut. No thinnings were included in the planning. The net revenues from the clearcut were

6 obtained by multiplying the total volume with the price 35 €/m^3. Thus, every produced cubic meter

7 was assumed to be equally valuable. Possible clearcut was assumed to be carried out at the end of

8 the period.

9

10 **3. Methods**
11

12 The problem here is a typical planning problem for a private forest owner, with the goal to get

13 incomes from the estate and also to save some cutting possibilities also for the future. We start with

14 a general multiobjective optimization problem that models the simultaneous optimization of harvest

15 schedule and data quality. The intention is to study how different trade-offs in the values of the

16 objectives affect the measurement and harvest decisions.

17

18 Let $x_{ijk} \in \{0,1\}$ be the harvest decision of stand $j \in J$, for scenario $i \in I$ and in period $k \in K$. Here,

19 sets J, I and K stand for the stands, scenarios and periods, respectively. Furthermore, let $(I_l)_{l \in L}$

20 (where L is an index set) be an exclusive partition of scenarios in I (i.e., $I_{l_1} \cap I_{l_2} = \varnothing$ for all $l_1 \neq l_2$

21 and $\cup_{l \in L} I_l = I$). The index sets I_l mathematically model the imperfect information, which was

22 explained previously. That is, the variance is smaller inside the index sets I_l than it is between

23 them. Finally, let $w_{jk} \in \{0,1\}$ be the decision of making a perfect inventory of stand j on period k

24 and $v_{jk} \in \{0,1\}$ be the decision of making an imperfect inventory of stand j on period k. In addition,

1   we fix an index $i_l \in I_l$ for each $l \in L$. Finally the model has the following parameters: parameter δ is

2   the interest rate, real values c_{jik} are the incomes (€/ha), q and r are the costs of full and partial

3   inventories (€/ha) and a_j are the areas of stands.

4

5   Then, we formulate the multiobjective optimization problem as

6     $$\max\left(\left(\sum_{k\in K}\sum_{j\in J}\left(x_{jik}c_{jik}-w_{jk}q-v_{jk}r\right)a_j\left(\frac{1}{1+\delta}\right)^{5k}\right)_{i\in I},\left(\sum_{j\in J}c_{ji\#K}\left(1-\sum_{k\in K}x_{jik}\right)\right)_{i\in I}\right)$$

7   (1)

8   subject to

9     $\sum_{k\in K}x_{jik}\leq 1,\, \textit{for all } j\in J, i\in I$ (2)

10    $x_{jik'}-x_{ji_lk'}-\sum_{k=1}^{k'}v_{jk}-w_{jk}\leq 0,\, \textit{for all } j\in J, i\in I, k'\in K$ (3)

11    $x_{ji_lk'}-x_{jik'}-\sum_{k=1}^{k'}v_{jk}-w_{jk}\leq 0,\, \textit{for all } j\in J, i\in I, k'\in K$ (4)

12    $x_{jik'}-x_{ji_lk'}-\sum_{k=1}^{k'}w_{jk}\leq 0,\, \textit{for all } j\in J, i\in I_l, l\in L$ (5)

13    $x_{ji_lk'}-x_{jik'}-\sum_{k=1}^{k'}w_{jk}\leq 0,\, \textit{for all } j\in J, i\in I_l, l\in L$ (6)

14    $x_{jik}, w_{jk}, v_{jk}\in\{0,1\},\,$ for all $j\in J, i\in I, k\in K$,

15   where #K refers to the number of periods. In the above problem, the objectives are to maximize the

16   discounted incomes in all scenarios and, at the same time, to maximize the value of the forest left

17   unharvested. The problem has $2\#I$ objectives. The first constraint (2) determines that each stand

18   can be clearcut at most once during the planning horizon. Constraints (3) and (4) make sure that if

19   nothing is measured, then all the harvest decisions in all the scenarios must be the same. It means

20   that if no measurements at all are carried out, all the decisions for the given stand are first-stage

1   decisions. These constraints can be interpreted as non-anticipavity constraints. Constraints (5) and

2   (6) make sure that if full inventory is not taken, then the harvest decisions must be the same within

3   the partitions I_l. When imperfect measurements are made, we can interpret the decisions as second-

4   stage decisions, where we know to which group of scenarios the realized scenario belongs. The

5   perfect information gives the possibility for third-stage decisions that can be optimal for each

6   scenario separately.

7

8   Unlike in typical multi-stage problems, here the uncertainty is not resolved in time, when moving

9   from one planning period to next, unless measurements are carried out. It also means that unlike in

10 typical multi-stage problems, the stands can be at different stages: it is possible that for some stands

11 stepping from first stage to second or third stage never happens, and it is likewise possible that in

12 some stands there is a need to step to the third stage right in the beginning of the planning horizon.

13 The movement from one stage to another is thus a result of a decision. In this context, the possible

14 consequences of uncertainty are that the harvest timing decisions are sub-optimal, i.e. produce less

15 incomes than would have been optimal. In addition, the uncertainty may cause the solution to be

16 infeasible (end-inventory value too small) or inefficient (part of the harvest possibilities are left

17 unused).

18

19 The optimization problem (1) - (6) means that both the incomes and the value of the forest left

20 unharvested must be maximized, whatever happens, i.e., whatever scenario realizes. The scenarios,

21 however, are not real but, instead, merely technical means to model the uncertainty. When one takes

22 into consideration the fact that there is a large number of scenarios and, thus, the problem

23 formulated has a very large number of objectives, it is desirable to find a formulation of the

24 problem with a lower number of objectives and formulation that handles the uncertainty in the

25 scenarios. For this, we formulate a biobjective optimization problem

1 $$\max\left(\sum_{j=1}^{\#J}\sum_{k=1}^{\#K}\left[\left(\frac{1}{\#I}\sum_{i=1}^{\#I}x_{jik}c_{jik}-w_{jk}q-v_{jk}r\right)a_j\left(\frac{1}{1+\delta}\right)^{5k}\right],\min_{i\in I}\left(\frac{\sum_{j=1}^{\#J}c_{ji\#K}(1-\sum_{k=1}^{\#K}x_{jik})}{\sum_{j=1}^{\#J}c_{ji1}}\right)\right) \qquad (7)$$

2 subject to

3 constraints $(2)-(6)$

4 $x_{jik}, w_{jk}, v_{jk} \in \{0,1\}$, for all $j \in J, i \in I, k \in K$.

5 The objectives of the above problem are to maximize the mean income over all scenarios and to

6 maximize the minimum value (over all scenarios) of the forest left unharvested relative to the value

7 at the beginning of the planning horizon. That is, if the value of the second objective is, for

8 example, greater than 0.7, then the reduction in the forest value is less than 30% in all scenarios.

9

10 The objectives of the optimization problem (7) and (2)-(6) cannot be combined directly into a single

11 objective, because it is not completely clear how they should be combined. By studying the above

12 biobjective optimization problem with different preferences between the objectives, we can get

13 better picture about what should be the most preferred way of cutting and measuring the stands.

14

15 Finally, we solve the biobjective optimization problem using the augmented ε-constraint method

16 (Movrotas 2009). That is, we solve optimization problems of the type

17 $$\max\sum_{j=1}^{\#J}\sum_{k=1}^{\#K}\left[\left(\frac{1}{\#I}\sum_{i=1}^{\#I}x_{jik}c_{jik}-w_{jk}q-v_{jk}r\right)a_j\left(\frac{1}{1+\delta}\right)^{5k}\right]+\rho\min_{i\in I}\left(\frac{\sum_{j=1}^{\#J}c_{ji\#K}(1-\sum_{k=1}^{\#K}x_{jik})}{\sum_{j=1}^{\#J}c_{ji1}}\right) \qquad (8)$$

18 subject to

19 constraints $(2)-(6)$

1
$$\frac{\sum_{j=1}^{\#J} c_{ji\#K}(1-\sum_{k=1}^{\#K} x_{jik})}{\sum_{j=1}^{\#J} c_{ji1}} \geq \varepsilon \text{ for all } i \in I \tag{9}$$

2     $x_{jik}, w_{jk}, v_{jk} \in \{0,1\}$, for all $j \in J, i \in I, k \in K$,

3

4   where the parameter ε ranges from 0.7 to 1. This means that we allow a drop from 0 to 30% in the

5   value of the forest when compared to the value at the beginning of the planning horizon. The

6   parameter ρ is given a very small value (in the computational test we have used 0.0001). Without

7   the additional term

8
$$+\rho \min_{i \in I} \left(\frac{\sum_{j=1}^{\#J} c_{ji\#K}(1-\sum_{k=1}^{\#K} x_{jik})}{\sum_{j=1}^{\#J} c_{ji1}} \right)$$

9   in the objective function (8), the optimal solutions could only be guaranteed to be weakly Pareto

10   optimal to the optimization problem (8), (2)-(6), (9), but with the complete objective function (8)

11   the optimal solutions of this problem are guaranteed to be Pareto optimal solutions to the

12   optimization problem (see Movrotas 2009). Finally, the above problem can be formulated as a

13   mixed-integer linear problem using the standard way of linearizing max-min type optimization

14   problems (see e.g., Posner and Wu 1981).

15

16   When solving the problem, all the mixed-integer linear optimization problems were solved using

17   the IBM ILOG CPLEX Optimization Studio V 12.5.1.0 (IBM ILOG CPLEX), which is commonly

18   known merely as Cplex. Cplex was running on a computing server with 64 cores and approx.

19   1032GB of Random-access memory (RAM). Cplex was able to take advantage of extensive

20   parallelization and was simultaneously running in parallel on up to 32 cores.

4. Results

The biobjective optimization problem was solved with different costs of perfect and imperfect measurements and with different interest rates. By using the ε-constraint method, we could study trade-offs between the expected income over all the scenarios and the end inventory value. According to the results, a fairly stable number of perfectly measured stands (being 8-10) is the best, while the number of imperfectly measured stands sharply diminished as a function of interest rate (Figure 2). With a 10€/ha cost, the number of stands profitably measured with imperfect data dropped already when the interest rate rose to 2 per cent, while with a 5 € cost/ha the biggest drop happened when the interest rate rose to 3. Yet, some stands were always measured with perfect information, even though the associated cost was set to 100€/ha.

When the proportion of measured area was shown, again the perfectly measured area seemed to be pretty constant, but the area of imperfectly measured stands dropped fast as a function of interest rate, and clearly faster with the 10€/ha cost. With the 1% interest rate, the proportion of the area measured was clearly over 40%. This result is due to the fact that the number of uncertain harvest decisions was greatest with the interest rate 1%, as the value growth percent in many of the stands was below 1% in part of the scenarios and above it in others. However, even with a 2% interest rate, the benefits of imperfect measurements were too low to measure much over 10% with 10€/ha costs, while with a 5€/ha cost it was profitable to measure about 20% with the interest rate 3. Besides this obvious difference, the two cost options produced fairly similar results and, thus, the results shown in the rest of the figures mostly concern 5€/ha costs.

When the measurement period and the quality of information is plotted as a function of stand age, it can be seen that with the interest rate 1%, it was profitable to measure mainly old stands during periods 3-6 with imperfect information, and mainly young stands during periods 7-9 with perfect

1   information (Figure 3). This suggests that imperfect measurements were mainly made for harvest

2   timing decisions, and perfect measurements for meeting the desired end-inventory level. When the

3   area measured with perfect information was studied more closely, it turned out that the probability

4   of a stand to be measured with perfect information rapidly decreased as a function of stand area.

5   The two smallest stands were measured with perfect information in all studied problems with

6   different parameters (Figure 4), but stands larger than 2 hectares were never measured with perfect

7   information. That also indicates that the perfect information was practically always just used to

8   meet the end-inventory requirement.

9

10   The use of imperfect and perfect data for different types of decisions could be seen from the

11   distributions of the harvest timings across scenarios and the quality of information required (Figure

12   5). The stands where no measurements were profitable were harvested either in the beginning

13   (period 1) or not at all (N). In both cases, the decision was clear (upper left, lower right). In one

14   stand (middle right) the decision to harvest was certain, and the main problem was to decide upon

15   the harvest timing. In the last three stands, both the decision of whether the stand should be

16   harvested or not was uncertain (in part of the scenarios the stand was left unharvested). In one of

17   these cases, the imperfect measurement was sufficient. In two stands, where perfect data was

18   needed, the decisions were more uncertain, for instance, the timing options varied more.

19

20   Varying the relative reduction allowed in the end-inventory from 0 to 30%, the area of measured

21   stands varied, but not in a linear fashion. With a 1% interest rate, the measured area was lowest

22   between 0.09 and 0.15 relative reduction, with 2% there seemed to be an increasing trend in the

23   imperfectly measured area up to 0.18 relative reduction, and with a 5% interest rate the area to be

24   measured was highest from 0.09 to 0.21 relative reduction (Figure 6). The variation was quite

25   modest, though. It depended on how uncertain the timing decisions and cutting decisions actually

1   were, and in the range studied the relative reduction did not affect much to the uncertainties (Figure

2   7). With high interest rates, there could be seen a slight increasing trend with certain harvest timing

3   decisions and a slight decreasing trend in certain unharvest decisions. This means that as the

4   relative reduction increased, part of the certain unharvest decisions changed to uncertain unharvest

5   decisions, part of the uncertain unharvest decisions changed to uncertain or certain timing decisions,

6   and all this affected the measured area. Instead, increasing the interest rate sharply increased the

7   number of certain decisions, i.e., decisions which were the same for all scenarios (Figure 7). With a

8   1% interest rate, no harvest timing decisions were certain, i.e., same in all scenarios.

9

10   The same could also be seen from the stand-level decisions. In the same stands presented before,

11   but with a higher relative reduction allowed, the main change was that in two stands the

12   measurement options was changed, but to opposite directions, one perfect measurement changed to

13   an imperfect measurement and vice versa (Figure 8). This indicated that the optimum in the

14   problem was quite flat, i.e., minor changes in the measurements of single stands could be

15   exchangeable, if the decisions to be made and the uncertainties of these decisions were fairly

16   similar.

17

18   The flatness of the optimum and the uncertainties of the decisions could be clearly seen from the

19   optimal measurement costs and gaps in the optimal solutions (gap = difference between the best

20   integer solution found and upper bound for the optimal solution) were in the same level of

21   magnitude. The gap was highest for the 1% interest rate, where the decisions were most uncertain,

22   and the optimal measurement cost levels as well (Figure 9). It means that the measurement

23   decisions in single stands were uncertain, and many quite good solutions could be found.

24

1   We also calculated the net value of information with 10€/ha cost for different interest rates and

2   relative reductions varying from 0 to 15%. It means that we subtracted from the objective function

3   values where the measurements we allowed (minus the measurement costs) the objective function

4   value assuming that no measurements were possible. At minimum, the net VOI was 4477€ (about

5   65€/ha, 1% interest rate), and at its maximum it was 8144€ (about 121 €/ha, interest rate 3%). The

6   net VOI for a large part is due to being able to harvest in all scenarios all the allowed volume, while

7   if measurements are not allowed the decisions need to be made in all stands according the worst

8   case scenario. The difference between 1% and higher interest rates is clear (Figure 10 left). With

9   1% interest rate there are a little more certain unharvest decisions, and therefore the problem is not

10 as bad as with the higher interest rates. The net VOI with the higher interest rates is also gradually

11 decreasing as a function of relative reduction (Figure 10 right). This means that with more

12 demanding constraints, the net VOI is higher.

13

14 The Pareto optimal solutions with the five different interest rates are shown in Figure 11. In the

15 studies case, the set of Pareto optimal solutions demonstrates an almost linear behavior, as the

16 objectives are very closely related.

17

18 **5. Discussion**
19

20 The traditional forest harvest scheduling problem with an end-inventory constraint can be

21 interpreted to involve two types of decisions: the decisions of timing of harvests during the planning

22 horizon, and decisions of selecting the stands that are left unharvested for the next planning horizon.

23 It can also be interpreted as a multiobjective planning problem with two objectives: the incomes

24 obtained during the planning horizon and the value of the growing stock at the end of the planning

25 horizon.

26

1   In the studied case, we added the decisions of measuring time and measurement quality to the

2   traditional forest harvesting scheduling problem. By using optimization methods to find optimal

3   solutions to the problem, we found out that imperfect information was mainly used for deciding the

4   timing of the harvest, while perfect information was mainly used for selecting the stands that are

5   left unharvested. There are two reasons for this. First, the perfect data in this study was assumed

6   quite expensive, and the benefits obtainable from better harvest timing decisions using perfect data

7   were in most cases lower than this cost. However, since the measurement costs as well as the

8   incomes are discounted, the perfect information becomes cheaper and more profitable towards the

9   end of the planning horizon, when the unharvest decisions are made. Obviously, had the perfect

10   data been cheap enough, all the measurements would have been made with perfect information.

11

12   The second reason is that we required in each case a specific end-inventory value. The loss of

13   incomes if a stand that can be harvested is left unharvested is quite high when compared to the

14   measurement cost even with perfect information. On the other hand, the specific end-inventory

15   demand means that if no information is available, the stand needs to be left unharvested if there is a

16   risk that even in one scenario the resulting solution will not meet the end-inventory demand. Thus,

17   the only way to ensure that all the harvesting possibilities can be used is to measure perfect

18   information. This was the case even if the imperfect information was fairly good in the studied case,

19   i.e., it explained 70% of the prior variation with only a fraction of the cost of the perfect

20   information. However, the need for perfect information can be covered in small stands. In fact, the

21   two smallest stands were always measured with perfect information.

22

23   The situation with a long planning horizon and a specific end-inventory demand is not entirely

24   realistic. In real life, after the first period decisions have been implemented, a new plan would be

25   made, meaning that a rolling horizon is used. Then, a state where the expensive perfect information

1   would be needed is never actually reached. It is probable that if the problem was reformulated using

2   an infinite planning horizon, where the aggregated incomes from an infinite series of planning

3   periods, the profitability of the imperfect information would further improve compared to perfect

4   information. It also means that strict demands are not the best way to model the decisions in

5   stochastic settings.

6

7   Another option to formulate the problem would have been to formulate the end-inventory demand

8   using soft constraints, i.e., penalizing the violation of the constraint but not requiring that the

9   constraint needs to be met in all cases. Then, the profitability of the imperfect information

10 compared to perfect information should depend on the penalty set to the violations: the more costly

11 the violation would be, the higher the profitability of perfect information would be. However, in

12 this approach the difficulty lies in setting the penalty. We would need to define how much the

13 decision maker is willing to pay in the objective function value to meet the constraints. It means

14 that we would need to model what the consequences of violating the constraints are to the decision

15 maker as a simple recourse problem (King and Wallace 2012). Furthermore, using other methods

16 than the ε-constraint method (e.g., involving the achievement scalarizing function (Wierzbicki

17 1982)) for solving the stochastic multiobjective optimization problem (1)-(6) might have led to

18 other results, since here the method used led to dealing with the end-inventory demand as a hard

19 end-inventory constraint. Thus, the effects of the multi-objective problem formulation on the

20 decisions and VOI should be studied more carefully in the future.

21

22 In this study, we assumed two different levels of measurement quality, perfect and imperfect data,

23 where imperfect resolves 70% of the prior uncertainty, and perfect data 100%. The procedure can

24 easily be generalized to more than two data quality options by introducing more variance

25 components, but the number of scenarios needed for the analysis would then necessarily be higher,

1 and the problem more complicated to solve. Then, the formulation of the scenario tree and also the

2 method used to solve the problem need to more carefully analysed. We also assumed that the same

3 proportion of prior variance was resolved for each of the three variables (height, age and basal

4 area), but it would be possible to consider options for improving just one variable at time.

5

6 If the whole area had been measured with perfect information in the beginning, the total cost would

7 have been 6729 €. Measuring the whole area with imperfect information would have cost either

8 672.9€ (10€/ha), or 336.45 (5€/ha). The optimal measurement cost levels varied from 170€

9 (measuring in total 4.22 ha, interest rate 5%) to 378€ (measuring in total 34.14 ha, interest rate 1%)

10 with a 5€/ha cost and from 136€ (measuring in total 3.78 ha, interest rate 5%) to 495€ (measuring in

11 total 23.22 ha, interest rate 1%) with a 10€/ha cost. The net profit from these investments to

12 measuring varied from 4477€ to 8144€ with 10€/ha cost, i.e. at least about 10 times the

13 measurement investment and with interest rate higher than 1% even higher than measuring the

14 whole area with perfect information would have been. The total optimal measurement costs were

15 surprisingly low compared to the measurement of all stands in the beginning, but it was clear that

16 the measurements were postponed to the last possible point (Kangas et al. 2014), and then perfect

17 information was in many stands measured rather than imperfect information. The difference

18 between the two cost levels of imperfect information was also fairly low, but with the higher cost

19 level in some stands the harvest timing decisions were rather made without information. The

20 optimal amount of money to invest in measurements obviously would be dependent on the prior

21 variance, and the number of uncertain decisions.

22

23 We assumed that in a multiobjective optimization situation, different objectives may require

24 different kinds of information, and that different preferences over various trade-offs would affect

25 the profitability of different measurements. In this study, the trade-off was measured with the

1   relative reduction allowed in the end-inventory. Its effect was relatively small, but it can be seen

2   that emphasizing incomes more meant that unharvest decisions could be turned into harvest timing

3   decisions or certain unharvest decisions to uncertain unharvest decisions, and both of these

4   phenomena would affect the measured area and the net VOI. It is probable, however, that if the end-

5   inventory demands were high enough, it would be certain that no stands should be harvested, or if it

6   was low enough, it would be clear that most stands would be harvested. Thus, having different

7   emphasis on trade-offs would have a more profound effect if the ranges analyzed were larger. This

8   remains to be studied in the future. In addition, since the value of unharvested forest depends on

9   similar uncertainties as the other objective, the incomes from the harvested forest, emphasizing

10   them in trade-offs did not affect much the quality of information needed. This might be different, if

11   the problem had more different objectives, e.g., recreational value of the forest that depends on

12   uncertainty in a different fashion.

13

14   On the other hand, also the interest rate can be seen to reflect the trade-offs between the objectives:

15   since the end-inventory value of forests was not discounted, increasing the interest rate decreased

16   the importance of the income objective relative to that of the end-inventory value. In this case, the

17   increasing interest rate also decreased the uncertainty concerning the harvest timing decisions and,

18   therefore, the usefulness of imperfect information decreased.

19

1   **References**
2

3   Birchler, U. and Bütler, M. 2007. Information Economics. Routledge advanced texts in economics

4   and finance. London. 462 p.

5   Birge, J.R. and Louveaux, F. 2011. Introduction to stochastic programming. Second edition.

6   Springer, New York. 499 p.

7   Duvemo, K., Lämås, T., Eriksson, L.-O. and Wikström, P. 2014. Introducing cost-plus-loss analysis

8   into a hierarchical forestry planning environment. Annals of Operations Research 219:415-

9   431. DOI 10.1007/s10479-012-1139-9.

10   Gustavsen, H. and Fagerström, H. 1983. Männyn, kuusen ja koivun muotolukujen vaihtelu. Folia

11   Forestalia 576. 32 p. Metsäntutkimuslaitos. (In Finnish).

12   Haara, A. & Leskinen, P. 2009. The assessment of the uncertainty of updated stand-level inventory

13   data. Silva Fennica 43: 87–112.

14   Haimes, Y.Y., Lasdon, L.S. and Wismer D.A. 1971. On a Bicriterion Formulation of the Problems

15   of Integrated System Identification and System Optimization. IEEE Transactions on

16   Systems, Man, and Cybernetics 1:296–297.

17   Hirshleifer, J. and Riley, J.G. 1979. The analytics of uncertainty and information – an expository

18   survey. Journal of Economic Literature Vol. XVII pp. 1375-1421.

19   IBM ILOG CPLEX Optimization Studio V12.5.1 documentation 2011. URL http://www-

20   01.ibm.com/support/knowledgecenter/SSSA5P_12.5.1/maps/ic-homepage.html.

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

22   Kangas, A., Hartikainen M. and Miettinen K. 2014. Simultaneous optimization of harvest schedule

23   and measurement strategy. To appear in Scandinavian Journal of Forest Research.

24   King, A.J. Wallace, S.W. 2012 Modeling with Stochastic Programming, Springer.

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

26   Miettinen, K. 1999. Nonlinear multiobjective optimization. Kluwer.

1 Movrotas G., 2009, Effective implementation of the ε-constraint method in Multi-Objective
2   Mathematical Programming problems. Applied Mathematics and Computation 213: 455–
3   465.
4 Mustonen, J. Packalén, P. and Kangas, A. 2008. Automatic segmentation of forest stands using
5   canopy height model and aerial photograph. Scandinavian Journal of Forest Research
6   23:534-545.
7 Packalén, P. & Maltamo, M.2007. The k-MSN method for the prediction of species-specific stand
8   attributes using airborne laser scanning and aerial photographs. Remote Sensing of
9   Environment 109:328–341. doi:10.1016/j.rse.2007.01.005
10 Posner, M. E. and Wu, C.-T. 1981. Linear max-min programming. Mathematical Programming 20:
11   166-172
12 R Development Core Team 2011. R: A language and environment for   statistical computing. R
13   Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
14 Vuokila, Y. and Väliaho, H. 1980. Viljeltyjen havumetsiköiden kasvatusmallit. Metsäntutkimus-
15   laitoksen julkaisuja 99.2. 271 pp. (In Finnish)
16 Wierzbicki, A. P. 1982. A mathematical basis for satisficing decision making, Mathematical
17   Modelling 3(5): 391–405
18
19

1   Table 1. The properties of the estate stands

Variable	Age (T)	Basal area (G)	Volume (V)	Mean diameter (Dgm)	Dominant height (Hdom)
Min	10.1	1.5	4.2	2.2	2.4
Mean	63.7	17.0	111.0	14.3	13.3
Max	103.3	37.5	274.1	23.4	21.3
Sd	27.36	9.98	78.12	5.41	5.42

2

3

1
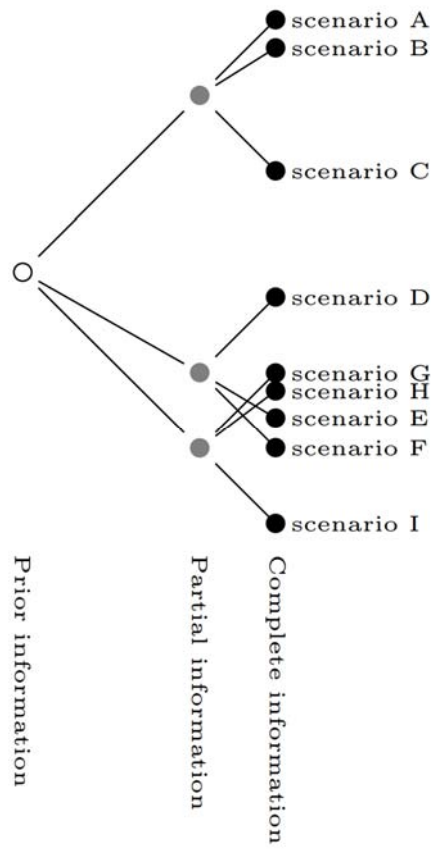
Figure 1. An example of a scenario tree, where there are three scenarios for the partial information and each scenario in partial information has three scenarios in complete information. The lengths of the lines represent the proportion of total variation.

2

3

1
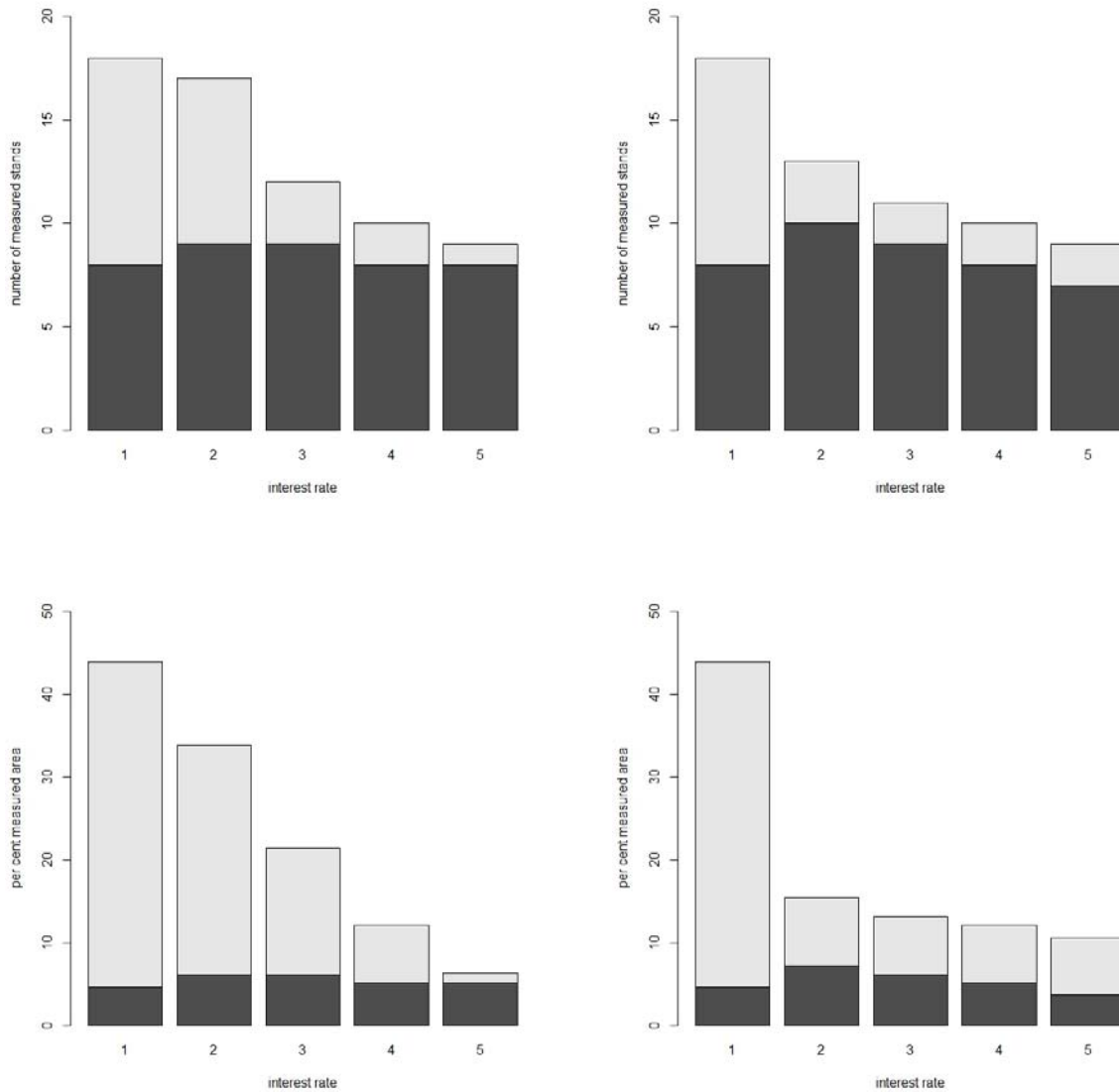
Figure 2. The upper row figures describe the number of stands measured with different interest rates, when the cost of imperfect inventory is 5€/ha (left column) or 10€ (right column), and the inventory cost of the perfect information is 100€/ha. Black color represents perfect information, and grey imperfect information. In the lower row is the proportion of the area measured with the same inventory costs and interest rates. The relative reduction allowed is 0.

2

3
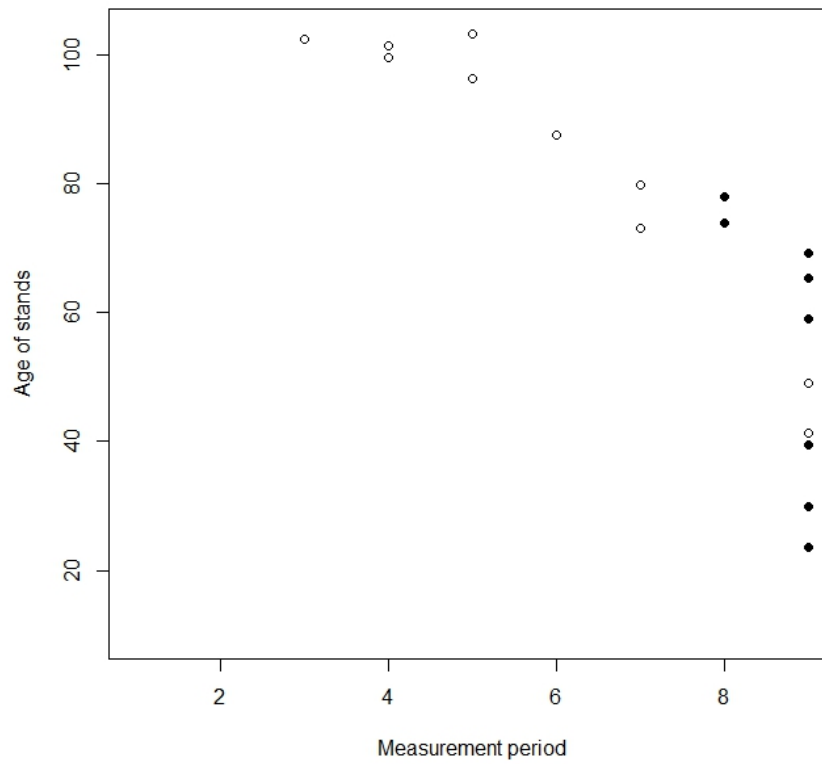
Figure 3. The measurement period as a function of stand age in the beginning of the planning

horizon. Black dots represent perfect measurements, circles imperfect measurements. The interest

rate is 1% and the relative reduction allowed is 0.

1

2
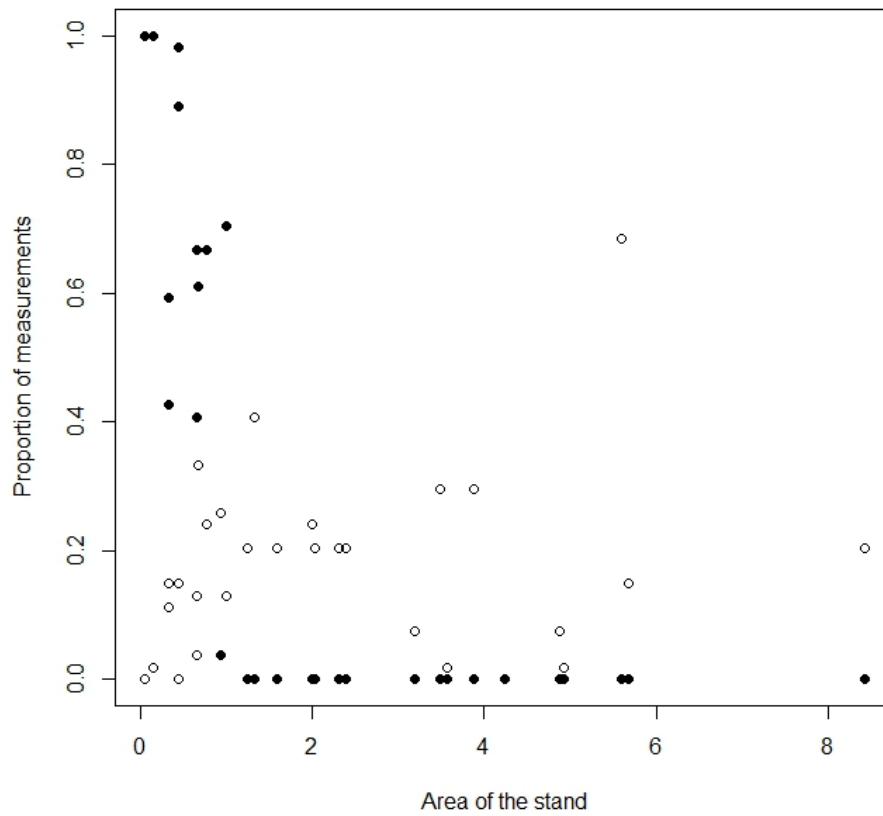
Figure 4. The proportion of cases where each stand was measured as a function of stand area.

Black dots represent perfect measurements, circles imperfect measurements.
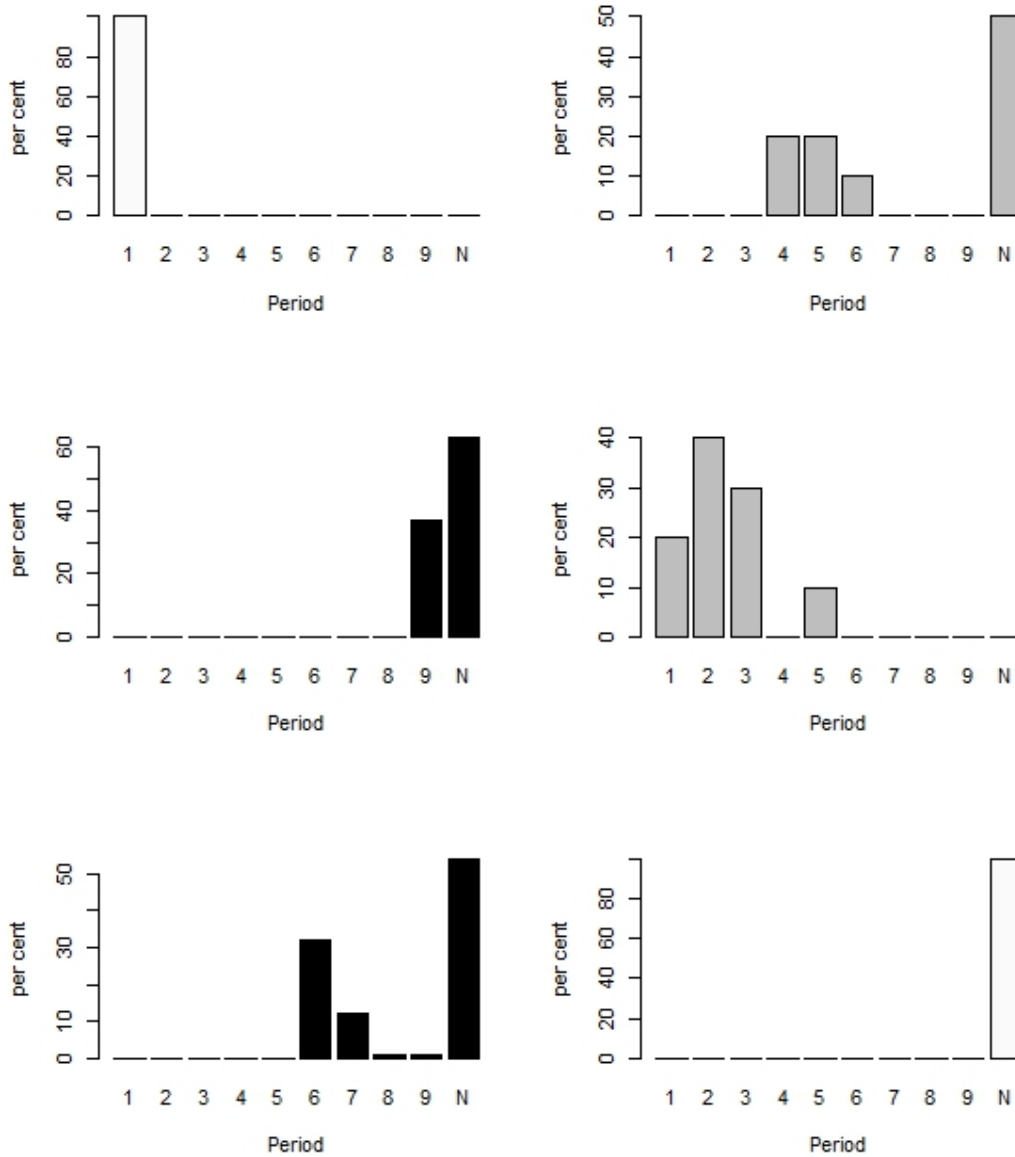
1

2
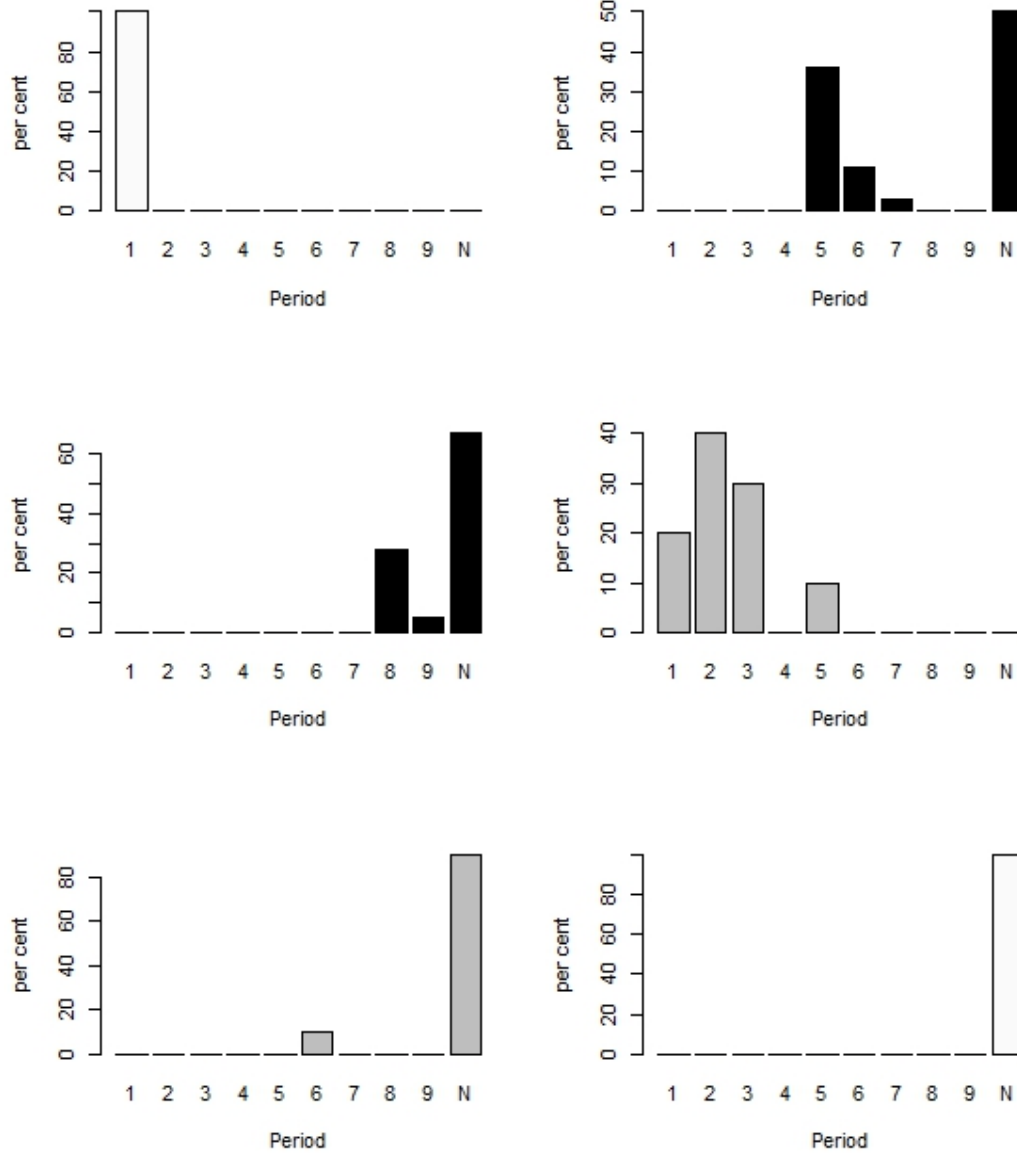
Figure 5. The harvest decisions for stands 2-7. Bars at periods 1-9 mean proportion of harvest decisions among the scenarios at that period. Period N means no harvest. Light grey denote no measurement, grey imperfect information and black perfect information. Interest rate is 2% and the maximum relative reduction allowed is 0.

1

2
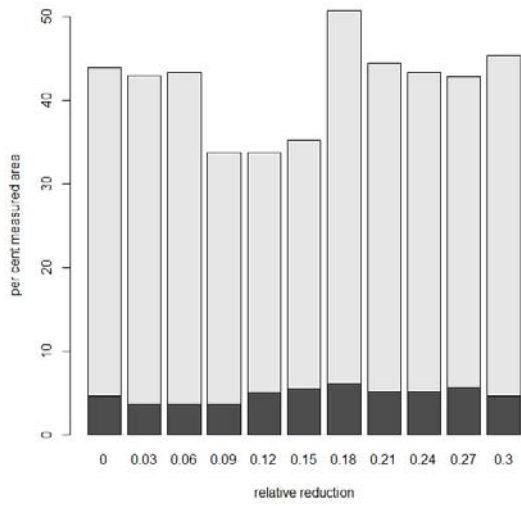
1%
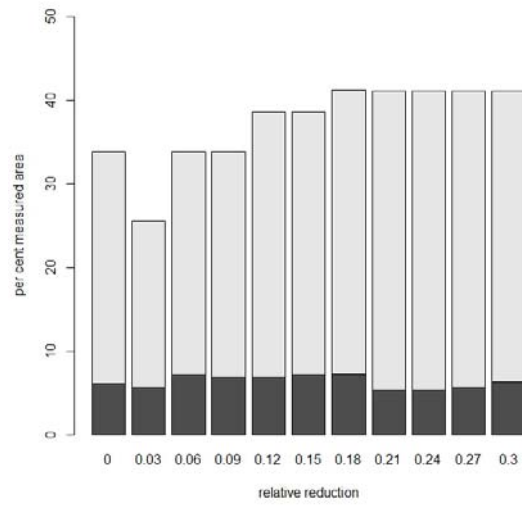
2%
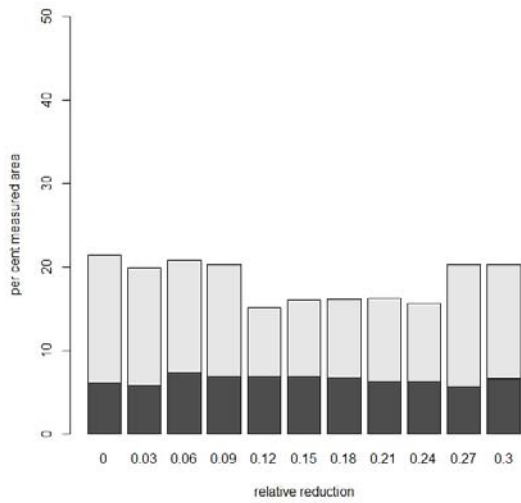
3%
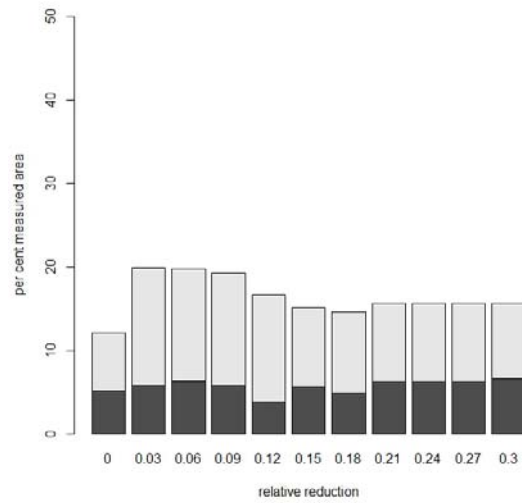
4%
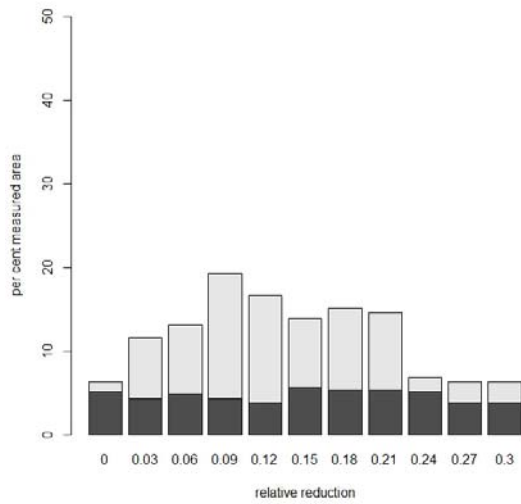
5%

Figure 6. The proportion of measured area when the maximum allowed relative reduction varied from 0 to 30%, with different interest rates. Black color means perfect information, grey imperfect information. Measurement cost of imperfect information 5€/ha.

1

2
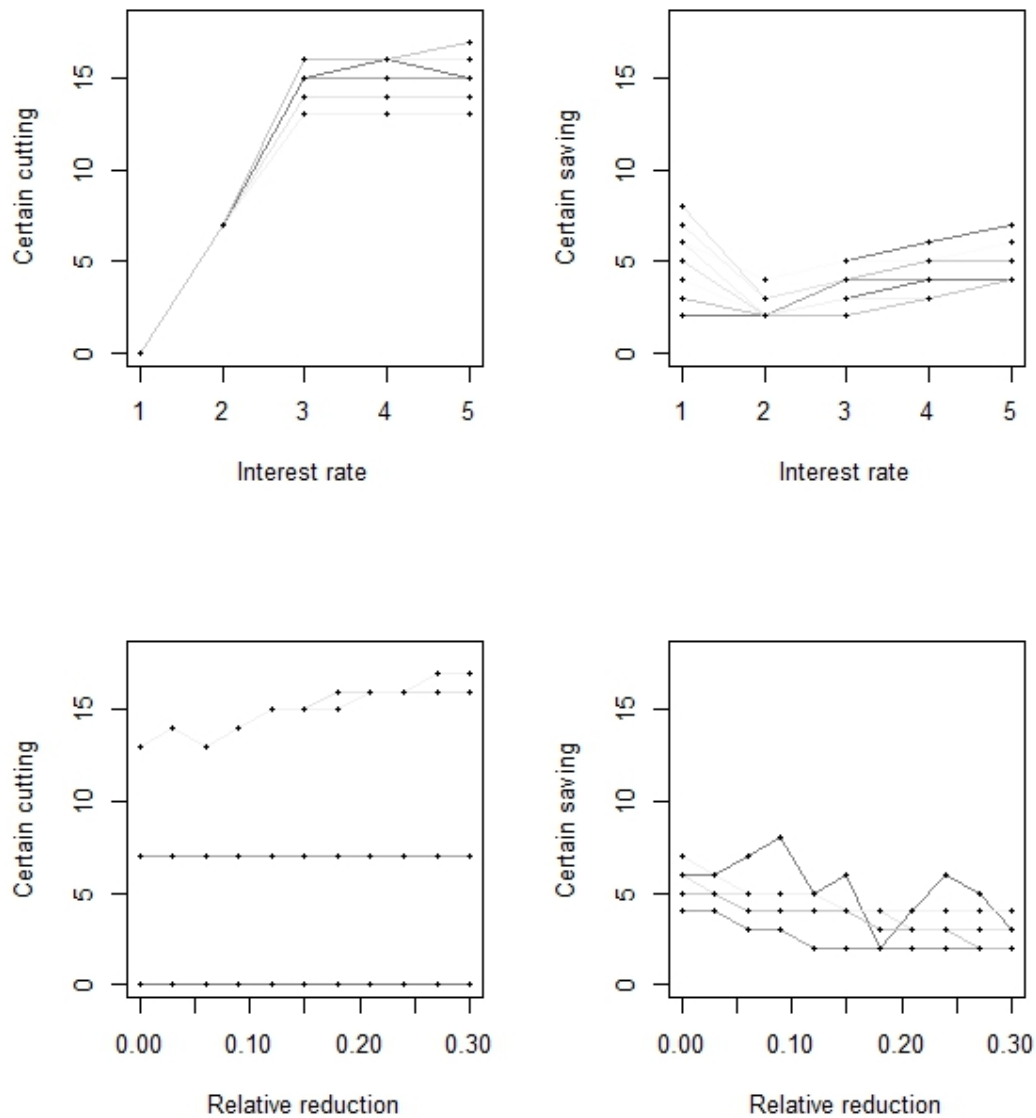
Figure 7. The upper left figure shows the number of certain harvest timing decisions as a function of interest rate. The different lines represent different values for maximal allowed relative reduction. The upper right figure shows the certain unharvest decisions. The lower left figure shows the certain harvest timing decisions as a function of relative reduction. The different lines describe different interest rates (interest rate is increasing from darkest to lightest line). The lower right figure shows the number of certain unharvest decisions.

1

1

Figure 8. The harvest decisions for stands 2-7. Bars at periods 1-9 mean proportion of harvest decisions among the scenarios at that period. Period N means no harvest. Light grey denote no measurement, grey imperfect information and black perfect information. Interest rate is 2% and the maximum relative reduction allowed is 15%.
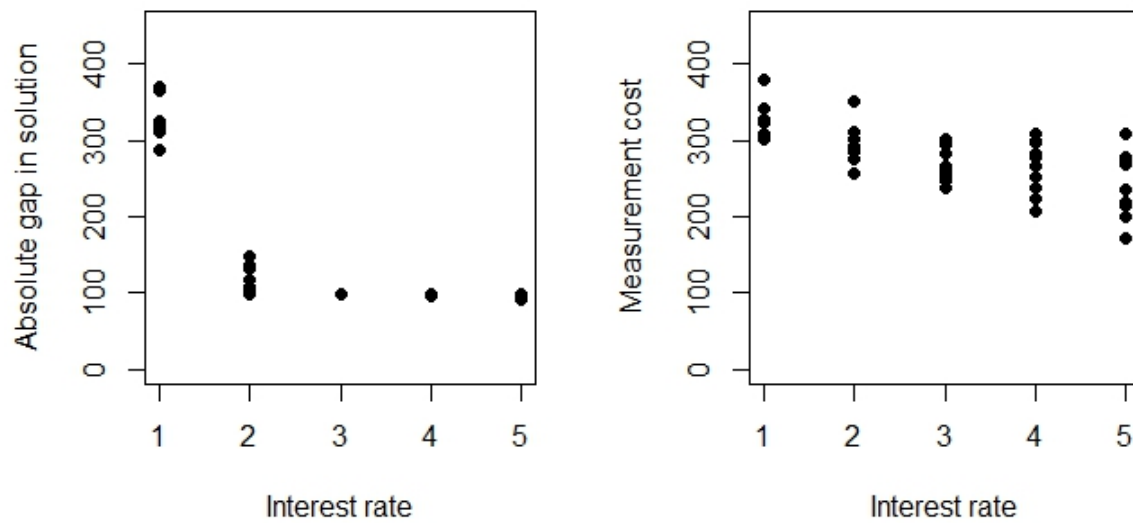
2

3

Figure 9. In the left hand figure is the absolute gap in the solution as a function of interest rate, and in the right hand figure the optimal total measurement costs. The dots describe the different end-inventory requirements.
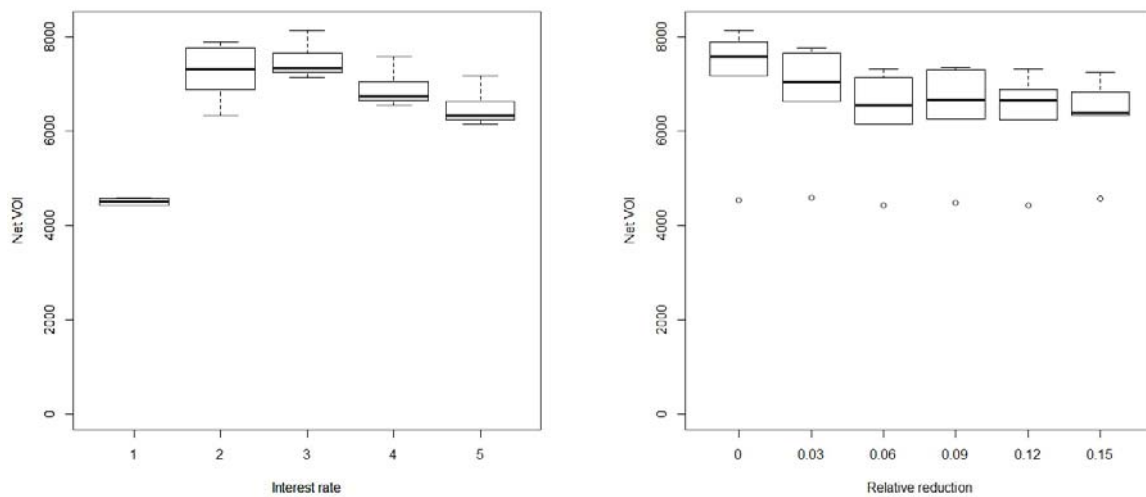
1

2

Figure 10. The net VOI as a function of interest rate (left) and as a function of relative reduction (right). The boxplots describe the variation between different values of relative reduction (left) and different interest rates (right).
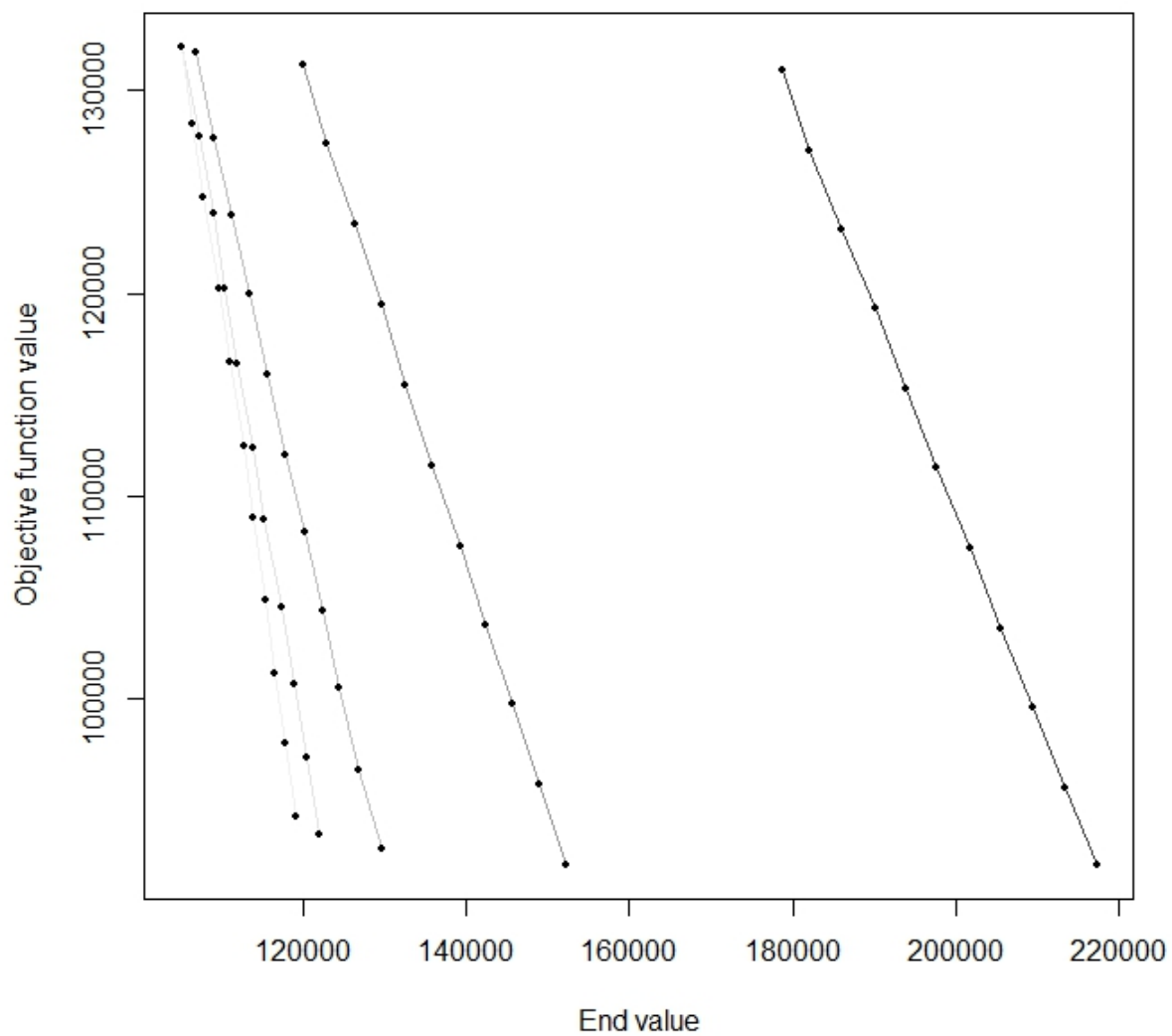
1

2

Figure 11. The Pareto front with the five different interest rates (increasing from the darkest line to the lightest line).

1