The Influence of Data Uncertainty on Planning and Decision Processes in Forest Management

Karl Duvemo
Faculty of Forest Sciences
Department of Forest Resource Management
Umeå

Doctoral Thesis
Swedish University of Agricultural Sciences
Umeå 2009

Acta Universitatis Agriculturae Sueciae
2009:6

ISSN 1652-6880
ISBN 978-91-86195-53-3

Print: Arkitektkopia 2009

The Influence of Data Uncertainty on Planning and Decision Processes in Forest Management

Abstract

This thesis focuses on how uncertainty in forest data affects the outcome of management planning and decision making.

In a review article (paper I), previous research aiming to evaluate forestry data are described. The methodology used and the results presented were discussed. A general conclusion was that previous studies concerned highly simplified planning situations, leaving some doubts concerning their real world applicability. In papers II and III, two quantitative approaches of data evaluation were applied to data from sample plot imputations. In paper II, a cost plus loss analysis of using the data for forest management decision making is presented. The usefulness of the data in forestry scenario analysis were scrutinized in paper III as errors in the predictions of future forest states, harvest levels and net income flow. In both papers it was concluded that improvements in the methodology of assessing data would be required.

In paper IV, an advance of the cost-plus-loss methodology was suggested by developing a simulation system that aims to capture the hierarchical structure and iterative nature of forestry planning. The simulation system included the tactical and operational levels of a continuous planning process at a specific corporate forest owner. It was characterized by annual re-planning with the option to reassess data of selected stands prior to operational planning. The planning simulation system was used in paper V for evaluation of current practice data. It was concluded that high decision losses occurred as a result of errors in the studied data. The introduction of holding level wood-flow considerations and incitements to cluster harvest activities reduced decision losses compared to stand wise planning without such considerations.

Keywords: Forest inventory, data evaluation, cost plus loss analysis, tactical planning, operational planning, decision support system, simulation of data errors, optimization.

Author's address: Karl Duvemo, Department of Forest Resource Management, SLU Skogsmarksgränd, SE-901 83 Umeå, Sweden.
E-mail: karl.duvemo@srh.slu.se

.

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 Duvemo, K., Lämås, T. (2006). The influence of forest data quality on planning processes in forestry. *Scandinavian Journal of Forest Research* 21, 327-339.

II Duvemo, K. Barth, A. Wallerman, J. (2007). Evaluating sample plot imputation techniques as input in forest management planning. *Canadian Journal of Forest Research* 37, 2069-2079.

III Barth, A., Duvemo, K., Wallerman, J. Evaluation of sample-plot imputations in sub-national forestry scenario analyses (manuscript).

IV Duvemo, K., Eriksson, L.O., Lämås, T., Wikström, P. Introducing cost plus loss analysis into a hierarchical forestry planning environment (manuscript).

V Duvemo, K. Wikström, P., Eriksson, L.O., Lämås, T. Cost plus loss analysis of current practice data in a large scale hierarchical forestry planning situation (manuscript).

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

1 Introduction

Investigating the value of forest data for forest management planning purposes means exploring the economics of uncertainty. This thesis focuses on how uncertainty in forest data affects the outcome of management planning and decision making. A large share of the work described herein has been devoted to explore the impact on the economical return of forestry, which in turn determines what value forest data will have.

1.1 Brief on decision analysis

To set the scene for this thesis, we need not only to describe forestry planning, but also the fundamentals of decision analysis on which forestry planning rely. The overview of decision analysis will provide some useful terminology introduced into a forestry context later on in this paper.

Decisions analysis deals with human decisions, i.e. how we use our freedom. More specifically, situations treated by decision theorists require that there are options to choose between and that we choose in a non-random, goal-directed way. Hence, decision theory is concerned with goal-directed behavior in the presence of options (e.g. Hansson, 2005). A further division is made into *normative* theories suggesting how decision should be made to be rational, and *descriptive* theories describing how decisions are actually made. It should be pointed out that norms of rationality are not the only norms important in decision making. It is however customary to consider ethical or political norms as external to decision theory.

Decision analysis does not include only rules of choice but the entire decision process. Numerous models of the decision process have been suggested over the last century, often with only small differences. The division of the process into 5 distinguishable steps by Brim et al. (1962) will serve as an illustrative example.

i) Identification of the problem
ii) Obtaining necessary information
iii) Production of possible solutions
iv) Evaluation of such solutions
v) Selection of a strategy for performance

It can be argued (Witte, 1972) whether the decision process is sequential as in the example above or if the steps are parallel or even circular (e.g. Mintzberg et al., 1976).

When making decisions, or choosing between options, humans try to obtain as good an outcome as possible according to some standard of what is good or bad. Decision theory assumes that such a standard is at hand, and proceeds to express this standard in a precise and useful way for relating options to each other. Most of the literature on decision making has focused on the actual choices of solutions (step v according to Brim) (Hansson, 2005). However, in several empirical studies (e.g. Mintzberg et al., 1976, Simon, 1960), it has been found that more time is spent in the steps corresponding to ii-iv) in Brims models and little time in doing the actual choices (making the decision). This is obvious in cases when an efficient value-standard is in place and the preceding steps have ranked the solutions unambiguously according to this standard. The choice is then reduced to simply picking the solution with the highest value according to the value-standard. This stresses the importance of the decision process steps preceding the actual choice.

In the real world, decision making is often affected by some degree of uncertainty influencing e.g. the actual feasibility and outcome of a solution. In analytical literature on uncertainty and information in economics, a division is made into *market* uncertainty and *event* uncertainty. The former focuses on uncertainties of demand-supply offers and behavior of other economic agents while the latter focuses on, for an individual, uncertainties considering its own endowment and production opportunities (Hirshleifer and Riley, 1979). Parts of these uncertainties can be reduced by acquiring more/better information thus enabling better decisions. Individuals or organizations may be willing to pay for information depending on how uncertain they are and on what is at stake (Macauley, 2006). Rationally, they can pay for the information as long as the expected gain from it exceeds the cost. The possible gain is what is commonly referred to as the concept of *value of information* (VOI), one of the most important notions of decision analysis.

For an economic activity, here referred to as a "project", the following definition and properties of VOI can be stated; *The VOI is the difference between the project value with particular information and the project value without that information, minus the cost of the information*. This implies that there must be an alternative outcome of a project; otherwise no information could add value, which is the same as to say that there must be uncertainty. If there is uncertainty there must be choices, if there is no choices – there is no decisions to be made, and information is in fact worthless.

The value of information is case specific and depends on many circumstances. From the work of Hirshleifer and Riley (1979) and McCall (1982) it can be concluded (Macauley, 2006) that VOI depends on several factors including;

i) How uncertain decision makers are
ii) What is at stake as an outcome of decisions
iii) What is the cost associated with accessing and using the information
iv) What is the price of the next best substitute for the information

Elaborating on i), it is obvious that if little uncertainty surround the decision prominent, little can be gained by adding information and, hence, VOI is small. VOI is small also if the decision makers options are limited, i.e. that even with better information available the decision makers can only make small or no adjustments of the decision. A third situation when VOI is low could be that although the information is important for the outcome and options are available, other aspects influencing the decision are of supreme importance and, thus ultimately decisive of the choice. If the situation is the opposite of the cases above, VOI can be large. The second factor (ii) depends on the value of the output of a decision on the market. Factors iii) and iv) are preferably treated together for comparison purposes as iv) reflect the possible existence of other sources of information than the one currently analyzed. An important note to iii) is that there may be other uses of the information or other users, hence the cost of the information may be shared.

Thus, if what can be gained by acquiring information prior to a decision exceeds the cost of the information, i.e. VOI is positive, then information acquisition is rational. Unfortunately, the content of the information acquired can never be known in advance and hence, its effect on the decision and outcome of the decision remain unknown. The decision to seek additional information must necessarily be made *ex ante* –before an event (Hirshleifer and Riley, 1979). Had the resulting information been known in advance, the information acquisition would be superfluous and

VOI would be negative (due to the cost of the information). VOI can be determined only after the information is acquired and the decision outcome is clear. So, when considering acquiring more/better information for decision making we must estimate the *expected* VOI.

Another aspect of uncertainty affecting decision making is the decision makers' attitude towards risk. That is, after making a decision under uncertainty, a decision maker may discover, on learning the relevant outcomes, that another alternative would have been preferable. This may create a sense of regret (Bell, 1982). Analogous, if the outcome of a decision made under uncertainty has a lower value than expected, a sense of disappointment can be created (Bell, 1985). A risk aversive decision maker is prepared to tradeoff financial return to avoid regret or disappointment. The rational decision in this case is made based on an expression of expected value in which regret and disappointment is explicitly included. Also other aspects of human behavior and preferences, such as willingness to play lotteries with negative expected values, affect real world decision making. In the following, no aspects of such behavior or preferences are included, i.e. the decision criterion is maximum expected value without considering risk preferences.

1.2 Forestry planning

The concept of *planning* in a general context has been defined by numerous authors. A definition well suited for, and previously cited in, a forestry planning context, was formulated by Cohon (1978);

> "Planning is the process by which analysts perceive a problem, define it, collect data about it, formulate it (perhaps mathematically as a model), and generate and evaluate alternatives for solving it, leading to the end of the process when decision makers choose an alternative for implementation".

This definition is well suited to view forestry planning as a decision process.

Forestry planning does not differ from planning for other purposes in any fundamental way, it does however have several distinctive features (Öhman, 2001). First, the forests are complex systems of which we have incomplete knowledge. Uncertainties stem from e.g. incomplete descriptions of the current state and effects of stochastic processes affecting the forests. Predicting many economic and ecological variables essential in the planning process is therefore associated with a great deal of uncertainty. Second, forestry is a long term economic activity requiring long term planning horizons. This increases the risk of unpredictable events (see e.g. Davis et al., 2001). Finally, forestry planning often includes different (and sometimes conflicting) goals such as maximizing the economic value of timber production, production of mushrooms, berries and game, increased/maintained biodiversity, carbon storage etc. (see e.g. Lämås, 1996).

To grasp the forestry planning problem, especially the implications of long time horizons, planning methods are frequently divided into a hierarchical structure. Typically, forestry planning is divided into strategic, tactical and operational levels (Weintraub and Cholaky, 1991, Nelson, 2001, Davis et al., 2001). Strategic planning is long term and aims to decide general strategies that can be applied and which output are possible. An important task of most strategic planning in forestry is to decide annual harvesting levels weighing present and future production potentials (e.g. Jonsson et al., 1993). Tactical planning is medium term, typically 5-10 years in the Nordic countries and creates object-specific plans. The management units are often treated in a spatial context. Operational planning is short term and consists of the administrative and steering functions for implementing the objectives set at higher levels.

The actual forestry planning ultimately aims to schedule operations on the level of management units. The most frequently used management unit, the forest *stand*, is a geographically contiguous parcel of land which is considered homogenous about tree vegetation (Davis et al., 2001). Also other features, such as different ground conditions, can influence the delineation of stands. The purpose of the delineation is to create areas for which specific management operations can be prescribed. Each stand can thus be treated separately according to the objectives of the manager. However, stands exist as parts of holdings and as parts of the forest landscape. The objectives of the decision makers and the state of a holding or landscape thus affect the management of the individual stand.

Within the framework provided by legislation and established practices, and considering public opinion and individual objectives, it is ultimately the landowner who makes decisions on forest management operations. To be rational, the decision maker needs to be able to envision the outcome and consequences of different decisions (Simon, 1976). Especially in large scale forestry and considering the long time horizon, this is typically aided by computerized decision support systems.

1.2.1 Forestry decision support systems

The main objectives of using a forestry decision support system is to produce and evaluate possible management alternatives (corresponding to steps iii–iv) in the decision process according to Brim et al. (1962)). In some cases, also step ii), obtaining the necessary information, is an integrated part of the system (e.g. Jonsson et al., 1993). Many forestry decision support systems exists worldwide, adapted to local conditions and requirements. An overview of models worldwide were compiled in 1996 (Nabuurs and Päivinen, 1996). Since then, more has been introduced. In the Nordic countries, the most widely used systems includes the Forest Management Planning Package (FMPP) (Jonsson et al., 1993), Hugin (Lundström and Söderberg, 1996), Avvirk 2000 (Eid and Hobbelstad, 2000), GAYA-JLP (Hoen and Eid, 1990, Lappi, 1992), MELA (Siitonen, 1995), Monsu (Pukkala, 1999, Pukkala, 2004) and SIMO (Tokola et al., 2006). A distinction is made between simulating and optimizing systems. A simulating system is used to simulate the effect of predefined decisions while an optimizing system searches for the best management decision according to the stated objective.

A simulating forestry scenario model is typically used by policy makers at national or sub-national level. For example, in Sweden there is a since long established tradition of estimating long term sustainable harvesting levels at the national level using simulating models (e.g. Anon., 1978, Bengtsson et al., 1989, Anon., 2000, Gustafsson and Hägg, 2004, Anon. 2008). Optimizing models are more commonly used by forestry companies for the scheduling of forestry operations.

The core of a forestry decision support system is a forest simulator. Given the initial state of each stand and possible treatment operations, the simulator computes the expected development of each forest stand (Eid and Hobbelstad, 2000, Lämås and Eriksson, 2003) taking into account e.g. tree growth, tree mortality and the effect of applied treatments. Typically, the state of a stand is computed in discrete time steps, each step representing a planning period. The period length is usually between one and ten years. For each planning period, the output of resources, e.g. harvest volumes, are predicted and presented.

The decision support system typically includes some framework for defining what kind of alternatives to be considered and computed by the forest simulator. Results at stand and forest level are compiled to make them interpretable for the decision maker. For example, in a strictly timber-oriented forestry situation, the objective may be to maximize net present value (NPV). In the decision situation (step v according to Brim et al.

(1962)), NPV then serves as an efficient value-standard to compare and choose among different alternatives.

Currently, forestry decision support systems are becoming more complex, especially as objectives other than economical are having a large impact on how forestry operations are planned and executed. Not only the objectives of the landowner, but also those of public interest, affects forestry. The public may appreciate the aesthetic and recreational value of forests, to pick berries and mushrooms or for hunting (e.g. de Vries and Goossen, 2002, Ihalainen et al., 2003, Lindhagen and Hörnsten, 2000, Pukkala et al., 1995). Societal demands on the forests also includes securing a high biodiversity (e.g. Angelstam and Andersson, 2001) and to play a mitigating role against climate change either through substitution effects (bio-fuel vs. fossil fuel) (Eriksson et al., 2007, Gustavsson et al., 2006) or by storing carbon (Backéus et al., 2005, Dean et al., 2004).

1.2.2 Multi-purpose forestry

To date, accounting for aspects or objectives other than timber production has mostly been formulated as restrictions put on timber production. Beside the challenges of incorporating other resources in model form, one reason for this is the known difficulties in estimating their exact monetary values (Boman and Mattson, 1999) thus making them possible to relate to the timber value. The emerging Swedish decisions support system Heureka (Lämås and Eriksson, 2003) will have the capacity to run in optimization mode with other objectives than NPV of timber production. Simultaneous optimization of several different resources though is still not possible. It is likely that future users will use the capacity to model and optimize these other resources to investigate their potential and possibilities before creating constraints applied on an optimized timber production plan.

When using a decision support system the need for input data is defined by the resources considered and properties of the system, such as the included growth and yield models. Data requirement concerns both what variables are needed and the format of these variables necessary to run the included models. To portray forest development, growth models using different approaches are typically available. Whole stand (or plot) models use mean age, basal area per area unit, etc. Diameter class models are based on data on the average tree within each class, while individual tree models uses data for each tree in the stand or in a sample (Davis et al., 2001). In the decision support system, results from the forest simulator are processed to become easier to interpret for the forest manager. Typically, the analysis of the development of a resource is simplified by introducing resource indicators which provide useful information on the resource, yet can be calculated based on the output from the forest simulator. Depending on what indicator to calculate, different additional information is required beside the tree data. Many aspects of the resource *biodiversity* depend on landscape patterns and functions. It is thus necessary to use data in a spatial context. Resource indicators on biodiversity that requires spatially comprehensive data (wall-to-wall data) is typically habitat suitability models for different species (Edenius and Mikusinski, 2006, Mikusinski and Edenius, 2006, Ricotta and Avena, 2003) or models for creating continuous areas of specific forest types (e.g. Öhman, 2000). Resource indicators concerning recreational values are other examples for which spatially comprehensive data at the landscape level is required (e.g. de Vries and Goossen, 2002, Lindhagen and Hörnsten, 2000). Also the traditional timber oriented analysis can be improved by the introduction of spatially comprehensive forest data as it enables, for example, harvest clustering and improved logistics (e.g. Gustafsson et al., 2000, Öhman and Lämås, 2003).

1.2.3 Data acquisition

Data for forestry planning can be collected in many different ways. A distinction between the methods are if the data are collected by a surveyor in the field (field inventory), or if data are collected from the air (remote sensing). Irrespective of this distinction, methods can either be subjective or objective. When using a subjective method, the surveyor typically estimates a variable directly by ocular assessment or possibly supported by measurement in what he believes to be representative areas. The accuracy of such a method depends heavily on the surveyors skills. Both systematic and random errors occur in such data and the size of the errors cannot be calculated. An objective method is based on statistical sampling theory. Measurements are carried out in a predefined manner and in areas or on objects selected in advance through random sampling procedures. An objective method should be independent of the surveyor's skills and produce unbiased estimates. Another advantage is that the precision can be estimated based on the data acquired. A more thorough description of data acquisition methods were recently provided by Barth (2007).

The data acquired will have different properties depending on which method that were used. Properties such as bias and standard deviation in estimated forest variables are used for comparing forest data from different sources. However, since the main objective of forest inventories is to provide information that can be used for decision making about management operations, these measures are insufficient in providing an unambiguous ranking of data acquisition methods (Eid et al., 2004).

1.3 Previous research on the influence of forest data uncertainty on planning and decision processes in forest management

In addition to the stated bias and standard deviation of a forest data set, a data assessment strategy can be further evaluated if a link between the errors, the consequential incorrect treatment decision, and the corresponding economic losses is established. Researchers previously pursuing this idea of data evaluation have used different approaches. These can be grouped in two categories; analytical approaches and simulation approaches. In the typical analytical approach, functions are derived to find the optimal intensity/timing of an inventory. Only a few examples of analytical studies exist. In a simulation approach, the effects on forestry planning with some erroneous data are compared with results from planning with data considered free from errors in repeated calculations. In both categories of approaches, cost-plus-loss analyses have been frequently applied. The basic theory of cost-plus-loss analysis is provided below, using the work of Hamilton (1970) as an example.

1.3.1 Analytical approaches

Hamilton (1970) first applied cost-plus-loss analysis in a forestry context following ideas from Yates (1960) and Cochran (1963). The total cost of an inventory effort (C_T) is formulated as:

$$C_T(n) = C_0 + C \times n \qquad (1)$$

where C_0 is a fixed cost, C is the cost per sampling unit and n the number of sampling units. The C term is specific for every sampling design. It may depend on the time to measure each unit, the transportation time between units, labour cost, etc.

The second part of this analysis is the loss function (L) that describes the losses that occur when decisions are based on inventory information that deviates from the true population values. This approach assumes that decision makers would make decisions that optimize their objectives if perfect data were available. In this setup, a forest variable (θ) to measure and upon which decisions are based, is identified. The loss function (L) is stated in some different formats, e.g:

$$L = \left|\hat{\theta} - \theta\right| \times \lambda \text{ or} \qquad (2)$$

$$L = (\hat{\theta} - \theta)^2 \times \lambda, \qquad (3)$$

where $\hat{\theta}$ is the estimated value of the variable, θ is the true value of that variable and λ is the parameter connecting error and loss. A generalized cost-plus-loss function can be expressed as $E[C_T + L]$. An expression for the optimal inventory intensity is then found by minimizing the expected value. The generalized results with loss expressed as in (2) and (3) are:

$$n = \left(\frac{\lambda^2 N^2 \sigma^2}{2\pi C^2}\right)^{\frac{1}{3}} \qquad (4)$$

and

$$n = \sqrt{\frac{\lambda \sigma^2}{C}} \qquad (5)$$

respectively, where σ^2 is the population variance, N is population size and other symbols as defined previously.

This setup is appealing as a simple formula for the optimal inventory intensity is provided. However, it contains some very simplifying assumptions and its validity for real world planning exercises must be questioned (Borders et al., 2008). First, and most important, it is assumed that decisions are made based on a single surveyed variable. Burkhart et al. (1978) addressed this problem by deriving the same formulas for a case of multiple decisions based on several variables from different inventories in a multiobjective management situation. In this complex case, individual error terms (λ) for each decision was expressed as the weighted (w) sum of errors in the variables estimated in the different independent inventories that affect the decision. The analyses relied heavily on the subjectively set parameters λ and w. Also, in the studies of Hamilton (1970) and Burkhart et al. (1978) the term for fixed inventory cost (C_0), as included in the formula for total cost (C_T), is not included in the formulas for optimal intensity. The fixed cost would definitely influence the decision to do the inventory at all.

Ståhl et al. (1994) choose an entirely different approach when trying to include the questions of *whether* and *when* to do an inventory. Here, the aim was to decide simultaneously optimal treatment and measurement actions in a stand. Thereby, decisions on inventory actions were brought into the forest management planning process. Ståhl et al. (1994) suggested that decisions be made using probability distributions of values in the calculations rather than point estimates. A planning framework was set up in which, using Bayesian theory, posterior distributions could be calculated when an inventory was carried out. Also the probability of obtaining different posterior distribution could be calculated. The value of different decisions was calculated for both prior and posterior distributions enabling evaluation of the inventories. Ståhl et al. (1994) then solved a multitemporal stand level management problem, in which inventory was one possibility in each planning period, using dynamic programming. Several simplifying assumptions were made in this study including that all errors were normally distributed and that the decision was based on only one variable.

1.3.2 Simulation approaches

Simulation approaches is computing intensive and for this reason they have been feasible only a couple of decades. Sprängare (1975) pioneered this field when setting up a framework for analysing the effects of using subjectively collected (erroneous) data for the selection of stands for final felling. In a small case study, a planning system was used that assigns a value on final felling priority and selects stands for the next 10 year period with values higher than a predefined limit. Planning effects were evaluated as the difference in lists of stands assigned to final felling when true and erroneous data were used. Sprängare's (1975) results indicate that planning with erroneous data leads to higher than optimal harvesting levels. As a starting point, Sprängare (1975) used subjectively inventoried data from all stands within the case study area. Results from an objective sample plot inventory of a sample of stands were used as reference data. By establishing a relation between the subjectively and objectively inventoried data in the sample, "true values" for all stands could be simulated using the established data type relation. It was pointed out that simulated values are not to be considered as true but rather that they have the same relation to the evaluated dataset as true values would have had. Planning outcome differences are thus considered to have realistic properties.

Jacobsson (1986) used a simulation approach for studying the optimal number of sample stands to be inventoried in a two-phase inventory for strategic planning with the Forest Management Planning Package (FMPP) (Jonsson et al., 1993). Besides finding the optimal inventory intensity, Jacobsson (1986) also showed that decreasing marginal utility of annual net income actually had little effect on the optimal number of sampled stands as compared to a case where marginal utility was not decreasing. It was also concluded in the study that strategic planning based on a small sample of stands yields results that are practically free from systematic errors. When creating the data for the study, Jacobsson (1986) followed the ideas of Sprängare (1975). The approach was reversed, however, as high quality sample plot data from a sample of stands were used to simulate data from a cheaper inventory method.

Eid (1991) studied planning on forest holding level by adding systematic errors to inventory data. Planning simulation was carried out using AVVIRK3 (Hobbelstad, 1988) in which the "average tree" in a stand is the basic unit for calculations. To the reference data, errors were added or subtracted systematically. For ten model forests, plans were based on the data with systematic errors in site index, average height, basal area or stand age. Errors in site index estimations were concluded to have the largest effect on the plans as they strongly affect both the predicted state of a stand and the treatments suggested for a specific state.

Eid (1993) also studied the effect of random errors in stand variables in strategic planning. Uncorrelated errors were generated to the stand data variables of site quality, basal area and dominant height corresponding to a 10% coefficient of variation. A strategic plan was set up in a version of GAYA-LP (Hoen, 1990, Hoen and Eid, 1990). Eid (1993) concluded that relative deviations in calculated net present value and final felling volumes were of lesser magnitude than the error level of the input data. However, the initial error levels were lower than those normally observed in practical forestry. Also the introduction of correlation between errors as well as errors in all variables would affect the outcome.

The aforementioned study was later extended to comprise the input variables basal area, mean height, site quality and age and errors levels of 10, 15 and 20%, (Eid, 2000). This study only considered maximization of NPV when scheduling final fellings without the requirement of a non-declining yield. The resulting NPV losses were small for errors in basal area and mean height, whereas they were noticeable for errors in stand age and site quality. The largest NPV losses were, not surprisingly, incurred for the combination of errors in all variables mentioned. Eid also showed dependencies between stand characteristics, the erroneous variables and NPV losses. For stands close to the economically optimal rotation age, errors in the estimated age had significant effects on NPV. Errors in site quality had large effects on NPV also in young stands.

Ståhl (1994a) investigated the optimal intensity in a standwise sample plot inventory. The problem was confined to determining the optimal time for final felling in mature stands where no other decision was to be made.

Twenty reference sample plots per stand were available holding treewise information. Ståhl used a bootstrapping technique to select subsamples of plots to represent the outcome of fictitious inventories of different intensities. Ståhl (1994a) showed that the sum of inventory costs and decision losses decreased rapidly when the number of plots increased towards the optimal number. When the sample size was increased beyond the optimal intensity, the sum of cost-plus-loss increased at a more moderate pace. In a later study, Ståhl (1994b) added complexity to the same basic problem, as stumpage prices were considered either deterministic or stochastic. In the latter case, the optimal inventory intensities were slightly reduced when compared with the case of deterministic prices.

Larsson (1994) considered the timing of final felling in a planning system in which average values of stand properties are used as input in growth predictions. The result from this planning approach was compared with a reference solution obtained in a planning system that used sample plot and single tree data as input. In a first case, the average values were based on the sample plot inventory results, i.e., they were given the values of the reference data. A second case considered planning with random errors in single variables. Simulated error levels of 10, 20 and 30% in age and volume were tested. In a third case, correlated random errors were introduced where the error structure was derived from empirical studies of existing stand registers. It was concluded that error levels of 10–20% in stand level data yield tolerable losses toward the cost of acquiring data of such quality. Larger errors quickly lead to high losses. Larsson used empirical data on error levels and structures (e.g. Ståhl, 1992). The same method as Sprängare (1975) was used for the generation of multivariate errors.

Holmström et al. (2003) presented a cost-plus-loss study comparing four different data acquisition strategies in strategic planning. One strategy included subjectively inventoried stands to which objectively inventoried sample plots were imputed with the kNN-method (Muinonen and Tokola, 1990, Tomppo, 1990). The second method was similar except that subjective aerial photograph interpretation data supported the kNN-assignment. The last two strategies were simulations of field inventories of objectively measured sample plots. It was concluded that the best choice of inventory method depended on stand type. No straightforward way of identifying the best method for a particular stand was pointed out. It was, however, indicated that large mature stands, preferably with a relatively short time to next treatment, should be inventoried with accurate methods. Small and/or less valuable stands may be inventoried with simpler methods. When choosing the optimal inventory method in each stand rather than the methods that did best on average, the total cost-plus-loss could be reduced by 15–50%.

Eid et al. (2004) compared one data acquisition strategy based on photo interpretation with one based on laser scanning by means of cost-plus loss analysis. The costs for the laser scanner data set were estimated to be more than twice as high as the cost of the photo interpreted data. However, the losses incurred when planning with the photo interpreted data were 3-4 times larger than decision losses when planning was based on the laser scanner data. When added together, the cost-plus-loss of the laser scanner data were only half of the equivalent for photo interpretation data. The authors suggested that the relation between the methods had a high general validity even if the respective method may be improved by altering their intensity.

Holopainen & Talvitie (2006) compared data acquisition methods based on laser scanner data, 2 or 3-dimensional measurements from digital aerial photographs and traditional, partly subjective, compartmentwise field inventories by means of cost plus loss analysis. The data was used for the scheduling of thinnings and final cut. The results showed that when inventory costs were not considered there were no significant differences between the expected NPV losses in 3D measurements of digital aerial photographs, laser scanning and the compartmentwise method. When the costs of the data were added, the traditional compartmentwise inventory produced the lowest cost plus loss.

Borders et al. (2008) investigated the impact of forest data errors on the expected NPV from simulated management of a fictive holding in the southern United States. They concluded that losses in excess of 170 US\$ ha^{-1} are likely to occur, stemming from the level of errors frequently found in current practice data.

In the cited work, uncertainty in the predicted development of the tree layer has mostly been assumed to stem from erroneous description of the initial state of the forest. Note that that errors in predictions may stem also from other sources such as model misspecification, errors caused by the randomness of nature, or errors in the estimation of model parameters due to uncertainties in the data for model development (Gertner, 1987, Kangas, 1996, Kangas, 1997, Kangas and Kangas, 1999, Nyström and Ståhl, 2001, Ståhl, 1994b).

Another common simplification is the assumptions of deterministic prices and a market insensitive to shifting demand supply offers. In fact, most work has disregarded all aspects of *market* uncertainty. Thus, evaluations to date are, with few exceptions (e.g. Ståhl, 1994b), limited to dealing with *event* uncertainty.

1.3.3 Value of information vs. cost plus loss analysis

In the discussion on decision analyses (section 1.1) the concept *value of information* was stressed. This concept is closely linked to the cost plus loss approach. In the value of information context, information acquisition is rational if the expected value of information (EVOI) is positive, that is if the expected gain in project value (EG) is greater than the expected cost (EC). EG is here expected project value with new information ($EPV_{posterior}$) minus expected project value without new information (EPV_{prior}).

$$EVOI = (EPV_{posterior} - EPV_{prior}) - EC \qquad (6)$$

The rational decision maker wants to maximize EVOI. This concept relates the outcome of information collection to the outcome without new information. It provides a measure on whether or not the collection is profitable. In this simple form –whether or not to acquire new information– it does however not provide any information if even larger gains are possible. In all forestry related research cited, the outcome of a data acquisition project is instead related to the maximum project value ($EPV_{perfect}$), which would be obtained using perfect information. Thus, the difference between $EPV_{perfect}$ and $EPV_{posterior}$ is the expected loss E[L] as used in the cost plus loss formula presented above. Using these notations the cost plus loss function to be minimized can be expressed as;

$$E[C+L] = EC + EPV_{perfect} - EPV_{posterior} \qquad (7)$$

What might be an advantage with the cost plus loss approach is that it provides the decision makers with information on how far from optimum they are operating. However, the approach does not give information on whether a certain acquisition policy is actually profitable. To do so, analysis of the outcome without new information also needs to be studied and the results compared. Regardless of what analysis approach is used, maximizing the EVOI and minimizing E[C+L] ultimately leads to the same data acquisition solution.

1.4 Objectives of the thesis

The main objective of this thesis was to improve and test methods for evaluating the appropriateness of forestry data sources as input in forestry planning. This included compiling, structuring and synthesizing experiences from previous research (Paper I) as well as improving evaluation methodology and doing case studies of specific forestry data. The evaluations comprised both traditional forestry data as currently used and data from recently developed remote sensing methods (paper II, III, V). Method development focused on i) measuring economic consequences of using erroneous data in decision making in timber oriented forestry (paper II and IV), and ii) measuring, in terms of forecasted resource indicators, effects of erroneous data in sub-national scenario analysis (paper III). The specific objectives of papers I-V were:

Paper I. Review the research area concerned with evaluating the appropriateness of forestry data in forest management planning. The included research were delineated to include research on planning effects caused by errors in forestry data even if no particular data acquisition strategy was evaluated. Research on data acquisition were included only if data evaluation included effects on planning and decision making.

Paper II. To evaluate forestry planning data obtained from recently developed sample plot imputation methods based on laser scanning and satellite image data. The consequences of data quality in forest management planning in terms of decision loss and inventory cost on forest stand level were considered.

Paper III. To evaluate the quantitative consequences of using spatially comprehensive data obtained from recently developed sample plot imputation methods based on laser scanning and satellite image data in a sub-national forestry scenario analysis. The evaluation focused on the errors in forecasted resource indicators, such as net income, cutting volume and stand volume.

Paper IV. To introduce cost plus loss analysis into a hierarchical forestry planning situation by mimicking specific tactical and operational planning procedures in a simulation approach. The consequences of data uncertainty in forest management planning in terms of decision loss and inventory cost on forest holding level were to be considered.

Paper V. To evaluate forestry planning data currently used in practice and the effects of updating data prior to harvest operations. Analysis to sensitivity in error levels was included. The consequences of data quality in forest management planning in terms of decision loss and inventory cost on forest holding level were considered.

2 Summary of papers

2.1 Reviewing the research area (Paper I)

Paper I is a review of the research area of this thesis. The aim was to compile previous findings, categorize and scrutinize methodologies and to identify needs for future research. The review was delineated to include studies relating the cost of a data acquisition strategy to the anticipated losses occurring when decisions are to be based on the (erroneous) data. Studies where the losses of using data of different quality were evaluated in a planning context were included even when a distinct connection to a data acquisition strategy was missing. Studies of data acquisition strategies without connections to the effects on planning were excluded.

The compilation of previous research is summarized in section 1.3. Findings and conclusions from paper I are included in the discussion (section 3).

2.2 Evaluating sample plot imputation techniques (Papers II-III)

In papers II and III, a recently developed technique for imputing sample plots to forest stands were evaluated from two different viewpoints. In Paper II the perspective was that of a forest owner engaged in timber oriented forest management. The timing of harvesting activities and the resulting losses in NPV were considered. In Paper III the perspective was instead that of national or sub-national policy makers with focus on the outcome of wood and revenues in specific planning periods. In the first case (Paper II) the imputation technique provides an alternative spatially comprehensive (wall-to-wall) data source to the traditional forest map and stand register. As implemented here it also provides individual tree data to each stand contrary to the average stand values available in traditional stand registers. In the second case (Paper III), it enables, contrary to analyses based on NFI plot data, the use of spatially comprehensive data in, e.g., regional analyses based on a sample of landscapes. It thereby enables the use of models requiring spatially comprehensive data such as habitat models.

2.2.1 Study area and forest data

The forest data used in papers II and III were collected as a sample of the stands at the Remningstorp estate in southern Sweden (lat. 58°30'N, long. 13°40'E). The estate is privately owned and dominated by Scots pine (Pinus sylvestris), Norway spruce (Picea abies), and Birch (Betula spp.).

Field data were collected by surveying 10-m-radius plots using the methods and models incorporated in FMPP (Jonsson et al., 1993). The position of each plot centre was measured using the averages of differentially corrected GPS measurements. The field plots were measured from 1998 to 2003. The condition of each plot was forecasted up to 2003 to match the acquisition date of the remote sensing datasets. A total of 64 stands were included in Paper II while 67 stands were included in Paper III. In all, 870 sample plots were used in both studies.

Satellite image data for the field plot centers were extracted from a SPOT-5 HRG scene, acquired at 10:05 PM on 3 June 2003 and geometrically precision corrected. Laser scanner data were acquired by the airborne TopEye system on 9 August 2003 at a flight altitude of 430 m, resulting in 1.5–2.0 pulses m^{-2}.

The results of sample plot imputations from a previous study (Wallerman and Holmgren, 2007) using the above mentioned data sources, provided the main target data of the evaluation. In these imputations, spatially comprehensive *carrier* data (SPOT-5 HRG, TopEye laser scanner) carried data from the primary data observations (870 reference sample plots) to locations missing primary data. This was accomplished by individually comparing carrier data, at 30 regularly spaced target locations in each stand, to carrier data from all locations for which primary data were available. Primary data from the location having the most similar carrier data observation were then imputed to the target location.

For Paper II, three carrier data combinations were selected for the evaluation; one using SPOT data as carrier data (ImpSp), one using laser scanner data (ImpLa), and one using both SPOT and laser scanner data (ImpLS). In Paper III only two combinations were used, one SPOT based and one laser scanner based.

For comparison purposes, simulated sample plot inventories in the field were created and used in Paper II. The simulation for each stand was carried out using bootstrapping techniques (Efron and Tibshirani, 1993). According to the standard assumptions of bootstrapping, the reference sample plots were assumed to capture the true variation of plot states in a stand. From the reference sample plots in each stand, 5 (Plot5) and 10 plots (Plot10) were sampled with replacement to simulate the outcome of field inventories of different intensities.

In both studies, the original sample plot data for each included stand were used as reference data against which all comparisons were made.

2.2.2 Methodologies

Paper II

Paper II was a traditional cost-plus-loss study. Costs of the imputation methods were estimated as the sum of remote sensing data, the field plot data and data processing for a holding and distributed to a per-hectare value. The cost corresponding to the simulated sample plot inventories were acquired from a leading contractor. This cost was sensitive to stand area (Table 1).

Table 1. *Cost of the different inventory methods for stands of different sizes as used in the cost-plus-loss analysis in Paper II.*

	Total cost of data (SEK)			
	Stand area			
Method	1 ha	2 ha	5 ha	10 ha
Plot10	-	1320	1450	1630
Plot5	-	1040	1130	1250
ImpLS	58	116	290	580
ImpLa	57	114	285	570
ImpSp	22	44	110	220

Based on the data sets to be evaluated, the optimal treatment program for each stand was identified using an adapted version of FMPP (Jonsson et al., 1993). It was assumed that decisions were based on the studied data for a period of ten years. From year ten and onwards, decision were based on the reference data set, with due regard to the possibly erroneous decisions already made. Ten years is a reasonable estimate of the elapsed time between data revisions, as currently used by forestry companies in Sweden. With this setup, we adapted the view that the loss in net present value (NPV) from using a particular dataset should be limited to the time within which it is actually being used in making decisions.

Stand wise loss of using the studied data were calculated as the difference between NPV of the treatment programs identified as optimal and NPV of the treatment program identified as optimal based on the reference data values.

Cost-plus-loss was calculated as the sum of costs and losses for each stand and each of the studied data sets. Alternatives with 2% and 4% real discount rates were tested.

Paper III

In Paper III, the analysis was set up to evaluate the effects of using sample-plot imputation data in national or sub-national forestry scenario analyses. Two imputation data sets were used, one based on SPOT data and one based on laser scanner data. The evaluation focused on predicted harvesting levels, net income, and standing volume in different planning periods. Different cases were specified to evaluate the consequences of different interest rates and forest landscape compositions.

The data from the 67 stands were evaluated in two groups of scenarios. The aim of group 1 was to evaluate the influence of the imputation data using different interest rates in the scenarios. In group 2, three landscapes with different age-class structures were evaluated. The aim was to assess the effect of data quality on the results when the age-class composition of the forest landscape varied. Different landscapes were constructed for the different scenarios. In group 1, all 67 stands were used and the area in each stand was set to 5 ha, so the total area of this landscape was 335 ha. In group 2, three landscapes were constructed, each having a different age-class composition. These landscapes were 5 000 ha in size, and the original 67 stands were given various area weights depending on their stand age. The stand register from the Remningstorp estate was used as the source of age information for each stand. The area weights of the stands in these landscapes are presented in Table 2.

Table 2. *Age structure of the landscapes used in group 2 scenarios and area per stand.*

		Landscape 2a		Landscape 2b		Landscape 2c	
Age class (year)	Number of stands (n)	Total area (ha)	Area per stand (ha)	Total area (ha)	Area per stand (ha)	Total area (ha)	Area per stand (ha)
0–20	12	1 000	83.3	1500	125.0	1 500	125.0
21–40	16	1 000	62.5	2 000	125.0	750	46.9
41–60	6	1 000	166.7	500	83.3	500	83.3
61–80	20	1 000	50.0	500	25.0	750	37.5
81–	13	1 000	76.9	500	38.5	1 500	115.4

The suggested treatments in all scenarios were based on an optimal management regime, which include finding the highest net present value for each stand. The prognoses for each stand were calculated using the FMPP. No further restrictions were made in the scenarios and each stand was treated independently, that is, no landscape level restrictions were placed on even harvest flows, etc.

Analyses based on the reference landscapes (using the original sample plot data) were considered to provide an error-free description of the development. Results are presented about the deviation between the reference analyses and the analyses using imputation data. Deviations are presented either as absolute numbers or as percentages, to provide the most relevant basis for interpreting the results. Furthermore, as a means to aggregate results over longer periods, both a mean deviation,

$$\frac{\left(\sum_{p=1}^{10}(y_{ref_p} - y_{imp_p})\right)/10}{\bar{y}_{ref}} \qquad (8)$$

and a mean absolute deviation, in relative terms, was calculated.

$$\frac{\left(\sum_{p=1}^{10}|y_{ref_p} - y_{imp_p}|\right)/10}{\bar{y}_{ref}} \qquad (9)$$

where *y* is the variable of interest with subscripts *ref* for reference solution, *imp* for the tested imputation dataset and *p* is index for period. Ten five-year planning periods were considered in the evaluation. The mean deviation was used as a measure of systematic deviation and the absolute mean deviation as a measure of variability.

2.2.3 Results & Discussion

Paper II

The decision losses calculated in Paper II, when using the simulated sample plot inventory data were, on average for all stands, considerably lower than when imputed data were used (Table 3). The average decision losses were generally lower when using a 4% real interest rate compared to a 2% rate. When comparing the imputation methods, the method using both laser and satellite data (ImpLS) gave the lowest level of average decision loss, while using satellite data only, gave the highest average decision loss. When data acquisition costs were added in the cost-plus-loss analysis, sample plot inventories was still more competitive compared to the imputation methods. Only in the smallest stands and when the highest discount rate was applied, an imputation method was superior.

Table 3. *Average cost-plus-loss for each method. Bold indicates the best method in each case.*

	Cost-plus-loss (SEK ha^{-1})					
	Interest rate					
	2%			4%		
	Stand area					
Method	2 ha	5 ha	10 ha	2 ha	5 ha	10 ha
Plot10	746	376	**249**	678	308	181
Plot5	**653**	**359**	258	553	**260**	**159**
ImpLS	827	827	827	**404**	404	404
ImpLa	1085	1085	1085	813	813	813
ImpSp	1872	1872	1872	1947	1947	1947

The high average cost-plus-loss when using the imputation methods was the result of very high decisions losses in a few stands. These were a result of limitations of the imputation method. That is, sample plots are to be imputed to a stand from the set of all available sample plots, excluding the plots from the stand in question. Thus, the method work better for stands which share important state properties with other stands. For stands with properties not well represented among the other stands there may simply not be any plots with the right properties to impute. Another limitation, most pronounced when using satellite images, is that the relation between the spectral radiance in the image and forest properties such as crown closure and volume, is weak in dense forests (Franklin, 1986; Horler & Ahern, 1996). This is due to the fact that once the canopy closes the spectral radiance saturates. As the result of the imputation is depending on the aforementioned relation, the usefulness of the method decreases in stands with closed canopies. Thus, coarse errors occur after imputation in dense stands or stands with uncommon properties resulting in the high decision losses observed.

If the best method for each stand was to be selected, the imputation methods would be the choice in most stands. An appealing idea would be to be able to choose different data acquisition methods in different stands. However, this study has found no general method for determining, based on prior information, which type of stand should be inventoried with which method. The relation between costs and losses in the cases using imputed data suggests that it will be more productive trying to improve precision in the data source, hence decreasing decision losses, than to reduce data acquisition costs.

In the analysis setup it was assumed that all plans are executed without alterations. In practical forestry, it is reasonable to assume that really poor decision can be cancelled, for instance, after conducting a pre harvest inventory. This is of course possible only for operations scheduled ahead of the optimal point in time. In an additional analysis, it was assumed that all final felling operations that were scheduled 15 years or more too early were cancelled. Although a considerable decrease in the average decision loss was observed, the overall conclusions of the study remained the same.

Other aspects of forestry planning not covered in this evaluation may further affect the impact of data errors. One is that stands generally exist as parts of larger holdings. On the forest holding level, issues like logistics, delivery agreements and the availability of production resources affects the management of individual stands. Forestry data evaluations could be improved by including such aspects.

Paper III

In Paper III, results are presented about the deviation between the reference analyses and the analyses using imputation data. The scenarios in group 1 were done simply to evaluate the effect of using the imputation data with different interest rates. The results of the analyses did not provide any evidence of major difference in performance of different data sources due to interest rate (Table 4). Comparing the two data sources, the laser-based imputations tended to do better independent of interest rate. In general, greater mean and absolute deviations were obtained when SPOT-based data were used.

Table 4. *Mean deviation (Mean dev.) and Mean absolute deviation (Mean abs. dev.) for standing volume, harvested volume and net income in the group 1 scenarios for different real interest rates for ten five-year planning periods*

Method	Interest rate	Standing volume (%)		Harvested volume (%)		Net income (%)	
		Mean dev.	Mean abs. dev.	Mean dev.	Mean abs. dev.	Mean dev.	Mean abs. dev.
Laser-	2%	-3	4	-6	25	-1	3
based	3%	-2	4	2	25	1	27
	4%	-2	4	1	23	1	23
SPOT	2%	-7	9	-5	30	-1	4
based	3%	-8	11	3	42	1	45
	4%	-13	14	-5	30	-10	33

The harvesting volume was in most cases underestimated in the first planning period, followed by an overestimation in the second. This means that harvesting was delayed compared to the reference case. The initial errors in the estimates for standing volume was in most cases small while systematic differences tended to occur over time. These trends are visualized in Figure 1 using the scenario 1b (real interest rate = 3%) as an example.

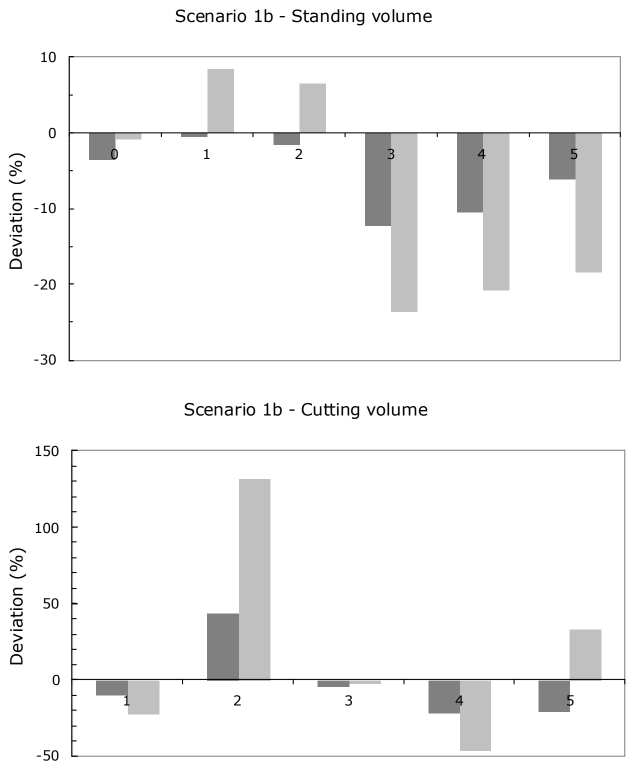

Figure 1. The upper diagram presents the deviation of standing volume estimated by laser-based (dark) and SPOT-based data (light) in Scenario 1b (3 % real interest rate). The lower diagram presents the deviation of cutting volume. Period 0 is the initial state and numbers 1–5 indicate the estimated states and harvests in future 5-year planning periods.

In the analysis of the group 2 scenarios where the effects of different age-class distributions where investigated, the cutting patterns over time was generally as in the group 1 scenarios. Also in these cases, the laser based data were usually better than the SPOT-based. Thus, as expected the accuracy of the carrier data and the imputation estimates clearly influence the accuracy of the scenario analysis. Also, with both imputation methods there was a tendency towards the mean in the initially estimated data. This may have profound effects on the scenario analysis results since the area of both young forests (low volume) and old forests (high volume) are underestimated. With a mean absolute deviation in cutting volumes of 25 % on average as observed for the laser based method, the following example may be illustrative; A deviation of 25 % in the test landscape correspond to approx 12 500 m^3 under bark on the 5000 ha estate used in scenario 2a. If the deviation is an underestimate, an industrial user may have to import the difference at a excess cost of EUR 14 for pulpwood and EUR 10 for sawtimber (Anon., 2005). The cost at estate level, assuming equal distribution between pulpwood and sawtimber would be EUR 150 000. Scaling this example to the national level of Sweden (22.7 mill. ha), costs for raw material would increase by EUR 210 million for a single year. Although hypothetical, it indicates the importance of acquiring relevant data for forestry scenario analyses.

One question is the generalization of the results from the case study. The test site is more homogenous about stand structure and tree species composition than Swedish forests in general. However, the tendency towards the mean as observed for the output of the data acquisition strategy is a general tendency. Similar consequences on scenario analysis are likely to occur also in other geographic areas. Thus, a conclusion is that imputation data should not be used without correction for the tendency towards the mean, in national or sub-national scenario analysis, since the results may be largely misleading.

2.3 Forest data evaluation in a hierarchical forestry planning environment (Paper IV-V)

One of the conclusions from Paper I and II was that the outcome of cost-plus-loss analysis are likely to be strongly affected by alterations of tactical plans in the operational planning step, which in previous research to a large extent has been ignored. In response to this and other deficiencies in data evaluation setups identified, the research for Paper IV-V were initiated to further develop and test cost-plus-loss analysis methodology in more complex forestry planning contexts. The idea for Paper IV rests on the assumption that the behavior of forest managers of an enterprise can be simulated including their planning and execution work. To achieve this, an attempt is made to, as closely as possible, model the continuous planning behavior of a large scale forest owner including tactical and operational planning as well as data acquisition routines. Due to differences in the operational planning routines between different companies it was decided to attempt to tailor it according to the behavior of a specific forest enterprise rather than to any generic forestry planning system existing in theory.

In Paper V the planning simulation system developed in Paper IV were used for the evaluation of empirical current practice data. Additional analyses were also made on data sets for which the error levels were reduced.

2.3.1 Material & Methods

The simulated planning process

The simulation of the tactical and operational planning routines was carried out to mimic the behavior of a large corporate forest owner. The starting point was that a strategic plan, decisive of annual harvest levels and the distribution of thinnings versus final fellings, were already in place. Planning at the selected company is carried out in the following manner. The strategic plan is based on a sample of stands which has been subject to objective sample plot inventories in the field. Both planning and data acquisition for strategic planning follows closely the procedures of the Forest Management Planning Package (FMPP) (Jonsson et al., 1993) Within the framework set by the strategic plan, the tactical planning aims at scheduling harvest operations for individual stands for the next ten years. Here spatial considerations are important e.g. clustering of harvest activities to improve logistics. Several issues which can be controlled through wood flow requirements are considered. Examples are, an even flow of different assortment to the industry, an even utilization of available harvest resources and a mix of stands with different ground conditions corresponding to the length of the season in which they are preferably harvested. The tactical planning step is today basically carried out manually based on data available in the stand database of the company. This data is acquired by quick and subjective field surveys and are hereafter referred to as *basic forestry data*. For further use in the planning process, stands selected in tactical planning for harvests within the first three years are included in a set of stands eligible for harvest (SEH set). As stands are added to the SEH set, they are subject to inventories including some measurements for updating stand volume, tree sizes etc. These are done when field staff also updates the delineation of a harvest area, mark paths for the forwarder to follow, etc. The data collected are used in operational planning and is referred to as *pre harvest data*. Operational planning aims at scheduling operations for the coming year by selecting stands from the SEH set and consequently, the operational plan includes actions to be executed the coming year. The same considerations as in the tactical step are valid also here. The tactical and operational planning steps are repeated annually, including pre harvest inventories of stands added to the SEH set. The company aims at keeping stands in the SEH set corresponding to three years of harvests continuously.

Due to unforeseen events, one-year plans are often partially overthrown during implementation. Re-planning in such cases is ad-hoc and aims at meeting new requirements or to adapt to effects of changes.

The simulation system of the planning process

The simulation system aims at mimicking the planning process at the selected company. The model of the annual and cyclic planning routine can be summarized as follows, starting with year t. Each tactical planning step covers 10 years while the operational planning step covers 2 years.

Start: Let iteration $t = 1$, and $T = 10$.

1. The tactical planning problem is solved based on a data set available generating a schedule for operations for each of the coming t to t+9 years. The solution is subject to restrictions from strategic planning.
2. The stands selected for operations the first three years are marked to belong to a set of stands eligible for harvest (SEH set).
3. The stands in the SEH set are subject to pre harvest inventories.
4. The operational planning problem is solved based on data from pre harvest inventories generating a schedule for operations for each of the coming t to (t+1) years.
5. The operations of the first year of the operational problem are assumed executed, cost and revenues from operations are evaluated using data considered free from errors.
6. The data set used for tactical planning is updated with the pre harvest data obtained for stands belonging to the SEH set. Both data sets are updated depending on executed harvest operations.
7. If $t \leq T+1$ then set $t = t+1$ and repeat annual planning cycle by going to 1. Else (if t = T+1), go to 8.
8. Evaluate the state of the forest at t= T+1 by solving the tactical problem for $t = T+1$ to $t=2T$ using a data set considered free from errors.

In this approach, T represents both the tactical planning horizon (in years), and the number of years for which a tactical problem is solved, by simulating ten years of implemented planning. To let the simulation run for ten iterations, creating operative plans considered executed for the coming ten years was chosen as this is a reasonable estimate of the elapsed time between *basic forestry* data revisions. It was believed that the effects of using erroneous data should be studied for the entire period during which it is involved in the decision making process.

Three different data sets are needed to form a complete evaluation in the system; first, the data available from the outset, the *basic forestry data*; second, the data acquired by *pre harvest* inventories and; finally, a *reference data* set considered free from errors.

The actual evaluations of data were carried out by comparing the outcome from simulation runs using different combinations of data in the tactical and operational planning steps. The outcome compared was the sum of net present values of costs and revenues incurred when plans are executed. Net present value of executed harvest activities year 1-10 are included as well as the capitalized value of the forest at the end of year 10. Costs include a combined cost of hauling harvest equipment and maintaining forestry roads affected by harvest activities. The combined cost is hereafter denoted *road access costs*. As the capitalized value of the forest at the end of year 10 was evaluated by solving a tactical problem with a ten-year planning horizon (step 8), road access costs were included also for years 11–20. The costs of pre harvest inventories were also included.

The simulation system of the planning process was formulated as a loop over a series of procedures created in AIMMS 3.7 (Bisschop and Roelofs, 1999). The optimization problems were solved using ILOG™CPLEX 11.0, (integrated solver in AIMMS). Further details of the system are given in paper IV.

The optimization models

The tactical and operational planning models were formulated as mixed integer programs as described in detail in Paper IV. The objective of the planning models was to maximize net present value (NPV) of stand harvest operations minus the present value of road access costs (Gustafsson et al., 2000). The revenues from harvest activities were introduced into the objective function as the NPV of treatment programs in which the specific harvest activity was included. Thus the selection of a harvest activity meant adding to the objective function, the value of a treatment program calculated, in GAYA (Eriksson, 1983) to the end of the rotation.

An important issue in forestry planning is the potential cost savings that can be obtained by coordinating harvesting in nearby stands. To recognize this in the model, the road network was divided into segments of approximately equal length and each stand was linked to the nearest road segment. If a stand is harvested, a road access cost occurs. Further harvesting activities in stands assigned to the same road segment shared this cost if they were carried out in the same season of the same year and by the same type of harvesting crew. Two types of harvesting crews were available, one for thinning and one for final harvest. Adaptation to seasonal changes of ground conditions was considered by allowing forestry operations in a stands only in designated seasons, depending on the actual stand ground conditions. To satisfy regular deliveries to industry, constraints were set for the year-to-year fluctuation of delivered volumes of each assortment. The harvest volume in each season was constrained to satisfy a certain relative interval of total harvests that year corresponding to the expected length of each season.

Forest data in Paper IV

The model was tested on a data set from a large holding situated in mid Sweden. For the purpose of this study, one administrative area was selected comprising nearly 4 000 stands covering approximately 30 000 ha of forest land. Stands in this area with a mean height of less than seven meters were removed from the analysis since no harvest operations could be anticipated for the next ten years. This resulted in 2 411 stands with a total area of 17 200 ha used for the planning problem.

Basic forestry data were available for all stands from the forest owner's database. For a sample of 80 stands, data from a sample plot inventory were also available. Field data were collected by surveys of 10-m-radius plots using the methods and models in the Forest Management Planning Package (FMPP) (Jonsson et al., 1993). This type of data is hereafter referred to as *reference data*. Although the true state of the forest stands is not perfectly captured by the sample plot inventory, regarding such data as accurately capturing reality has become standard procedure in cost-plus-loss analysis (Eid et al., 2004, Holmström et al., 2003, Holopainen and Talvitie, 2006, Larsson, 1994, Ståhl, 1994a) as it is assumed to have the same relationship with the studied data as true reference data would have had.

The relationship between the *basic forestry data* and the *reference data* were studied for the sample of stands and then used to simulate artificial *reference data* for all stands. The simulation aimed at preserving the multivariate dependency structure of the errors in different forest variables.

As no data from pre harvest inventories could be made available for this case study, the *reference data set* is used in place of *pre harvest data* for operational planning. This did not affect the possibilities to run the simulation system but must be considered when analyzing the results.

Forest data in Paper V

Basic forestry data was acquired from the same holding as in Paper IV and used to form a study area also in Paper V. Due to minor differences in area selection (in a GIS) the study area consisted of 2499 stands, with a mean tree height over 7 m, covering an area of 16 872 ha.

The data needed to create the data sources relationships were obtained from a different holding in mid Sweden. At this area, *basic forestry data* were extracted from the forest owners' database for a sample of 237 stands which had recently been subject to sample plot inventories in the field. Also, another sample of 70 stands was selected as the stands had recently been subject to pre harvest inventories as well as sample plot inventories in the field. Thus, two relationships, *reference data* ↔ *basic foresty data* and *reference data* ↔ *pre harvest data* could be created in the same way as in paper IV (Figure 2). The pre harvest inventory is carried out as a part of the *field planning* of the harvest operation. The cost of the entire field planning activity, including the pre harvest inventory, is estimated at 235 SEK ha^{-1} by the forest owner. No further division of the cost to different parts of the field planning was provided by the forest owner but a reasonable assumption is that the pre harvest inventory represents less than half of the cost.

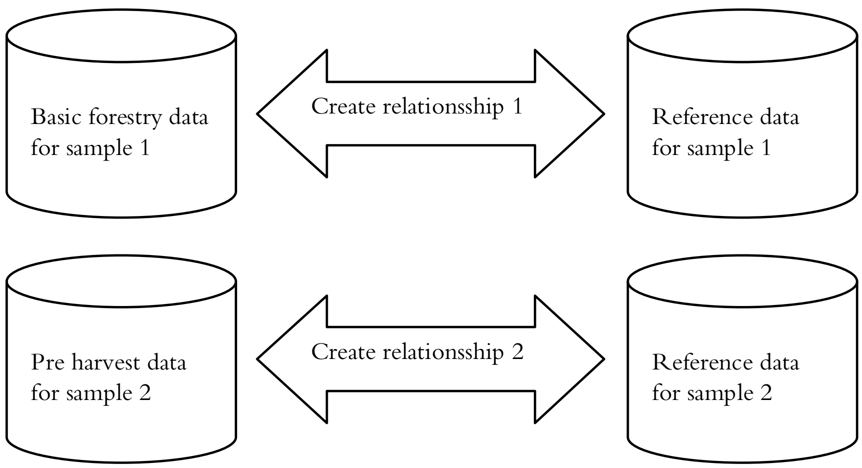

Figure 2. The relationship between the data types were studied on two different samples of stands.

Both relationships created involved *reference data*. It was thus assumed that the most efficient simulation of synthetic data would be to start with a *reference data* set from which the two other data typed could be simulated directly. From the outset, the only complete dataset for the study area was the *basic forestry data*. Hence, this data set was assigned to be used as the *reference data* set of the holding (Figure 3). The alternative to first simulate *reference data* from *basic forestry data* (as in Paper IV) and then simulate *pre harvest data* from the newly created *reference data* was abandoned as additive effects of errors could cause unrealistic correlation between the *basic forestry* and *pre harvest* datasets.

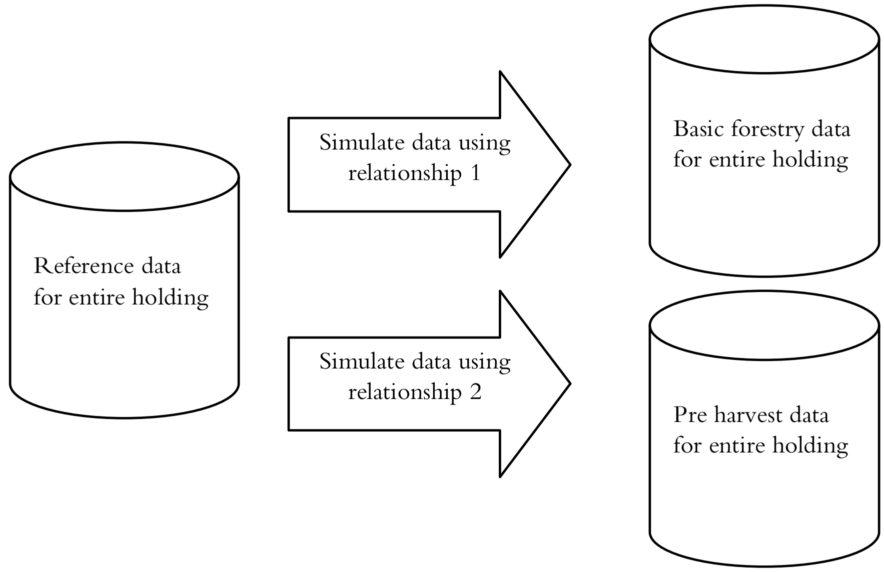

Figure 3. The basic forestry and the pre harvest data sets were simulated separately from the reference data set for the holding.

The datasets created were primarily used in two simulation runs; first a baseline analysis, in which each of the simulated data sets were used as intended; and second, the reference analysis. In addition, alternatives were run with different combinations of input data to enable several different analyses. Analysis of sensitivity to error levels were carried out by decreasing the errors simulated onto the basic forestry and pre harvest data sets. Both systematic and temporary errors were adjusted -35% before simulation to create new data sets. The adjusted datasets are denoted as *improved* and marked in tables with a "+". The combinations of datasets used in the simulations are shown in Table 5 including main analysis possibilities.

Table 5. *Data sets used in different simulation runs and the main analysis objectives of each run. The "+" sign indicates improved data quality.*

Simu-lation	Tactical step	Operational step	Primary analysis
1	Reference	Reference	Reference
2	Basic forest	Reference	Decision loss as incurred by basic forestry data in tactical step
3	Basic forest	Basic forest	Assess the value of pre harvest data
4	Basic forest +	Basic forest +	Assess value of improving basic forest data
5	**Basic forest**	**Pre harvest**	**Baseline, decision loss with data as currently used in practice**
6	Basic forest	Pre harvest +	The value of improved pre harvest data

2.3.2 Results & discussion

Paper IV

In the case study of paper IV, three combinations of data types were used in the tactical and operational planning steps of the simulation system. Summarized results of the simulations are presented in Table 6.

Table 6. *Results of the three simulations as NPV of forest management, the cost of accessing and maintaining roads, the sum of these, the difference in total NPV compared with the reference solution (SEK million), and the difference in NPV ha^{-1} compared to the reference solution (SEK ha^{-1}).*

	Data for		Net present value (NPV)				
Simu-lation	Tactical planning	Operat. planning	Forestry	Road access	Sum	Loss Sum	**Loss (SEK ha^{-1})**
1	Reference	Reference	345.7	–6.8	338.9	**-**	-
2	Basic for.	Reference	343.8	–7.1	336.7	–2.2	**126**
3	Basic for.	Basic for.	321.4	–6.7	314.7	–24.2	**1409**

A comparison between the reference case and the case using basic forestry data in both planning steps indicated what can be lost if the basic forestry data are used for both tactical and operational planning. This can also be expressed as the *expected value of perfect information* (EVPI) setting the upper bound on how much can be spent on acquiring new data (Kangas, 2008); in this context, the obtained loss of SEK 1 409 ha^{-1} is quite large.

When including the results of the simulation using basic forestry data in the tactical planning step and reference data in the operational planning step, it could be concluded that most of this loss stems from the operational planning step. In other words, little is lost in the tactical planning step from using *basic forestry data* if higher-quality data are collected in the operational step. If these complimentary data are "perfect" (*reference data* as in simulation 2), the loss is limited to SEK 126 ha^{-1}.

The results from the test case indicated that the cost-plus-loss simulation system could be run on a holding large enough to make wood flow restrictions and harvest clustering meaningful. Regarding computing capacity, the MIP model used for solving the planning problems was memory intensive, strongly limiting current possibilities of increasing the number of stands in the problem. The studied case, comprising 2 410 stands, resulted in AIMMS using approximately 0.65 GB of memory. Attempts to double the number of stands in the problem failed, as the total required memory (including that needed by Windows XP) exceeded the capacity of the PC (>2 GB).

Paper V

The outcomes of the simulations in paper V were primarily compared to the reference run to assess the loss in NPV stemming from suboptimal management decisions. The losses presented in Table 7 are thus related to the reference run.

Table 7. *Results of the simulations as the loss in NPV compared to the reference solution. Simulation 5 (base line) mimics the present practice at the selected company.*

	Data set entered into:		NPV loss	
Simulation	Tactical planning	Operational planning	Mill. SEK	SEK ha^{-1}
1	Reference data	Reference data	0.0	–
2	Basic forest	Reference	5.1	304
3	Basic forest	Basic forest	16.5	986
4	Basic for. improved	Basic for. Improved	13.7	814
5	Basic forest	Pre harvest	15.2	907
6	Basic forest	Pre harvest, improved	13.2	787

The results of the simulations indicated, as in paper IV, that decision losses from using *basic forestry data* in the tactical step were moderate and that they increased after the operational step. This observation was however not as pronounced as in paper IV. The decision loss of 304 SEK ha^{-1} in simulation 2 stems from the tactical planning step, that is, from erroneous selection of stands to the SEH set due to errors in the *basic forestry data*. This can be compared to the decisions loss of 986 SEK ha^{-1} when *basic forestry data* is used in both planning steps (simulation 3). The difference in decision loss of 682 SEK ha^{-1} may be interpreted as what is lost in the operational step from using the *basic forestry data*.

Results from the baseline run (simulation 5) show that the decision loss caused by errors in current practice data was 907 SEK ha^{-1}. This can be compared to the decision loss of 986 SEK ha^{-1} in the case when basic forest data is used in both planning step (simulation 3). The difference, 79 SEK ha^{-1} is what was gained in terms of reduced decision loss by acquiring pre harvest data. However, in the course of planning, more than half of the stands, corresponding to an area of 11 123 ha, were subject to field planning of operations (incl. pre harvest inventory) at a total cost of 2.61 mill. SEK. This means that the actual cost was 156 SEK ha^{-1} on average for the holding. When comparing this additional cost to the gain in reduced decision loss of 79 SEK ha^{-1}, it is obvious that the profitability of acquiring pre harvest data is dependent on how costs are distributed between different parts of the field planning activity. The assumption that the cost of the pre harvest inventory is less than half of the total cost of the field planning activity, indicates that it may be profitable. Regardless of this distribution, the gain or loss is small.

The decision loss of 787 SEK ha^{-1} when using the improved pre harvest data represents an improvement compared to using the current pre harvest data by 120 SEK ha^{-1} on average for the entire holding. Considering the area actually subject to pre harvest inventories in the simulation run (11 191 ha), it means that an additional cost of 180 SEK ha^{-1} or less can be accepted for the improvement of pre harvest data.

In simulation 4, planning was based on the *improved basic forestry data* in both planning steps. The decision loss of 814 SEK ha^{-1} can be compared to that of 986 SEK ha^{-1} when using the original *basic forestry data*, showing a reduction in decision loss of 172 SEK ha^{-1}. The size of the reduction sets the limit to what additional cost can be accepted for a 35 % reduction in basic forestry data error levels, in cases when pre harvest inventories are skipped.

The studied data sets were also evaluated as input in stand level management planning, that is, simply selecting the treatment program per stand with the highest NPV without considering holding level wood flow restrictions or implications of harvest clustering. The selection was done in a single planning step. The decision loss when using the basic forestry data was increased by 153 SEK ha^{-1} compared to the corresponding outcome from the simulated planning system (simulation 3). Also when using the pre harvest data set as input in stand level planning, the average decision loss appears to increase (by 72 SEK ha^{-1} compared to simulation 5). This result is however not completely comparable as basic forestry data affects the outcome in simulation 5. The reduction of decision losses when adding complexity to the planning problem is an important finding.

The basic approach taken in this study to overcome the simplifications common in cost-plus-loss studies is to set the analysis within the context of the planning procedure of the forest company. Thus the hierarchical structure of the simulated planning system acknowledges the introduction of pre harvest data as currently practiced in Sweden. Also, several issues important when planning on holding level is included. However, even though the proposed method represents an advance in the realism of cost-plus-loss analysis in forestry planning, there are still some important issues to address. The tactical planning steps of the simulated planning process follow the procedures of Swedish forest companies, the only difference essentially being that a more sophisticated method is used, the mixed integer optimization model, than in actual planning. The operational planning step is more artificial. Experience shows that even carefully prepared one-year plans are likely to be, at least to some extent, overthrown by unforeseen events. A model of manager behavior when handling unforeseen events has yet to be implemented in a cost-plus-loss analysis. Moreover, in the operational planning step, the optimization model represents greater sophistication than found in actual practice. Also, the implications of using simulated, instead of real data in the analysis is unknown. Hence, one possible improvement to the proposed advancement of cost plus loss analysis is to explore the effects of using simulated data. Another is to fathom the effects of stochasticity at the operational planning stage.

3 Discussion & Conclusions

This thesis summarizes results from five papers about evaluation of data for forest management planning and decision making. The thesis deals with two areas related to forest data collection: evaluation of data sources and development of evaluation methodology.

3.1 Evaluations of forest data

Papers II and III were evaluations of the outcome from new data collection methods in which sample plots were imputed to forest stands (Wallerman and Holmgren, 2007). The imputations were based on SPOT medium-resolution satellite data and laser scanning data from the airborne TopEye sensor. The main conclusions are:

i) As input for forest management planning at the stand level, the data imputation methods were inferior to field based sample plot inventories. This was due to high decision losses.

ii) Using imputed data in national scenario analysis resulted in an underestimation of future standing volume. Simulated harvest levels were underestimated or delayed in time. The proportions of these deviations imply that the data were not well suited to be used in scenario analysis.

iii) Imputations based on laser scanner data was much more useful than imputations based on satellite data.

In Paper V, an evaluation of current practice data at a corporate forest owner was done with the planning simulation system developed in paper IV. The simulation of a hierarchical planning system including simulation of pre harvest inventories, and other procedures in operational planning improves the cost plus loss methodology. Conclusions from paper V are:

i) The size of the decision loss (907 SEK ha^{-1}), as found in the analysis of using current practice data, represents a high loss of revenues for a corporate forest owner.
ii) Errors due to incorrect selection of stands to the set of stands eligible for harvest (in tactical planning) causes a minor part of the total decision loss incurred. The larger part stems from erroneous decisions on the operational level.
iii) The current practice of doing inventory just before harvest (pre harvest inventory) is barely profitable with regard to the reduction in decision loss. Other uses of the pre harvest data, not included in the present analysis, may increase the value of inventory.
iv) The introduction of holding level wood-flow considerations and incitements to cluster harvest activities reduced decision losses compared to stand wise planning without such considerations.

3.2 Development of evaluation methodology

The underlying question of this thesis is about the forest manager's investment in data for forestry planning, how good is good enough?

Attempts to provide answers through the analytical approaches cited above are based on some very simplifying assumptions on the planning and decision processes as well as reliance on subjectivity, and leave many question unanswered about the real world validity of the results (Borders et al., 2008). The last published effort in this area was made in 1994 (Ståhl et al., 1994). Difficulties in formulating the problem in a way that are both complex enough to produce meaningful results and yet solvable may explain the seemingly low interest to tackle these issues analytically.

Instead, most recent research on data acquisition strategy, including this thesis, has used the simulation approach. An obvious advantage with this is that the relation between data errors and decision loss is a result of the simulation. Hence, the derivation of a function that connects error and loss (Burkhart et al., 1978, Hamilton, 1978) which is considered a weak part of those studies (Holmström et al., 2003, Ståhl, 1994a) is not needed. Also, including several forest variables, upon which the decisions can be based, is rather straight forward in a simulation approach while it increases complexity dramatically in the analytical approach. An aspect that needs to be considered concerning the simulation approaches is that they are, or should be regarded as, case studies. This raises some questions about their general validity.

One important conclusion from paper I was that cost plus loss analysis in previous research concerned highly simplified planning and decision situations. How the results would change if these simplifications could be overcome had not been studied. The simplifications identified in paper I and II as likely to affect cost plus loss results raised important question to be dealt with. Some of these questions, their implications and, in some cases, suggested solutions are presented below:

i) Does existing information have a value?

Kangas (2008) argues that ignoring the existence (value) of prior information is a short-coming of previous cost plus loss analyses in forestry research. While cost plus loss analysis is an efficient way to compare different data acquisition strategies it is usually done without considering the alternative of using prior information. However, prior information always exists in forestry (Kangas, 2008). In principle, including prior information as one data source in cost plus loss analysis is possible. A probable reason why this is usually not done is that data acquisition usually has been carried out periodically for an entire holding or parts thereof. This is reflected in most cost plus loss analysis by the assumption that the decision to acquire new data has already been made. How prior information is accounted for in papers IV and V depends on the focus of the analysis. If the target of the evaluation is the pre harvest inventory, the basic forestry data constitutes the prior information. If the target of the analysis is the acquisition of basic forestry data for tactical planning, the problem of not considering prior information remains.

In theory, this problem can be overcome using the proposed simulation system if information on how data deteriorates over time can be made available. By comparing the increase in decision loss from simulation runs with ageing data with the data cost, the optimal life length of data can be found. As a consequence, the question of when to acquire new data at all can be answered.

ii) Have all uses, for which the data has a value, been included?

The value of forestry data depends on for what it is used. If the information can be used for several purposes or by other users the cost of the data can be shared. In the evaluations in papers II, IV and V, the most important silvicultural decisions in timber oriented forestry, thinnings and final felling, are included. At a corporate forest owner, the same data is likely to be used also for other purposes, e.g. in decision making related to issues concerning environmental considerations. No such decisions are accounted for in studies included in this thesis.

Another aspect not dealt with when developing the methodology for papers IV and V is the information value for other users. An obvious example concerns the effect of erroneous predictions of harvested volumes. Deviations in harvested volumes may have an impact on transportation and wood processing planning.

For example, if erroneous forest data lead to deviations in predicted wood delivered to a sawmill, then the sawmill may have increased costs for altering production plans or buying additional (or selling surplus) timber. Therefore, the value of the information would increase. In the settings of papers II, IV and V, this could be seen as an example of what Ketzenberg et al. (2007) terms *marginal use* of information. If the needs of the timber-users were to be included as a marginal use, the total value of the information would increase and hence motivate a more intensive data acquisition strategy. In principle, it does not matter if the timber user is the same as the forest owner or if it is another economic agent; that concerns only how the economic rent accruing from better data should be distributed and is a question dealt with in transaction cost theory (see e.g. Williamson, 1985). For instance, the cost of better data of one agent is compensated by incorporating the value of precision in wood flows in the form of price increases or bonuses. That is, a higher price is paid if the actual deliveries meet agreed requirements. Analogous, a lower price is paid if deliveries differ from agreements. This is often found in practice and agreement may apply to the volume, quality and timing of deliveries.

iii) For how long is the data used, and what happens thereafter?

Commonly in cost plus loss studies, the studied data is used for decision making for the remainder of the rotation of a stand. In studies limited to deciding the timing of final harvest in mature stands, this may be appropriate. In analyses where younger stands are included this is more problematic as it is unlikely that decision on harvest activities in a distant future will be based on data available today. To avoid this, another approach has been to limit the time within which decision are actually made based on the studied data, and to evaluate the state of the forest thereafter using data considered free of errors. The latter approach appears viable if periodical reassessments of data, as often practiced, are assumed. In papers II, IV and V ten years was assumed to be a reasonable estimate between data revision in current practice. Thus the analysis was restricted to decisions based on erroneous data for ten years. After that, the state of the forest were evaluated by simulating optimal management with due regard to erroneous decisions already made.

As suggested in i), extending the analysis to include the effect of data ageing would enable a better understanding of economically viable length of life of the data.

iv) How does operational planning and planning on holding level affect the decision loss?

A common simplification in previous cost plus loss analyses is the assumption of total *responsiveness* to the data. This means that decision are based entirely on the studied data, as is the case in management planning on stand level. Planning on holding level introduces other considerations which may influence the decisions for individual stands. Thus, in planning on holding level, the assumption of total responsiveness to the data is no longer valid. In general, the value of information decreases whenever the decision responsiveness to the information decreases (Ketzenberg et al., 2007). Thus, when developing the planning simulation system in paper IV the holding level considerations assumed most important, harvest clustering and even-flow requirements were included. The outcome of the study was in compliance with the general relation between information responsiveness and value, affirming the importance of including holding level considerations in the analyses.

3.3 Future research

There are several possibilities to develop the suggested cost plus loss approach. The theoretical framework could for instance be improved by incorporating functionality for considering aspects of how stochastic exogenous events affect the feasibility of forestry operations. Also aspects of *market* uncertainty could be added. Explicitly, these may include uncertainties in demand and prices. However as long as the nature of the price process is in doubt it is difficult to appreciate how market uncertainty interacts with data uncertainty. For evaluation of data in practical forestry, the applicability could be improved by elaborating on the functionality accounting for how delivery agreements are set up. Especially in cases were a value is put on the precision in predicted (delivered) harvest volumes, the value of information is likely to be affected.

Collection of real data sets needed for a cost plus loss study is associated with considerable cost. It is thus likely that also future evaluations will be carried out using simulated data. The effect on the results from using simulated data as opposed to using real data should be scrutinized.

3.4 Concluding remarks

When dealing with the issues considered problematic in cost plus loss analysis, it was often assumed more important to closely mimic observed manager behavior than generic planning procedures. Thus, thriving to improve cost plus loss methodology throughout the work in this thesis has led in the direction of analyses in more complex but also more specified planning environments. The high level of specification in papers IV and V includes the details of the evaluation system as well as corporate specific study data. Hence, the outcome from the simulations presented should be viewed as cases studies and their general validity cannot be guaranteed. However, the evaluation system developed in paper IV has an overall design that makes it, with only minor alterations, usable for most of the large corporate forest owners in Sweden.

It is thus suggested that evaluations of data of interest are done regularly in different versions of evaluation systems tailored according to the behavior of different forest owners. As well as providing results relevant in each specific case, the accumulated experience from many such evaluations will provide valuable knowledge on how data with different properties affect the sizes of decision losses and also how they are interconnected with the properties of forestry planning. With such experiences at hand, also the general validity of the cases can be better judged.

References

Angelstam, P. & Andersson, L. (2001) Estimates of the needs for nature reserves in Sweden. *Scandinavian Journal of Forest Research*, 38-51.

Anon. (1978) *Skog för framtid. Betänkande av 1973 års skogsutredning,* Stockholm.

Anon. (2000) *Skogliga konsekvensanalyser 1999 -skogens möjligheter på 2000-talet.* Jönköping, Skogsstyrelsen. Rapport 2/2000. ISSN 1100-0295.

Anon. (2005) *Swedish statistical yearbook of forestry 2005.* Jönköping, National Board of Forestry.

Anon. (2008) *Skogliga konskvensanalyser 2008 –SKA-VB 08.* Jönköping, Skogsstyrelsen. Rapport 25/2008. ISSN 1100-0295.

Backéus, S., Wikström, P. & Lämås, T. (2005) A model for regional analysis of carbon sequestration and timber production. *Forest Ecology and Management,* 216, 28-40.

Barth, A. (2007) Spatially Comprehensive Data for Forestry Scenario Analysis. *Acta Universitatis Agriculturae Sueciae 2007:101.*

Bell, D. E. (1982) Regret in decision making under uncertainty. *Operations Research,* 30, 961-981.

Bell, D. E. (1985) Disappointment in decision making under uncertainty. *Operations Research,* 33, 1-27.

Bengtsson, G., Homlund, J., Lundström, A. & Sandewall, M. (1989) *Long-term forcasts of timber yield in Sweden, AVB 85.* Umeå. Swedish University of Agricultural Sciences. Dept of Forest Survey. Report 44.

Bisschop, J. & Roelofs, M. (1999) *AIMMS The user's guide.,* Haarlem, The Netherlands, Paragon Decision Technology B.V.

Boman, M. & Mattson, L. (1999) Evironmental economics: a young and vital branch on the Swedish forest economics tree. *Scandinavian Forest Economics,* 37, 1-28.

Borders, B. E., Harrison, W. M., Clutter, M. L., Shiver, B. D. & Souter, R. A. (2008) The value of timber inventory information for management planning. *Canadian Journal of Forest Research,* 38, 2287-2294.

Brim, O. G., Glass, D. C. & Goodman, N. (1962) *Personality and Decisions Processes, Studies in the Social Psychology of Thinking,* Stanford, Stanford University Press.

Burkhart, H. E., Stuck, R. D., Leuschner, W. A. & Reynolds, M. R. (1978) Allocating inventory resources for multiple-use planning. *Canadian Journal of Forest Research,* 8**,** 100-110.

Cochran, W. G. (1963) *Sampling techniques,* New York, Wiley.

Cohon, J. L. (1978) *Multiobjective programming and planning,* New York, Academic Press.

Davis, R. G., Johnson, K. N., Bettinger, P. S. & Howard, T. E. (2001) *Forest management - To sustain ecological, economic and social values,* New York, McGraw Hill.

de Vries, S. & Goossen, M. (2002) Modelling recreational visits to forests and nature areas. *Urban Forestry and Urban Greening,* 1**,** 5-14.

Dean, C., Roxburgh, S. & Mackey, B. g. (2004) Forecasting landscape-level carbon sequestration using gridded, spatially adjusted tree growth. *Forest Ecology and Management,* 194**,** 109-129.

Edenius, L. & Mikusinski, G. (2006) Utility of habitat suitability models as biodiversity assessment tools in forest management. *Scandinavian Journal of Forest Research,* 21**,** 62-71.

Efron, B. & Tibshirani, R. (1993) *An introduction to the bootstrap,* New York, Chapman & Hall.

Eid, T. (1991) *Consequences of erroneous data basis for planning and management of forest holdings.* Ås, Skogforsk. Rapport Skogforsk 9/91.

Eid, T. (1993) *Random errors and strategic planning in forestry*. Ås, Norwegian Forest Research Institute. Medd. Skogforsk 46/7.

Eid, T. (2000) Use of uncertain inventory data in forestry scenario models and consequential incorrect harvest decisions. *Silva Fennica,* 34**,** 89-100.

Eid, T., Gobakken, T. & Naesset, E. (2004) Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analyses. *Scandinavian Journal of Forest Research,* 19**,** 512-523.

Eid, T. & Hobbelstad, K. (2000) AVVIRK-2000: A large-scale forestry scenario model for long-term investment, income and harvest analysis. *Scandinavian Journal of Forest Research,* 15**,** 472-482.

Eriksson, E., Gillespie, A. R., Gustavsson, L., Langvall, O., Olsson, M., Sathre, R. & Stendahl, J. (2007) Integrated carbon analysis of forest management practices and wood substitution. *Canadian Journal of Forest Research,* 37**,** 671-681.

Eriksson, L. O. (1983) *Timber class formation by cluster analysis*. Garpenberg. Department of forest technology, Swedish University of agricultural sciences. *Report 157.*

Franklin, J. 1986. Thematic Mapper analysis of coniferous forest structure and composition. *International Journal of remote Sensing* 7(10), 1287-1301.

Gertner, G. (1987) Approximating precision in simulation projections: an efficient alternative to Monte Carlo methods. *Forest Science,* 33**,** 230-239.

Gustafsson, K. & Hägg, S. (2004) Skogliga *konsekvensanalyser 2003 - SKA 03.* Jönköping, Skogsstyrelsen. Rapport 4/2004

Gustafsson, K., Larsson, M., Källman, U. & Eriksson, L. E. (2000) Logistical Determinants in Timber Harvesting Schedule. Accounting for geographical aspects in medium term forest planning. In: *Proceedings of the First World Symposium on Logistics in Forest sector. Logistics in the Forest Sector.* Helsinki, Finland, 2000. University of Helsinki and IUFRO.

Gustavsson, L., Madlener, R., Hoen, H. F., Jungmeier, G., Karjalainen, T., Klöhn, S., Mahapatra, K., Pohjola, J., Solberg, B. & Spelter, H. (2006) The role of wood material for greenhouse gas mitigation. *Mitigation and Adaptation Strategies for Global Change,* 11**,** 1097-1127.

Hamilton, D. A. (1970) *Precision requirements for some information in timber management decisions.* Ph.D. dissertation, Iowa State University.

Hamilton, D. A. (1978) Specifying precision in natural resource inventories. In *Proceeding from the workshop: Integrated inventories of renewable natural resources.* Tucson, Arizona, USDA Forest Service.

Hansson, S. O. (2005) *Decision Theory -A Brief Introduction.* [Online] Department of Philosophy and the History of Technology, Royal Institute of Technology (KTH) Stockholm. Available from: http://www.infra.kth.se/~soh/decisiontheory.pdf [Accessed 25th October 2008].

Hirshleifer, J. & Riley, J. G. (1979) The Analytics of Uncertainty and Information -An Expository Survey. *Journal of Economic Litterature,* XVII**,** 1375-1421.

Hobbelstad, K. (1988) *Planleggningsprogrammet AVVIRK3.* Ås, Norges Landbrukshögskole, Institutt for skogstaksjon. Melding 42.

Hoen, H. F. (1990) Theoretical and empirical studies of long range forest management planning. *Doctor Scientiarium Theses* 1990:23, Agricultural University of Norway.

Hoen, H. F. & Eid, T. (1990) En modell for analyse av behandlingsstrategier for en skog ved bestandssimulering og lineær programmering. Ås, Norwegian Forest Research Institute. Report 9/90.

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

Holopainen, M. & Talvitie, M. (2006) Effect of data acquisition accuracy on timing of stand harvests and expected net present value. *Silva Fennica,* 40**,** 531-543.

Horler, D.N.H. & Ahern, F.J. Forestry information content of Thematic Mapper data. *International Journal of Remote Sensing* 7(3), 405-428.

Ihalainen, M., Salo, K. & Pukkala, T. (2003) Empirical prediction models for *Vaccinium Myrtillus* and *V. vitis idaea* berry yields in North Karelia, Finland. *Silva Fennica,* 37**,** 95-108.

Jacobsson, J. (1986) *Optimization and data requirements - A forest management planning problem.* Ph.D. dissertation,Umeå, Swedish University of Agricultural Sciences.

Jonsson, B., Jacobsson, J. & Kallur, H. (1993) The Forest Management Planning Package. Theory and application. Studia Forestalia Suecicae 189.

Kangas, A. (1996) On the bias and variance in tree volume predictions due to model and measurement errors. *Scandinavian Journal of Forest Research,* 11**,** 281-290.

Kangas, A. (1997) On the prediction bias and variance in long-term growth projections. *Forest Ecology and Management,* 96**,** 207-216.

Kangas, A. (2008) Value of Forest Information. *European Journal of Forest Research,* (Accepted for publication).

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

Ketzenberg, M. E., Rosenzweig, E. D., Marucheck, A. E. & Metters, R. D. (2007) A framework for the value of information in inventory replenishment. *European Journal of Operational Research,* 182**,** 1230-1250.

Lappi, J. (1992) JLP. A linear programming package for management planning. Suonenjoki, The Finnish Forest Research Institute, Research paper 414

Larsson, M. (1994) *The significance of data quality in compartmental forest registers in estimating volume growth and non-optimal losses -A study of final felling compartments in northern Sweden.* Lic. Dissertation, Swedish University of Agricultural Sciences, Department of Biometry and Forest Management, Report 26.

Lindhagen, A. & Hörnsten, L. (2000) Forest recreation in 1977 and 1997 in Sweden: changes in public preference and behaviour. *Forestry,* 73**,** 143-153.

Lundström, A. & Söderberg, U. (1996) Outline of the Hugin system for long-term forecasts of timber yields and possible cut. *IN* Päivinen, T., Roihuvuo, L. & Siitonen, M. (Eds.) *Large-scale forestry scenario models: experiences and requirements.* Joensuu, Finland, 15-22 June, 1995, EFI.

Lämås, T. (1996) *Forest Management Planning for Biodiveristy and Timber Production.* Ph.D. Dissertation, Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics. Report 3

Lämås, T. & Eriksson, L. O. (2003) Analysis and planning systems for multiresource, sustainable forestry: the Heureka research programme at SLU. *Canadian Journal of Forest Research,* 33**,** 500-508.

Macauley, M. K. (2006) The value of information: Measuring the contribution of space-derived earth science data to resource management. *Space Policy,* 22**,** 274-282.

McCall, J. J. (1982) *The Economics of Information and Uncertainty,* Chicago, University of Chicago Press.

Mikusinski, G. & Edenius, L. (2006) Assessment of spatial functionality of old forest in Sweden as habitat for virtual species. *Scandinavian Journal of Forest Research,* 21**,** 73-83.

Mintzberg, H., Dury, R. & Théorêt, A. (1976) The Structure of 'Unstructured' Decision Processes. *Administrative Sciences Quarterly,* 21**,** 246-275.

Muinonen, E. & Tokola, T. (1990) An application of remote sensing for communal forest inventory. *IN* Sylvander, R. & Poso, S. (Eds.) *The usability of remote sensing for forest inventory and planning. Proceedings from the SNS/IUFRO workshop in Umeå 26-28 February 1990.* Umeå, Swedish University of Agricultural Sciences, Remote sensing Laboratory, Report 4.

Nabuurs, G. J. & Päivinen, R. (1996) *Large scale scenario models: a compilation and review.* Joensuu, Finland. European Forest Institute. EFI Working Paper 10.

Nelson, J. (2001) Assessment of harvest blocks generated from operational polygons and forest-cover polygons in tactical and strategic planning. *Canadian Journal of Forest Research,* 31**,** 682-693.

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

Pukkala, T. (1999) *Monsu-metsäsuunnitteluohjelma. Ohjelmistojen toiminta ja käyttö (User's Guide of the Monsu forest planning program).* Joensuu, University of Joensuu.

Pukkala, T. (2004) Dealing with ecological objectives in the Monsu planning system. *Silva Lusitana,* 12 Especial**,** 1-15.

Pukkala, T., Nuutinen, T. & Kangas, J. (1995) Integrating scenic and recreational amenities into numerical forest planning. *Landscape and Urban Planning,* 32**,** 185-195.

Ricotta, C. & Avena, G. (2003) On the relationship between Pielou's evenness and landscape dominance within the context of Hill's diversity profiles. *Ecological Indicators,* 2**,** 361-365.

Siitonen, M. (1995) The MELA system as a forestry modelling framework. *Lesnictvi-Forestry,* 41**,** 173-178.

Simon, H. A. (1960) *The New Science of Management Decision,* New York, Harper & Row.

Simon, H. A. (1976) *Administrative behavior,* New York, The MacMillan Company.

Sprängare, B. (1975) *A method for analysing the sensitivity of the long-range planning which depends on errors in stand data*. Stockholm, Royal college of forestry. Rapporter och uppsatser 87.

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

Ståhl, G. (1994a) Bootstrap-estimation of the cost-plus-loss minimizing number of plots in compartmentwise forest inventories. *IN: Ståhl, G. Optimizing the Utility of Forest Inventory Activities*. Ph.D. Dissertation, Umeå, Swedish University of Agricultural Sciences, Section of Forest Mensuration and Management, Report 27.

Ståhl, G. (1994b) Optimal Stand Level Forest Inventory Intensities under Deterministic and Stochastic Stumpage Value Assumptions. *Scandinavian Journal of Forest Research,* 9**,** 405-412.

Ståhl, G., Carlsson, D. & Bondesson, L. (1994) A method to determine optimal stand data acqusition policies. *Forest Science,* 40**,** 630-649.

Tokola, T., Kangas, A., Kalliovirta, J., Mäkinen, A. & Rasinmäki, J. (2006) *SIMO-SIMulointi ja Optimointi uuteen metsäsuunnitteluun*. Helsinki, Metsäntutkimuslaitos Metsätieteen aikakauskirja*,* 1**.**

Tomppo, E. (1990) Designing a satellite image aided National Forest Survey in Finland. *IN:* Sylvander, R. & Poso, S. (Eds.) *The usability of remote sensing for forest inventory and planning: proceedings from the SNS/IUFRO workshop*. Umeå, Swedish University of Agricultural Sciences, Remote sensing Laboratory.

Wallerman, J. & Holmgren, J. (2007) Estimating field-plot data of forest stands using airborne laser scanning and SPOT HRG data. *Remote Sensing of Environment,* 110**,** 501-508.

Weintraub, A. & Cholaky, A. (1991) A hierarchical approach to forest planning. *Forest Science,* 37**,** 439-460.

Williamson, O. (1985) *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, The Free Press, New York.

Witte, E. (1972) Field research on complex decision-making processes - the phase theorem. *International Studies of Management and Organisation,* 2**,** pp. 156-182.

Yates, F. (1960) *Sampling methods for censuses and surveys,* London, Charles Griffin and Co.

Öhman, K. (2000) Creating continuous areas of old forest in long-term forest planning. *Canadian Journal of Forest Research,* 30**,** 1817-1823.

Öhman, K. (2001) Forest planning with consideration to spatial relationships. *Acta Universitatis Agriculturae Sueciae, Silvestria* vol. 198.

Öhman, K. & Lämås, T. (2003) Clustering of harvest activities in multi-objective long-term forest planning. *Forest Ecology and Management,* 176**,** 161-171.

Acknowledgements

First and foremost, I thank my supervisors. Their guidance throughout the last 5 years of Ph.D. studies has not only been of great help but a crucial prerequisite for progress. My gratefulness towards my main supervisor, Associate Professor Tomas Lämås, for aid on scientific matters goes without saying. My gratitude towards your patience and your faith is however something that I have not often expressed, it has been of equal importance. I also want to express my gratefulness to my co-supervisor Dr. Peder Wikström. Our discussions have always been rewarding, at least on my part! I owe special thanks to Professor Ljusk Ola Eriksson who not only encouraged me to follow this path in the first place but who has been a constant source of inspiration.

In the everyday work, Dr. Andreas Barth was often the first to be exposed to my ideas and variously successful scientific efforts. I hope that you benefited from this as much as I did. My gratefulness towards Mikael Egberth can be given no simply expressed reason, they are just too numerous. Thanks for everything! I also extend my thanks to the many colleagues at SLU that I've cooperated with or who have simply contributed to creating a creative and pleasant workplace.

Finally, I want to take the opportunity to express my gratitude towards my growing family who have endured throughout this time. Although sad to leave the University, I'm happy that my commuting days are finally over.

Sundsvall 2009-01-15
Karl Duvemo