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

**Title:** Data-Based Forest Management with Uncertainties and Multiple Objectives

**Year:** 2016

**Version:**

# Data-Based Forest Management with Uncertainties and Multiple Objectives

Markus Hartikainen, University of Jyvaskyla, Department of Mathematical Information Technology, P.O. Box 35. FI-40014 University of Jyvaskyla, Finland, markus.hartikainen@jyu.fi

Kyle Eyvindson, University of Jyvaskyla, Department of Biological and Environmental Science, P.O. Box 35. FI-40014 University of Jyvaskyla, Finland, kyle.j.eyvindson@jyu.fi

Kaisa Miettinen, University of Jyvaskyla, Department of Mathematical Information Technology, P.O. Box 35. FI-40014 University of Jyvaskyla, Finland, kaisa.miettinen@jyu.fi

Annika Kangas, Natural Resources Institute Finland (Luke), Economics and Society Unit, Yliopistokatu 6, FI-80101 Joensuu, Finland, annika.kangas@luke.fi

**Abstract.** In this paper, we present an approach of employing multiobjective optimization to support decision making in forest management planning. The planning is based on data representing so-called stands, each consisting of homogeneous parts of the forest, and simulations of how the trees grow in the stands under different treatment options. Forest planning concerns future decisions to be made that include uncertainty. We employ as objective functions both the expected values of incomes and biodiversity as well as the value at risk for both of these objectives. In addition, we minimize the risk level for both the income value and the biodiversity value. There is a tradeoff between the expected value and the value at risk, as well as between the value at risk of the two objectives of interest and, thus, decision support is needed to find the best balance between the conflicting objectives. We employ an interactive method where a decision maker iteratively provides preference information to find the most preferred management plan and at the same time learns about the interdependencies of the objectives.

**Keywords:** Forest management planning, multiobjective optimization, interactive multiobjective optimization, Pareto optimality, uncertainty

## 1 INTRODUCTION

In forest management, a forest area is typically divided to decision units, so-called stands, which are relatively homogeneous with respect to the age structure and the species composition of trees. Forest management planning means selecting optimal harvest schedules including one or more treatment option(s) and their timing for each

of these stands. The treatment options may include harvesting all the trees (final felling) or a part of them (thinning) within any stand, and planting new seedlings and tending them after a harvest has been carried out. The timing of the treatment options in the schedules is described by dividing the planning horizon to several planning periods of one or more years.

For most decision makers, forest management planning is a multiobjective decision making problem. Harvesting implies incomes from forests but, on the other hand, it diminishes the recreational and esthetical values of the forest, and it may have adverse effects on the natural values of the forest area, for instance, the biodiversity within the area and the viability of wildlife populations living in the area.

Importantly, forest management planning involves a lot of uncertainty as it is not possible to measure all trees within a forest area. This means that there is uncertainty concerning the current state of the forests. Furthermore, measuring biodiversity is prohibitively expensive, so using proxy variables (biodiversity indices) is the only possibility. Finally, all forest decisions concern the future (typically the next 5-20 years), so that the state of the forest stands and the biodiversity as well as the consequences of the treatment options need to be predicted using statistical models. As we do not know the exact consequences of the management decisions, the decisions involve uncertainty which the decision makers may wish to manage.

In this paper, we present an application that accounts for the conflicting objectives in forestry (incomes and biodiversity) and manage the risk involved in them using the value at risk (often denoted by VaR) concept. We apply an interactive multiobjective optimization method to find the most preferred treatment option for each stand considered.

## 2    BACKGROUND AND RELATED WORK

Selecting the optimal treatment schedule for each stand is a combinatorial optimization problem, where the size of the problem depends on the number of separate stands and the number of possible treatment schedules for them. Linear programming has been widely used to solve this problem since the 1960s (Kangas et al. 2015). Most applications have multiple objectives, and have typically been solved using the $\varepsilon$-constraint method (Miettinen 1999). Heuristic optimization has also been used since the 1990s. Within a heuristic framework, multiobjective optimization has typically been based on a weighted (additive) utility function. The main reason for not utilizing multiobjective optimization more has been the lack of computational tools to do so.

Multiobjective optimization problems can in general be formulated as

$$\max\{f_1(x),\ldots,f_k(x)\}$$
$$\text{s.t } g_j(x) \geq 0 \text{ for all } j = 1,\ldots,J$$
$$h_k(x) = 0 \text{ for all } k = 1,\ldots,K$$
$$a_i \leq x_i \leq b_i \text{ for all } i = 1,\ldots,n$$
$$x = (x^i, x^r), \text{ where } x^i \in Z^a, x^i \in R^{n-z}.$$

In the above problem, $f_i$ are objective functions to be maximized simultaneously, $g_j$ and $h_k$ are inequality and equality constraints, respectively, bounds $a_i \leq x_i \leq b_i$ are called box constraints and, finally, the decision variable vector $x = (x^i, x^r)$ consists of integer-valued variables $x^i$ and real-valued variables $x^r$. All the constraints form a feasible set Q which is a subset of the decision space. The image of Q mapped with the objective functions is called the feasible objective set $f(Q)$ and elements of it are so-called objective vectors $z$.

In multiobjective optimization, there typically does not exist a single optimal solution but, instead, there exist multiple so-called Pareto optimal solutions, where none of the objective functions can be improved without impairing at least one of the others. For this reason, one needs additional preference information to decide which of the Pareto optimal solutions is the best one. The person giving this information is called a decision maker (DM).

One way of giving preferences is providing aspiration levels, which are values of objective functions that should desirably be achieved, and they constitute a reference point $z^{ref} \in R^k$. Aspiration levels are employed in so-called achievement scalarizing functions (Wierzbicki, 1986). The main idea behind achievement scalarizing functions is that they measure the preferability of a solution given a reference point in a way that is theoretically justifiable. This is defined as order-consistency by Wierzbicki (1986).

There exist many achievement scalarizing functions following the characterization of Wierzbicki (1986). In this paper, we use the following achievement scalarizing function to be maximized

$$s^{asf} : f(Q) \times R^k \to R,$$

$$(z, z^{ref}) \mapsto \min_{i=1,\ldots,k} (z_i - z_i^{ref}) / (z_i^{ideal} - z_i^{nadir})$$

$$+ \rho \sum_{i=1}^{k} z_i / (z_i^{ideal} - z_i^{nadir}).$$

In the above problem, $z$ is a so-called objective vector in the image space of the feasible set, $z^{ideal}$ is the ideal vector of the problem containing the maximum values of the individual objectives and $z^{nadir}$ is the nadir vector containing the minimum values of the individual objectives within the set of Pareto optimal solutions (see e.g. Miettinen, 1999). The summation term at the end is called an augmentation term and the constant $\rho$ is a small positive constant, e.g. 0.0001. The augmentation term guarantees that the solutions are indeed Pareto optimal.

Employing the achievement scalarizing function means solving the optimization problem

$$\max_{x \in Q} s^{asf} (f(x), z^{ref})$$

for any reference point $z^{ref}$ given by the DM. Given a reference point, the optimal solution of the above problem is a Pareto optimal solution to the original multiobjective optimization problem (Wierzbicki, 1986). The variant of the reference point method (Wierzbicki, 1986) employed utilizes the achievement scalarizing function.

A vast majority of applications within forest management planning assumes a decision situation under certainty. A couple of applications of stochastic optimization

have been published since the 1980s, and most of them in the last 10 years. To our knowledge, any kind of risk management has only been included in two published papers (Eyvindson and Kangas 2016a, Eyvindson and Chen 2016). The latter employed conditional value at risk (Rockafeller and Uryasev 2000). Applications utilizing value at risk (Duffie and Pan 1997) have not yet been published in forest management planning. To our knowledge, applications including two or more different value at risk concepts have not been published in any field. However, in forest planning, it is quite possible that the DM is willing to accept high risks for poor incomes but only a low risk for losing biodiversity or vice versa. The main reason why the uncertainties have been ignored so far is that the problem including uncertainties requires quite heavy calculations, which can be regarded too demanding for large areas with a lot of stands.

In interactive multiobjective optimization methods (e.g. Miettinen, 1999, Miettinen et al. 2008), the final Pareto optimal solution, to be called the most preferred solution, is identified by iterating the steps of re-defining the preferences and producing a solution fulfilling these preferences as well as possible, until the DM is satisfied. The idea is that in this way the DM learns about what kind of preferences are attainable and what kind of solutions are achievable. As mentioned, the preferences can be expressed, for instance by giving aspiration levels. Even though the benefits of interactive methods have been acknowledged in the forest management planning field (Pykäläinen 2000), only a few applications have been published.

## 3    MULTIOBJECTIVE OPTIMIZATION OF FOREST INVENTORY

One complication in forest management planning is that most decisions concern the future. Even when a decision is implemented immediately, the consequences of the actions in the future need to be predicted. This prediction is carried out using a forest growth simulator.

In the Finnish forestry, there are three different operational simulators that are used by all organizations making forest planning. In this study, we used the SIMO simulator, which includes more than 400 equations to predict, e.g., the diameter growth and the height growth of each tree, and the probability that a tree dies in the next years. The simulator also predicts the total volume of timber available from a harvest carried out at a specific time, and the income based on the harvested volume. Forest growth can be predicted fairly accurately for the next 1-5 years, but as the time period becomes longer, the uncertainties grow.

While forest growth can be simulated, biodiversity cannot really be measured in practical forestry. Biodiversity should include genetic variation within species, variation of species within each stand and the variation of habitats (types of forest stands) within the forest area. Of these, only the habitat variation is an operational objective for optimization. It is even more difficult to predict the consequences of different actions on the biodiversity. Therefore, the usual approach is to utilize a proxy variable, a so-called biodiversity (BD) index, in the optimization. The BD indices are

based on the characteristics of the stands, which enables analyzing the consequences of actions also from the biodiversity point of view (Kangas and Pukkala 1996). As these forest characteristics include uncertainty, and using a proxy in itself includes uncertainty, these estimates are highly uncertain.

The stochasticity involved can be dealt with by using a set of scenarios. Parametric distributions are utilized to describe the uncertainty in any one of the input variables of a system, and a parametric distribution can also be used to describe the uncertainty in the simulated growth. The variables of interest, like incomes from harvests, are a result of several input variables and statistical models, and therefore describing the uncertainty is easiest to carry out with a set of scenarios, each of them essentially describing one possible future development of the forest area. Then, the whole set of scenarios describes the uncertainty in the variables of interest. Using a set of scenarios also enables describing the stochastic optimization problem involving uncertainties in a way where linear optimization can be used (a so-called deterministic equivalent of the stochastic problem).

It is possible to evaluate the quality of the set of scenarios used to describe the stochastic problem, through the use of evaluation techniques. The so-called Sample Average Approximation (SAA, Kleywegt et al. 2001) method compares the solution generated by a smaller number of scenarios to the solution generated by a much larger number of scenarios. This is iterated several times to generate confidence intervals of the gap in optimality and expected objective value of the solution. One application of this method has been realized in forestry, and the number of scenarios required to effectively represent the uncertainty depended on the quantity of uncertainty and risk preferences of the DM (Eyvindson and Kangas 2016b).

## 4 MODELLING FOREST MANAGEMENT AS A MULTIOBJECTIVE MIXED INTEGER LINEAR PROBLEM

We assume that we have S stands, T treatment schedules for each stand, R scenarios and P periods of time to consider. We have simulated values $I_{t,s,r,p}$ for the income and $B_{t,s,r,p}$ for the biodiversity with indices denoting

- treatment schedules $t$ (including one or more timed treatments),
- stands $s$,
- scenarios of future development of the forest stand $r$
- and 5-year periods $p$.

The problem of choosing the best treatment schedules for all the stands can be formulated as a multiobjective optimization problem

$$\max\left\{\min_{p\in P, r\in R_I}\sum_{t\in T, s\in S} x_{t,s} I_{t,s,r,p}, \quad \min_{p\in P, r\in R_B}\sum_{t\in T, s\in S} x_{t,s} B_{t,s,r,p}, \quad \mathop{E}_{r\in R}\sum_{t\in T, s\in S} x_{t,s}\min_{p\in P} I_{t,s,r,p}, \quad \mathop{E}_{r\in R}\sum_{t\in T, s\in S} x_{t,s} B_{t,s,r,p}, \delta_I, \delta_B\right\}$$

s.t. $R_I \subset R, \#R_I \geq \delta_I \#R,$

$\qquad R_B \subset R, \#R_B \geq \delta_B \#R,$

$\qquad \sum_{t\in T} x_{t,s} = 1$ for all $s \in S,$

$\qquad x_{t,s} \in \{0,1\}$ for all $t \in T, s \in S.$

In the above problem, # denotes the number of elements in a set and we have six objectives to be maximized:

1. Minimum income in the scenarios that are in the set of scenarios $R_I$ (in euros). This objective will be denoted by $VaR_I$ in what follows. According to the first constraint, the set $R_I$ is a subset of the complete set of scenarios $R$ and the number of scenarios in this subset is greater or equal to $\delta_I$ times the number of scenarios in the complete set of scenarios. This means that this is the income at risk for the risk level $1 - \delta_I$.

2. Minimum biodiversity index in the scenarios that are in the set of scenarios $R_B$. This objective will be denoted by $VaR_B$. Being an index, this variable is unitless, and only the relative differences can be interpreted. According to the second constraint, the set $R_B$ is a subset of the complete set of scenarios $R$ and the number of scenarios in this subset is greater or equal to $\delta_B$ times the number of scenarios in the complete set of scenarios. This means that this is the biodiversity at risk for the risk level $1 - \delta_B$.

3. Expected minimum income across the periods p in the complete set of scenarios $R$ (in euros). This objective will be denoted by $E_I$.

4. Expected minimum biodiversity index across the periods in the complete set of scenarios $R$. This objective will be denoted by $E_B$.

5. Probability $\delta_I$ of the set $R_I$. The risk level for the value at risk for the income is, thus, $1 - \delta_I$.

6. Probability $\delta_B$ of the set $R_B$. The risk level for the value at risk for the biodiversity is, thus, $1 - \delta_B$.

Traditionally in forestry, the variation of income over periods has been handled by seeking a so-called even-flow solution. In an even-flow solution, the income is equal over all periods. With biodiversity, stability is even more important. We, however, maximize the minimum income and biodiversity over periods, instead of seeking the even-flow solution. The solution that maximizes the minimum income or biodiversity over the periods is sometimes the even-flow solution, if all the periodic incomes are the same as the minimum income, or it is better than the best available even-flow solution, if the income or the biodiversity in one of the periods is higher than the min-

imum income or biodiversity. For this reason, we believe that it makes more sense to maximize the minimum over the periods than to seek for the even-flow solution.

Our decision variables in the problem formulated are both the sets $R_I$ and $R_B$ and the treatment decisions $x_{t,s}$ for all treatment schedules t and stands s. The treatment decision $x_{t,s}$ is one, if the treatment schedule is chosen for the stand s and 0 if not. Because only one treatment schedule can be chosen for each stand, the third constraint allows only one of the $x_{t,s}$ values to be one, and the others must be zero for all the stands.

In order to solve the optimization problem efficiently, it can be converted into a mixed integer linear problem (MILP). Once we can re-formulate the problem as a MILP, we can use the efficient MILP solvers available e.g., IBM ILOG CPLEX Optimization Studio (see e.g., http://www-01.ibm.com/support/knowledgecenter/ SSSA5P_12.6.3/ilog.odms.studio.help/Optimization_Studio/topics/COS_home.html) or Gurobi Optimization (see e.g., http://www.gurobi.com/documentation/). Our problem can be re-formulated as a multiobjective MILP

$$\max\left\{ I_{min,subset}, \ B_{min,subset}, \ \sum_{r\in R}\sum_{t\in T, s\in S} x_{t,s} \frac{\min\limits_{p\in P} I_{t,s,d,p}}{\#R}, \ \sum_{r\in R}\sum_{t\in T, s\in S} x_{t,s} \frac{\min\limits_{p\in P} B_{t,s,d,p}}{\#R}, \ \delta_I, \delta_B \right\}$$

$$\text{s.t.} \ \sum_{r\in R} t^I_r \geq \delta_I \#R,$$

$$\sum_{r\in R} t^B_r \geq \delta_B \#R,$$

$$\sum_{t\in T} x_{t,s} = 1 \text{ for all } s \in S,$$

$$t^I_r \in \{0,1\} \text{ for all } r \in R,$$

$$t^B_r \in \{0,1\} \text{ for all } r \in R,$$

where

$$I_{min,subset} = \min_{p\in P, r\in R}\left( \sum_{t\in T, s\in S} x_{t,s} I_{t,s,r,p} + (1-t^I_r)M \right)$$

$$B_{min,subset} = \min_{p\in P, r\in R}\left( \sum_{t\in T, s\in S} x_{t,s} B_{t,s,r,p} + (1-t^B_r)M \right).$$

In the re-formulated problem, the constant M is a big number that allows for the minimum over the scenarios R in both the biodiversity and income to be the minimum over the scenarios, for which the variable $t^I_r$ or $t^B_r$ has the value one. Because of the two first new constraints, the new variables $t^I_r$ and $t^B_r$ must be one for the ratio of scenarios given by their respective probability variables $\delta_I$ and $\delta_B$.

The re-formulated problem is, however, computationally extremely expensive to solve when the number of stands is high. For this reason, we approximate it with a problem where the decision variables for the treatment schedules $x_{t,s}$ are allowed to take any real value between 0 and 1, instead of being binary variables. This is a common approximation in forest management. The interpretation of treatment schedules with non-binary values is that a part of the stand is treated with a different schedule. For this reason, we can do this approximation and still get solutions which can be implemented as treatment schedules for the stands. This leaves us with a multiobjec-

tive optimization problem, with both integer ($t^I_r$ and $t^B_r$) and real-valued ($x_{t,s}$) variables.

# 5 EXPERIMENT AND RESULTS

## 5.1 An overview of the experiment

We have conducted an experiment and Figure 1 shows the data flow in it. We first need to acquire the data (1), which is estimated for a set of 0.04 ha segments of a forest area. To make the data more manageable, this data is segmented (2) into fairly homogenous stands. To include uncertainty into the model, randomization (3) is included into the forest data. With each instance of the stand level data, the forest stand is simulated (4) to predict future forest resources according to different treatment schedules. Once the future resources are predicted and different scenarios are created, we approximate (5) the ideal and nadir vectors using optimization. Then we can start the (6) solution process with an interactive method using reference points from a DM. Once the DM is satisfied, we produce a list of treatment schedules for all stands, and we can implement the plan for the forest area.

In what follows, Section 5.2 describes the data, segmentation of the forest area, randomization of the data and simulation. Section 5.3 oulines single-objective optimization and interactive multiobjective optimization methods applied. The output of our experiment are preferred treatment schedules for the stands that take into account the conflict between income and biodiversity and the DM's preferences and handle the uncertainties inherent to the problem.

## 5.2 Forest Inventory Data with Uncertainties

The forest inventory data has been acquired through the combined use of remote sensing and statistical models based on a field plot measurement. The forest data, obtained from the Multi-source National Forest Inventory (MS-NFI) for 2011,was available in a raster dataset with a pixel size representing a 20 x 20 m footprint, provided by the Natural Resources Institute Finland (Luke), available from http://kartta.luke.fi/index-en.html. Our area of interest is a large forest area (8,415 ha) to the north of the city of Jyväskylä in Finland. The entire forest area has been segmented into a set of 2,842 stands where each stand represents a relatively homogenous forest area. The segmentation is based on a mainly cloud free Landsat image (LC81880162013155LGN00), utilizing an algorithm developed by the Technical Research Centre of Finland (VTT). In the experiment, we only used 300 stands, and 50 scenarios for the whole forest area. The forest-level scenarios were constructed by randomly selecting one of 25 stand-level scenarios for each stand, resulting in a problem with a size similar to a 15,000 stand problem. The selection of 50 scenarios was based on previous research (Eyvindson and Kangas 2016b).
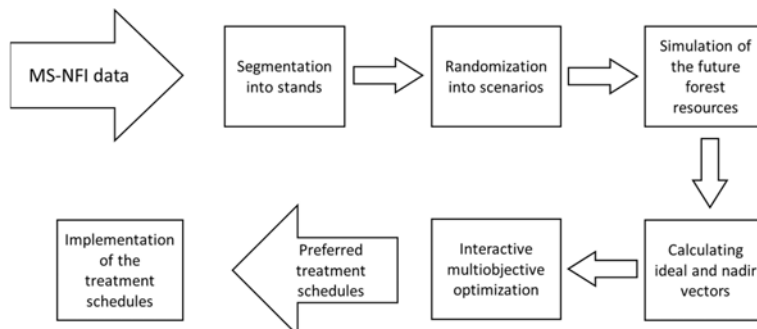


**Figure 1:** Data flow in the conducted experiment

As collected, the data are estimates of the expected value of the forest resources, and the actual forest resources may differ from the expectation. Two sources of uncertainty were included: initial inventory errors and forest stand age. We assumed that the inventory errors for the height, basal area (a measure of the forest density) and forest stand age were normally distributed, with a mean of 0 and a standard deviation equal to 10% of the mean of both variables. These estimates of errors reflect the current state-of-the-art inventory methods for Finnish forest conditions (Mäkinen et al. 2010).

The simulation of future forest resources was conducted using an open-source forest simulation and optimization software (SIMO, Rasinmäki et al. 2009). For each stand, SIMO generates a list of possible treatment scenarios which could be used in the forest during the planning period. They include silvicultural (i.e. planting, tending and fertilizing) and harvesting options (i.e. final felling or thinning). The maximum number of possible sets of treatment schedules for all stands was 34. For each stand, a total of 25 simulations were run for all treatment schedules. Each simulation was run with a random adjustment to the error estimates. As uncertainty is not spatially correlated (the estimate of one stand does not influence the estimate of another stand), one possible realization of what is contained in the forest is a random selection of all stands. With this number of simulations, there are $25^{300}$ possible combinations of different treatment schedules for the stands. Thus, going through all of them would not be possible and optimization is needed.

In this experiment, in order to estimate the biodiversity of the forest, a weighted combination of habitat suitability indices of red-listed saproxylic species for boreal forests was used (Tikkanen et al. 2007). These models require an estimate of dead-wood in the forest, which is especially difficult to measure using remote sensing techniques. The amount of dead-wood was estimated based on average quantities of dead-wood in Finland (Tomppo et al. 1999) and the age of the forest stands. Two functions were used to estimate the quantity of deadwood, one to represent the increase of dead-wood from a middle aged forest to an old-aged forest, while the other represented the decrease of deadwood from a recently harvested stand to a middle aged one.

### 5.3 Decision making

The decision making environment has been implemented using a Jupyter Notebook, which has been made freely available at https://github.com/maeehart/MOD2016. The data used in decision making is also available at the same repository. The data of the stands is represented as a text file with the format required by the IBM® ILOG® CPLEX® Optimization Studio. The DM involved was an expert in forest management planning. The experiment was run on a sixteen core Intel Xeon E5-2670 processor with 64 GB of RAM. The computer was running Ubuntu Linux 14.04.3 LTS.

The multiobjective optimization problem of forest management planning was modeled using the Optimization Programming Language (OPL) of IBM. The problem was not, however, directly converted into a text file using OPL but, instead, the essential components (i.e., objectives, constraints and decision variables) of the problem were included in a Python dictionary. This is because OPL does not directly support

multiobjective optimization, but the problems need to be scalarized (in this case with the achievement scalarizing function) in order to be completely modelled with OPL and then to be solved with IBM® ILOG® CPLEX® Optimization Studio.

Before starting the decision making process, one had to estimate the ideal and nadir vectors of the problem. This was done using the pay-off table approach (Miettinen 1999), where one first optimizes each objective individually, evaluates the values of all objectives at solutions obtained and forms the k-dimensional ideal vector from the best value of each objective and estimates the components of the nadir vector from the worst values obtained. The pay-off table for the problem considered is given in Table 1. In the pay-off table, the rows represent objective function values calculated at the solution where an objective function obtained the best value (ideal on the diagonal) and the components of the nadir vector are the worst values of each column. The biggest values in the table are written in bold face and the smallest values are underlined. The ideal vector is, thus

$$(VaR_I, VaR_B, E_I, E_B, \delta_I, \delta_B) = (5246601,\ 520,\ 4828656,\ 520,\ 1,\ 1)$$

and the nadir vector is all zeros.

**Table 1. Pay-off table for the problem considered**

| $VaR_I$ | $VaR_B$ | $E_I$ | $E_B$ | $\delta_I$ | $\delta_B$ |
|---|---|---|---|---|---|
| **5246601** | 358 | 4275987 | 358 | <u>0</u> | 1 |
| 0 | **520** | <u>0</u> | **520** | 1 | <u>0</u> |
| 3588782 | 398 | **4828656** | 399 | 0 | 1 |
| <u>0</u> | 519 | <u>0</u> | **520** | 0 | 1 |
| 0 | <u>0</u> | <u>0</u> | <u>0</u> | 1 | 1 |
| 0 | <u>0</u> | <u>0</u> | <u>0</u> | 1 | 1 |

The Pareto optimal solution corresponding to the reference point specified by the DM was obtained by solving the achievement scalarizing function introduced earlier. The achievement scalarizing function was implemented using a Python function, which compiles a string, which is then written into a text file. The text file was, finally, solved by calling the IBM® ILOG® CPLEX® Optimization Studio and attaching the data files to the call.

Before asking the DM for any reference point, a so-called neutral compromise solution (Wierzbicki 1999) for the problem was computed and shown to the DM together with the ideal and the nadir vectors. The neutral compromise solution was

$$(4595853, 456, 4631499, 456, 0.54, 0.68).$$

The neutral compromise solution gives information about the objective function values that are roughly in the middle of Pareto optimal solutions. This solution means that

- there is a $100\% - 54\% = 46\%$ chance that income is worse than 4595853€,
- there is 32% chance that the biodiversity is worse than 456; and that
- the mean income and biodiversity are 4631499€ and 456, respectively.

The DM found the solution quite good. However, she wanted to improve the security on the biodiversity and was willing to give up on the income. Thus, she gave a reference point

$$(2500000, \ 504, \ 2600000, \ 505, \ 0.7, \ 0.85).$$

In it, the confidence on the biodiversity is higher and both the mean and value at risk for the biodiversity are larger, and the values for income are worse than in the neutral compromise solution.

The solution of the achievement scalarizing function corresponding to the reference point was

$$(2545940, \ 509, \ 2642280, \ 510 \ 0.8, \ 1).$$

This Pareto optimal solution is better in all objectives than the reference point. Thus, the DM was really happy. In the reference point method, the DM specifies a new reference point as long as she wishes to continue iterating.

As the DM wanted to still see whether it would be possible to improve the expected income at the cost of the reliabilities, she gave a new reference point

$$(250000, \ 508, \ 3000000, \ 509, \ 0.7, \ 0.85),$$

which has a higher value for the expected income. However, this reference point led to a Pareto optimal solution

$$(2913004, \ 507, \ 2986231, \ 508, \ 1, \ 0.98),$$

which means that the higher expected income came at the cost of biodiversity instead of the reliabilities.

This was not what the DM wanted. For this reason, she wanted to make one more check whether the income could be improved at the cost of reliabilities and gave one more reference point

$$(2000000, \ 504, \ 3000000, \ 510, \ 0.7, \ 0.85),$$

where she had improved the expected biodiversity and allowed the value at risk for the income to get worse. This resulted in a Pareto optimal solution

$$(2755038, \ 507, \ 2982154, \ 508, \ 0.98, \ 1).$$

The DM was satisfied with this solution, because it is better than the reference point that she gave in all of the objectives except the expected biodiversity and income, and even in these objectives, the value is very close. In addition, the DM was really happy that the probability of the biodiversity being over the value at risk is 100%. For this reason, the DM chose this solution as the final, most preferred solution.

The interpretation of this final preferred solution is that

- expected values of the minimum income and minimum biodiversity (i.e., minima over the periods) are $2982154 \, €$ and $508$;
- there is only a two percent risk that the minimum income is smaller than $2755038 \, €$; and
- the minimum biodiversity is guaranteed to be over $507$ with 100% probability.

This was very much what the DM was hoping for, because the expected values are at satisfactory levels and risk levels in the value at risk are really low (i.e., 2% and 0%). Overall, the interactive solution process enabled the DM to get convinced of the

goodness of the final solution and gain insight about the interdependencies of the objectives.

## 6    CONCLUSIONS

We have presented an approach of using interactive multiobjective optimization to handle the conflict between both

- income and biodiversity and
- risk and expected outcomes

in forest management planning. Our approach is based on

- simulating forest growth using standard models,
- modelling uncertainty using a scenario-based approach,
- converting the decision problem into a six-objective optimization problem, and
- using a reference point based interactive method to find preferable treatment schedules for the forest stands.

In the solution process, we applied the reference point method of Wierzbicki (1986) in its simplest form. Thanks to the interactive nature of the method, the DM could learn about the interdependencies between the objectives and the attainability of her preferences. Thanks to this, she could gain confidence of the final solution selected.

From the DM's point of view, an interactive method is essential. As the biodiversity is a unitless index, interpreting the results is only possible in relative terms. Relating the current solution to the previous solutions is needed in order to be able to construct a set of preferences.

Having the risk level and the value at risk at that particular risk level both as objectives for both the incomes and biodiversity further emphasizes the need for an interactive approach. It would be very difficult to give pre-defined hopes (like weights) for these variables. It may be easy enough to set e.g. weights when there is only one expected value and the value at risk at stake, but the setting would require quite an abstract analysis unless the task can be carried out in iterative fashion enabling learning.

A drawback of the solution process was that the computation times with the given data for a single given reference point were rather long, ranging from less than one day to almost five days. This is a problem when using interactive methods, when the DM may get tired in waiting for new solutions to be computed (as argued e.g., in Korhonen and Wallenius 1996). This is a major challenge to extend this method to scale, which is needed with large forests. Large forests may contain up to hundreds of thousands of stands, while our data contained only 300 stands. In many cases, such forest properties can be divided into parts that can be handled independently from each other, but still there is a need to make the method less time-consuming. One way could be the use of hierarchical planning. For instance, Pantuso, Fagerholt and Wallace (2015) have developed a method of solving hierarchical stochastic programs for a maritime fleet renewal problem. They were able to solve large problems in a hierarchical framework, which could not be solved using CPLEX alone.

# 7    ACKNOWLEDGEMENTS

# 8    REFERENCES

Eyvindson, K. & Cheng, Z. 2016. Implementing the conditional value at risk approach for even-flow forest management planning. Can J For Res. 46(5): 637-644.

Eyvindson, K. and Kangas, A. 2016a. Evaluating the required scenario set size for stochastic programming in forest management planning: incorporating inventory and growth model uncertainty. Canadian Journal of Forest Science. 46(3):340-347.

Eyvindson, K. and Kangas, A. 2016b. Integrating risk preferences in forest harvest scheduling. Annals of Forest Science. 73:321-330.

Kangas, A., Kurttila, M., Hujala, T., Eyvindson, K. and Kangas, J. 2015. Decision support for forest management. Second edition. Managing forest ecosystems, Vol 30. Springer. 307 p.

Korhonen, P. and Wallenius, J. 1996 Behavioural Issues in MCDM: Neglected Research Questions. Journal of Multi-Criteria Decision Analysis, 5:178–182.

Miettinen, K. 1999. Nonlinear Multiobjective Optimization. Kluwer. 298 p.

Miettinen, K., Ruiz, F. and Wierzbicki, A.P. 2008. Introduction to multiobjective optimization: Interactive approaches, in "Multiobjective Optimization: Interactive and Evolutionary Approaches", Ed. by Branke, J., Deb, K., Miettinen, K., Slowinski, R., Springer-Verlag, Berlin, Heidelberg, 27-57.

Mäkinen, A., Kangas, A. and Mehtätalo, L. 2010. Correlations, distributions and trends of inventory errors and their effects on forest planning. Can. J. For. Res. 40(7): 1386–1396.

Kangas, J. and Pukkala, T. 1996. Operationalization of biological diversity as a decision objective in tactical planning. Canadian Journal of Forest Research 26:103-111.

Kleywegt, A, Shapiro, A. and Homem-de Mello, T. 2001. The sample average approximation method for stochastic discrete optimization. SIAM Journal on Optimization, 12(2):479-502.

Pantuso, G, Fagerholt, K. and Wallace S.W. 2015. Solving hierarchical stochastic programs: Application to the maritime fleet renewal problem. INFORMS Journal on Computing, 27(1), 89-102.

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

Rasinmäki, J., Mäkinen, A. and Kalliovirta, J. 2009. SIMO: an adaptable simulation framework for multiscale forest resource data. Comput. Electron. Agric. 66(1):76–84.

Tikkanen, O. P., Heinonen, T., Kouki, J. and Matero, J. 2007. Habitat suitability models of saproxylic red-listed boreal forest species in long-term matrix management: cost-effective measures for multi-species conservation. Biological Conservation, 140(3), 359–372.

Tomppo, E. 1999. Keski-Suomen metsäkeskuksen alueen metsävarat ja niiden kehitys 1967–96. Metsätieteen aikakauskirja 2B/1999: 309–387. [in Finnish]

Wierzbicki, A. P. 1986. On the completeness and constructiveness of parametric characterizations to vector optimization problems. OR Spectrum, 8(2):73-87.

Wierzbicki, A.P. 1999. Reference point approaches, in: Gal, T., Stewart, T. J., Hanne, T. (Eds.), Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory, and Applications, Kluwer Academic Publishers, Boston, 9-1–9-39.