The Effect of Forest Information Quality on the Planning and Decision Process in Forestry

Rami Saad

Faculty of Forest Sciences
Department of Forest Resource Management
Umeå

Doctoral Thesis
Swedish University of Agricultural Sciences
Umeå 2017

Acta Universitatis Agriculturae Sueciae
2017:11

Cover: Mid-Swedish forest landscape
(photo: T. Lämås)

ISSN 1652-6880
ISBN (print version) 978-91-576-8795-1
ISBN (electronic version) 978-91-576-8796-8

Print: SLU Service/Repro, Uppsala 2017

The Effect of Forest Information Quality on the Planning and Decision Process in Forestry

Abstract

Managing the forests is important for the provision of timber, environmental benefits, recreation, etc., to both forest owners and society. In order to manage the forests, information on the state of the forest is required. However, forest information is not free from errors, that is, it suffers from some degree of uncertainty leading to suboptimal decisions and potential economic losses.

Remote sensing data are frequently used to acquire forest information. In this thesis the use of remote sensing techniques was elaborated further. In Paper I airborne laser scanning (ALS) data was used to estimate stand stem diameter distributions and, thus improving forest information compared to traditional stand mean values. In Paper II ALS data was used first as auxiliary data when using the local pivotal method (LPM) and a micro-stand approach for locating reference plots. Second, ALS data was used as auxiliary data when imputing forest information to evaluation plots. The combined approached showed a potential for improvement and has the potential to be a competitive method when considering cost efficiency.

Improving forest information can be done through acquiring new information or through assimilating new with old information. Paper III presents the benefits of data assimilation process. It provides more accurate estimates as compared to traditional methods and it also provides the associated uncertainty. Paper IV presents a visual illustration method to incorporate estimates of uncertainty in forest planning. The method is applicable in current decision support systems (DSSs) for use in stand level decision making situations.

The results of this thesis may also shed light on the reason why uncertainty so far is typically ignored in forest planning. Taking the results and the potential benefits of this thesis forward could lead to the development of new DSSs considering uncertainty. Moreover, the data assimilation process should be further investigated, as it is a promising framework for assimilating different sources of information into a single useful source of information for forest planners and decision makers.

Keywords: forest management planning, suboptimal loss, Heureka, decision support system, local pivotal method (LPM), segmentation, most similar neighbor (MSN) imputation, remote sensing, uncertainty, data assimilation, risk preferences, stochastic optimization

Author's address: Rami Saad, SLU, Department of Forest Resource Management, Skogsmarksgränd, SE-901 83 Umeå, Sweden
E-mail: rami.saad@ slu.se

Dedication

I dedicate the fruit of my work to my parents whom I truly love. My virtuous father Emil, and my virtuous mother Fatina, you are a blessing in my life. (Psalms 128).

Then Samuel took a stone and set it up between Mizpah and Shen and called its name Ebenezer; for he said, "Till now the Lord has helped us."

1 Samuel 7:12

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Saad, R., Wallerman, J. & Lämås, T. (2015). Estimating stem diameter distributions from airborne laser scanning data and their effects on long term forest management planning. *Scandinavian Journal of Forest Research* 30(2):186-196.

II Saad, R., Lämås, T., Wallerman, J. & Holmström, J. (2016). Local pivotal method sampling design combined with micro stands utilizing airborne laser scanning data in a long term forest management planning setting. *Silva Fennica* vol. 50 no. 2 article id 1414.

III Saad, R., Eyvindson, K., Gong, P., Lämås, T. & Ståhl, G. (2017). Potential of using data assimilation to support forest planning. *Canadian Journal of Forest Research*. DOI: 10.1139/cjfr-2016-0439.

IV Eyvindson, K., Saad, R., & Eriksson, L.O. Incorporating stand level risk management options into forest decision support systems (submitted manuscript).

Papers I-III are reproduced with the permission of the publishers.

The contribution of Rami Saad to the papers included in this thesis was as follows:

I Planned the study with the co-authors. Carried out most parts of the calculations and analyses and wrote major parts of the manuscript.

II Planned the study with the co-authors. Carried out major parts of the calculations and analyses and wrote major parts of the manuscript.

III Planned the study with the co-authors. Compiled major parts of the content included and wrote parts of the manuscript.

IV Elaborated the initial idea, carried out parts of the calculations and analyses and wrote parts of the manuscript.

Abbreviations

ALS	Airborne laser scanning
DA	Data assimilation
DSS	Decision support system
NPV	Net present value

1 Introduction

Forests are important natural resources that provide the forest owner and society with timber and non-timber products and thereby forests involve large monetary and other social values. Therefore it is important to manage the forests to provide a high utility to forest owners and society.

Information on the present state of the forest serves as the basis for the forest planning process and therefore plays an important role in planning and decision processes. However, forest information is not free from errors; therefore these errors are considered a source of uncertainty in the forest planning process. Uncertainty stemming from forest information or other sources typically leads to suboptimal decisions in forest planning (Duvemo and Lämås 2006; Pasalodos-Tato et al. 2013). Therefore, considering and reducing uncertainty in forest information is of major importance for forest planners. Furthermore, the need for accurate (higher quality) forest information is increasing, in particular because forests may have multiple uses involving several spatiotemporal scales (Duvemo and Lämås 2006; Kangas 2010; Duvemo et al. 2014).

This thesis concerns the effect of forest information quality on planning and decision processes; hence, the effect of different forest alternatives comprised of field data and remote sensing data used in forest planning were evaluated in terms of timing of actions and suboptimal losses.

1.1 Forest planning and DSS

The process of planning includes determining certain goals, managing actions and evaluating in order to achieve these goals. Similar processes occur in forestry, where the goal might be, for instance, maximizing timber production (Davis and Liu 1991; Randhawa et al. 1996; Pukkala 1998; Edwards and Steins 1999; Kazana et al. 2003; Leskinen et al. 2009). However, forests pose a series of serious challenges to the forest planner that turns this kind of planning into a unique process that requires special attention. Such challenges would be, for example, complex growth models and inaccurate information of the current state. The complexities in forest planning can also potentially stem from the conflicting multiple goals, such as simultaneously maximizing economic value and forest conservation objectives for biodiversity purposes. Nevertheless, the most typical goal stated by forest owners is to achieve a high and sustained economic yield – often expressed in terms of net present value (NPV) – given restrictions based on nature conservation and other considerations. The result of the planning process is a series of management actions that satisfy the stated goal(s).

Forest planning generally considers at least two time horizons: long term and short term planning (Bettinger et al. 2009; Davis et al. 2001; Eriksson 2008). The long term planning process is an important step towards enabling sustainable forest management (i.e., that today's forest management does not restrict possibilities of forest use in the future). This step requires consideration of several aspects such as nature conservation, carbon sequestration and balancing the volume, species composition, and assortment distribution of harvests. Moreover, the results of long term planning for the near future are actually the targets of short term planning; therefore, there is a strong linkage between long term planning and short term planning. In short term planning, in contrast to long term planning, the focus is shifted to facilitate timber procurement and logging procedures and that can be achieved by planning the harvest schedule for single stands. The industrial demands are heavily dependent on this step since the forest industries decide on the potential delivery of wood products beforehand (Bettinger et al. 2009; Davis et al. 2001; Eriksson 2008). In this thesis the effect of forest information quality was tested only in a long term forest planning setting (100 years).

Since the 1970s (e.g. the US Timber RAM (Navon 1971) and FORPLAN systems (Alston and Iverson (1987))), computerized decision support systems (DSSs) became important tools for the support of forest planners in the planning process. The development of forest DSSs is an active research area, one example being the Heureka system (Gordon et al. 2013; Borges et al. 2014; Wikström et al. 2011) developed at the Swedish University of Agricultural Sciences (SLU). It enables forest planning for both long and short term, however, the system is developed more with long term planning in mind. Moreover, there are several successful systems similar to Heureka, for instance, the MELA system in Finland (e.g., Redsven et al. 2013), SGIS in Norway (e.g., Rørstad et al. 2010) and SADfLOR in Portugal (e.g., Garcia-Gonzalo et al. 2015).

In the planning procedure, the Heureka DSS is used to simulate and evaluate different possible treatments, then maximize a goal stated by the user, such as maximum NPV subject to economic and environmental restrictions. In order to compute the NPV as a consequence of different actions and to evaluate them, forest information (either in terms of stand mean values such as basal area, number of stems, mean diameter and height, etc. or as individual tree data) is required to be imported into the DSS. The Heureka DSS was used in the studies included in this thesis where computation of NPV was required.

1.2 Forest information (field survey and remote sensing)

Forest information is used in the planning process in order to predict future states and to determine optimal future actions. Typically forest information is often organized in databases, where relatively homogenous parcels of forested land are aggregated into stands. The stand is a concept used in forestry as both a description and treatment unit for a piece of homogenous forest land. Thereby forest information is typically contained in forest maps and can be found in forest stand databases. This type of information typically consists of stand level values in terms of mean values, for instance, mean age, mean stem number, and mean basal area per hectare. Nevertheless, it is not always straightforward to identify what information is most important for the forest planner; it is, however, common to assume that at least stem volume, variation of the stem diameter, stand age, and tree species composition are essential for the decision maker as these factors are closely related to the value of timber and pulp wood in a given forest stand (Bettinger et al. 2009; Davis et al. 2001; Eriksson 2008). Contrary to stand mean values, forest information can be in the form of individual tree data making up tree lists with stated species and size for each tree. This type of information is seldom available in stand databases. DSSs often use individual tree models in the calculations, therefore if only stand mean values are available these values are transferred to tree lists using built-in models in the DSSs, typically based on assumptions on stand stem diameter distributions. As these stem distribution models are quite rough a potential improvement is to provide diameter distributions based on field surveys or remote sensing.

Stand level forest information compiled in stand register databases is gathered using different methods, for instance, ocular methods (subjective choice of sample units and few guiding measurements) or field sample plots (objective approach). Estimation based on ocular methods typically contain systematic and random errors while estimates based on objective sampling, such as plot sampling, can be expected to contain random errors only and the data acquired are typically considered to be unbiased. Stand register databases are commonly updated on a 5 to 10 year rotation, or prior to planned management actions. In the last decade, the use of the remote sensing techniques, such as airborne laser scanning (ALS; McRoberts et al. 2010), for acquiring forest information is increasing rapidly and, subsequently, the challenges of using these data in an efficient way are also increasing. Remote sensing techniques, such as aerial photography or especially ALS, have the potential to capture high quality information for forest planning purposes (Gobakken & Næsset 2004; Næsset et al. 2004; McRoberts et al. 2010). This information is generally found to outperform traditional sources of information, such as data based on ocular methods and old and forecasted stand register data, for management planning.

Compared to traditional ocular estimates typically providing relative random errors of 20%, for example, for basal area and timber volume (Ståhl 1992), remote sensing techniques provide relative random errors of 3-15% (paper II). Recently, nationwide ALS campaigns in countries such as Denmark, Switzerland, the Netherlands, Finland and Sweden have been conducted and promoted in order to, among other purposes; provide remote sensing data to forest planners at a low cost (Nilsson et al. 2016).

ALS data can be used for estimating stand variables, either as stand mean values or values for individual trees for each stand. ALS data can be acquired at a low number of laser pulses per area unit (≤5 pulses per m^2) or at a higher number of laser pulses per area unit (typically >5 pulses per m^2). In general the low number of laser pulses is used for the area-based method (providing forest data per raster cell, Næsset 2002) while a high number of pulses enables the estimation of individual tree data (e.g., Persson et al. 2002; Solberg et al. 2006; Breidenbach et al. 2010).

Stand delineation is an important procedure aimed at reducing variation within the delineated units by aggregation of similar forest areas. Historically, forests have been delineated manually into stands, either by field survey or aerial photo interpretation. Since stand delineation is performed manually and subjectively, a high variation in forest characteristics within stands is often encountered. ALS data were found to be a data source for forest planning via automatic delineation of stands as (e.g. Olofsson and Holmgren (2014)). Stand delineation using ALS (Koch et al. 2009) potentially reduces the variation within the delineated stands, thus improving the forest information. Stand delineation using automatic delineation algorithms has also led to the introduction of the concept of "micro-stands" (Hyvönen et al. 2005; Pippuri et al. 2012). Micro-stands in general are homogenous areas, but could be smaller than traditional stands. However, in forest planning new challenges can appear if the treatment units become smaller than traditional sizes.

Until recently, when new forest information has been acquired, the old information was discarded. However, a new approach to merge data, termed Data assimilation (DA), allows the use of different sources of information acquired at different time points. In its essence, DA is a process which can merge data from different sources into a single usable source. DA has recently been advocated as a tool for improving the information from forest inventories (Czaplewski & Thompson 2008; Ehlers et al. 2013; Nyström et al. 2016). The DA procedure has the potential to increase the quality of forest information by assimilating all new sources of information (typically data from remote sensing) with existing information (typically forecasted data from field surveys).

1.3 Uncertainty in forest planning

It is difficult, if not impossible, to acquire forest information free from errors. These errors typically lead to suboptimal decisions and therefore are considered to be one of the sources of uncertainty in the forest planning process. Moreover, in a broader perspective, the sources of uncertainty in forest planning can be categorized as belonging to one of the following three sources: (i) uncertainty in forest information, (ii) uncertainty in projection models and thus the future forest state, and (iii) uncertainty in exogenous factors such as timber prices, preferences of the decision maker, climate change and risks associated with natural or human caused external events or disasters (Pasalodos-Tato et al. 2013; Yousefpour et al. 2012). The three sources of uncertainty have the potential to affect the forest planning process negatively. In the four studies which make up this thesis only uncertainty in forest information was addressed. However, uncertainty in forest information, such as forest inventory data, has the potential to affect other sources of uncertainty. That is, forest inventory data of high quality can make natural hazard risk assessment, such as risk of wind throw, more accurate. In addition, growth prediction results are negatively affected by low quality inventory data and other sources of uncertainty. Studies of uncertainty however, often consider a single source of uncertainty where in reality uncertainties are often related. Contrary to a single source of uncertainty Holopainen et al. (2010) consider different uncertainties (forest inventory, growth model and timber prices) simultaneously given different interest rates in order to determine the importance of different source of uncertainties on the NPV. Another example is a study by Mäkinen et al. (2012) where a cost-plus-loss analysis approach is applied to determine the optimal inventory interval where uncertainties stemming from both forest inventory and growth predictions are considered simultaneously. Considering different sources of uncertainties simultaneously was not addressed in this thesis.

Several studies have been performed on the topic of uncertainty in forest information, such as the review studies by Duvemo and Lämås (2006) and Kangas (2010), and this uncertainty has been found to have an important effect in the planning process. The uncertainty in forest information can cause inaccurate estimates to be made and therefore lead to wrong management actions and timing of actions. Wrong management actions have the potential to lead to economic losses. Nevertheless, the topic of uncertainty in forest information has so far not received proper attention (Duvemo and Lämås 2006) since this topic has been found to be complex. Kangas (2010) also emphasizes the complexity of the topic and suggests methods to improve the use of the available forest information (not to improve the forest information directly).

In forest information, one common way to control or reduce uncertainty is to simply carry out a new inventory (Ståhl et al. 1994). Carrying out a new inventory is a research topic in itself as there are several available methods to do this, such as by using remote sensing techniques combined with field survey (Næsset 2002, Gobakken and Næsset 2004). Remote sensing has become an essential part of forest inventory since it improves the accuracy of estimated forest variables compared to traditional data acquisition methods, such as ocular field surveys. Moreover, remote sensing can be utilized as auxiliary information in the sampling design and therefore improves the quality of forest information. However, it is not clear which inventory method is the optimal; therefore cost-plus-loss analysis could be used in this case where the direct cost of carrying out the inventory is added to the expected loss from imperfect information. The cost-plus-loss technique has been employed in several studies to evaluate the effects of uncertainty in forest information (Eid 2000; Holmström et al. 2003). Moreover, the mentioned DA process is a new approach to processing of forest information, and has the potential to reduce uncertainty in forest information (Ehlers et al. 2013).

In case no new inventory is carried out to improve data quality, securing an even harvest volume and balanced age class distribution are considered ways to handle uncertainty in forest planning (Hahn et al. 2014). Another way of controlling uncertainty in an indirect way is to add risk premium to the discount rate when maximizing the NPV since it reduces the effects of outcomes far into the future and therefore reduces uncertainty.

Another option is to use the errors in the available information; however, the errors are rarely known but can often be estimated or approximated through parametric distributions. If the errors are estimated then sensitivity analysis, scenario analysis or Bayesian decision theory could be used (Ståhl et al. 1994). Applications of stochastic methods for instance, where the uncertainty is integrated in the planning model, can be difficult due to intricate mathematical programming (Pukkala 1998; Pasalodos-Tato et al. 2013) and the ability of traditional mathematical programming methods, such as linear programming, to account for uncertainty. Nevertheless, in the last study of this thesis an attempt was made to develop a simple application that takes into account the estimated error in the forest planning.

1.4 Objectives of the thesis

The aim of this PhD thesis is to improve the forest planning process by improving forest information and by suggesting ways to use this information in a forest DSS. Improving forest information was conducted either through field surveys and remote sensing (Papers I and II) or through applying the DA procedure (Paper III). Improving the decision processes in the forest DSS was conducted through proposing methods to deal with uncertainty in the forest DSS (Paper IV). It will throughout be assumed that the forest owners' objective is maximum NPV.

The specific objectives of Papers I-IV were:

Paper I. The purpose of the study was to estimate diameter distributions using ALS information and to determine if these distributions notably improved decision making in terms of reduced suboptimal losses compared to traditional methods of simulating tree lists from stand mean values.

Paper II. The purpose of the study was to test the efficiency of (i) the new sample design referred to as the Local Pivotal Method (LPM; Grafström et al. 2012), combined with a micro-stand approach and then compared with a (ii) traditional layout of reference plots in a systematic grid over the forest area.

Paper III. The objective of this study is to explore, highlight and discuss the potential benefits in forest management planning of using DA processes in forest inventories. The information provided by the DA process contains novel features, but there are also challenges in applying DA.

Paper IV. The purpose of this study was to examine methods of incorporating risk and uncertainty in a forest DSS at stand level. The methods are illustrated by taking uncertainty stemming from measurement errors of the initial state as a case. For the stochastic approach, we highlighted two different approaches of managing risk besides the expected value approach.

2 Paper summary

2.1 Paper I - Improving data quality by utilizing ALS in forest information estimation

Data obtained from ALS are frequently used for acquiring forest data in terms of stand mean values. DSSs typically use individual tree models in their calculations in case stand mean values are the input data to the DSS. These individual tree models are required to simulate tree lists from the stand mean values.

In Paper I ALS information was used also to estimate stem diameter distribution and tree lists were then elaborated and imported to the DSS. Thereby the rough model within the DSS to simulate tree lists could be avoided. Moreover, the improvement in terms of reduced inoptimality losses of this approach compared to a mean value approach was analyzed in a long term (100 years) planning setting.

2.1.1 Field survey and laser data

The study was performed in a boreal forest area in northern Sweden. A field survey was performed in 2008 and 2009 in which all stands (in total 124) were surveyed using 2–15 (mean 7.33) circular sample plots in each stand. Plot radii for the stands included were 10 m (117 stands) and 5 m (7 stands). The stem diameter at breast height and species of all trees on the plots were registered. The height and age of at least three trees on each plot were also registered (Table 1).

Table 1. *Characteristics of the stands used in the study according to the field survey (124 stands, total area 1,135 hectares).*

Variable	Mean	Minimum	Maximum
Area (ha)	9	0.14	66.7
Age (year) 2)	59 1)	20	169
Stem volume (m^3 ha^{-1})	146 1)	24	569
Stem diameter 2) (cm)	19.72 1)	11.27	34.2

1) Area weighted mean, stand area as the weight.
2) Basal area weighted within stand.

The surveyed area was scanned using the ALS system TopEye (S/N 425) carried out using a helicopter platform (August 2008). Flying height was 500 m above ground and pulse intensity was approximately 5 pulses per m^2. The height above ground was determined for all returns, using a digital elevation model produced from the classified ALS data. A set of fundamental ALS metrics was then computed from the ALS data in accordance to the area-based method (Næsset

2002). These metrics correspond to the height above ground (i.e., the 10th, 20th, …, 90th, 95th and 100th height percentiles), mean and standard deviation of height above ground.

2.1.2 Methodology

In order to analyze the usefulness of diameter distributions estimated from ALS data three alternatives were used in this study. The first alternative was acquired through a sample plot field survey of 124 stands which is referred to as "the observed alternative". The second alternative was based on the ALS metrics. Stand mean values estimated from the second alternative (corresponding to traditional stand register information) made up the third alternative, termed later as "the mean values alternative".

Based on the observed data and the ALS data, functions estimating plot level forest variables including stem diameter distribution, in the form of a Weibull distribution (two parameters), were elaborated. Along with the ALS metrics the proportion of basal area of pine was also used as it turned out to be an important variable. In order to estimate the parameters of the Weibull distribution, linear regression was employed (after applying stepwise regression) where the dependent variables were the variables in the observed alternative and the independent variables were the ALS independent variables and the proportion pine.

An essential step in the processing of the ALS alternative was the generation of tree lists. This was achieved by using the fitted Weibull distribution parameters to generate a diameter distribution for each plot and incorporating the fitted number of stems per hectare (estimated for each plot separately).

2.1.2.1 Accuracy measurement

The estimated diameter distribution accuracy was determined using two error indices, computed for each stand separately using the diameter classes' absolute differences. The first error index (e) gives one measure of the degree of the diameter distribution errors, in which the total number of the trees is taken into account. Its value can range between 0 and 200, where 0 represents a perfect match between two compared distributions. The second error index (δ) termed the total variation distance index (Levin et al. 2009), measures the degree of the diameter distribution errors that are independent of the total number of trees. Each diameter class in each stand was divided by the total number of stand trees in order to obtain a diameter probability distribution. The value of index δ can range between 0 and 1, where 0 represents a perfect match of two compared distributions.

2.1.2.2 Suboptimal loss

The observed alternative and ALS alternative were imported as tree lists, while Heureka simulated tree lists for the mean value alternative. In Heureka, a set of potential management alternatives is generated. A management alternative is a sequence over time of management actions such as regeneration, thinning and final felling. Each action has a calculated net cost or income, and an NPV is calculated for each potential management alternative. Then for each stand the alternative providing the highest NPV is selected. The optimal management strategies selected for the ALS and mean values alternatives were applied to the forest information in the observed alternative. The differences between the NPV of the observed alternative and the NPV of the applied programs on the forest information in the observed alternative were considered to be the suboptimal losses. The applied treatment programs were fixed only for the two first periods (10 years) since it is expected that in the future new and better information is probable after a period of time (Holmström et al. 2003).

2.1.3 Results

The ALS alternative provided a better match to observed diameter distributions compared to the mean values approach. The ALS alternative had about 35% lower error indices and these results are in line with the previous studies (Table 2).

Table 2. *Summary of error indices indicating the accuracy of diameter distributions estimated using the ALS and mean values approaches compared to the measured diameter distributions.* e_{ALS} *and* $e_{Heureka}$ *are Reynold indices (range 0 – 200), while* δ_{ALS} *and* $\delta_{Heureka}$ *are total variation distance indices (range 0 -1) for the ALS and mean values approaches, respectively. The index value 0 in both indices present perfect matches of the compared distributions.*

	Error indices			
	Reynolds index		Total variation distances index	
	e_{ALS}	$e_{Heureka}$	δ_{ALS}	$\delta_{Heureka}$
Mean	50.896	79.160	0.251	0.388
Maximum	123.529	159.191	0.542	0.777
Minimum	23.348	39.021	0.090	0.145
Standard deviation	17.454	25.262	0.088	0.122

Using the Heureka DSS the suboptimal losses in terms of net present value due to erroneous decisions were compared (Table 3). Although no large difference was found, the ALS approach showed smaller suboptimal loss than the mean values approach. These suboptimal losses were calculated based on two different pulpwood and sawn timber price lists. The first price list was the default price list used by Heureka and second was the hypothetical price list. The latter price list was constructed to emphasize the increase in the sawn timber as the log diameter size increased, where in the default price list the sawn wood price was not very sensitive to the size of the log diameter.

Table 3. *Calculated NPVs (3% real discount rate).* $NPV_{Observed}$ *is the NPV of the observed alternative.* NPV_{ALS} *and* NPV_{Mean} *are the NPV based on the forest information in the observed alternative where the two first period's management alternatives from the ALS and mean values alternatives were applied on the observed alternative, respectively. The difference between* NPV_{ALS} *and* NPV_{Mean} *is considered to be the suboptimal loss when ALS information is utilized.*

	NPV results (SEK ha^{-1})			
	$NPV_{Observed}$	NPV_{ALS}	NPV_{Mean}	Decrease in suboptimal loss utilizing the ALS information compared to the mean values alternative
Default price list	38,824	38,778	38,712	66
Hypothetical price list	34,139	34,090	33,979	111

2.1.4 Discussion

The ALS derived tree lists yielded smaller - although not large - suboptimal losses than the lists generated from stand mean values. Thus, in addition to providing robust estimates of stand characteristics such as tree height and basal area, ALS can provide valuable estimates of diameter distributions, thereby improving forest planning. Furthermore the use of error indices also showed that the stand level ALS-based tree lists were closer to the observed diameter distributions than the Heureka derived tree lists.

A potential way to further improve the approach is to use non-parametric methods to estimate plot level diameter distributions, as described by Gobakken and Næsset (2005) and Maltamo et al. (2009). In such a case no parametric diameter distribution is assumed (in contrast to our assumption of Weibull distributions), and in operational applications today, imputation techniques, based for instance on kMSN methods (Maltamo et al. 2009), are usually applied. In this approach, predictions are made using the actual diameter measurements in the reference data and no smoothing or distribution assumptions are needed. Such methods can be further evaluated in future studies to assess their potential for improving data to be used in forest DSSs.

In conclusion, the results of the study indicate that ALS-based estimates of diameter distributions have the potential to further improve the planning process, although in this study the gain in NPV was not very high. Use of ALS data should reduce losses from suboptimal decisions, but the level of reduction depends on, for example, the design of timber price list.

2.2 Paper II - Improving data quality by utilizing ALS in the sampling process

Estimation of forest variables obtained through the use of remote sensing requires field reference plots. The allocations of the reference plots are a major challenge since these reference plots are supposed to represent the field in an efficient way. Traditionally allocations of reference plots were performed systematically within the field or within pre-stratified traditional stands. However, a high variability in the forest variables within the forest area and stands is often encountered. For this reason, acquiring a representative sample through reference plots is a difficult task. Recent developments in automatic segmentation using ALS data for forest delineation (Olofsson and Holmgren 2014) have led to introduction of the micro-stand approach which can potentially reduce variation within the forest units. Methods that can potentially reduce the variation within the forest treatment units (i.e., stands) are of major interest to test and therefore increase the forest information quality.

In Paper II a new sampling design, named the local pivotal method (LPM) (Grafström et al. 2012), was combined with the micro-stand approach and compared with the traditional systematic sampling design for estimation of forest stand variables. The LPM uses the distance between units in an auxiliary space to obtain a well spread sample, i.e., it is unlikely that two similar units (in the auxiliary space) will be selected in the sample and therefore guarantee a well spread sample.

The ALS data were used in this study for three main steps. First for delineating the stands (Olofsson and Holmgren 2014), second the ALS were used as the auxiliary space in order to obtain a well spread sample once the LPM (Grafström et al. 2012) was utilized, and third for estimating the forest information using the non-parametric method MSN (Moeur and Stage 1995).

The effect of the new sampling design on long term forest planning (100 years) was evaluated through relative root mean square error (RelRMSE), stem diameter distribution error index and suboptimal losses.

2.2.1 Field survey and laser data

This study was performed using data obtained from the Remningstorp estate located in southern Sweden. This forest holding covers about 1200 ha of productive forest land. ALS data acquisition was performed by scanning from a helicopter on August and September 2010 with a Riegl LMS-Q560 system. The area was scanned from a flight height of approx. 400 m above ground level and the pulse density was at least 10 measurements per square meter. The metrics derived using the height distribution of laser returns from ALS were as follows: (1) average height (AH), (2) standard deviation of heights (SDH), (3) vegetation ratio (VR), (4), average crown height (ACH), (5) 10th height percentile (P10), (6) 50th height percentile (P50), and (7) 90th percentile (P90).

Two sets of surveyed reference plots were acquired by the two sampling designs and used for imputing data to evaluation plots. The first set of reference plots, acquired by LPM, made up four imputation alternatives (with varying number of reference plots) and the second set of reference plots, acquired by systematic sampling design, made up two alternatives (with variable plot radius).

In the LPM sampling design, the forest area was first delineated into micro-stands based on ALS data using the segmentation algorithm developed by Olofsson and Holmgren (2014). Five meter radius plots were allocated to a sub-sample of 100 micro-stands selected using the LPM (Grafström et al. 2012). In this study, the LPM was applied using the standardized ALS metrics AH, SDH, VR, ACH, P10, P50, and P90 for measuring the distance between units, i.e., the selected micro-stands were spread in the auxiliary data (ALS metrics) space. The probability of selecting a particular micro-stand was set proportionally to the sum of ACH for all raster cells within the micro-stand because this sum was correlated to the total stem volume. The reason for this selection was to ensure that a sufficient amount of reference data were obtained in micro-stands with a high economic value. The number of 5 m circular sample plots surveyed in each selected micro-stand was selected proportionally to the variation of P90 within the micro-stand in order to allocate more plots in micro-stands with a high variation of tree height. Thus, the number of 5 m plots was not necessarily evenly distributed among the micro-stands but depended on the variation of P90 in each micro-stand. The LPM was also utilized in a second step to locate (geographically) plots inside each micro-stand using the same ALS metrics as used in the first step to ensure that a high variation was captured. Within the LPM sampling design all trees with a diameter at breast height larger than 4 cm were calipered and the species recorded, but no other data were collected. Heights of the calipered trees on the plots were estimated using allometric models relating tree height to diameter (Söderberg 1992). The plots (originally 981 plots in total) serving as the ALS training data were allocated over the original 100 micro-stands. The number of plots actually surveyed was 881 because nine micro-stands were excluded from study. This number was later reduced to 856 as some plots were considered to be outliers when analyzing the field survey data and the ALS data relationships (Table 4).

The systematic sample included 10 m radius plots surveyed in 2010 and located in a 200 m × 200 m grid over the major part of the estate, resulting in 263 plots located on forest land. The plots were surveyed using the methods and models estimating the state of the forest available in the Heureka system by employing the modules Ivent and PlanStart (Wikström et al. 2011). All trees greater than 4 cm in diameter at breast height were callipered and a sub-sampling of trees to measure height and age was performed. Heights of the remaining calipered trees on the plots were estimated using allometric models relating tree height to diameter (Söderberg 1992). In total, the final data set used consisted of 216 plots. Table 4 provides a summary of the data collected.

The 40 m radius evaluation plots, measured between 2010 and 2013, were allocated subjectively in mature and old forest using stand data from the existing forest management plan. This was done in order to obtain data for several tree species compositions and a range of mean stem volume per hectare. The allocation also ensured that each 40 m plot was placed well inside the boundaries of the selected stand in order to avoid any influence of edge effects. In total, 30 40 m radius plots were surveyed (Table 4).

Table 4. *Summary statistics of the datasets (min; mean; max)*

Dataset	Number of plots	Tree height [m]	Stem diameter [cm]	Stem volume [m3/ha]	Age [yr]
Micro stand and LPM approach	856*	3.5;19.2;34	4;27.1;67.9	1;284;2058	10;50;168
Systematic design	216**	4.9;18.2;31.6	5.2;24.1;51.9	2;229;655	14;51;160
Evaluation plots	30***	15.8;24;32.4	20;30.6;42.3	157;360;685	33;60;112

*5 m radius
** 10 m radius
*** 40 m radius

2.2.2 Methodology

Based on the ALS data, rasters of estimated forest data were created by imputation of survey data (i.e., calipered diameters, tree height measurements and recorded tree species) from the reference survey plots to each cell in the evaluation plot raster. The raster cell sizes were chosen to approximate the size of the reference plot data used. The plots having a 5 m (LPM sampling design) radius were imputed to the raster cells having a size of 10 m × 10 m, and the plots having a 10 m radius (systematic sampling design) were imputed to the raster cells having a size of 18m × 18 m. Data for each 40 m radius evaluation plot were generated by aggregation from the estimated rasters, either 10 m × 10 m or 18 m × 18 m, covering the evaluation plot and imported into the Heureka DSS as if it were data surveyed on site.

Seven alternatives were examined in the study following the imputation technique (non-parametric method). The first alternative, termed the observed alternative, comprised the field survey observations of the 30 evaluation plots with 40 m radius. In the second to the fifth alternatives the 5 m radius reference plots from the micro stand and LPM approach were used. The number of reference plots was reduced from the original 856 to 500, 250 and 91, respectively, to test how the sampling intensity affected the accuracy. The alternatives are hereafter termed S500, S250, and S91, respectively. The fifth alternative – termed subS91 –also contained 91 reference plots sampled using the LPM, but given the condition that one plot was sampled from each micro stand. The sixth and seventh alternatives used 216 reference plots from the systematic grid design sampling, the alternatives hereafter termed syst10m and syst5m, respectively. For the syst10m alternative, the original 10 m radius plots were used, from which 5 m radius plots were constructed and served as reference plots for the syst5m alternative. Table 5 provides a summary of the data in the six alternatives.

2.2.3 Validation, accuracy measurement and suboptimal loss

The imputations (forest variables) were validated using the relative root mean square error (RelRMSE) and relative bias (RelBias) for each forest variable.

The accuracy of the estimated stem diameter distributions was assessed using the total variation distance index (Levin et al. 2009) computed for each of the evaluation plots (40 m radius) using the diameter classes' absolute differences. The value of the error index can range between 0 and 1, where 0 represents a perfect match of two compared distributions.

Suboptimal losses of the imputed alternatives were computed in the same way as it is described in the summary of the I paper using the Heureka system (Wikström et al. 2011).

Table 5. *Summary statistics of the different alternatives (min; mean; max)*

Dataset	Number of plots	Tree height [m]	Stem diameter [cm]	Stem volume [m3/ha]	Age [yr]
S500	500*	3.5;19.2;33.7	4;27.1;62.3	2;287;2058	12;50;131
S250	250*	3.5;19.1;32.3	4;26.9;59.7	2;276;930	13;49;131
S91	91*	3.5;19.3;32.4	4;27.7;59.4	2;301;1510	12;50;131
subS91	91*	3.5;19.6;33.7	4;27.7;62.3	2;303;2058	13;50;100
syst5m	216**	4.4;18;37.1	4.6;23;51.4	3;226;748	14;51;160
syst10m	216***	4.9;18.2;31.6	5.2;24.1;51.9	2;229;655	14;51;160

*5 m radius
** 5 m radius extracted from the 10 m radius plots
*** 10 m radius

2.2.4 Results

When comparing the LPM alternatives with the systematic design alternatives, the latter gave a lower RelRMSE for three (BWH, BWD and MA) out of the five variables. When comparing the approaches with a similar number of plots and the same plot radius, i.e., S250 vs. Syst5m, the RelRMSE was roughly equal except for mean stem volume per hectare (MSV), for which S250 showed a much lower RelRMSE (18.6% vs. 27.3% for Syst5m; Table 6).

Table 6. *Summary of the relative root mean square error (RelRMSE) and relative bias (within parenthesis) values, for the forest variables basal area weighted mean tree height (BWH), basal area weighted mean tree diameter (BWD), mean stem volume per hectare (MSV), basal area (BA) and stand mean age (MA), indicating the accuracy of the airborne laser scanning based imputation of the reference plot data to the 40 m radius evaluation plots for the six different estimated alternatives.*

	RelRMSE (%)				
Alternative	BWH	BWD	MSV	BA	MA
S500	9.6(-6.8)	8.7(5.7)	24.7(9.9)	26.8(17.1)	21.6(-7.7)
S250	10.8(-7.8)	9.0(4.7)	18.6(-2.1)	16.2(8.1)	20.4(-9.6)
S91	13.6(-9.5)	10.7(3.4)	23.3(-3.1)	20.5(8.5)	28.7(-13.2)
subS91	10.6(-7.0)	11.9(5.8)	38.8(11.9)	38.2(19)	25.0(-7.9)
Syst10m	8.4(-5.6)	8.0(0.0)	23.8(-10.6)	18.3(-8.0)	20.0(-4.0)
Syst5m	10.1(-8.0)	8.7(-4.2)	27.3(-14.9)	18.3(-8.0)	23.1(-6.2)

Error indices of the estimated stem diameter distributions are presented in Table 7. For the LPM alternatives, the mean error indices increased with decreasing number of reference plots; S91 and subS91 showed equal mean error indices. Among the systematic sampling design, Syst10m gave a lower mean error index than Syst5m, and Syst10m also performed better than the micro-stand and LPM

Table 7. *Summary of error indices of the estimated diameter distributions indicating the accuracy of the estimated diameter distributions using the two different sampling methods, local pivotal method and systematic, with different reference plot intensity. The index value 0 indicates a perfect match of the compared distributions.*

	Error indices of the estimated diameter distributions - total variation distance index			
Alternative	Mean	Minimum	Maximum	sd*
S500	0.27	0.12	0.64	0.11
S250	0.29	0.15	0.62	0.1
S91	0.35	0.13	0.69	0.12
subS91	0.35	0.2	0.54	0.1
Syst10m	0.24	0.12	0.48	0.08
Syst5m	0.28	0.15	0.53	0.09

* sd: standard deviation

alternatives. The alternatives with roughly the same number of plots and plot radius (i.e., S250 and Syst5m) gave similar mean error indices.

The suboptimal losses were calculated for each imputation alternative separately. The suboptimal losses using the LPM sampling design were 47, 62, 39 and 28 euros ha^{-1} for the S500, S250, S91 and subS91 alternatives, respectively. The suboptimal losses using the systematic sampling design were 61 and 63 euros ha^{-1} for the syst10m and syst5m alternatives, respectively. The alternative subS91 yielded the lowest suboptimal loss but the highest error index of the estimated diameter distributions (along with S91) and the highest mean stem volume RelRMSE.

2.2.5 Discussion

According to the results – RMSE, bias (Table 6), stem diameter distribution error indices (Table 7) and suboptimal losses – none of the sampling designs (i.e., LPM vs. systematic) showed a clear advantage over the other. Both the random and systematic errors showed large variations among the imputation alternatives. This was, to some degree, expected because the inference method does not utilize weighting (i.e., smoothing of several observations), unlike in the kMSN method ($k > 1$), to reduce the prediction errors. Thus, the few extreme reference observations may have had a very large influence.

Among the systematic sample design alternatives (i.e., syst10m and syst5m), the RelRMSE, stem diameter distribution error index and suboptimal loss were smaller for the 10 m radius reference plots than for the 5 m radius plots, indicating higher accuracy for syst10m then expected. Among the alternatives using the LPM sampling design, i.e., S500, S250, S91 and subS91, the stem diameter distribution error index increased as the number of reference plots decreased, indicating decreasing accuracy. This pattern of lower estimation accuracy (as for the error index) as the sample size was reduced did not hold for all estimated variables regarding the values of the RelRMSE, RelBias (Table 6) and suboptimal loss. It is likely that the obtained results were a consequence of the small evaluation dataset used in the study ($n = 30$). If the number of evaluation plots had been larger, the effect of a few extreme values on the RelRMSE and RelBias for some imputation alternatives would have been smaller.

Overall, the results of this study showed the great potential for improving data acquisition methods by employing the new sampling design LPM and micro-stand delineation techniques. One reason for choosing the LPM sampling design is its cost efficiency compared to the systematic sampling design.

2.3 Paper III- Improving data quality by assimilating different sources of forest information

Paper III is a concept paper where the idea of DA is highlighted as a novel idea for the utilization of forest information and thus for improving forest planning and decision making. The potential benefits of DA information were explored and the challenges were discussed for different forest planning contexts.

Uncertainty in forest information typically results in economic losses, ecological, and social values as a consequence of suboptimal management decisions. Incorporating uncertainty into the planning process can be difficult due to intricate mathematical algorithms and the ability of traditional mathematical programming methods, such as linear programming (Pukkala 1998; Pasalodos-Tato et al. 2013) to account for uncertainty. If the results of planning and decision processes are complicated, managing uncertainty may be hard to interpret and apply (Mowrer 2000). Therefore forest planners in practice ignore uncertainty for the sake of simplicity. Moreover, uncertainty, for instance, in inventory information increases through time as growth models are used to update forest information (Nyström & Ståhl 2001; Fig 2). While the growth models are of high quality, predictions are simplifications, and there are no techniques available to remove the uncertainty of predictions of the future forest state (Pietilä et al. 2010). The tool for controlling this uncertainty is to collect new information. Traditionally once new forest information is acquired the old forest information would no longer be used in the forest planning process, and the potential value of the old information was ignored.

DA is an approach to merge data acquired at different time points, as well as data acquired using different acquisition techniques. In the realm of forest inventory data, DA has the potential of improving accuracy of the information as well as providing an estimate of the uncertainty of the data (Figure 1).

Recent developments in remote sensing have allowed for the possibilities of acquiring forest information from distance at reduced cost (Næsset 2002; Gobakken & Næsset 2004). Regardless of how forest information is acquired, it is not free from errors and these errors are one of the many sources of uncertainty in forest planning. By using DA as new information arrives, the DA process can update the existing information, and provide new estimates of forest information variables, and their estimated uncertainty. Through this process the quality of the information will be improved by assigning less importance to the information with lower quality, updating both the estimate and the estimates of uncertainty are also updated. This provides the forest planner with information on both the point estimate of the study variable and its corresponding uncertainty. In this way, DA can be a cost efficient process of producing forest information. Therefore, DA can be seen as a continuous procedure where the information in any point of time will be up-to-date (either by forecasting or by new information) which improves the planning possibilities.

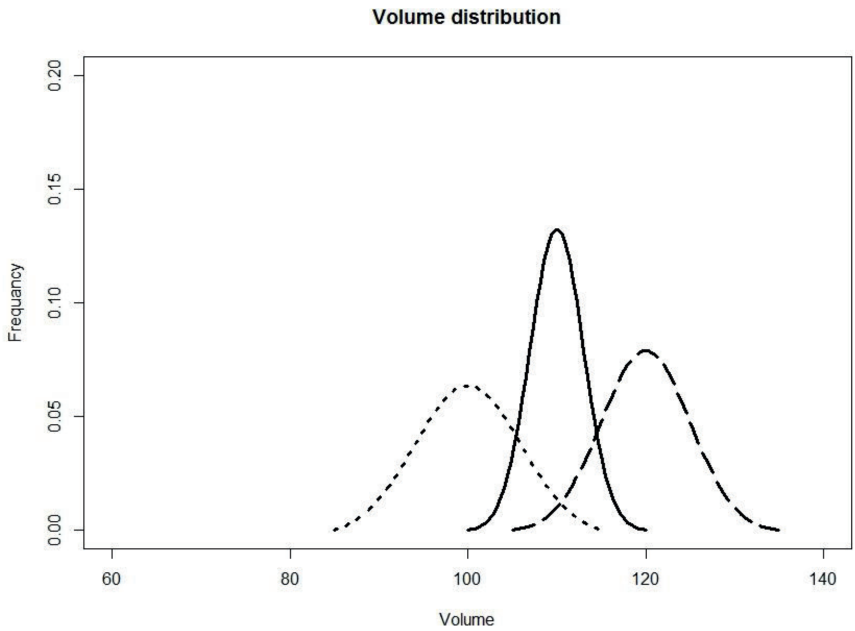

Figure 1. The forecasted information of the timber volume (dotted line), i.e., prior distribution, is combined with the new information (dashed line) in order to obtain the posterior distribution (solid line), which results in an updated estimate of the timber volume. As shown the posterior distribution is narrower compared to the prior distribution.

2.3.1 Potential of using DA in forest planning

DA can improve all optimization methods; however the information generated by DA can be more fully utilized with optimization methods which incorporate estimates of uncertainty. For instance, robust optimization (Bertsimas and Sims 2004) can integrate estimates of uncertainty through the utilization of a two-point error distribution while stochastic programming can model any distributions through a Monte Carlo process (Birge and Louveaux 2011). Thus by improving the estimation of uncertainty and integrating this information into the decision process using stochastic optimization models, improvements can be expected and calculated (Chapter 4 of Birge and Louveaux 2011).

Five major potential benefits in forest management planning when using DA procedures were identified: (i) The accuracy of the information will be improved, (ii) The information will be kept up-to-date even though no new measurement is made, (iii) The DA procedure will provide information with estimated accuracy, (iv) Stochastic decision making can be applied, which can integrate the estimated uncertainty of the information into the decision making process and (v) DA data allow for the analysis of optimal data acquisition decisions.

2.3.2 Conclusions

Improving forest information through the DA process offers several benefits to forest planners. The primary benefits are the improved accuracy of the current forest information and the provision of uncertainty estimates surrounding this information. To realize the benefits of DA, current DSS tools require the ability to explicitly incorporate information about the uncertainty of forest information and make modifications so that stochastic optimization tools can be used. There are several techniques applied in research which can handle uncertainty, but that implementation in DSSs in practice seems to be missing except in the Finnish SIMO DSS (e.g., Rasinmäki et al. 2009); however, the application that consider uncertainty in SIMO is not yet widely used. Thus there is a need to develop DSSs that can incorporate uncertainty in the decision making process, for instance, Bayesian approaches where the probability distribution of true values can be utilized in the decision making. Furthermore, DA systems in forestry need to be further investigated and developed in order to be implemented properly in forestry. Only a few empirical studies of using DA for forest information (e.g., Nyström et al. 2015) have been conducted so far and it is recommended to further assess the benefits of DA in forest inventories.

2.4 Paper IV - Incorporating risk and uncertainty through a forest DSS at stand level

This paper is designed to be a methodological paper where it aims to study the incorporation of uncertainty and managing risk through a forest DSS at stand level. The majority of the DSS designs to support forest planning are done through point estimate of the forest information in the initial state. That implies that even if the quality of information is available through empirical studies it is not used in the current DSS systems. Therefore, we highlight two options for portraying how differing decisions can relate to decision makers with varying risk preferences and then incorporate estimates of uncertainty into current forest DSSs.

2.4.1 Methods

Estimates of uncertainty can be evaluated from empirical data provided by specific forest inventory methods. Additionally, new methods (e.g., data assimilation) are being introduced to provide additional information regarding the uncertainty of forest information. The description of uncertainty in most forest applications assumes that the forest information consists of continuous variables. To be able to use these data in the DSS structure, the underlying estimate of uncertainty needs to be approximated through scenarios. Each scenario represents a realization of possible forest attribute values. The probability for each scenario (posterior) is calculated given the estimate of uncertainty (prior).

Given that the error refers to the initial state, each scenario can be represented as a stand in the DSS. As the data are in the same format as with a single point estimate, this does not require fundamental changes to the structure of the stand management program generated by the current DSS. Then, a range of management programs over the planning horizon is computed for each pseudo stand.

The most proliferous measure involving risk and uncertainty is probably the expected value; in this case the maximum expected NPV, which is defined as

$$E(NPV)=\max[j \in J, E(NPV(j))]$$

where J is the set of management programs and E(NPV(j)) is the expected NPV of management program j, which in turn is defined as

$$E(NPV(j)) = \text{sum}\ [n \in N, p(n)*NPV(j,n)]$$

where N is the set of scenarios, p(n) the probability of scenario n, and NPV(j,n) the NPV of program j under scenario n. The maximum expected NPV can of course be complemented by information of variance as an assessment of its uncertainty.

Two different methods of incorporating uncertainty and managing risk through a forest DSS at stand level are proposed:

1. The first option relates to the quantifiable risk measures, namely the conditional value at risk and the value at risk (VAR/CVAR). Being able to manage downside risk may be of value for risk averse decision makers. At the stand level it is possible to analyze the trade-offs between managing risk and maximization of NPV utilizing the Value-at-Risk (VaR) or the Conditional Value-at-Risk (CVaR) in combination with the expected NPV.

2. The second option is based on the decision maker's view of what the forest inventory could be. A visual representation is provided to highlight which decision would be most beneficial dependent on the

decision maker's opinion of the forest inventory results. By focusing on the stand level, risk preferences can be rather easily incorporated into the current forest decision support software. One benefit would be that it gives the forest owner a visual illustration of the relation between erroneous variable and optimal management program.

2.4.2 Case study

Eight stands were used in this small case study from the Jönköping area in southern Sweden. These stands were inventoried in 2011, using a plot sampling inventory method. The original stands' variables consist of mean values (corresponding to the stand register's mean values). The selection of the stands was done subjectively, with an aim of having different stands to cover the variety in stand age, different species and site index (Table 8).

In this study errors were simulated for the basal area value of each stand. The simulated errors were replicated with 100 scenarios of a normal distribution with a mean corresponding to the point estimate and variance 30 % of the corresponding stand's basal area. The 100 different pseudo stands were imported to the Heureka DSS and a maximum of 100 management programs were created for each scenario.

For the VAR/CVAR measures the results (please refer to Table 3 and 4 in the paper) show a tendency to favor longer rotations with increased risk aversion. This shift to longer rotations allows for the stand to grow, allowing the possibility for harvesting a stand with rather low BA to be less unfavorable in comparison with the stand being more heavily stocked.

Table 8. *Characteristics of the eight stands used in the study.*

Stand	Basal area	Site index	Age	Dominant species
1	30	27	37	Pine
2	20	24	53	Pine
3	16	28	73	Pine
4	27	24	36	Spruce
5	18	31	50	Spruce
6	27	24	53	Spruce
7	43	35	58	Spruce
8	24	22	87	Spruce

Figure 2 presents the management programs yielding maximum NPV for different scenarios (i.e., different BA values). For each stand the 100 BA scenarios were divided into broader intervals where in each interval the dominant management plan is presented. The BA intervals were here selected subjectively based on the list of best programs over the scenarios. Additionally, from Figure 2 it can be noted that for almost all stands, the same management programs are suggested by the VaR/CVaR measure (refer to Table 3 and 4 in the paper). The results from the E(NPV) method (refer to Table 2 in the paper) and the results from the CVaR and VaR can be seen to provide similar management suggestions. The tendency that increased risk aversion implies longer rotations observed for VaR/CVaR is here associated with smaller BA.

2.4.3 Conclusions

The results of this study highlight the potential of incorporating estimates of uncertainty into forest DSSs. To be of practical use, tools to manage uncertainty should be integrated into the application of forest DSSs. The advantage of incorporating estimates of uncertainty, such as the proposed method in this study, is that decision makers will be able to consider his/her risk preferences, such as risk-neutral or risk-averse, and adjust forest management decisions to reflect these attitudes toward risk. For instance, if we examine the decisions taken for the second stand, the E(NPV) and point estimate decision is to conduct a thinning in period two and a final felling in period six. If the decision maker is risk averse the option to conduct the final felling in period 7 or 8 may be preferable. Even though the E(NPV) is lower, during that delay the BA will increase and the probability of harvesting the forest with a low NPV is reduced.

To make forest planning under risk and uncertainty among practitioners more prevalent will require that analytical tools are part of the same DSS where other planning tasks are performed. To add the functionality of conducting this type of analysis within the DSS, an applicable software package should be developed. This package should allow for running the same management program for a set of scenarios which incorporates a variety of uncertainty estimates, and should integrate this information for ease of analysis. The proposed methods are chosen in order to be easily implementable in existing DSSs. From a programming point of view, a visual illustration method is probably the easiest approach since each scenario, or pseudo stand, can be treated separately and only the maximum value program needs to be kept. Probably, the most demanding part to arrange is the visual illustration, where a mechanism for aggregating programs is needed. With the E(NPV) and the CVaR and VaR methods the maximum value program for each scenario is kept and the only subsequent operation is to sort the kept management programs according to the criteria.

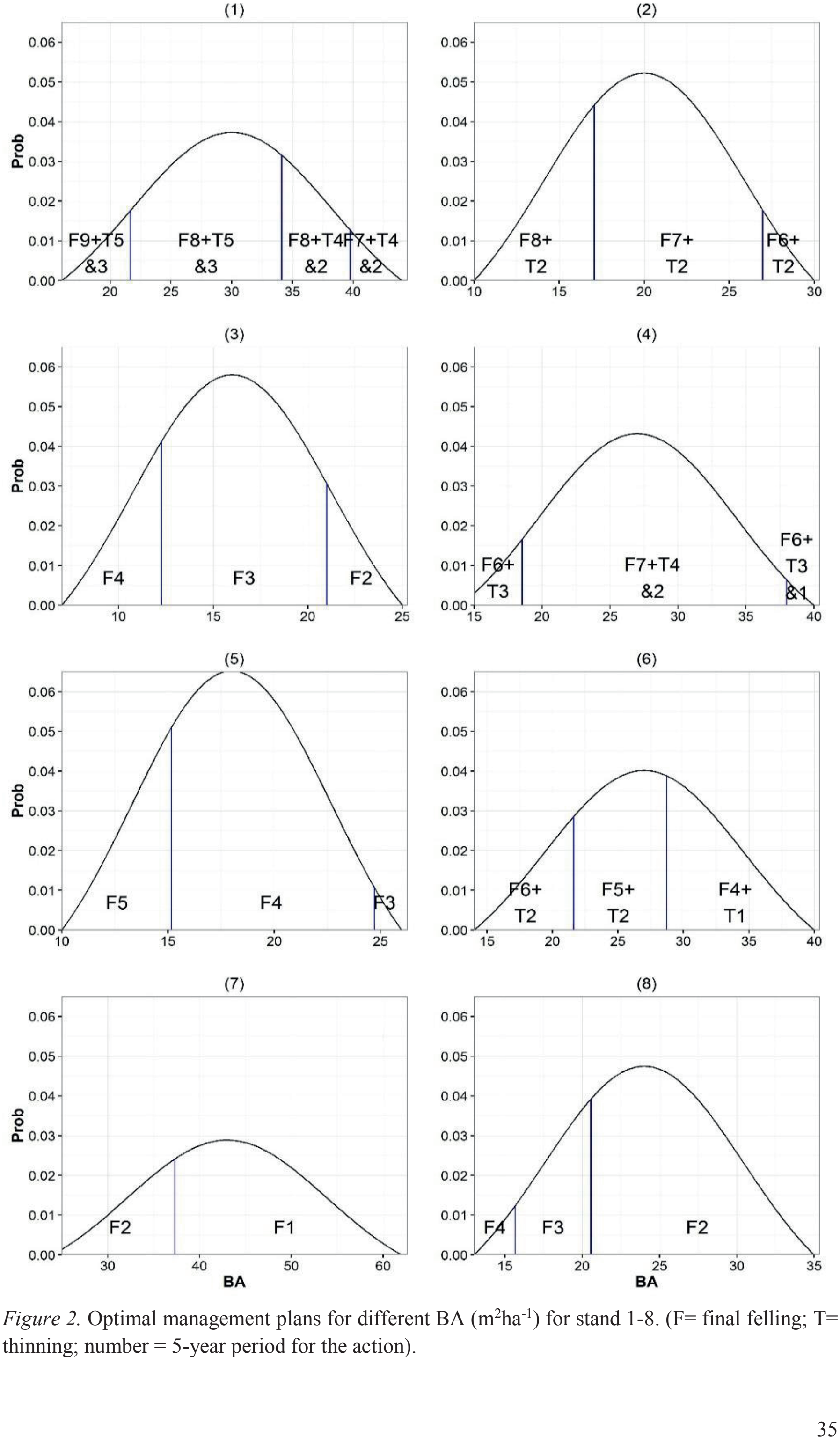

Figure 2. Optimal management plans for different BA (m^2ha^{-1}) for stand 1-8. (F= final felling; T= thinning; number = 5-year period for the action).

3 Discussion

Managing forest holdings relies to a large degree on the quality of forest information. Typically, forest information is uncertain; therefore suboptimal decisions are expected to occur even with recently acquired forest information. In general, the uncertainty in forest information can be addressed from two different fronts. The first front can be seen as a direct intervention where forest information quality is improved by collecting new information, utilizing ALS data in the sampling design and in the estimation and/or by combining new information with existing (old) information. The second front can be seen as indirect intervention where the existing planning methods and decision making processes that account for uncertainty are improved. Both approaches can be studied simultaneously to compare their efficiency and suitability. This kind of study was, however, not considered in this thesis. Papers I to III address the uncertainty directly by improving forest information quality where Paper IV addresses indirectly the uncertainty by incorporating an estimate of the uncertainty in the existing planning and decision making process.

The second front which was just mentioned can be seen as an alternative to acquiring higher quality information by improving existing planning processes and tools by incorporating and considering estimates of uncertainty in the DSS. The outcome of such an approach is presented in Paper IV where a visual illustration method is proposed to account for the risk preference of the decision maker. This idea of visual illustration is of interest for further studies as it potentially helps decision makers and forest owners to understand and observe the effect of risks. As it is hard to explain as well as to understand the tradeoff between the E(NPV) and what in the worst case might happen (VaR and CVaR), a visual illustration may be of interest. To the best of my knowledge visual illustration methods have not been considered in previous literature in forestry, and it is therefore likely to be an innovative approach in forest planning.

To combine both fronts to handle uncertainty is of interest but was, however, not covered in this thesis. Nevertheless there is evidence (e.g., Ståhl et el. 1994) that combining different methods for handling uncertainty can increase the chance of reducing the magnitude of the suboptimal loss. One example of this is from Bayesian decision theory where the error, lower though, of the new information combined with the old information can be utilized in a system that takes decisions based on information in term of distributions and not point estimates.

The DSSs nowadays rely on the deterministic assumptions regarding the simulator and optimization models. Therefore, giving estimates of uncertainty cannot be utilized in the planning process in such systems since they cannot handle estimates of uncertainty. In Papers III and IV this problem was addressed and discussed and scenario analysis (deterministic equivalent) was proposed as a solution. The method can potentially be implemented in existing DSSs. It seems that such solutions are in their early stage of development. One example is the implementation of scenario analysis in the Finnish SIMO DSS (Rasinmäki et al. 2009); however, actual use of the solution seems so far not to be fully utilized by the forest planners.

In Papers I and II ALS data were utilized in the empirical analysis and shown to have a high potential as using these data reduced suboptimal losses due to suboptimal decisions. The information stemming from ALS data has been proven to provide robust and valuable estimates of forest information, for example, stand mean values or stem diameter distributions (McRoberts et al. 2010). Incorporating stem diameter distribution in the planning process was shown in Paper I to improve the planning process compared to stand mean values only.

In Paper II ALS data were also used in data acquisition as they were used as auxiliary information in a LPM approach to design a field sample. ALS data were used in mapping (delineation of micro-stands) and for selecting micro-stands to be sampled. Moreover, ALS data and LPM were used to allocate field plots with selected micro-stands. Micro-stands have the potential to reduce within stand variation and therefore potentially improve the planning process. An automatic segmentation technique was used to delineate the micro-stands. For example, spectral data or mean tree height within raster elements can be used for segmentation. The idea of micro-stands is to provide a good description of the forest, where after the micro-stand should be aggregated into treatment units (e.g., harvest units) at some stage in the planning process. Hence, the optimal micro-stand size is of interest and the effect of stand delineation is also of interest for future studies. Approaches for clustering of micro-stands into treatment units should be developed further (Packalen et al. 2011). It is of interest also to study the effect of incorporating estimates of uncertainty stemming from growth forecasts (Nyström and Ståhl, 2001) in the same way as was done in Paper IV for the initial state of forest information.

In this thesis data quality was considered in forest planning without any restrictions or interactions between stands in a forest holding which is the simplest approach to manage forest. Using this approach the best treatment for each stand is optimized in order to maximize the NPV of each individual stand within the forest. Nevertheless, the results of this thesis have the potential to affect forest level planning including restrictions and/or interactions between stands. In such forest level planning, a best combination of treatments for the stands is optimized in order to reach a general goal for the landscape as a whole.

However, the potential to do this was not assessed in this thesis. For instance, the micro-stand approach might have an impact if the planning were to be performed at the landscape level since the clustering process performed at that level to generate a new layout for harvest operations may result in different treatment units as compared to traditional stands. This has to be further analyzed and tested to see whether new challenges would arise.

4 Concluding remarks

For future studies and in order to utilize the benefits of this thesis, new DSSs that are able to account for estimates of forest information uncertainty have to be developed. These kind of DSSs should be built in a dynamic way that will allow new information to be input into the system at any point of time. Thus, a Bayesian decision process can then be applied.

Nowadays, enormous amounts of forest information stemming from remote sensing techniques and harvesting machines are available; however, the use of these different sources of information is poor. Rather than improving or investing in one source of information, assimilating these sources seems to be a promising approach. The DA process should be further investigated as it seems to be a promising framework that can make use of many different sources of information, and provide output that can be useful for forest planners.

As discussed in Paper III, the DA process has potential to provide several main benefits to forest planning by improving performance both in short as well as long term planning setting. Without modifying the existing DSSs, the improvements provided by DA include better information accuracy, as well as estimates of the uncertainty of the information. These factors are important, but in order to benefit fully from DA there is a need to develop DSSs towards incorporating procedures which will allow them to benefit from Bayesian decision making in the later stages of the decision process. This also involves making growth forecasts which account for the uncertainty of the projected information (e.g., Nyström & Ståhl 2001).

In regards to practical forestry the use of remote sensing data and consideration of uncertainty in forest information is essential. Remote sensing data improve the quality of forest information through increasing the accuracy and as a consequence have the potential to reduce suboptimal losses. The use of the remote sensing data in practical forestry is presently applied on a broad scale, contrary to the use of information about uncertainty. Typically, uncertainty in forest information was ignored for simplicity; however, uncertainty can lead to unwanted or unexpected results and subsequently to the forest planner retrospectively wishing they had considered uncertainty at an earlier stage. Therefore, while the current DSSs is still using the deterministic optimization methods, the use of visual illustration (Paper IV) through scenario analysis (deterministic equivalent) can be utilized in practical forestry in order to facilitate for forest planners to better understand and account for uncertainty.

References

Alston, RM. and Iverson, DC. (1987). The road from Timber RAM to FORPLAN: How far have we traveled? Journal of Forestry 85: 43-49.

Bertsimas, D. and Sim, M. (2004). The price of robustness. Oper. Res. 52:35-53.

Bettinger, P. Boston, K. Siry, JP. and Grebner, DL. (2009). Forest Management and Planning. Academic Press: page 67.

Birge, J. and Louveaux, F. (2011). Introduction to stochastic programming. Second edition. Springer, New York.

Borges, J. G. Nordstrom, E. M. Garcia-Gonzalo, J. Hujala, T. and Trasobares, A. (2014). Computer-based tools for supporting forest management. The experience and the expertise world-wide. Report of Cost Action FP 0804 Forest Management Decision Support Systems (FORSYS). Dept. of Forest Resource Management, Swedish Univ. of Agricultural Sciences. Umeå. Sweden.

Breidenbach, J. Næsset, E. Lien, V. Gobakken, T. and Solberg, S. (2010). Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. Remote Sens. Environ. 114:911–924.

Czaplewski, RL. and Thompson, MT. (2008). Opportunities to improve monitoring of temporal trends with FIA panel data. In Forest and Analysis (FIA) Symposium 2008., 21–23 October 2008, Park City, Utah. Edited by W. McWilliams, G. Moisen, and R.L. Czaplewski. USDA For. Serv., Rocky Mountain Research Station, Fort Collins, Colorado, USA.

Davis, LS. Johnson, KN. Howard, T. and Bettinger, P. (2001). Forest Management. E. Bartell (4th Edition).

Davis, L. S. and G. Liu (1991). Integrated forest planning across multiple ownerships and decision makers. For. Sci. 37: 200-226.

Duvemo, K. and T. Lämås (2006). The influence of forest data quality on planning processes in forestry. Scand. J. For. Res. 21: 327-339.

Duvemo, K. Lämås, T. Eriksson, L. O. and Wikström, P. (2014). Introducing cost-plus-loss analysis into a hierarchical forestry planning environment. Ann. Oper. Res. 219: 415-431.

Edwards, V. M. and Steins, N. A. (1999). A framework for analysing contextual factors in common pool resource research. J. Environ. Policy Plan. 1: 205-221.

Ehlers, S. Grafström, A. Nyström, K. Olsson, H. and Ståhl, G. (2013). Data assimilation in stand-level forest inventories. Can. J. For. Res. 43: 1104-1113.

Eid, T. (2000). Use of Uncertain Inventory Data in Forestry Scenario Models and Consequential Incorrect Harvest Decisions. Silva Fenn. 34(2): 89-100.

Eriksson, L. O. (2008). The forest planning system of Swedish forest enterprises: A note on the basic elements. SLU, Dept of Forest Resource Management, Work report 232.

Garcia-Gonzalo, J. Bushenkov, V. McDill, M. E. and Borges, J. G. (2015). A decision support system for assessing trade-offs between ecosystem management goals: An application in Portugal. Forests 6:65-87.

Gobakken, T. and Næsset, E. (2004). Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. Scand. J. For. Res. 19: 529-542.

Gobakken, T. and Næsset, E. (2005). Weibull and percentile models for lidar-based estimation of basal area distribution. Scand. J. For. Res. 20: 490-502.

Gordon, S. N. Floris, A. Boerboom, L. Lämås, T. Eriksson, L. O. Nieuwenhuis, M. Garcia, J. and Rodriguez, L. (2013). Studying the use of forest management decision support systems: an initial synthesis of lessons learned from case studies compiled using a semantic wiki. Scand J For Res. doi:10.1080/02827581.2013.856463

Grafström, A. Lundström, N. L. P. and Schelin, L. (2012). Spatially Balanced Sampling through the Pivotal Method. Biometrics 68: 514-520.

Hahn, W. A. Härtl, F. Irland, C. L. Kohler, C. Moshammer, R. and Knoke, T (2014). Financially optimized management planning under risk aversion results in even-flow sustained timber yield. For. Policy Econ. 42: 30-41.

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

Holopainen, M. Mäkinen, A. Rasinmäki, J. Hyytiäinen, K. Bayazidi, S. & Pietilä, I. (2010). Comparison of various sources of uncertainty in stand-level net present value estimates. For. Policy Econ. 12: 377–386.

Hyvönen, P. Pekkarinen, A. and Tuominen, S. (2005). Segment-level stand inventory for forest management. Scand. J. For. Res. 20: 75-84.

Kangas, A. S. (2010). Value of forest information. Eur J Forest Res 129: 863-874.

Kazana, V. Fawcett, R. H. and Mutch, W. E. S. (2003). A decision support modelling framework for multiple use forest management: The Queen Elizabeth forest case study in Scotland. Eur. J. Oper. Res. 148: 102-115.

Koch, B. Straub, C. Dees, M. Wang, Y. and Weinacker, H. (2009). Airborne laser data for stand delineation and information extraction. Int. J. Remote Sen. 30: 935-963.

Leskinen, P. Hujala, T. Tikkanen, J. Kainulainen, T. Kangas, A. Kurttila, M. Pykäläinen, J. and Leskinen, L. (2009). Adaptive decision analysis in forest management planning. For. Sci. 55: 95-108.

Levin, D. A. Peres Y. and Wilmer E. L. (2009). Markov Chains and Mixing Times. Providence, RI: American Mathematical Society: Page 48.

Maltamo, M. Næsset, E. Bollandsås, O. M., Gobakken, T. and Packalén, P. (2009). Non-parametric prediction of diameter distributions using airborne laser scanner data. Scand. J. For. Res. 24: 541-553.

Mäkinen, A. Kangas, A. and Nurmi, M. (2012). Using cost-plus-loss analysis to define optimal forest inventory interval and forest inventory accuracy. Silva Fenn. 46:211-226. https://doi.org/10.14214/sf.55.

McRoberts, R. E. Cohen, W. B. Erik, N. Stehman, S. V. and Tomppo, E. O. (2010). Using remotely sensed data to construct and assess forest attribute maps and related spatial products. Scand. J. For. Res. 25: 340-367.

Moeur, M. and Stage A. R. (1995). Most Similar Neighbor: An Improved Sampling Inference Procedure for Natural Resource Planning. For. Sci. 41: 337-359.

Mowrer, H. T. (2000). Uncertainty in natural resource decision support systems: sources, interpretation, and importance. Comput. Electron. Agric. 27:139-154.

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sen. Environ. 80: 88-99

Næsset, E. Gobakken, T. Holmgren, J. Hyyppä, H. Hyyppä, J. Maltamo, M. Nilsson, M. Olsson, H. Persson, Å. and Söderman, U. (2004). Laser scanning of forest resources: the nordic experience. Scand. J. For.Res. 19: 482-499.

Navon, D. I. (1971). Timber RAM: A long range planning method for commercial timber lands under multiple use management. USDA Forest Service Research Paper PSW-70.

Nilsson, M. Nordkvist, K. Jonzén, J. Lindgren, L. Axensten, P. Wallerman, J. Egberth, M. Larsson, S. Nilsson, L. Eriksson, J. and Olsson, H. (2016). A nationwide forest attribute map of Sweden predicted using airborne laser scanning data and field data from the National Forest Inventory. Accepted for publication in Remote Sen. Environ.

Nyström, M. Lindgren, N. Wallerman, J. Grafström, A. Muszta, A. Nyström, K. Bohlin, J. Willén, E. Fransson, J. E. S. Ehlers, S. Olsson, H and Ståhl, G. (2015). Data assimilation in forest inventory: first empirical results. Forests 6:4540-4557.

Nyström, K. and Ståhl, G. (2001). Forecasting probability distributions of forest yield allowing for a Bayesian approach to management planning. Silva Fenn. 35:185-201.

Olofsson, K. and Holmgren, J. (2014). Forest stand delineation from lidar point-clouds using local maxima of the crown height model and region merging of the corresponding Voronoi cells. Remote Sen. Letters 5: 268-276.

Packalén, P. Heinonen, T. Pukkala, T. Vauhkonen, J. and Maltamo, M. 2011. Dynamic treatment units in eucalyptus plantation. For. Sci. 57: 416-426.

Pasalodos-Tato, M. Mäkinen, A. Garcia-Gonzalo. J. Borges, J. G. Lämås, T. and Eriksson, L. O. (2013). Review. Assessing uncertainty and risk in forest planning and decision support systems: review of classical methods and introduction of innovative approaches. For. Syst. 22: 282-303.

Persson, Å. Holmgren, J. and Söderman, U. (2002). Detecting and measuring individual trees using an airborne laser scanner. Photogramm. Eng. Remote Sens. 68: 925-932

Pippuri, I. Kallio, E. Maltamo, M. Peltola, H. and Packalen, P. (2012). Exploring horizontal area-based metrics to discriminate the spatial pattern of trees and need for first thinning using airborne laser scanning. Forestry doi:10.1093/forestry/cps005.

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

Pukkala, T. (1998). Multiple risks in multi-objective forest planning: Integration and importance. For. Ecol. Manage. 111: 265-284.

Randhawa, S. U. Olsen, E. D. and Lysne, D. H. (1996). A decision analysis approach to forest resource management. Int. J. Ind. Eng-Appl P. 3:95-101.

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

Redsven, V. Hirvelä, H. Härkönen, K. Salminen, O. and Siitonen, M. (2013). MELA2012 Reference Manual (2nd edition). The Finnish Forest Research Institute. 666 p. ISBN 978-951-40-2451-1 (PDF).

Rørstad, P. K. Trømborg, P. K. Bergseng, E. and Solberg, B. (2010). Combining GIS and forest modelling in estimating regional supply of harvest residues in Norway. Silva Fenn. 44:435-451.

Solberg, S. Næsset, E. Bollandsas, O. M. (2006). Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. Photogramm Eng Remote Sens. 72:1369-1378.

Ståhl, G. Carlsson, D. and Bondesson, L. (1994). A method to Determine Optimal Stand Data Acquisition Policies. For. Sci. 40:630-649.

Ståhl, G. (1992). A study on the quality of compartmentwise forest data acquired by subjective inventory methods. Swedish University of Agricultural Sciences, Department of Biometry and Forest Management Report 24 (In Swedish).

Söderberg, U. (1992). Functions for forest management. Height, form height and bark thickness of individual trees. Report 52. Department of Forest Survey, Swedish University of Agricultural Sciences, Umeå, Sweden. 87 p. [In Swedish with English summary].

Wikström, P. Edenius, L. Elfving, B. Eriksson, L.O. Lämås, T. Sonesson, J. Öhman, K. Wallerman, J. Waller, C. and Klintebäck, F. (2011). The Heureka forestry decision support system: an overview. Mathematical and Computational Forestry & Natural-Resource Sciences 3:87-94.

Yousefpour, R. Bredahl Jacobsen, J. Thorsen, B. J. Meilby, H. Hanewinkel, M. and Oehler, K. (2012). A review of decision-making approaches to handle uncertainty and risk in adaptive forest management under climate change. Ann. For. Sci. 69:1-15.

Acknowledgements

Now that my PhD studies are coming to an end, I would like to take the opportunity to thank a couple of people.

This thesis could not have been completed without the support of my main supervisor Tomas Lämås. Tomas, for everything, and especially for keeping an open-door policy to any question I had or help I needed during my entire PhD program, thank you. I thank my co-supervisor Jörgen Wallerman for his endless support, and for having taken the time to discuss both work and personal stuff. I would like to thank and mention my co-authors Johan Holmgren, Kyle Eyvindson, Ljusk Ola Eriksson, Peichen Gong and Göran Ståhl. I thank Anton Grafström for his help in statistical analysis throughout my PhD studies and Hampus Holmström for his help with the Heureka analysis. I thank Heather Reese and Neil Cory for help with checking the language and for general comments on the thesis. All of my colleagues from the forest planning section, thank you.

Finally, I would like to thank everyone else that in one way or another have given their time to help me succeed in my PhD studies.

January 10th, 2017