Spatially Comprehensive Data for Forestry Scenario Analysis

Consequences of Errors and Methods to Enhance Usability

Andreas Barth

Faculty of Forest Sciences
Department of Forest Resource Management
Umeå

Doctoral thesis
Swedish University of Agricultural Sciences
Umeå 2007

Acta Universitatis Agriculturae Sueciae
2007: 101

ISSN 1652-6880
ISBN 978-91-85913-00-8

Tryck: Arkitektkopia, Umeå 2007

Abstract

Barth, A. 2007. *Spatially comprehensive data for forestry scenario analysis: consequences of errors and methods to enhance usability.* Doctor's dissertation.
ISSN 1652-6880, ISBN 978-91-85913-00-8.

This thesis focuses on the use of forest data for national level policy making. Three major issues were considered: (i) to determine typical requirements of data in forestry scenario analysis, (ii) to evaluate and further develop methods to determine data requirements, and (iii) to develop methods that improve data usability in forestry scenario analysis. Increasingly, the trend is to use spatially comprehensive data as a basis for forestry scenario analysis. Compared to traditional approaches, often limited to sample data, this allows for a broader scope. This is needed since sustainable forestry today must encompass economical and ecological, as well as social perspectives. Different approaches to linking data acquisition strategies with decisions that typically are based on forestry scenario analyses were used in the determination of data requirements.

In Paper I, a qualitative framework was developed and applied. The conclusions were that none of the currently used Swedish data acquisition strategies were able to provide data for adequate multi-resource forestry scenario analysis at national level. In Papers II and III, two quantitative approaches were used for the evaluation of sample-plot imputations; using a decision support system the quantitative consequences of errors and cost-plus-loss with simulations were considered. From Paper II it was clear that traditional approaches to acquiring spatially comprehensive data may lead to severe errors in scenario analyses. Both papers concluded that improvements are required in the methodology of assessing the data. In Paper IV, an analytical cost-plus-loss approach was used to address the issue of decision-making at the national level linked to national forest inventories. The conclusion was that the current level of Swedish national forest inventory is motivated fully by the role of the inventory to provide information for national level timber harvesting planning, whereas the inventory serves many other purposes as well. In Papers V and VI, methods were developed and tested regarding how the usability of spatially comprehensive data for national level forestry scenario analysis can be enhanced. In Paper V an algorithm for spatially consistent imputation within forest stands was developed and found to deliver good results in a case study. In Paper VI, a framework for landscape level imputation aiming at preserving overall composition while enhancing spatial configuration was outlined and tested. A core component of the framework was a restricted imputation algorithm that ensured that the classical imputation problem of data "tending towards the mean" was avoided. Case studies showed promising results, but it is clear that the methodological tool-kit must be further developed before it can be applied in practice.

Keywords: forest inventory, data acquisition, forest management, decision support systems, forestry scenario analysis, data requirements, national level forest planning, decision-making, policy-making, cost-plus-loss analysis, medium-resolution satellite data, laser scanner data, national forest inventory, data usability, optimisation, imputation

Author's address: Andreas Barth, Department of Forest Resource Management, SLU, SE-901 83 Umeå, Sweden.

Content

Appendix

Papers I-VI

The present thesis is based on the following papers, which will be referred to by their Roman numerals:

I. Barth, A., Lind, T., Petersson, H. & Ståhl, G. 2006. A framework for evaluating data acquisition strategies for analysis of sustainable forestry at national level. *Scandinavian Journal of Forest Research* 21: 94-105.

II. Barth, A., Duvemo, K. & Wallerman, J. 2007. Evaluation of sample plot imputation in sub-national forestry scenario analysis. Manuscript.

III. Duvemo, K., Barth, A. & Wallerman, J. 2007. Evaluating sample plot imputation techniques as input in forest management planning. *Canadian Journal of Forest Research* (In press).

IV. Barth, A. & Ståhl, G. 2007. Determining sampling size in a national forest inventory by cost-plus-loss analysis. Manuscript.

V. Barth, A., Wallerman, J. & Ståhl, G. 2007. Spatially consistent nearest neighbor imputation of forest stand data. Manuscript.

VI. Barth, A., Lind, T. & Ståhl, G. 2007. Improving spatial consistency in landscape level data for forestry scenario analysis. Manuscript.

Papers I and III are reproduced with kind permission of the publishers.

Introduction

Forest management planning today involves decisions where a wide variety of resources and objectives must be considered simultaneously (cf., Davis *et al.*, 2000). Thus, planning is becoming more complicated. This involves the entire process from assessing the input data to making final decisions based on results from a decision support system. Analyses made with these systems are typically termed “forestry scenario analyses.” This thesis focuses on the requirements on forest data to be used in such decision support systems in general, and specifically on requirements on spatially comprehensive data for forestry scenario analysis at national and sub-national level. Spatially comprehensive data are determined as a lattice of units, where one unit is linked to adjacent units, and then to the second adjacent units, and so on (cf., Cressie, 1993). Units can be linked to each other in regular or irregular patterns. The data can be a complete coverage (i.e. “wall-to-wall” data) or only a partial coverage of an area.

Decision-making in forestry

Human demands for different forest resources, which in turn affect the objectives of forestry, tend to change over time (Davis *et al.*, 2000; Ekelund & Hamilton, 2001). During the twentieth century timber production had a unique role in forest management. Even today, the main objectives for forestry are also economical, and the production of timber substantially contributes to the common wealth. However, society currently has a strong interest in other benefits from the forest as well. Many societal objectives are common to the objectives of individual landowners, but priorities vary. Forest resources include a variety of good and services, and not only the production of timber. For example, the public appreciates forests for its recreational values as well as for picking mushrooms and berries, and hunting and fishing (e.g., Pukkala, Nuutinen & Kangas, 1995; Lindhagen & Hörnsten, 2000; de Vries & Goossen, 2002; Ihalainen, Salo & Pukkala, 2003). In addition, the aesthetic values of forests are of importance for humans’ well being. Environmental issues such as securing high biodiversity (e.g., Gustafson, 1998; Guisan & Zimmermann, 2000; Angelstam & Andersson, 2001; Ricotta & Avena, 2003; Larson *et al.*, 2004) and using the forest for storage of carbon (e.g., Dean, Roxburgh & Mackey, 2004; Backéus, Wikström & Lämås, 2005; Petersson & Ståhl, 2006) are other examples of environmental services that affect the objectives of forestry. Furthermore, in the Nordic countries another important use is for reindeer herding (e.g., Proceviat, Mallory & Rettie, 2003; Sandström *et al.*, 2003).

Decisions in forestry are made at many levels in society, from landowners managing their forest properties to governmental policy-makers providing the legal framework for forest management (Davis *et al.*, 2000; Cubbage & Newman, 2006). The international society makes agreements between governments by using international conventions. Governmental bodies and politicians take decisions to develop forest policies at the national and sub-national level. These decisions are often influenced by these international conventions. Legislations, subsidies, and information campaigns are tools that typically can be used to implement policies. Also, non-governmental organisations can influence decisions through such practises as lobbying or certification schemes (Cashore, Auld & Newsom, 2003).

As in many Nordic countries, a large proportion of the productive forest land in Sweden is owned by private individuals whom are often called non-industrial private forest owners (NIPF). In Sweden the share of land owned by NIPF is more than 50% (Anon., 2006). Furthermore, a limited number of large forest companies possess 25% of the forest land. Less than 20% of the forest land is owned by the public. In Sweden the final decisions on forest management in practise rely on the 350 000 individual forest owners and the 240 000 registered forest companies that own forest (Anon., 2006). However, both the public and the forest industry directly and indirectly affect the forest owners' decision making process.

Concrete decisions about forest management at levels such as estate and stand-level are made by landowners. These decisions are made in agreement with the objectives of the forest landowner and within the framework of national forest policy. Typically, international conventions consider biodiversity and environmental issues, for example the Convention on Biological Diversity (CBD) and the United Nations Framework Convention on Climate Change (UNFCCC) (Holmgren, 2002). A decision at the national level could typically involve policies concerning reasonable levels of conservation areas or legislation on regeneration measures after cuttings. Typically, planners at large forest enterprises determine sustainable cutting levels to guarantee the supply of timber for their pulp- and sawmill or make decisions on which stands to fertilize within the next years. Private forest landowners with small properties typically consider when to do the next cutting, what tree species to plant, and what other silvicultural measures to apply.

Decisions made by forest owners are based on anything from pure intuition to complex scenario analysis. In forestry a distinction between formal and incremental planning can be made (Saaty, 1985). Whereas formal planning typically is based on mathematical models predicting future scenarios based on specific assumptions, incremental planning is based on the experience and intuition of the decision-maker. In reality a combination of the two methods is preferred and often applied.

Forestry planning processes are typically applied hierarchically, divided into different levels based on the time perspective. In forestry three planning levels are often distinguished: strategic, tactical, and operational planning (Weintraub & Cholaky, 1991; Davis & Martell, 1993; Martell, Gunn & Weintraub, 1998; Tittler, Messier & Burton, 2001). The strategic long-term planning consists of functions for goal formulations and the aim is typically to find silvicultural programs or to determine sustainable harvest levels (cf., Martell, Gunn & Weintraub, 1998). The time perspective here is almost unlimited but in practice it is around one hundred years under boreal conditions. This planning level is important for the policy-makers and the decision-makers in large scale forestry. Based on the strategic planning results, tactical planning involves decisions at an intermediate time scale to determine an optimal configuration for the harvesting tracts. The tactical planning considers issues such as timber assortments, fertilisation and road systems with a time perspective of one to ten years (Martell, Gunn & Weintraub, 1998). The operational planning is the implementation level used for short term forest operations. Decisions concerning the usage of available resources of harvest capacity, the final selection of harvesting units, and timber logistics by linking short term industrial demands with forest management activities are typically made at this level (Martell, Gunn & Weintraub, 1998).

The implementation of different decisions taken by the forest landowner and the policy-maker at the national level differs. The forest planning process at the forest landowners' level often aims to identify treatments at the level of single stands, while decisions made by national-level policy-makers are general, and often affect all forest land.

Decision support systems for forestry scenario analysis

Strategic planning in forestry considers the effects of today's decisions over decades, which is rather unique in business management. One reason for these long time perspectives is the long rotation periods in forestry. Another reason is the possibility of making reliable long term prognoses of forest development (cf., Söderberg, 1986; Peng, 2000; Lämås & Eriksson, 2003; Kangas & Kangas, 2004). The effect of many decisions can not be evaluated within short time horizons. Whether or not the forest owner planted the right tree species will in many cases remain unknown until final felling, which in boreal forests can be one hundred years after the decision was made. Similarly, if society finds the right strategy to protect a threatened species will not be known after several decades, and if the answer is negative, it may be too late to change the strategy.

In formal forest planning, forestry scenario analysis can be applied to evaluate alternative management strategies. Forestry scenario analysis decision support systems are used to simulate the effect of different alternative decisions. The simulations are typically used by policy-makers evaluating the effects of different scenarios (e.g., Gustafsson, 2000). In many applications mathematical programming is used to find optimal strategies for forest management (von Gadow & Puumalainen, 2000; Hoen, Eid & Økseter, 2001). Many of today's systems can be used both in optimisation and simulation mode.

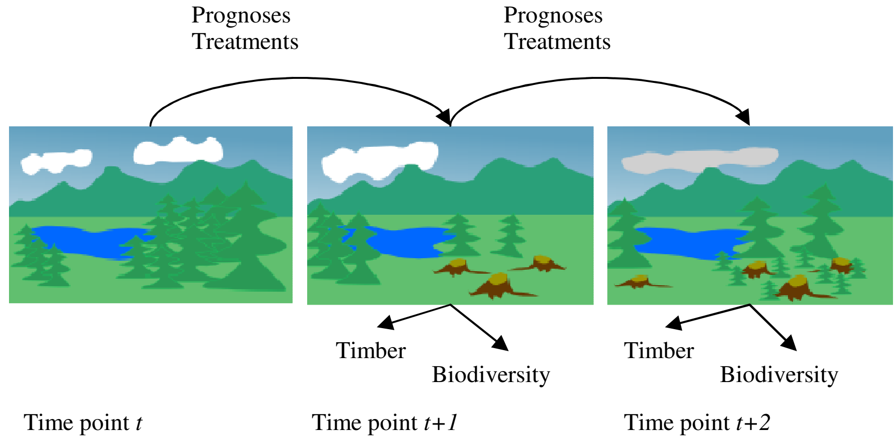

Fig. 1. An outline of the prediction of forest ecosystem development and the derivation of resource indicators. Each resource (e.g. biodiversity) is defined in terms of one or more resource indicators (e.g. area of suitable habitat or volume of coarse woody debris).

A forest simulator constitutes the basis in a decision support system (Fig. 1). Given a description of the current state of the forest, models are used to forecast the ecosystem development (Eid & Hobbelstad, 2000; Lämås & Eriksson, 2003; Gobakken, Lexerød & Eid, 2004). The models typically consider growth, mortality, and the effect of different treatments. Climate change and soil nutrient status may also be considered. Examples of models used to forecast the development of a forest ecosystem are presented in Table 1. Based on these prognoses the outcomes of different resources are simulated and can be optimised. One or more indicators can be used to evaluate the outcome of a resource. Volume of harvested timber is an example of an indicator that can be used to quantify the outcome of timber harvest. Similarly, the amount of coarse woody debris or one or more habitat suitability indices can be used as indicators of biodiversity.

Table 1. Examples of models used for forecasting the forest ecosystem

Models	Reference
Tree growth	Agestam (1985), Andreassen & Tomter (2003), Huang & Titus (1999), Lexerød (2005), Monserud & Sterba (1996), Nyström & Kexi (1997), Porté & Bartelink (2002), Söderberg (1986)
Mortality	Achim *et al.* (2005), Blennow & Sallnäs (2004), Eid & Tuhus (2001), Fridman & Ståhl (2001), Pukkala *et al.* (2005), Talkkari *et al.* (2000), Thor, Ståhl & Stenlid (2005), Valinger & Fridman (1997)
Treatments	Karppinen (1998a), Karppinen (1998b), Karppinen (2005), Pesonen (1995), Pettersson & Högbom (2004), Pukkala, Ketonen & Pykäläinen (2003), Raulier, Pothier & Bernier (2003)
Climate change	Andalo, Beaulieu & Bousquet (2005), Zheng *et al.* (2002)
Soil nutrient status	Rolff & Ågren (1999)

Among Swedish forestry companies the Forest Management Planning Package (FMPP) (Jonsson, Jacobsson & Kallur, 1993), developed in the 1970s and 1980s, is widely used. For policy making, a simulation system called HUGIN (Lundström & Söderberg, 1996) has been used to evaluate different forest management strategies at the national and sub-national levels (Gustafsson & Hägg, 2004). These systems are today being replaced with a new Swedish system called Heureka (Lämås & Eriksson, 2003). The advantage of Heureka is that it includes the evaluation of multiple forest resources and that the same system can be used both by forest managers and by policy-makers. Four applications are included: national and sub-national analyses, long-term forest planning, operational planning, and planning for individual forest land owners.

Internationally, there are large numbers of decision support systems. In Finland, MELA (Siitonen, 1995) has been the main system since the 1970s. Initially it was developed for national forest scenario analyses but today it is also widely used for forest management optimisation. Another Finnish system is Monsu (Pukkala, 1999; Pukkala, 2004), which is a simulating and optimising system which includes spatial analysis of biodiversity, scenic beauty, and recreation scores in combination with timber production. SIMO (http://www.mm.helsinki.fi/MMVAR/SIMO/; 16 Aug. 2007) is an ongoing project in Finland which aims to develop modules that can be used in future software for simulation and optimisation of forest development. Other European systems are GAYA-SGIS (Næsset, Gobakken & Hoen, 1997) in Norway, and EFISCEN (Pussinen *et al.*, 2001) at the European level. In North America systems such as FORPLAN (Johnson, Stuart & Crim, 1986), LANDIS II (Scheller *et al.*, 2007), and CLAMS (Johnson, Duncan & Spies, 2007) are being used. Forest sector models are another example of decision support systems used in the forest industry involving production and marketing, and the use of capital, labour and raw material (e.g., Andersson, Kallio & Seppälä, 1986; Lönnstedt, 1986; Kallio, Moiseyev & Solberg, 2004). Such models have a different problem structure and different approaches to data acquisition; this type of model is not considered further in this thesis.

In optimising systems there has been a number of different approaches to solve forest management problems (Dykstra, 1984; von Gadow & Puumalainen, 2000; Pukkala & Kurttila, 2005). Optimising methods include an objective function that is minimised or maximised depending on the objective. A typical objective includes maximising the net present value or minimising the transportation cost of harvest machinery. Linear programming is one of the first methods used in forestry applications (Dykstra, 1984; von Gadow & Puumalainen, 2000). Other optimising solution methods such as integer programming and mixed integer programming have also been used in forestry (Dykstra, 1984; Jones, Meneghin & Kirby, 1991; Hof & Joyce, 1993). Dynamic programming has been used to optimise a sequence of interrelated decisions (e.g., Dykstra, 1984; Ståhl, 1994; Lohmander, 2000). The size of the problem is often a limiting factor for these methods, and also their inability to solve problems with spatial details (e.g., Murray & Church, 1995; Bettinger & Chung, 2004). As a consequence, heuristic solution methods are often used, such as simulated annealing (e.g., Lockwood & Moore, 1993; Öhman & Eriksson, 1998), threshold accepting (cf., Bettinger *et al.*, 2002), genetic algorithm (e.g., Lu & Eriksson, 2000; Tomppo & Halme, 2004), and tabu search (e.g., Bettinger, Sessions & Boston, 1997; Wikström & Eriksson, 2000). The heuristic solution methods do not necessarily find the globally optimal solution, but they do generally find a relatively good solution within a reasonable time.

Requirements for data in forestry scenario analysis

The requirements for data in forestry scenario analysis depend on what kind of decision will be made based on the analysis. Data requirements need to be determined by what resources and indicators are to be included in the analysis and furthermore by what details of the scenarios are to be considered. Thus, the complexity of the decision support system will affect the data requirements. Summarising data requirements for the included models will result in the overall requirements. In a typical forestry scenario analysis these models have a wide range of requirements. Some models require data on characteristics of single trees while others only need average tree data at the stand level. At the same time, many models also work within geographical windows, requiring information on adjacent units. Independent of the scale at which data are to be assessed there are cases when spatially comprehensive data are required. In some models the spatial arrangement of forest trees is of importance and in other models information on the arrangement of forest stands is necessary. Examples of models that typically can be included and their requirements on data are given in the next sections.

Data requirements of the forest simulator

In the forest simulator empirical stand models or individual tree models are used to forecast the development of the forest state (Peng, 2000). The basic stand models utilise growth and yield equations for the forest stand, while more advanced stand models also consider the distribution of tree size (Peng, 2000; Kangas & Kangas, 2004). These models are called size class models, or often diameter distribution models (e.g., Bailey & Dell, 1973; Kangas & Maltamo, 2000), and have better capability to estimate the outcome of different timber assortments than the more basic stand models. The stand models require data on the forest stand level, such as basal area number of trees. The size class models also require data at the level of size classes, such as the number of trees and basal area in specific size classes. Single tree models forecast the growth of single trees, and require data such as diameter, height, and species. Empirical validation studies have been made to assess the accuracy of different growth and yield models. Kangas & Kangas (2004) assert that the stand level models have generally performed better in these tests than the tree level models, due to cumulating errors on the tree level. However, the use of single tree growth and yield models has several advantages compared to more aggregated models used on the stand level. The experience from the Swedish systems such as HUGIN and FMPP are that single tree data enable reliable projections of growth (Söderberg, 1986; Lämås & Eriksson, 2003). Detailed information may be provided in the prognosis and increased flexibility of evaluation of treatment alternatives is possible (Peng, 2000; Wikström & Eriksson, 2000; Gobakken, Lexerød & Eid, 2004; Kangas & Kangas, 2004). Some growth models also require spatial data on the level of single trees, such as tree position or distance between trees. Such distance dependent growth models are applied to describe the competition of growth between single trees (e.g., Biging & Dobbertin, 1992).

The data requirements between decision support systems differ. Some systems, such as AVVIRK-2000 (Eid & Hobbelstad, 2000), are based on forest stand data. Each forest stand is represented by an average tree, and the forecast is based on basal area mean diameter, mean height weighted by basal area and the number of stems. Other decision support systems, such as FMPP (Jonsson, Jacobsson & Kallur, 1993), use data from single trees. Prognoses for a sample of stands are based on a list of single tree data, where each sampled tree diameter is registered. Age, height, and timber quality are also measured for a sample of trees, and are estimated for the residual trees. There are also decision support systems that require spatially comprehensive data, for example, Heureka (Lämås & Eriksson, 2003). For some applications a list of trees are required for all stands in the landscape.

Details and tools in the decision support system that affects data requirements

There are a number of details that can be considered in a forestry scenario analysis which will affect the data requirements. If the analyses aim to evaluate different forest management strategies in a forest stand considering the risks of wind damages, data for adjacent stands will be required (e.g., Blennow & Sallnäs, 2004; Zeng, Pukkala & Peltola, 2007). The exposure of wind depends both on the stand structure and the surrounding terrain. Another example of improvement that affects the requirements of forest data is when the behaviour of non-industrial private forest owners is considered in a forestry scenario analysis. Information about the forest estate in combination with preferences of the forest owner enables analyses that consider the behaviour of the forest owners under different conditions (Pesonen, 1995; Karppinen, 1998b; Karppinen, 1998a; Lönnstedt, 1998). For example, the willingness among NIPF to cut when implementing different management policies or fluctuation of timber prices can be included in such analyses. This could be valuable for the policy-makers when evaluating forest policies at the national level or for a forest industry analysing potential harvesting levels in a timber catchment area. A further advantage of having data for every stand is a better connection between strategic and tactical planning (Lämås & Eriksson, 2003; Andersson & Eriksson, 2007). While strategic planning aims to establish long-term harvesting levels, tactical planning seeks the right configuration of forest stands available for cutting. Data from all stands are required to identify which areas are available for different treatments and to optimise the spatial distribution of the cutting areas. Spatially comprehensive data enhance the possibilities for dynamic treatment of units in forest planning (Holmgren & Thuresson, 1997; Lind, 2000; Heinonen, Kurttila & Pukkala, 2007).

Data requirements of resource indicators

The requirement of details in the forest simulator depends to some extent on which resources are considered in the analysis. Timber is one resource which is strongly dependent on the forecast of the tree layer. Here, net income in different planning periods may be used as an indicator for timber. This indicator would be correlated with wood quality properties and outcome of different timber assortments. These models often require data on the level of single trees (Wilhelmsson *et al.*, 2002; Moberg, 2006; Wilhelmsson, 2006). Also spatially comprehensive data may improve the utilisation of the timber resources, for example, when harvesting activities are being clustered (Öhman & Lämås, 2003).

Biodiversity is another resource that depends on the landscape patterns and functions. Landscape metrics may be used to evaluate the ecological value of the landscape; these metrics both determine the composition of a landscape and its spatial configuration (Riitters *et al.*, 1995; Gustafson, 1998). The composition of the landscape depends on its mixture of different patch types, independent of spatial location, while spatial configuration metrics characterize the arrangement of these characteristics. As an example, the composition provides the area of deciduous stands in a landscape, while the spatial configuration metrics are used to determine the arrangement of these deciduous stands, such as the size, shape and connectivity between the deciduous stands. These metrics can be used as indicators in the forestry scenario analysis, but also to evaluate, for example, habitat suitability for a certain species (e.g., Hirzel, Helfer & Metral, 2001; Ricotta & Avena, 2003; Edenius & Mikusinski, 2006; Mikusinski & Edenius, 2006). These landscape metrics often require spatially comprehensive data. However, studies do exist where landscape metrics have been derived from sampling data (e.g., Kleinn, 2000). In forest management, a further advantage of using spatially comprehensive data is the possibility of creating continuous areas of old growth forests for biodiversity purposes (e.g., Öhman, 2000). All indicators of biodiversity do not require spatially comprehensive data, and some may be simulated with forest stand or single tree data (e.g., Kolström, 1998; Kruys *et al.*, 1999; Lähde *et al.*, 1999; Kruys, Jonsson & Ståhl, 2002; Bollmann, Weibel & Graf, 2005). As an example of this use of stand data, Kolström (1998) used diameter distribution of dead and living trees to describe stand structures. Kruys, Jonsson & Ståhl (2002) introduced a method for forecasting the decay-class distribution of dead trees over time.

Recreational values are another example of a resource that typically would require spatially comprehensive data at the landscape level (Pukkala, Nuutinen & Kangas, 1995; Lindhagen & Hörnsten, 2000; de Vries & Goossen, 2002). Spatial configuration metrics are also of importance in some of these models. As an example, Pukkala, Nuutinen & Kangas (1995) integrate recreational values into forest planning by determining the variety and recreational values of each forest stand. The variety is described by the total length of boundaries between different forest stands.

Data acquisition for forestry planning and analysis

Data acquisition for forestry scenario analysis can be conducted in many different ways, either using field inventories where a surveyor visits the forest to conduct measurements, or with remote sensing where data primarily are gathered from aerial surveys. In practical applications the methods often are combined. Independent of whether the inventory is conducted in field or from air, methods may be either subjective or objective (Ståhl, 1992). Typical for subjective methods is that the surveyor directly estimates the variables or makes supporting measurements in representative areas. The accuracy of such methods is strongly dependent on the personnel's experience and skills (Kangas, Heikkinen & Maltamo, 2004). Errors will contain both a systematic and a random component, and cannot be estimated unless check assessments are conducted. Objective methods typically are conducted based on statistical sampling methods or by total tallies. Here, assessments are performed in areas selected by random sampling and by using repeatable methods for measurements. Advantages of objective methods are independency of surveyors, that estimates normally are unbiased, and that the precision can be determined based on data acquired (Ståhl, 1992).

Field inventories

In subjective field inventories, variables are acquired directly by the surveyor with ocular methods or based on subjective measurements (Ståhl, 1992). Trees that the surveyor finds typical for a stand are measured with instruments such as a relascope (Bitterlich, 1984) or a calliper. These methods are quick and can cover relatively large areas at a low cost. In objective field inventory statistical sampling theory is applied. Often field-plots are used to sample trees within a stand or over a larger area, independent of stand boundaries; trees can also be selected using a relascope. A time efficient method, but with a slightly biased estimator, is the point-to-tree sampling method, where a fixed number of trees is measured at each sampling point. Model based assumptions about the underlying process or empirical approximation can be used to improve the estimator (Kleinn & Vilčko, 2006). For rare objects other sampling methods may be used, such as line intersect sampling for assessing downed coarse woody debris (e.g., Ringvall & Ståhl, 1999) or line transects sampling to assess wildlife populations (e.g., Ringvall, Patil & Taillie, 2000). Advantages of the field inventory methods are that a large number of variables can be measured. Often field-inventories are planned based on some prior information, such as previous inventories, maps, or remote sensing data.

Some examples of typical field-based inventories in Sweden are:

- At the forest estate level a stand register is often available for the landowner. Data are assessed for every stand, most often with subjective methods in field inventories aided by aerial photo interpretation (Ståhl, 1992). The method used is quick and effective for tactical and operational forest planning. The data will contain both systematic and random errors, and are not suitable for long-term planning.
- For long-term planning at the level of forest companies, a second-phase sampling is used to assess data for the forest scenario analysis with FMPP

(Jonsson, Jacobsson & Kallur, 1993). Stands are stratified based on the stand register. In each stratum a sample of forest stands are selected proportional to stand area. For each sampled stand an objective inventory is performed with approximately 10 field-plots and data at the level of single trees are assessed.

- For national level planning and reporting, forest field-plot data from the national forest inventory (NFI) (Ranneby et al., 1987) have been used. In the NFI more than 10 000 field plots are inventoried every year. The field-plots are sampled using a systematic grid of tracts and are independent from stand borders and land use. Normally half of the field-plots are on forest land. For the national forestry scenario analysis, field-plots from five years of inventory are used (Gustafsson & Hägg, 2004).

In recent years, technical developments have resulted in new devices for ground-based surveys. Devices such as laser range finders, GPS, electromagnetic compasses, and electronic clinometers have improved productivity and are used in several surveys. Furthermore, with terrestrial laser scanning, ground-based measurements of forest variables, such as tree diameter, density, and upper stem diameters can accurately be assessed (Thies *et al.*, 2004; Watt & Donoghue, 2005). However, the latter method is not yet a viable alternative for practical inventories.

Remote sensing

The main advantage of remote sensing is the ability to obtain information over large areas at low cost per area unit. Thus, remote sensing is often applied to attain spatially comprehensive data. Both airborne and space-borne sensors are used to acquire data (Lillesand & Kiefer, 2000). Space-borne satellites are commonly used to provide forestry with remote sensing data. Sensors can either use active or passive energy techniques. The passive sensors (e.g. optical sensors) measure the reflection of naturally available energy while the active sensors (e.g. radar and laser sensors) supply their own source of energy (Lillesand & Kiefer, 2000).

Both airborne and satellite-borne optical sensors have been used in forestry for decades. Typically, these sensors depend on weather and light conditions, due to the measurement of natural available reflections. Film-based aerial photograph systems have been widely used in forestry for more than half a century (cf., Congalton & Green, 1999; Hauska, 1999). Delineation of forest stands and identification of tree species are some examples of applications commonly used in boreal forests. Measurements of tree heights and crown closure have been used to estimate stand volume. These interpretations are made visually, in some cases with equal accuracy as relascope measurements in field (Eid & Næsset, 1998). During the last years digital sensors have been developed and improved the resolution and usability of aerial photographs. In digital images more automatic image-processing can be performed (Pitkanen, 2001; Olofsson *et al.*, 2006). During the last decades optical images have also been available from space-borne sensors (Tomppo *et al.*, 2002). The resolution varies due to sensor, and medium resolution sensors such as SPOT and LANDSAT provide images with a resolution around 10-30 meters (cf., Magnusson & Fransson, 2005). Fine resolution sensors such as Ikonos and

Quickbird are also available for forest applications (e.g., Kayitakire, Hamel & Defourny, 2006). These sensors have a resolution of one meter or less and are detailed enough to detect single trees; the quality of these images are comparable with aerial photographs. Low resolutions images are also available and are most often used in global assessments.

The active sensors supply their own source of energy and many acquire data regardless of cloud cover and light conditions. Laser scanning and radar are two techniques based on active sensors that are used in forestry (Nilsson, 1996; Næsset, 1997; Fransson & Israelsson, 1999; Lefsky *et al.*, 1999; Fransson, Walter & Ulander, 2000; Holmgren, 2004). Both airborne and space-borne sensors are available with both systems, but primarily airborne applications have been used so far in forestry. The radar systems emit radio waves from a fixed antenna mounted below the aircraft (Hyyppä *et al.*, 1997; Fransson & Israelsson, 1999; Fransson, Walter & Ulander, 2000; Lillesand & Kiefer, 2000). The most commonly used technique is the synthetic aperture radar systems. These systems are equipped with a physically short antenna but use the velocity of the aircraft to synthesize the effect of a long antenna. Radar operates in the microwave portion of the electromagnetic spectrum and the wavelengths used in forestry applications are from centimetres up to meters. In Sweden airborne radar was used to estimate the volume of storm damaged timber after the storms of 1999 (Fransson et al., 2002) and 2005. Another active sensor is the laser scanning system that has been introduced to Scandinavian forestry during recent years (Nilsson, 1996; Næsset *et al.*, 2004). Laser scanning systems use either pulses or continuous waves of near infrared or green light to measure the distance to a target object on the ground (Wehr & Lohr, 1999). The continuous wave sensors constantly return signals reflected from the ground, while the discrete pulse systems either receive a first and a last return or multiple returns. The pulses are either projected directly on the ground or distributed over the ground during the flight. The most commonly used technique to distribute the pulses is with a scanner, which distributes the pulses over the ground perpendicular to the flight direction. The laser system measures coordinates of targets on the ground and the vegetation in three-dimensions. A digital terrain model is produced and tree height and tree cover is measured, so that basal area and volume can be estimated (Næsset et al., 2004). Research has been done to identify single trees and their properties such as position, height, crown width, stem diameter, and species (Hyyppä *et al.*, 2001; Holmgren & Persson, 2004). Airborne laser scanning provides planning data for individual forest owners in Norway (Næsset, 2004). In North America a profiling laser was used to assess multi-resource forest data at a sub-national level (Nelson et al., 2003).

With remote sensing data, different methods for the estimation of forest variables can be applied and most often field data are required. Methods such as visual interpretation, digital photogrammetry, classification and single tree detection are some commonly used methods. Regression and non-parametric methods have also been used to predict forest variables. In applications based on regression (e.g., Hyyppä *et al.*, 1997; Næsset, 2004; Magnusson & Fransson, 2005), the relationship between the variable of interest and a number of independent variables in the remote sensing data is modelled. The variables of interest are often estimated independently and are obtained as interpolated values, which may result in an unnatural relationship between the variables (e.g., Holmström, 2001). Another approach for predicting forest variables is by using non-parametric methods. There are a large number of different methods, such as k Nearest Neighbour (*k*NN) (Tomppo, 1990; Tokola *et al.*, 1996; Nilsson, 1997), Most Similar Neighbour (MSN) (Moeur & Stage, 1995), and Gradient Nearest Neighbour (Ohmann & Gregory, 2002). The independent variables in the remote sensing data are used to simultaneously predict a number of forest variables. Units with a complete list of variables of interest supply reference data. The linkage is provided by carrier data (cf., Holmström, Nilsson & Ståhl, 2001) that must be available from all units. Typically, carrier data comprise only a few variables, but are variables that can be inexpensively assessed for all units in the target population, as well as in the reference data set. The units may be entire stands, but more often they are plots or pixels (e.g., Holmström, Nilsson & Ståhl, 2002; LeMay & Temesgen, 2005; Wallerman & Holmgren, 2007). Similarities in carrier data between reference and target units are used to determine what reference data set to be imputed to a certain target unit. Similarity generally is expressed in terms of some suitable distance metric, for example, in Euclidean distance (Holmström, Nilsson & Ståhl, 2002).

Data quality

Accuracy and precision are considered as two important properties of an estimator (Tamhane & Dunlop, 2000). Accuracy is used to determine the deviation between estimated and true values. Estimators that produce estimates close to true values are considered accurate (Schreuder, Ernst & Ramirez-Maldonado, 2004). Often root mean square error (RMSE) is used as estimator to determine accuracy. Precision is closely related to accuracy but determines the deviations between individual measurements and their mean value. The precision is often characterised with the standard deviation, and estimated with the standard error (SE). In surveying, random errors express the random variability in the measurements and bias is the systematic non-random error. The sampling errors are typically random errors whereas measurement and judgement errors are typically both (Ståhl, 1992). An accurate estimate is obtained if precision is high and the estimate is unbiased (Schreuder, Ernst & Ramirez-Maldonado, 2004).

The accuracy of the estimates varies due to the use of different techniques and methods for estimation. An overview and evaluations of different remote sensing techniques are given by Magnusson (2006). However, assessing the accuracy of spatially comprehensive data for forestry scenario analysis is more complex. Remote sensing techniques contain spatially auto correlated errors (Congalton, 1988; Foody, 2002) which affect the accuracy in, for example, a forestry scenario analysis. Studies have shown the sensitivity of error patterns in spatially comprehensive data, both in scenarios using landscape metrics (Wickham *et al.*, 1997; Langford *et al.*, 2006) and estimating habitat-suitability indices (Fleming *et al.*, 2004). To assess the errors in classified data, an error matrix can be used (Congalton & Green, 1999). In spatially comprehensive data, data quality also includes the consistency between variables in adjacent units. In this thesis spatial consistency is denoted to spatial data when the natural variability between units is accurately described.

Planning sampling surveys

When planning forest inventories, a number of aspects have to be considered. What parameters should be assessed, when should the inventory be conducted, and how accurate data are required? These questions are difficult to answer and, as a result, planning of inventories is often based on tradition and previous experience. However, the parameters to assess can partly be determined based on what decisions will be made and the requirements of data in the forestry scenario analysis system to be used. When to conduct an inventory is also dependent on what type of decision will be made based on the data and time for next treatment (Ståhl, 1994; Ståhl, Carlsson & Bondesson, 1994). To determine an appropriate target of accuracy of a forest inventory, considering the relationship between cost and precision is one possible approach (Thompson, 2002). The trade-off between inventory cost and precision for different intensities of inventories can be studied. An example is presented in Fig. 2, but to determine a reasonable trade-off is not simple. With fixed economical budgets it is possible to determine the expected accuracy when using different methods and techniques. Often the accuracy of different inventory methods is known and can be compared with other inventory methods. Mehtätalo & Kangas (2005) developed models for the expected error of the total volume and saw timber volume due to sampling errors. For a given inventory budget, optimisation was used to find the inventory strategy that minimised the expected error.

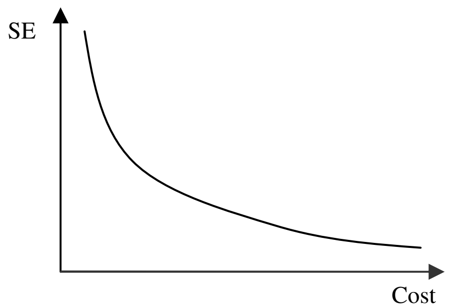

Fig. 2. An example of the trade-off between inventory cost and precision.

Cost-plus-loss analysis

When a forest inventory is planned, the aim of inventory and the ability to make adequate decisions has to be considered as well. The link between decision and inventory data can in many cases be difficult to establish (Duvemo & Lämås, 2006). When the link between data and decision-making is clear, and when the loss due to poor decisions can be assessed in monetary terms, cost-plus-loss analysis (e.g., Hamilton, 1978; Ståhl, Carlsson & Bondesson, 1994; Eid, 2000; Holmström, Kallur & Ståhl, 2003; Eid, Gobakken & Næsset, 2004) can be applied in the planning of forest inventory. In a cost-plus-loss analysis the sum of the inventory cost and the loss due to poor decisions based on the inventory data is minimized. The economical loss of a non-optimal decision is the difference between the income of decision based on perfect data and the income of the decision based on the data available. A general view of cost-plus-loss analysis is presented in Fig. 3. Formal cost-plus-loss analysis has so far generally only been applied to optimising the net present value at the forest stand level. The net present value is the revenue of all future treatments of the forest discounted to present value. Cost-plus-loss analysis can be performed using either an analytical approach (e.g., Ståhl, Carlsson & Bondesson, 1994) or by using a simulation approach (e.g., Eid, 2000). A review of the literature on cost-plus-loss was recently provided by Duvemo and Lämås (2006).

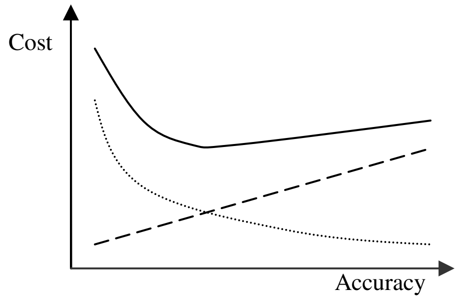

Fig. 3. A general view of cost-plus-loss analysis showing the inventory cost (dashed line) and the non-optimality loss (dotted line). The solid line is the cost-plus-loss which should be minimised.

Analytical cost-plus-loss

The analytical approach of cost-plus-loss analysis could be described as follows (Hamilton, 1978; Ståhl, Carlsson & Bondesson, 1994; Duvemo & Lämås, 2006). The objective of a cost-plus-loss analysis is to minimize the sum of inventory cost and expected loss due to non-optimal decisions. The inventory cost is typically specified as:

$$C = c_0 + c_1 n \quad (1)$$

Here c_0 is the fixed cost of the inventory and c_1 is a variable cost per sampling unit. The loss function can take many forms: linear, quadratic, one-sided, and discontinuous functions may approximate the loss function (Hamilton, 1979). Here, the linear (Eq. 2) or quadratic (Eq. 3) loss function are given as examples (Hamilton, 1978; Duvemo & Lämås, 2006).

$$L = \lambda|\varepsilon| \quad (2)$$

$$L = \lambda\varepsilon^2 \quad (3)$$

Here, λ defines the relationship between error and loss. The error (i.e., the deviation between true value and estimated value) is given by ε. In a general case, knowing that sampling errors typically are normally distributed, the expected value of the absolute deviation in the linear case is $(2/\pi)^{1/2}std(\varepsilon)$ and in the quadratic case $E(\varepsilon^2)=S^2/n$. The expected cost-plus-loss, assuming simple random sampling of units can then be expressed as:

$$E(C+L) = c_o + c_1 n + \lambda\sqrt{\frac{2}{\pi}}\frac{S}{\sqrt{n}} \tag{4}$$

$$E(C+L) = c_0 + c_1 n + \lambda\frac{S^2}{n} \tag{5}$$

where S is the population variance.

By minimizing the expected value, the optimal inventory intensity can be determined in the linear case as Eq. 6 and in the quadratic case as Eq. 7.

$$n_{opt} = \left(\frac{\lambda S}{2c_1}\sqrt{\frac{2}{\pi}}\right)^{2/3} \tag{6}$$

$$n_{opt} = \sqrt{\frac{\lambda S^2}{c_1}} \tag{7}$$

Cost-plus-loss analysis using simulation

In the simulation approach of cost-plus-loss analysis, data to be evaluated are used in forest planning. An example of this is in a forestry scenario analysis which optimises net present value (e.g., Holmström, Kallur & Ståhl, 2003; Eid, Gobakken & Næsset, 2004). Treatment schedules based on the evaluation data are applied to the analysis with perfect data. The scenario analysis is also applied using the perfect data and the deviation of net present value between the two analyses is then determined to be the decision loss.

Planning sampling surveys for national forestry scenario analysis

Means to evaluate the consequences of data acquisition strategies are limited and seldom used. Cost-plus-loss analysis is a rather unique tool for evaluating forest data in decision-making. However, in many relevant situations cost-plus-loss is to some extent limited in the ability to evaluate data acquisition. One such situation is in decision making at the national level, where the connection between decisions and data are not that obvious. For example, when national data are reported to international conventions, the path from data to decision is long and unclear. Another challenge is the multiple purposes of national forest data and it can be extremely difficult to determine a complete loss function. In planning a multi-resource inventory, not all resources can be expressed in monetary terms. In this case, a cost-precision approach would be achievable, but a true cost-plus-loss analysis would be difficult. The requirement of spatially comprehensive data also limits the use of analytical cost-plus-loss analysis; however, a simulation approach would be conceivable.

The development of more complex forestry scenario analysis at national level and a wider spectrum of available data acquisition methods necessitate further progress of means with which to evaluate the consequences of data quality. However, when data quality is linked to decision making, an important note is that there are many other uncertainties beside data errors which affect the outcome of forest planning (Kangas, 1997).

Objectives

The main objectives of this thesis were to evaluate different data acquisition methods and develop tools to enhance data usability within national and sub-national level forestry scenario analysis. Special considerations were given to situations where resource indicators required spatially comprehensive data. Three major issues were considered: i) to determine typical requirements of data in forestry scenario analysis, ii) to evaluate and further develop methods to determine data requirements, and iii) to develop methods that improve data usability in forestry scenario analysis. Means of linking data acquisition strategies with decisions that typically are based on forestry scenario analyses were used in the determination of data requirements in Papers I-IV. In Papers V and VI, methods to improve the usability of spatially comprehensive data were developed. The specific objectives were

Paper I. To provide a framework for evaluating data acquisition strategies for national forestry scenario analysis. A qualitative approach was used to determine which data quality characteristics are of importance and what data acquisition strategy should be applied.

Paper II. To evaluate the quantitative consequences of using spatially comprehensive data based on airborne laser scanning and medium resolution satellite images in a sub-national forestry scenario analysis. The evaluation focuses on the errors in forecasted resource indicators, such as net income, cutting volume and stand volume.

Paper III. To apply cost-plus-loss analysis in a simulated approach for evaluating the quantitative consequences of using spatially comprehensive data based on airborne laser scanning and medium resolution satellite images in decision-making at the forest stand level. The consequences of data quality in forest management planning in terms of decision loss and inventory cost were considered.

Paper IV. To apply cost-plus-loss analysis for determining an appropriate sample size of a national forest inventory for estimating sustainable harvest levels at a national level. An analytical cost-plus-loss analysis approach was used.

Paper V. To develop a method whereby the within-stand spatial consistency was considered in the estimation of spatially comprehensive stand data. A non-parametric method for estimation of forest characteristics and a heuristic optimising approach to improve the quality characteristics of data were used. The method was then evaluated in a simple case study.

Paper VI. To provide a framework for improving composition and spatial consistency in spatially comprehensive data at the landscape level. Core parts of the framework were evaluated in a case study.

Summary of papers

Framework for determining forest data requirements in forestry scenario analysis (Paper I)

In Paper I a framework for evaluating data acquisition strategies for national forestry scenario analysis was provided. Here, a qualitative approach was used to determine which data quality characteristics are of importance and what data acquisition strategy should be applied.

Evaluation of data acquisition strategies

Planning a forest inventory to acquire data for national forestry scenario analysis has become more complex due to the multi-resource objectives in forestry and the development of new inventory techniques, mainly in the field of remote sensing. Routines for evaluating different data acquisition strategies are needed. Connecting the forest inventory with decision-making is, however, difficult. More analytical approaches for evaluating data acquisition strategies can be performed with the cost-precision approach or by cost-plus-loss analysis (Hamilton, 1978; Ståhl, Carlsson & Bondesson, 1994). However, the cost-precision approach only determines the accuracy given certain cost and does not consider the use of data in decision-making. In cost-plus-loss analysis the decisions are also considered. In multi-resource forestry this approach is not directly applicable due to the difficulty of expressing many of the resources in monetary terms.

Multi-resource forestry scenario analysis requires an expansion of the traditional cost-plus-loss analysis. Here, two approaches could be possible.

- One approach would be to define all considered resources into monetary terms along principles outlined in environmental economics (Mattsson & Li, 1993; Boman & Mattsson, 1999; Boman, Bostedt & Persson, 2003). Then, cost-plus-loss analysis is applied. However, it is known from environmental economics that it is very difficult to estimate the exact values of different resources (Boman & Mattsson, 1999). It is likely that it is even more difficult to estimate the effects of non-optimal decisions in monetary terms.
- Another approach would be to evaluate the effect of every single resource using different norms for each resource. In this case there are no straightforward means to compare the overall effect of data for different resources. Thus, making a decision about a forest inventory strategy will be largely subjective, but at least the consequences for including different resources will have been evaluated.

As an alternative to these two quantitative approaches, a more generic approach regarding the choice of an inventory strategy could be applied. This approach is conducted in two steps. First, the type of indicators that can be applied when different data acquisition strategies are used is identified. Secondly, for a given set of indicators, an assessment is be made of the likely consequences of using data with the specific quality that can be expected from a certain inventory strategy. To accomplish this type of analysis, some concepts are needed. First, the indicators typically included in forestry scenario analyses and their data requirements have to be determined. Secondly, a conceptual way of characterizing forest inventory data quality is also proposed for use in the evaluation.

Resource indicators and their data requirements

As stated in the introduction the data requirements for forestry scenario analysis systems are highly dependent on the models used in making forecasts of forest ecosystem development and resource indicators. The values of the indicators can be derived based on the management scenario assumed and the forecasted ecosystem states (cf. Fig. 1). To derive the values of the indicators, there is a need for models that link data regarding the forecast ecosystem state to values for the specific indicator. A restriction when developing this type of model, therefore, is that it can only be based on data that can be satisfactorily forecasted. When determining what data are needed as input to a decision support system, one needs to consider what data are required to forecast the ecosystem state and the indicators. Summarising the data requirements of the models presented in the introduction, there is a wide range of demands. Some of the indicators require crude data at landscape or stand level, while other indicators demand more detailed descriptions at the single tree level.

Characterising forest inventory data quality

The next step is a generic quality assessment to identify the likely consequences of different types of errors. The trade-off between inventory cost and data quality is an important issue since an exact description of the current state is never possible to obtain. However, data quality is a complex property that cannot be quantified appropriately with any single measure. As a basis for the evaluation of different data acquisition strategies, five different features of data quality are proposed. To present these features, a distinction is made between (1) the descriptions made within a single description unit, and (2) the relationships between the description units in the forest landscape. A description unit is the smallest area described in a data set, for example, a pixel, a field plot or a forest stand (depending on the analysis set-up). Following an inventory, the landscape is described in terms of a set of description units allocated over the landscape with or without geographically determined locations. The features distinguished within a single description unit are

- degree of detail, in terms of how many variables are assessed,
- accuracy of the variable estimates, and
- consistency between the variables.

The features distinguished regarding the relationships between the description units are

- spatial completeness and
- spatial consistency of errors.

Each of the five features is now described in more detail.

Degree of detail

The degree of detail may differ within a description unit; some data acquisition strategies generate a long list of variables, while others are only able to provide a few (Fig. 4). A high degree of detail may be obtained when a plot is inventoried in the field. In this case, it is simple to add additional variables to the data set. In contrast, a pixel in a satellite image is generally described with only a few digital numbers. Here, forest variables have to be predicted. In general, a description with a high degree of detail provides better opportunities for forest analyses than a description with a low degree of detail. With a detailed description, normally it is possible to account for more resources and indicators in the analyses compared with the case where only a very crude description of a forest is available.

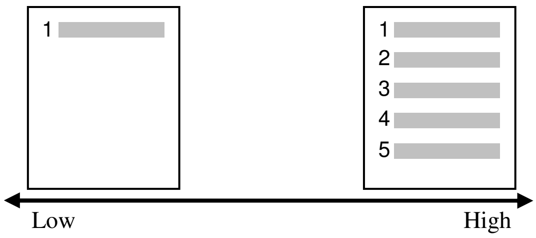

Fig. 4. Degree of detail depends on the number of variables that describes the unit. Many parameters indicate a high degree of detail.

Accuracy of variables

Accuracy is a measure of how well an estimated value corresponds to the true parameter value. Some data acquisition strategies are more accurate than others in describing forest variables (Fig. 5). For example, a field measurement of the basal area of a stand generally results in a more accurate value than if aerial photo-interpretation is used for that purpose (Ståhl, 1992). Descriptions with high accuracy are preferred in forestry analyses.

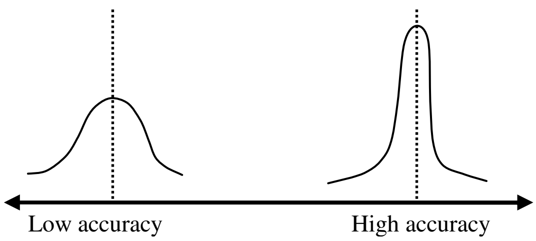

Fig. 5. The accuracy of variables is a measure of the relationship between an estimated value and the true parameter value (dotted line). In the figure, this is illustrated in terms of distribution functions for estimated values around some true parameter value.

Consistency between variables

Another data quality feature to be considered within description units is the consistency of the error structure for the estimated variables (Fig. 6). Consistency in errors means that if one variable, at random, overestimates the true value for a certain plot, variables that are logically connected to the first variable should also be overestimated at that plot for the errors to be consistent if the correlation is positive. The consequences of an inconsistent error structure might be severely erroneous forecasts since the models are usually derived using consistent data. For example, the predicted or measured tree volume in a description unit needs to be consistent with stand age otherwise growth predictions for the unit may be severely biased. Variables that are measured or predicted independently from each other run a larger risk of obtaining low consistency. The correlation structures of error distributions have been studied by Kangas & Kangas (1999).

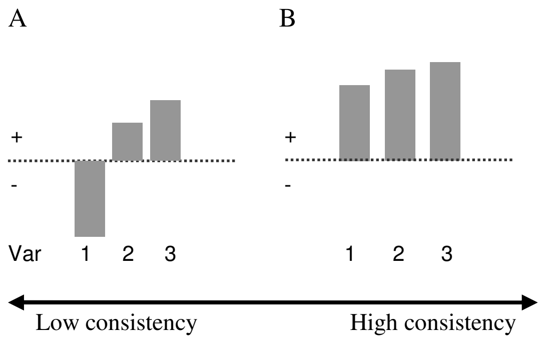

Fig. 6. Examples of high and low consistency of the error structure for three positively correlated variables assessed within two different description units (A and B). The dotted line symbolises the true value.

Spatial completeness

A landscape can be described with different numbers of description units, and thus the assessed proportion of the landscape will vary (Fig. 7). For example, spatially comprehensive data for a landscape can be provided by satellite images, while owing to cost and practical issues, field measurements will seldom cover more than a small fraction of the landscape. To meet modern modelling requirements a spatially complete description is sometimes preferred (e.g., Lämås & Eriksson, 2003). With a high proportion of the landscape assessed, indicator models that require this type of data can be applied.

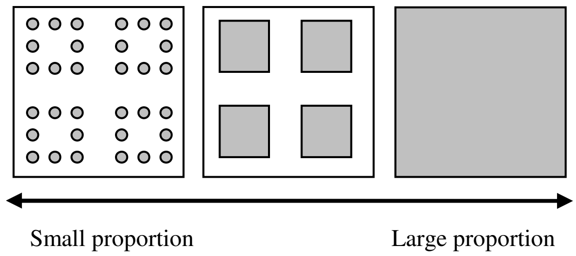

Fig. 7. Spatial completeness, expressed as the ratio of the sampled area to the total area. In the illustration, grey represents the area that has been sampled and white is non-sampled areas.

Spatial consistency of errors

The last quality characteristic of forest data is the spatial consistency of errors. When compiling data at the landscape level, different description units are linked to each other in some type of grid system to set up a partial or complete cover description. In doing this, one needs to ensure that the large-scale patterns in the landscape are realistic, for example that forest stands are represented by realistic features as well as having realistic within-stand variability. This consistency might not be of importance for predicting the total or average value of some variable in a region or a stand, but when models that require data from a larger neighbourhood are applied (e.g. habitat models) the output will depend on the realistic way in which patterns are described in the landscape. Management decisions depend on both within- and between-stand variability. High consistency of errors is preferred when accurate landscape metrics are important, for example, in a national level scenario analysis. Unless the spatial localisation of different features in the dataset is perfect, it would be preferable to have spatially correlated errors rather than completely random distributed errors (Fig. 8). Random distributed errors, so called white noise, are however advantageous in some applications, such as in a tactical and operational level planning when the exact position of different resources is important.

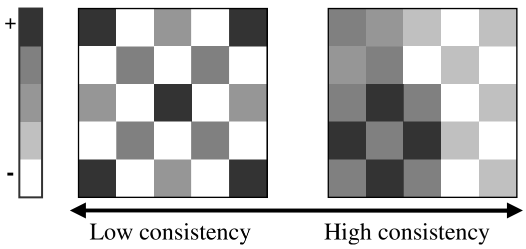

Fig. 8. High spatial consistency of errors will form a realistic landscape pattern. If the consistency is low, it will not be possible to delineate forest stands. Black colour denotes description units with overestimated variables, and white colour with underestimated variables. A map of gray values would also indicate a landscape of high consistency due to almost perfect data.

Application of the evaluation framework: an example

In Paper I a case example was conducted to illustrate how to use a generic approach when making decisions on data acquisition strategies for national forestry scenario analysis based on the proposed evaluation framework. Three commonly used data acquisition strategies were selected: (1) objective sample plot inventory similar to a NFI in a Nordic country (‘‘field plots’’); (2) stand level inventory using a combination of aerial photographs and ocular assessment in the field (‘‘stand-level’’); and (3) use of field sample plots in combination with satellite images and *k*NN imputation of plots (‘‘imputation’’). The cost for each method was in Paper I determined to be € 0.30 ha^{-1}, € 7.00 ha^{-1}, and € 0.30 ha^{-1}, respectively. The evaluation of the three strategies follows the qualitative generic alternative that has been previously described. The assessment is done in two steps: the first step considers what type of indicators can be used based on the available data from the evaluated strategy. The second step considers the likely quality using the same data to estimate the outcome of each resource.

Results and discussion

Quality characteristics for each of the three data acquisition strategies were considered, these were linked to the data requirements on single resource indicators (Table 2). Scores were given in the table as a simple assessment of both how well a resource can be described in terms of indicators (“Content”) and the likely quality of the indicator predictions (“Quality”).

In the Field plot strategy, single tree data are available and the content for resources such as timber, forest fuels, and carbon storage can be predicted at an acceptable level. The diameter distribution of the trees is known, and the value of the timber and different management strategies can be considered (e.g., Kangas & Kangas, 2004). However, the lack of spatial data limits the use of many indicators of resources such as biodiversity and recreation (e.g., Edenius & Mikusinski, 2006). With the Field plot strategy, the quality of predictions would likely be generally good due to the use of objectively sampled data. In the Stand level strategy, only average values are known for each stand. The spatially comprehensive data would increase the content of resources such as biodiversity and recreation, and to some extent also timber and forest fuel. For example, the risk of wind damage and the behaviour of the landowners can be considered in the scenarios (e.g., Blennow & Sallnäs, 2004). However, not knowing the diameter distribution of trees would limit many important indicators such as timber value. The subjective inventory method contains a bias and the quality of predictions would be poor (e.g., Ståhl, 1992). The Imputation strategy generates both spatially comprehensive data and a list of single trees, which would provide a good basis for good content with most of the resources. However, the complexity of the data might affect the quality negatively.

Table 2. *Overview of the evaluation in the case study of the three strategies, and how well a resource can be described in terms of indicators ("Content") and the likely quality of the indicator predictions ("Quality"). One score (●) = poor, two scores (●●) = acceptable, and three scores (●●●) = good.*

Resource	Field plots		Stand level		Imputation	
	Content	Quality	Content	Quality	Content	Quality
Timber	●●	●●●	●●	●	●●●	●●
Forest fuel	●●	●●●	●●	●	●●●	●●
Carbon storage	●●	●●●	●	●	●●	●●
Biodiversity	●	●●●	●●	●	●●	●●
Recreation	●	●●●	●●	●	●●	●●
Reindeer herding	●	●●	●●	●	●●	●
Berries and mushrooms	●	●●	●●	●	●●	●

In a forestry scenario analysis, where all the resources are of interest, it is clear that either the field plot or the imputation strategy would be preferred. The cost of the stand-level strategy is simply too high to make this method appropriate. It is obvious that none of the methods would be optimal in all cases, and thus a conclusion is that it should be relevant to combine methods when setting up suitable data acquisition procedures for national-level analyses. A complete database of all required data and information for a forestry scenario analysis at the national level would grow unnecessarily large. The use of overly large data sets would make the costs of data acquisition high and affect the efficiency of the analysis in a negative way. To simplify the analysis and to reduce the cost of inventory it is probably relevant to include only a sample of the landscape in the analysis. In this limited area, all the necessary spatially comprehensive data can be integrated.

Using the suggested framework is largely subjective and the results exemplified in Table 2 will not give the same result upon repetition, but nonetheless should follow a similar pattern. Thus, the result will depend on the knowledge and experience of the users. However, the framework provides some support when making decisions about data acquisition strategies and will hopefully assist the decision-maker in finding arguments for a proper decision.

Evaluating data acquisition strategies (Paper II-IV)

In Paper I a generic qualitative reasoning for evaluating forest data acquisition provided a framework for support when making decisions. However, these results are likely to be strongly dependent on the knowledge and experience of the users. More quantitative methods can be applied theoretically to learn more about the effects of different data sources in different decision situations. Papers II and III study the effects that data have on decisions with a simulation approach in a decision support system. In Paper III, the data are evaluated by means of cost-plus-loss analysis. The focus of the decision making is from the view of a forest owner optimising treatments in a forest stand. However, at national and sub-national level policy-making, a forestry scenario analysis is dependent on the quantitative outcome of different resources. Consequently, Paper II focuses on the accuracy of outcomes in specific planning periods. In Paper IV, analytical cost-plus-loss analysis is applied to determining an appropriate sampling size for a national level sample-plot inventory. Decisions at the national level are considered in the loss function.

Material

An overview of the material and methods used in Papers II-IV are presented in Table 3. In Papers II and III imputation based on SPOT medium-resolution satellite data and laser scanning data were used in the evaluation. These data were available from a previous study and offered a unique opportunity to test cost efficient sample-plot data in forestry scenario analysis. A comprehensive description of the data are available in Wallerman & Holmgren (2007). In Paper IV, data from the Swedish NFI were used (Ranneby *et al.*, 1987). More details of the data and imputation method are given in the following section.

Table 3. *A general overview of forest data and methods used in Paper II-IV.*

		Paper II	Paper III	Paper IV
Evaluation data	SPOT	●	●	
	Laser	●	●	
	Laser & SPOT		●	
	Field-plots		●	
	Swedish NFI			●
Method	FMPP	●	●	
	Error consequences	●		
	Cost-plus-loss		Simulation	Analytical

Forest data (Papers II and III)

The forest data used in Papers II and III were collected as a sample of stands in a 1 200 ha estate, Remningstorp, 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 assessed by surveying 10 m radius field plots using methods and material developed for the FMPP (Jonsson, Jacobsson & Kallur, 1993). A systematic grid was used to sample approximately 10 field plots in each sampled stand. In addition, 16 stands inventoried using a cluster of 4 by 4 adjacent field plots were available. In total 67 and 64 stands were available for the simulations in Paper II and Paper III, respectively. Satellite image data for the field plot centres were extracted from a geometrically precision corrected SPOT-5 HRG scene, acquired at 10:05 AM on 3 June 2003. 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}.

Evaluation data (Papers II and III)

In total three remote sensing data sets were evaluated as carrier data. As reference data 870 field-plots from the inventoried stands were used. Estimation of forest variables in each forest stand was made using MSN (Moeur & Stage, 1995;

Temesgen *et al.*, 2003). Approximately 30 field plots with data at the level of single trees were assigned to each forest stand. The distance metric was the Euclidean distance weighted by a vector of the squared canonical correlation (Moeur & Stage, 1995; Wallerman & Holmgren, 2007). The independent variables were transformed to define an efficient measure of similarity between forest variables of the sample-plot data, accounting for the different forest information content, scale, and correlation of independent variables. Different independent variables were used in the three data sets. In the SPOT-based data, the XS1, XS2, XS3, and XS4 bands were used. In the Laser-based data, measures used were all 10th percentiles, the 95th percentile, mean height, standard deviation of height, and semi-variogram parameters. These semi-variogram parameters corresponded to nugget, sill, and range (cf., Cressie, 1993). In the combined SPOT- and Laser-based data, the same satellite and laser data were used, with the exception that the laser data did not contain all the 10th percentiles. In Paper II, the field-plot inventoried stands were considered to provide the true description of the forest estate. In Paper III a field inventory of 5 and 10 field-plots was simulated for each stand using bootstrapping (Efron & Tibshirani, 1993). In the data evaluation in Wallerman and Holmgren (2007), the mean volume was estimated with an RMSE of 18% using the laser-based data, and of 33% using the SPOT-based data. In contrast, the mean volume estimated with data assessed according to the FMPP instructions had an RMSE of 12% (cf., Ståhl, 1992). These data are normally used as input data in the FMPP.

Forest data (Paper IV)

In Paper IV data from the Swedish NFI for the period 2003-2006 were used to derive the empirical variances needed for the cost-plus-loss analysis. In Sweden NFI data are acquired in permanent and temporary tracts (or, according to the nomenclature in some countries, plots). This tracts consist of circular plots (or sub-plots) with 10 or 7 meters radii (Ranneby *et al.*, 1987). The country is divided into a total of six regions. Every year about 1 400 tracts are inventoried which corresponds to more than 10 000 plots. Half of the plots are located on forest land and two thirds of these are permanent plots. In Table 4, a basic summary of the inventory design and forest characteristics is given.

Table 4. *Basics about the Swedish NFI sample size in different regions and summary statistics about the total land and fresh water area of Sweden, forest area and volume according to the forest definition of FAO.*

	Tracts (n yr^{-1})		Plots (n yr^{-1})		Total area (1000 ha)	Forest area (1000 ha)	Volume (m^3 ha^{-1})
	Perm.	Temp.	Perm.	Temp.			
Region 1	109	58	859	678	11 813	5 232	74.2
Region 21	96	53	749	612	6 355	4 844	83.3
Region 22	94	49	736	562	6 555	4 538	114.7
Region 3	125	67	982	787	6 865	5 166	125.0
Region 4	269	226	2 079	1 301	10 717	6 395	160.1
Region 5	180	86	709	493	2 735	1 367	178.0
Country	873	539	6 113	4 433	45 040	27 542	117.1

Methods

Different approaches were used in the evaluation of data (Table 3). In Papers II and III the consequences of errors in forest data were evaluated using a decision support system. In Paper II, the results are analysed in detail, which was done by considering the effect of the outcome of several indicators in specific planning periods. In Paper III, the decisions concerning treatments were studied with a simulation approach of cost-plus-loss analysis. Here, the decisions are applied to the true state of the forest, while in Paper II the probable consequences of using poor data are considered. In Paper IV an analytical approach of cost-plus-loss was used to determine an appropriate accuracy level for a NFI.

In Papers II and III, the evaluation data were used as input for strategic level planning in the FMPP (Jonsson, Jacobsson & Kallur, 1993). Detailed growth projections and economic yield calculations were performed. Data enters the planning system at the level of single trees on sample plots. The trees on each plot in a stand are projected five years at a time, and different treatment options are applied at the stand level. The FMPP simulates a large number of different treatment schedules for each individual stand and calculates the net present value of each treatment schedule. The treatment schedule resulting in the highest net present value was chosen as the optimal schedule.

Consequences of errors in data (Paper II)

In Paper II, the outcome in terms of the indicators *harvesting volume*, *net income*, and *standing volume* for each planning period were compared between simulations based on a certain data source and the true description. The mean deviation was estimated as a measure of systematic deviation and the absolute mean deviation as a measure of the average variation for ten 5-year planning periods.

In Scenario 1, one landscape consisted of the 67 stands with each stand having an area of 5 ha. Three scenarios were calculated with different interest rates (2%, 3%, and 4%) and denoted Scenario 1a, 1b, and 1c, respectively. In Scenario 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 was used as the source of age information for each stand. The area weights of the stands for these landscapes are presented in Table 5.

Table 5. *Age structures of the landscapes used in Scenario 2 and area per stand.*

		Landscape 2a		Landscape 2b		Landscape 2c	
Age class (yrs)	Number of stands	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	1 500	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
> 80	13	1 000	76.9	500	38.5	1 500	115.4

Cost-plus-loss using simulation (Paper III)

In Paper III the simulations in FMPP were used to calculate the loss due to non-optimal decisions based on errors in input data. Only decisions from the first ten years were considered in the analysis. From ten years and onwards correct data were used. The evaluation data were used as input in the forest scenario analysis. Identical evaluations were carried out using data with 2, 5, and 10 ha large stands. Analysis with two different interest rates, 2% and 4%, were carried out. The suggested optimal treatment schedule based on the evaluation data were then used during a ten year time period with true input data. The difference in net present value for each stand was determined as the loss (Fig. 9). The loss and the inventory cost of each data set were then summarised.

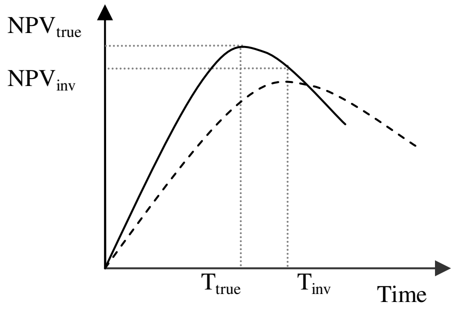

Fig. 9. The solid line is the true Net Present Value (NPV), and the dashed line is the NPV based on inventory data. The decision loss is calculated as the deviation between true NPV and the NPV based on the inventory data. T is the optimal time for cutting based on the true and inventoried state.

Analytical cost-plus-loss (Paper IV)

In Paper IV an analytical approach of cost-plus-loss was applied. By determining the relationship between error and loss in a NFI, an appropriate sample size could be determined. The analysis was only considered when the data were to be used to make a prognosis about sustainable harvesting levels at the national level. The losses varied depending on whether harvesting levels were underestimated or overestimated. In most applications the loss is determined with a quadratic loss function. These are appropriate when losses decrease at decreasing rate of error, such optimising net present value. However, these loss-functions would not be a realistic scenario in determining sustainable harvesting levels at national level. Thus, an linear loss function is appropriate when losses are proportional to the absolute value of the error in the inventory estimate (Hamilton, 1979). Harvesting estimate errors were determined linearly related to losses and thus a linear loss function of the following kind was used:

$$L = \begin{cases} \beta \lambda_1 |\varepsilon_v| & if \ \varepsilon_v > 0 \\ \beta \lambda_2 |\varepsilon_v| & if \ \varepsilon_v \leq 0 \end{cases} \tag{8}$$

Here, L is the loss, ε_v is the deviation between the correct total volume and the estimated total volume based on calculations using Swedish NFI data, and λ_1 and λ_2 are constants relating harvesting level error with loss. Harvesting levels are not directly estimated within the Swedish NFI; instead, data on trees and site conditions on plots are entered into the HUGIN system (Lundström & Söderberg, 1996). With this system, estimates of current growth increments form the basis for the estimation of sustainable harvesting levels. HUGIN uses simulation to make a prognosis on future growth. Thus, contrary to FMPP used in Paper III, there is no optimisation of net present value. For the purpose of this study we believe that it is a reasonable, and simplifying, approximation to assert that the output harvest level is the same as the estimated current net increment. In turn, the net increment can be expressed as a proportion, β, of the estimated total volume (Eq. 8). The expected cost-plus-loss can then be expressed as:

$$\begin{aligned} E(C+L) &= c_0 + c_1 n + \beta \frac{\lambda_1+\lambda_2}{2}\sqrt{\frac{2}{\pi}}\frac{Std\,(V)}{\sqrt{n}} = \\ &= c_0 + c_1 n + \beta \frac{\lambda_1+\lambda_2}{2}\sqrt{\frac{2}{\pi}}\frac{A\,Std\,(V_{ha})}{\sqrt{n}} \end{aligned} \tag{9}$$

Here, $Std(V)$ is the population (of infinitely many tracts) standard deviation of total volume, $Std(V_{ha})$ is the corresponding standard deviation of per-hectare volume, and A is the area of the area studied. To find the cost-plus-loss minimal number of tracts, Eq. 9 was differentiated with respect to n and the derivative set to zero (after verifying that the second order derivative was positive). The following expression for the optimal number of clusters was obtained:

$$n_{opt} = \left(\frac{\beta(\lambda_1+\lambda_2)\sqrt{2/\pi}\,AStd(V_{ha})}{4c_1} \right)^{2/3} \tag{10}$$

This formula was applied to each type of plots (permanent and temporary) in each region of the Swedish NFI, as well as for the country as a whole, by specifying that the conditions in mid-Sweden would represent average conditions at the country level.

The c_1 coefficients were € 1 700 and € 850, for full-day and half-day tracts, respectively. The analysis was assumed to be valid for a five year period. Error-losses were generalised based on a scenario about what would happen if data were over- or underestimated. These are further discussed in Paper IV; basically, an underestimation would lead to increasing import of timber, with an extra cost of € 38 m^{-3} over five years. If the errors of data lead to an overestimation, investments in new industry capacity and decreasing imports were considered in determining a decision-loss. In the case of overestimation, the extra cost was € 20 m^{-3} for five years. The β coefficients ranged from 0.027 to 0.039 and the standard deviation $Std(V_{ha})$ for the volume per hectare ranged from 34.6 to 94.0 m^3 ha^{-1}.

Results and discussion

Paper II

Focusing on specific planning periods in Scenario 1 provides information regarding how data quality influences the scenarios in the short term. In general, for all three scenarios (a, b, and c), the cuttings in the first planning period were delayed to the second planning period. A generalized pattern regarding standing volume was that initial estimates were rather accurate, followed by an underestimation in later planning periods. In the SPOT-based scenarios this is probably due to an overestimation of cuttings in planning period 2. The delayed cutting levels were a result of a slightly underestimated area of mature forest available for cutting. Deviations in the estimated standing and cutting volumes for each planning period, using a 3% interest rate, are presented in Fig. 10. In the laser-based scenario the cutting of 1 300 m^3 was delayed to the second planning period. This error is more than 3.5 m^3 ha^{-1}, which on a larger scale would affect decision-making considerably. Also here, the descriptions of the initial standing volume were similar, independent of the data source, and closely match the true description. Due to differences in optimal treatment schedules between the scenarios, however, the estimated standing volume differs in subsequent periods.

The results of the scenarios do not provide any evidence of major effects due to data quality with different interest rates. Comparing the two data sources, the laser-based imputations tended to perform better in the scenario analyses. In general, the scenarios predicted greater mean deviations and mean absolute deviations when using the SPOT-based data. Overall, the mean deviation indicated an underestimation of standing volume independent of the effect in the harvested volume. More detailed results, summarizing the complete planning period, are presented in Table 6.

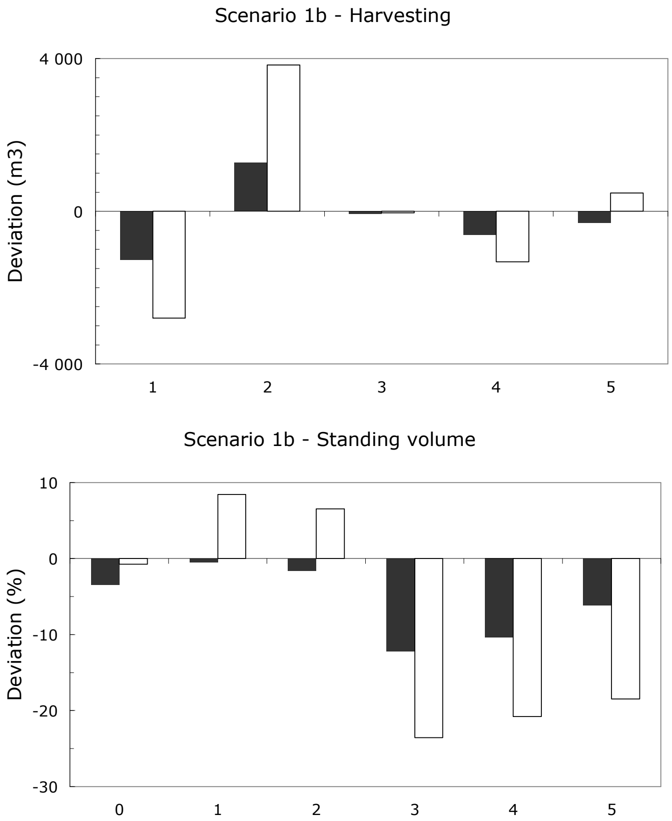

Fig. 10. The upper diagram presents the deviation of cutting volume, estimated using laser-based (dark) and SPOT-based data (light) in future 5-year planning periods. The lower diagram presents the deviation of standing volume for the same landscape. An interest rate of 3% was used in the scenario analyses.

Table 6. *Mean deviation (Mean dev) and Mean absolute deviation (Mean abs dev) for standing volume, harvested volume and net income in Scenario 1 for different interest rates for ten 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-based	2%	-3	4	-6	25	-1	3
	3%	-2	4	2	25	1	27
	4%	-2	4	1	23	1	23
SPOT-based	2%	-7	9	-5	30	-1	4
	3%	-8	11	3	42	1	45
	4%	-13	14	-5	30	-10	33

The harvesting levels in the landscapes with different weights to different age-class scenario analyses are presented in Fig. 11. Due to optimisation of net present value, typically most harvesting occurs in the first planning period. In Scenario 2a, the 5 000 ha area had an even age-class structure. The laser-based data initially overestimated the harvesting levels in the first planning period, but overall the cuttings were underestimated. A similar pattern could also be seen in the young landscape in Scenario 2b. In Scenario 2c, which mainly consists of both old and young forest, laser-based data underestimated the harvesting. However, summed over ten planning periods, the harvesting was overestimated with a mean deviation of 3% and 8% for the laser- and SPOT-based data, respectively. This is due to overestimation of stand age in the younger forests.

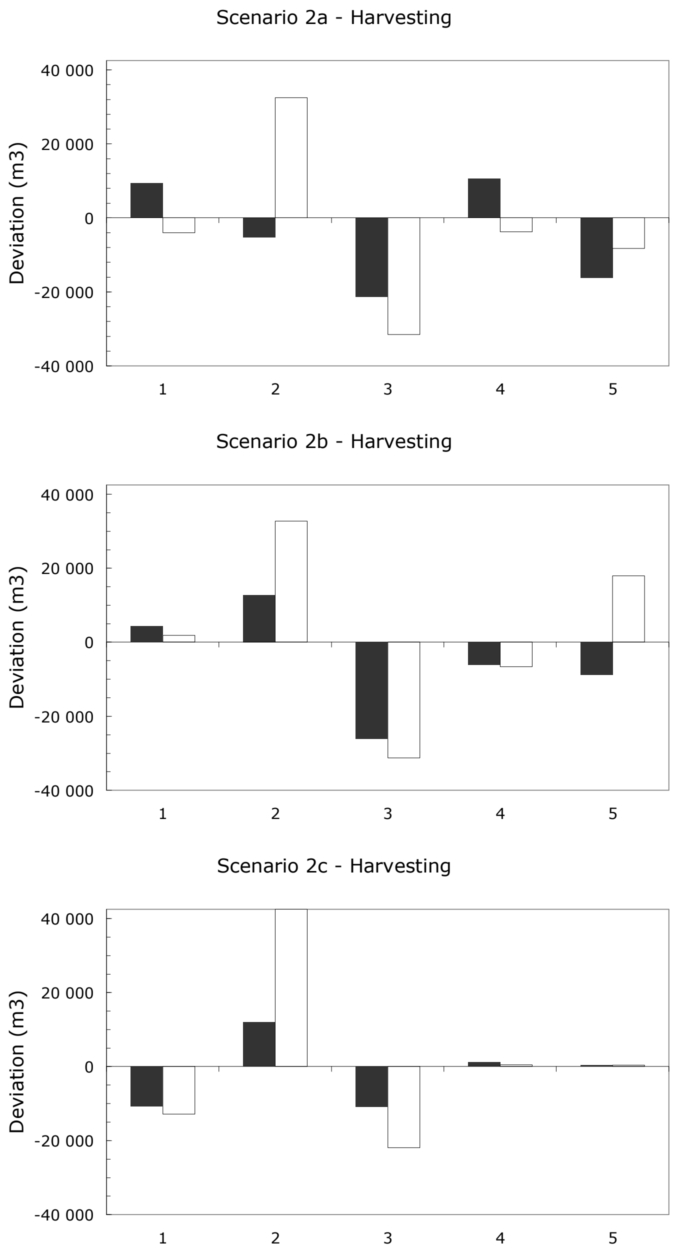

Fig. 11. The deviation of cutting volume for scenarios on Landscape 2a, 2b, and 2c. Estimated using laser-based (dark) and SPOT-based data (light) in future 5-year planning periods. An interest rate of 3% was used in the scenario analyses.

It is clear however, that there are large deviations looking at single planning periods. It can be stated that the laser-based data performed better than the SPOT-based data. Independent of the data, errors of this size would most probably have an effect on the decision-makers. This is not surprising; in Wallerman & Holmgren

(2007) the poor composition of reference data are discussed. At the extremes of the reference data set, sparseness in the distribution of reference data might cause bias (cf., McRoberts *et al.*, 2007). Maltamo *et al.* (2006) found that high volume plots were underestimated and low volume plots were overestimated due to the methodology using a similar method for imputation. Thus, the landscapes become averaged with an overrepresentation of the middle-aged forests. In the scenarios based on the imputation data, standing volume is underestimated and decreases over time. This is independent from harvesting levels which mostly are underestimated or delayed. This would probably also affect the cost of brushing and the proportions of the harvest whether timber came from final felling and thinning.

Paper III

The average decision losses are considerably lower when the simulated field-plot data are used instead of the imputation based data (Table 7). With a 4% interest rate the loss is generally lower than with a 2% interest rate. It is only in the case of using SPOT-based data that the loss is higher with the 4% of interest rate. Among the imputation based methods, the combination of laser- and SPOT-based data performed best, while using only SPOT-based data resulted in the highest average decision loss.

Table 7. *The average decision loss (SEK ha^{-1}) for all stands and each method*

	Interest rate	
Method	2%	4%
Field-plots (n=10)	86	18
Field-plots (n=5)	133	33
Laser & SPOT	769	346
Laser	1028	756
SPOT	1850	1925

Summarising the average cost-plus-loss results, the field-plot method in general resulted in the lowest cost-plus-loss (Table 8). The cost-plus-loss tended to be lower with a higher interest rate. Furthermore, an increasing stand size decreased the average cost-plus-loss per hectare.

Table 8. *The average cost-plus-loss (SEK ha^{-1}) for each method. Boldface indicates the best-performing method in each case.*

	Interest rate					
	2%			4%		
	Stand area					
Method	2 ha	5 ha	10 ha	2 ha	5 ha	10 ha
Field-plot (n=10)	746	376	**249**	678	308	181
Field-plot (n=5)	**653**	**359**	258	553	**260**	**159**
Laser & SPOT	827	827	827	**404**	404	404
Laser	1085	1085	1085	813	813	813
SPOT	1872	1872	1872	1947	1947	1947

If only one inventory method is used, the average cost-plus-loss analysis favours the field-plot inventories. In Sweden, sample-plot inventory is considered too expensive for large-scale forest inventory. The decision loss based on the imputation based data is 5 to 40 times higher than the cost of inventory. In further developing the imputation methods, the improvement of data quality is of higher importance than reducing the inventory cost. However, when comparing the field-plot inventories and the imputation-based data, caution is needed. The imputation-based data are real data with only 64 observations and the average decision loss is affected by a few high values. The field-plot data are represented by 50 data sets for each stand and average decision loss will be an average value for all 50 repetitions. A final conclusion is, however, that it would be more productive to improve data precision, and decrease decision loss, than to cut data acquisition costs.

Paper IV

To obtain an estimate of the optimum number of plots at the national level, given a certain tract type, the conditions within Region 3 (mid-Sweden) were assumed to hold for the entire country although the area of the country was substituted for the area of the region. In Table 9, the optimum number of tracts is presented based on calculations using Eq. 10.

Table 9. *Optimum number of tracts (5 yrs) per stratum and at the national level, separately for each tract type (if applied uniquely).*

	Region						
Tracts	1	21	22	3	4	5	Country
Perm.	1 653	1 334	1 707	1 860	3 046	2 502	6 520
Temp.	1 663	1 240	1 648	1 739	5 108	2 429	6 094

Selecting mid-Sweden as a typical part of Sweden, results indicated that the Swedish NFI sample size should be about 1 219 using temporary tracts and 1 304 for using permanent tracts if expected cost-plus-loss was minimised. This can be compared with the current level of 1 412 tracts. In Paper IV, an alternative approach was used to determine a “worst-case” cost-plus-loss, indicating that the sample size should be in the order of 2 500 tracts annually. As a rough conclusion, it could be stated that the current inventory sample size is in the right order of magnitude. However, it should be stressed that this assessment only considered using the NFI for determining harvesting level, whereas in reality it serves a large number of different purposes.

Some of the simplifications made in the study may have overestimated the optimum sample size. For example, it is not likely that past data and conclusions are fully disposed of when new data and a new analysis are made. Indeed, it is likely that new results indicating very different harvesting levels would be treated with great caution. To some extent this was accounted for in the study, but in reality decisions may not follow the calculated harvesting levels as strictly as was assumed in this study.

Methods to enhance usability of spatially comprehensive data for forestry scenario analysis (Paper V-VI)

As stated in Paper I, detailed data about the structure within a forest stand are sometimes required, often to the level of single trees. Furthermore, not only stand level data are required but also the composition and spatial configuration of stands within a forest landscape are essential information (e.g., Gustafson, 1998; Shifley *et al.*, 2006). Knowledge about spatial patterns allows for more detailed forestry scenario analysis and several resource indicators even demand such data. Habitat suitability models are one example of a resource indicator that is strongly correlated with the structure of the landscape. Non-parametric methods for imputation can preserve between-variable consistency within a unit, but do not consider the consistency between variables in geographically nearby units. In Papers II and III the consequences of errors in data used in forestry scenario analysis were evaluated. Poor decisions were not only due to low accuracy in the sample-plot imputations but also the poor composition of data had an effect. Thus, in spatially comprehensive data, spatial consistency has to be considered when acquiring data for forestry scenario analysis, both within a forest stand and between forest stands. This issue is seldom stressed in data acquisition.

Material and methods

In Papers V and VI, two frameworks on how spatial consistency can be improved are suggested. Paper V provides an approach that can be used within a forest stand while Paper VI is a further development and provides an approach that can be used to capture spatially consistent data at the landscape level, also including the within stand variability.

Details of the method and material for Paper V

An overview of the framework suggested in Paper V is presented in Fig. 12. First, a definition of data quality and determination of target values are needed. Secondly, an initial description of the forest stand provides a starting point. Finally, an optimising search algorithm is used to modify the description to meet the variability targets.

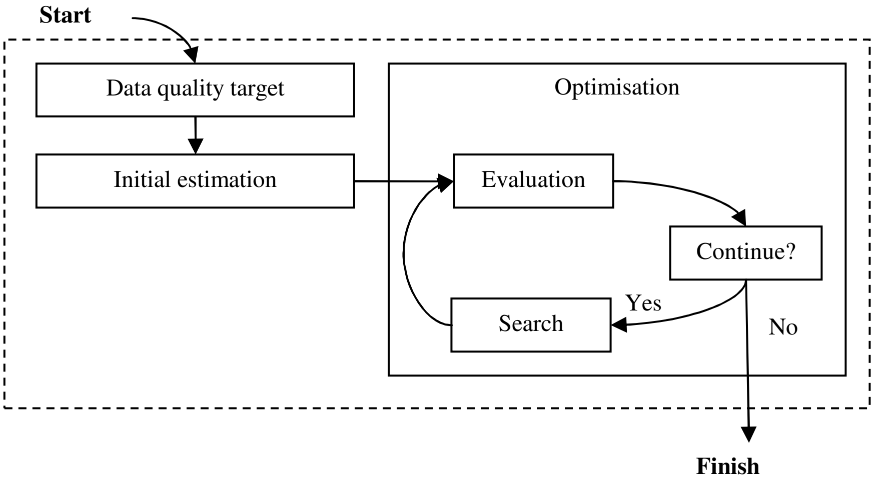

Fig. 12. An overview of the method used in Paper V.

Non-parametric methods can be used for acquiring spatially comprehensive data in a forest stand. Reference data assessed in field are imputed to the forest stand using carrier data. The carrier data contain at least one independent variable in every cell of the forest stand. In Paper V, the initial estimation was based on *k*NN (Tomppo, 1990; Nilsson, 1997). In this study, *k* was set to one. This is typically done when variable consistency within a unit should be preserved. Euclidean distance is used to determine similarity based on carrier data, *C*, between the target unit, *t*, and the reference unit, *r*. The Euclidean distance between points in *n*-space is defined as:

$$d_{t,r} = \sqrt{\sum_{i=1}^{n} (C_{i,t} - C_{i,r})^2} \tag{11}$$

In the objective function, targets for data quality characteristics were determined. Three quality characteristics were determined in the objective function:

- Correlation (C): pair-wise correlation between adjacent units within the forest stand;
- Short range variance (SRV): average variance of units within a 3x3 units moving window; and
- Accuracy (A); a measure of the accuracy of the estimated stand average value.

Correlation and short range variance were used as metrics of spatial variability, whereas accuracy was used to fix the initial estimation. A target was set for these three quality characteristics. The accuracy target was determined by the initial estimation while the correlation target and short range variance were determined from the true state of the forest. The objective function was specified in the following way:

$$O = w_C\left(\left|C - t_C\right|/t_C\right) + w_{SRV}\left(\left|SRV - t_{SRV}\right|/t_{SRV}\right) + w_A\left(\left|A - t_A\right|/t_A\right) \tag{12}$$

Here, w is the weight for the different components and t is the predetermined target value. In this study the weights were set to 1.0 for correlation and short range variance, and 0.5 for accuracy.

Simulated annealing (SA) was used as an optimising search algorithm (Lockwood & Moore, 1993; Öhman & Eriksson, 2002; Pukkala & Kurttila, 2005). SA is typically simple to apply in complex problems, but does not necessarily find the global optimal solution. However, in general, a relatively good solution can be found within a reasonable amount of time. In brief, SA replaces a current solution with a randomly selected nearby solution. Better solutions (in terms of objective function value) are always accepted, whereas worse solutions are accepted with a probability, p, which depends on the corresponding objective function value (O) and the global parameter T (temperature):

$$p = e^{(o_{new} - o_{best})T_i^{-1}} \tag{13}$$

With decreasing values of T, the probability of accepting worse solutions decreases. Initially, a high value of T is specified; then a cooling schedule is applied so that the temperature decreases until it is close to 0. The reason for this procedure is to allow the algorithm to escape from local optima.

To find an alternative solution within the SA algorithm, the same basic principle as in the initial estimation was applied, although only for a few randomly selected units at each iteration. The *k*NN algorithm did not have to select the most similar units but instead a random number was selected in order to exclude a certain number of most similar reference units from being included in the imputation procedure. In the case study the random numbers were selected with uniform probability within the range of 1 to 20. This choice was made subjectively considering the variability in the carrier data set.

Details of the method and material for Paper VI

In Paper VI, a framework for improving the spatial consistency at a landscape level is proposed; the framework consists of four steps:

1. The spatial configuration at landscape level is determined with remote sensing data using methods available for segmentation and classification.
2. Total forest area is calibrated using NFI information.
3. The composition at landscape level is captured by a newly developed restricted imputation technique.
4. The composition and spatial consistency within stands is improved by rearranging imputed sample-plots (a) between stands and (b) within each stand.

Landscape spatial configuration is considered as being accurately captured with available remote sensing data and methods for segmentation (e.g., Hagner, 1990; Pal & Pal, 1993; Pekkarinen, 2002) and classification (e.g., Lu & Weng, 2007). After the initial step, the landscape has been segmented into patches and each segment has been assigned a class label. Furthermore, with the aid of an error matrix, the forest area has been adjusted so that the area of forest land corresponds to the area estimate based on NFI data (Czaplewski & Catts, 1992; Congalton & Green, 1999). Further details on the two initial steps are provided in Paper VI, however, the main focus in this study is on Steps 3 and 4.

In Step 3, imputation is used to assign field-plot data to each pixel in the stands. This procedure preserves within-pixel consistency between variables but does not control within-stand variability or spatial consistency. Here, an algorithm similar to the way in which *k*NN was used in Paper V was used, but differs in some important aspects. Satellite data are used as carrier data, and sample-plot data from the NFI are imputed into the patches classified as forest stands. Contrary to ordinary imputation, each sample-plot in the reference data set can only be imputed into the target units for a limited number of times; each sample plot is represented in the reference set as many times as it should be found in the landscape according to the NFI data. This approach secures that the composition of the "wall-to-wall" landscape at the pixel level will be the same as the composition of the NFI data.

The above kind of restricted imputation can be performed with different algorithms, each having its specific implication for the remaining parts of the imputation framework. The ambition is to obtain a close-to-final distribution of reference plots within stands. With the suggested method, the satellite image digital numbers are sorted in descending order both in target and reference units. Then, target and reference units are matched pair-wise in descending (or ascending) order. Following these steps of imputation, a landscape with forest data is obtained having the same composition, at the plots scale, as the NFI. However, the restricted imputation cannot assure that the composition in terms of stand level mean values or within-stand variability and spatial consistency are appropriately determined.

According to the assumptions in the problem, the landscape composition in terms of stand level values is unknown, but within-stand variability is assumed to be known from case studies. The last step of the methodological framework, contain two parts: (a) the position of reference plots are exchanged between stands in order to improve within-stand variability (and indirectly, the composition of stand level mean), (b) the position of reference plots are exchanged within each stand to improve spatial configuration of stand data.

Similar to the method in Paper V, optimisation is used to improve the within stand variability. However, Threshold Accepting (TA) was used for optimisation, which is a similar method to the SA algorithm (Dueck & Scheuer, 1990). The TA method was used because of the capability to stop the search when no more improvements are taking place. TA examines a single adjustment to a current solution, and accepts every new solution that *is not much worse* than the previously accepted solution (Bettinger *et al.*, 2002). The initial solution was provided from the restricted imputation in step 3. Current solutions are then changed by rearranging the reference plots two by two. The difference between the last accepted solution and the new solution is determined ΔE. This value is computed with the deviation between the values of the objective function in the two solutions. An initial threshold level T_{TA} is set by the user, and only if ΔE is less than T_{TA}, the new solution is accepted. The process continues until no more improvements occur during a user defined number of iterations. The threshold value is then made smaller (T_{TA}=T_{TA}-ΔT_{TA}). The process finally ends when one of the following three criteria are fulfilled: 1) the number of non-improving iterations exceeds a maximum level *C*; 2) the total number of search iterations exceeds a maximum level *S*; or 3) *T* reaches a user defined stopping point.

The composition for each stand is expressed in terms of variance. An objective function for the landscape is then determined:

$$O_L = \sum_{i=1}^{N} \left(\frac{\left| Var_i - t_{Var_i} \right|}{t_{Var_i}} \right)^2 \qquad (14)$$

Here, Var_i is the variance in each stand and t_{var} the target variance for the current stand. In an optimal solution the distance between target variance and variance is 0, thus, the TA was used for minimisation. The imputed reference plots are allowed to be rearranged within the complete landscape. The target values of each stand are typically determined by empirical data of typical forest stands.

Once the previous optimisation algorithm has ended, a last step is to rearrange the plots within each stand to improve spatial consistency. TA was used also for this purpose, using pair-wise correlation, *Corr*, between adjacent units and short-range variance, *SRV*, to determine the quality characteristics of spatial variability, as done in Paper V. The objective function is specified in the following way (Eq. 15):

$$O_S = \left(\frac{\left| C - t_C \right|}{t_c} \right)^2 + \left(\frac{\left| SRV - t_{SRV} \right|}{t_{SRV}} \right)^2 \qquad (15)$$

Here, *t* is the predetermined target values. Typically, empirical data from a sample of forest stands is used to determine the target values in the objective function.

Case study with the method developed in Paper V

The method for spatially consistent imputation was tested in a case study. Spatial consistency was considered for one forest variable: mean stem volume at plot level. A forest was simulated using a semi-variogram (Cressie, 1993) describing the variability of a northern Swedish forest stand (Fig. 13). In order to obtain stands with various characteristics three alternative stands were simulated by using different values for range in the semi-variogram. For each cell in the forest stand a digital number was simulated using a regression model; these data were used as carrier data (Fig. 14). The errors in the carrier data were assumed to be spatially independent. An independent set of reference data was also simulated, using 1 000 uniformly distributed random numbers between 0 and 550 m^3ha^{-1}. Based on these simulated reference volumes, carrier data were simulated according to the procedures described above. Target values for correlation and short-range variance were determined by the stands, while the target for accuracy was determined by the initial volume estimate.

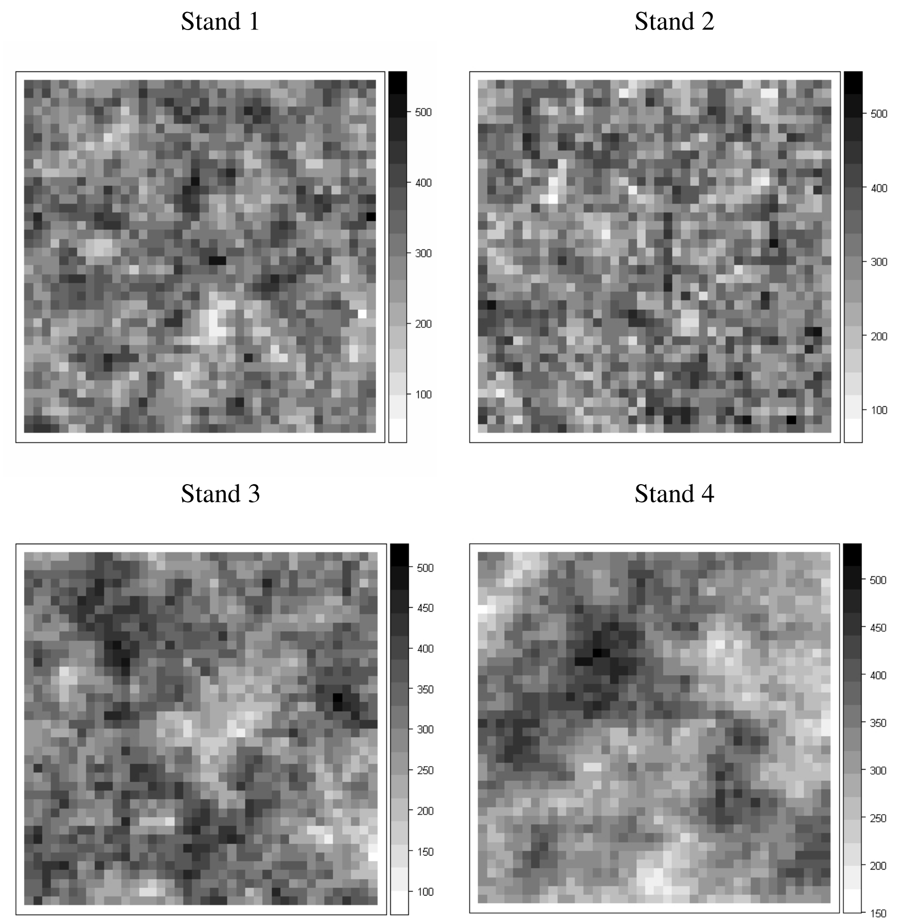

Fig. 13. Maps of the stands used in the case study of Paper V. The volume for each pixel is displayed according to the scale bar on the right side.

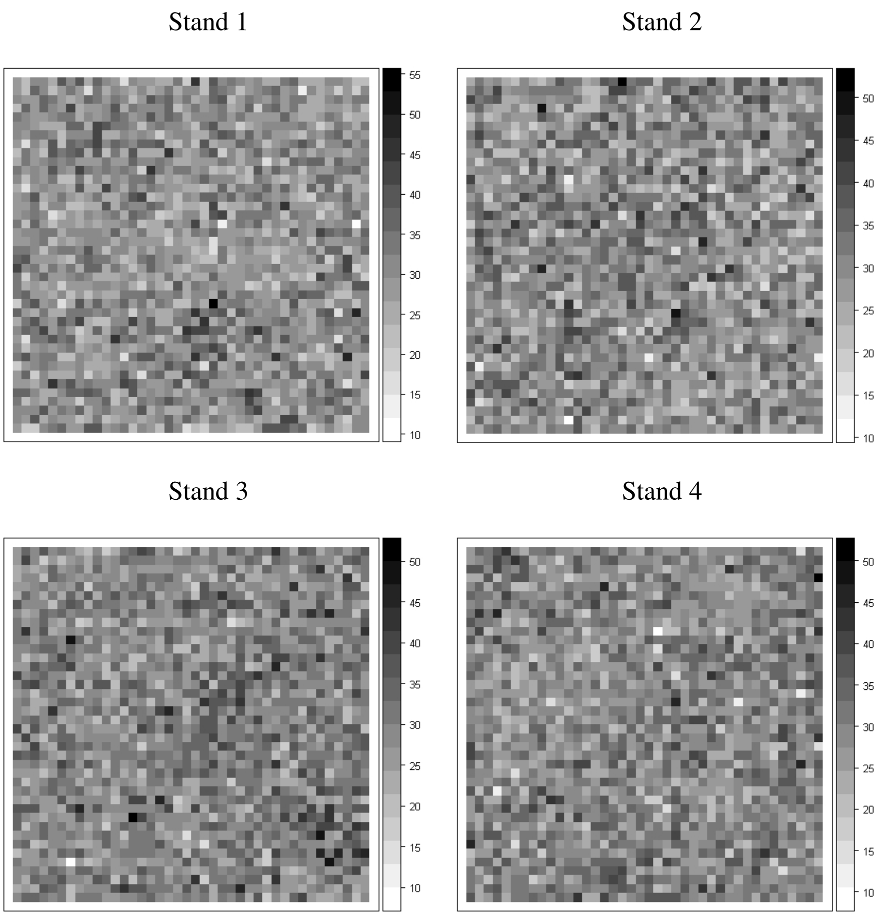

Fig 14. The simulated carrier data used for the stands in the case study. The digital number in each pixel is displayed according to the scale bar on the right side.

Case study with the method developed in Paper VI

The two last steps of the framework were evaluated in a case study using similar routines as in Paper V to provide test data. Nine stands were included in the case study, each of the stands were assumed to be 16 hectares large, square shaped, and each comprising 400 pixels (20x20 m pixel size). For stands with high volumes the same semi-variogram was used as in Paper V. By changing the sill (the maximum semi-variance), four alternative semi-variograms were used for forest stands with lower volume. The sill was changed in such a way that non-negative values of forest volume were attained in each stand. The semi-variograms and the simulated standing volume in each stand are presented in Table 10. Furthermore, a map of the simulated landscape is provided in Fig. 15.

Table 10. *Overview of the semi-variograms and the target standing volume in the simulation of the forest stands.*

	Sill	Nugget	Range	Volume (m^3 ha^{-1})
Stand 1	10	0	31	10
Stand 2	100	0	31	50
Stand 3	1000	0	31	100
Stand 4	1000	0	31	150
Stand 5	2500	0	31	200
Stand 6	2500	0	31	250
Stand 7	4753	0	31	300
Stand 8	4753	0	31	350
Stand 9	4753	0	31	400

Carrier data were simulated using the same formulas as in Paper V. A sample of reference data were acquired with simple-random sampling without replacement (Thompson, 2002). In total, 10% of the pixels were selected with stand data and carrier data. Since the sampling proportion was 10%, each sample plot and carrier data value was replicated ten times in the reference data set used for the restricted imputation.

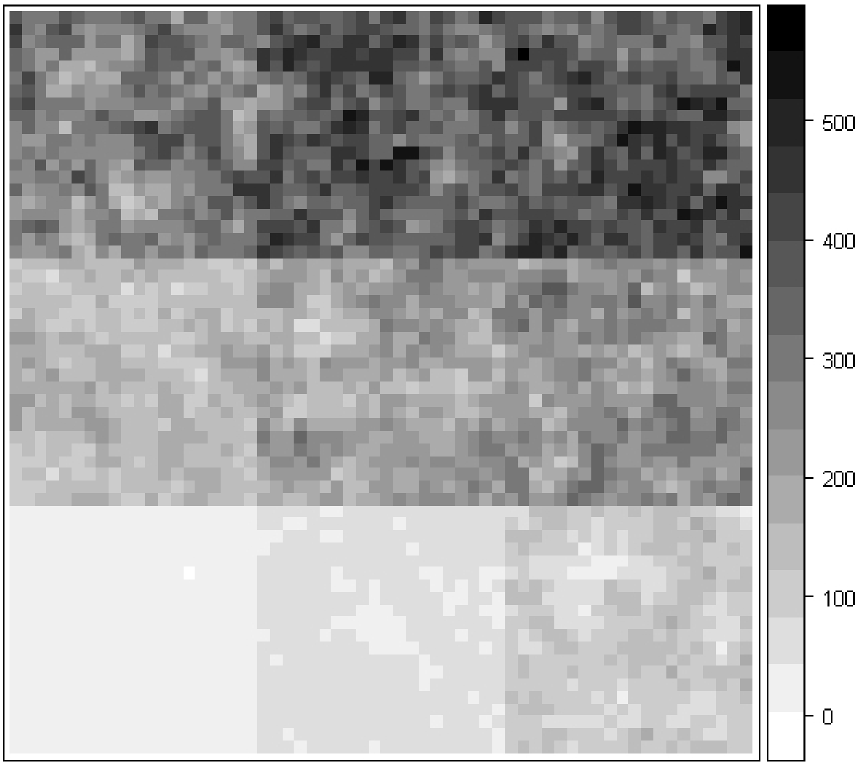

Fig. 15. Map showing the nine simulated forest stands in the landscape, the average standing volume according to the scale bar at the right side (m^3 ha^{-1}).

Results and discussion

Case study results and discussion for Paper V

The results show that it is possible to improve the spatial consistency in the estimations (Table 11). In 3 out of 4 stands the algorithm reaches the targets values in the objective function. In Stand 4, which had the most demanding targets, accuracy and short-range variance were achieved. However, the correlation could only be improved to 0.75 compared to the target of 0.81.

Table 11. *Results from Paper V. Boldface indicate successful results and italic non-optimal results. SCI is the spatially consistent imputation.*

		Stand 1	Stand 2	Stand 3	Stand 4
Accuracy	Target	308	246	249	253
	Initial	308	246	249	253
	SCI	**308**	**246**	**249**	**253**
Correlation	Target	0.51	0.36	0.67	0.81
	Initial	0.02	-0.01	0.04	-0.01
	SCI	**0.51**	**0.36**	**0.67**	*0.75*
Variance (short-range)	Target	3 033	3 873	1 973	1 056
	Initial	20 110	10 021	9 527	9 094
	SCI	**3 033**	**3 873**	**1 973**	**1 056**

Maps of Stand 1 and 2 are presented in Fig. 16 and Stand 3 and 4 in Fig. 17. Based on the random structure from the initial imputation, spatially consistent imputation improves the variability of the forest stand. The pattern in the maps may look smooth although the variability between some pixels is very high.

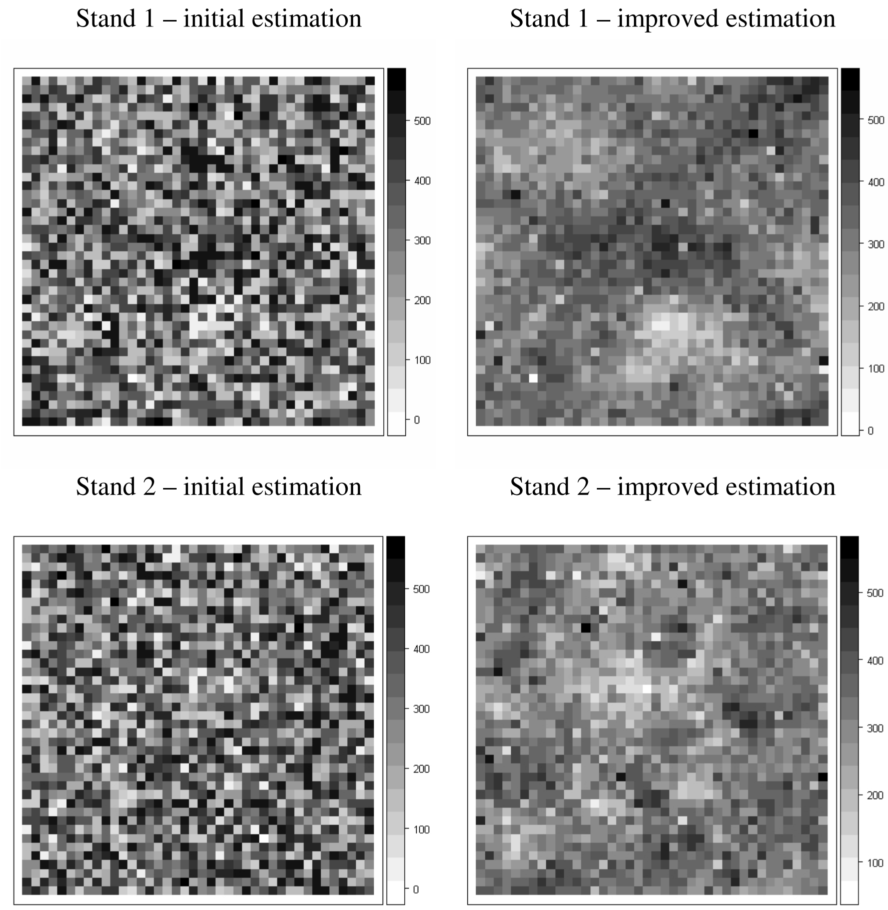

Fig. 16. Left column maps show the stand after the initial imputation based on *k*NN. The right column maps are the same stands but when spatially consistent imputation has been used. Volume in each pixel is displayed in the scale bar at the right side of each map.

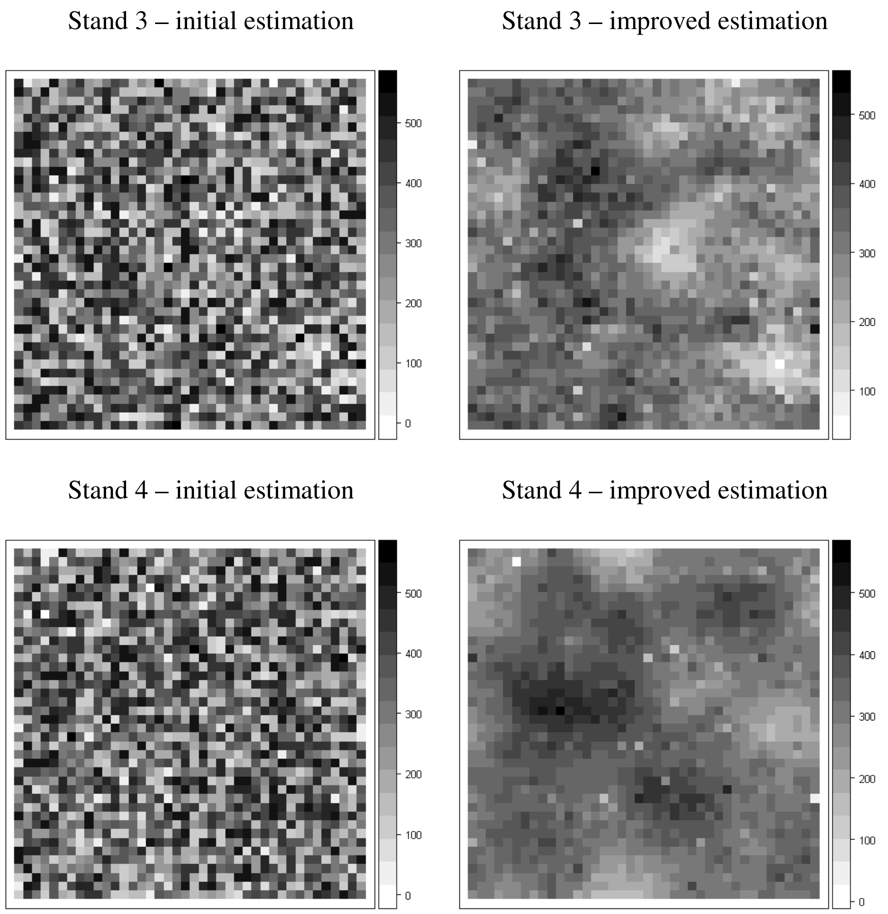

Fig. 17. Left column maps show the stand after the initial imputation based on *k*NN. The right column maps show the same stands when spatially consistent imputation has been used. Volume for each pixel is displayed in the scale bar at the right side of each map.

A positive effect of the improvement of spatial consistency was that variability, in terms of distributions of estimated stem volume in each pixel, was clearly improved (Fig. 18). No quality characteristics to enhance the variability were directly included in the objective function. For many variables a correct distribution of data would be valuable in the simulation of stand development.

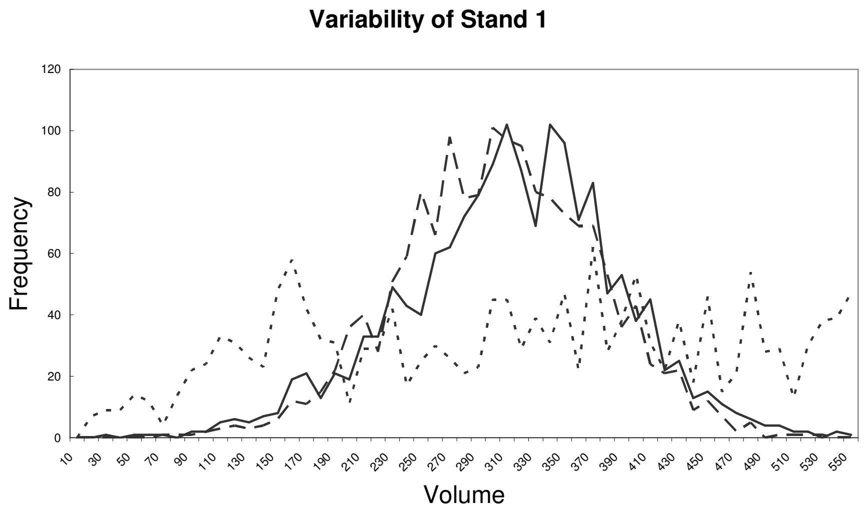

Fig 18. Variability within Stand 1, in terms of distributions of estimated stem volume in each pixel. The simulated reference data are given by the dashed line, the initial *k*NN imputation by the dotted line, and the result of the spatially consistent imputation by the solid black line.

The case study results illustrate that the proposed algorithm is useful for spatially consistent imputation within forest stands. Some possible improvements can be concluded. The carrier data do not consider the spatial dependency of errors that often occur in remote sensing data (Lillesand & Kiefer, 2000). When using more realistic carrier data, more information on the spatial patterns would probably improve the possibility of deriving a spatially accurate estimation. In that case the time efficiency would also be improved. Another possible improvement would be to further evaluate metrics to determine spatial consistency. Whether or not this problem could be solved using a semi-variogram approach was not tested, and further studies are needed in order to evaluate the spatial variability metrics which would be most appropriate to use. However, the spatial consistencies of the estimations were considerably improved.

For practical use of the algorithm there are some issues that have to be considered. In the case study, target values were assumed to be known from pilot studies, however, this method could probably also be used in practice. Solution time is another critical issue, although efficiency was not a priority in the development of the currently used software. Most critical, however, is the performance when spatial consistency of multiple variables is considered. In the case study, only one variable was considered. In operational applications, the demand for spatial consistency in general would involve several variables in the reference data set, such as, tree age, tree species composition, and site conditions. Different solutions to this problem are possible. One simplistic approach would be to consider spatial variability only for one variable, but select a variable known to have high correlation with other plot variables. Volume actually could be such a variable. Another solution would be to extend the algorithm with variability targets for more than one variable. In principle this is straightforward, although it is likely that the algorithm would encounter problems to reach all the targets. Solution times are likely to be extended and it would probably also be important to have a large reference dataset.

Case study results and discussion for Paper VI

The average variance for the nine stands in the forest landscapes was 2 193 $(m^3\ ha^{-1})^2$. The initial solution based on the restricted imputation resulted in an average variance of 12 284 $(m^3\ ha^{-1})^2$. Rearranging the reference plots in the landscape and by minimising Eq. 14, an average variance of 2 230 $(m^3\ ha^{-1})^2$ was reached. The results of the rearrangement in each forest stand are presented in Table 12. In nearly all cases the final solution was close to the true value. The main problems were in the two stands with the lowest volume, Stand 1 and Stand 2. Here the final *C* and *SRV* were too high. The required number of iterations in the optimisation of each stand was between 1 200 and 2 500.

Table 12. *Results from the improvements of the spatial configuration within each forest stand. Number of required iterations is presented as well as initial values from the restricted imputation and the final values after the optimisations. Boldface values indicate where the target was reached and italic values where it was not reached.*

	Iterations	*C*			*SRV*		
		Initial	Final	True	Initial	Final	True
Stand 1	2 400	0.00	*0.66*	0.49	2 565	*54*	6
Stand 2	2 564	-0.01	*0.65*	0.45	10 391	*94*	66
Stand 3	1 835	0.05	**0.50**	0.50	15 550	**612**	611
Stand 4	1 275	0.09	**0.47**	0.47	14 271	**661**	662
Stand 5	1 377	0.07	**0.54**	0.54	13 081	**1 410**	1 410
Stand 6	1 376	-0.06	**0.41**	0.41	14 768	**1 768**	1 769
Stand 7	2 006	-0.06	**0.49**	0.49	14 342	**3 120**	3 118
Stand 8	1 297	0.04	**0.41**	0.41	12 327	**3 152**	3 152
Stand 9	1 375	0.01	**0.43**	0.43	11 961	**3 123**	3 214

A map of the initial solution provided by the restricted imputation is presented in Fig. 19. These data serve as the initial solution for the optimisation algorithm. Due to the saturation effects in the carrier data and the random errors in the simulations of carrier data, stand borders are difficult to distinguish.

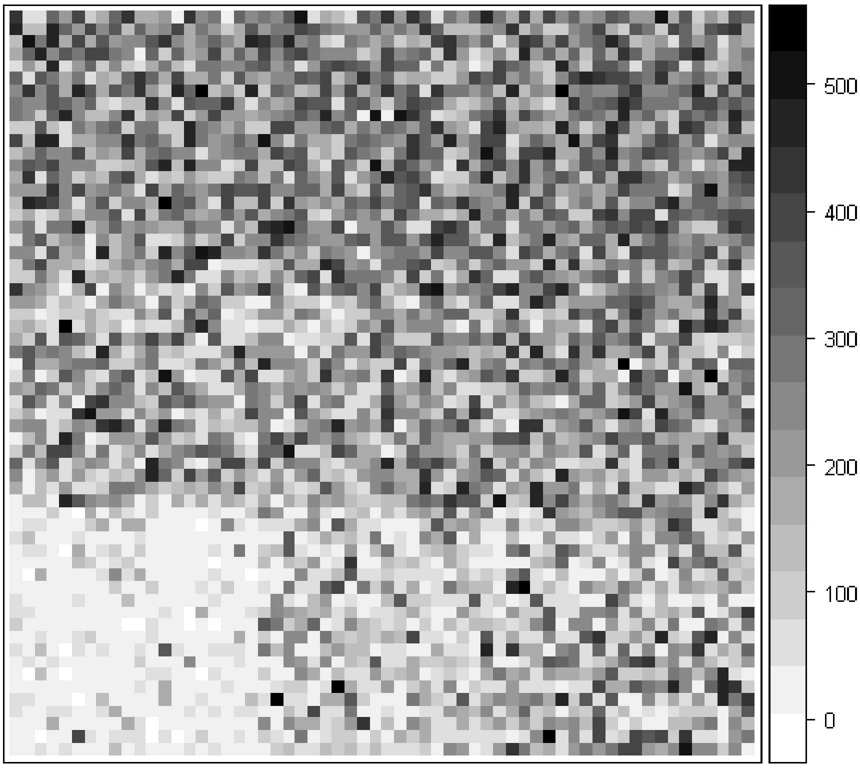

Fig. 19. Map showing the average standing volume in the nine forest stands in landscape after the restricted imputation in Step 3. Average standing volume in each pixel is according to the scale bar at the right side (m^3 ha^{-1}).

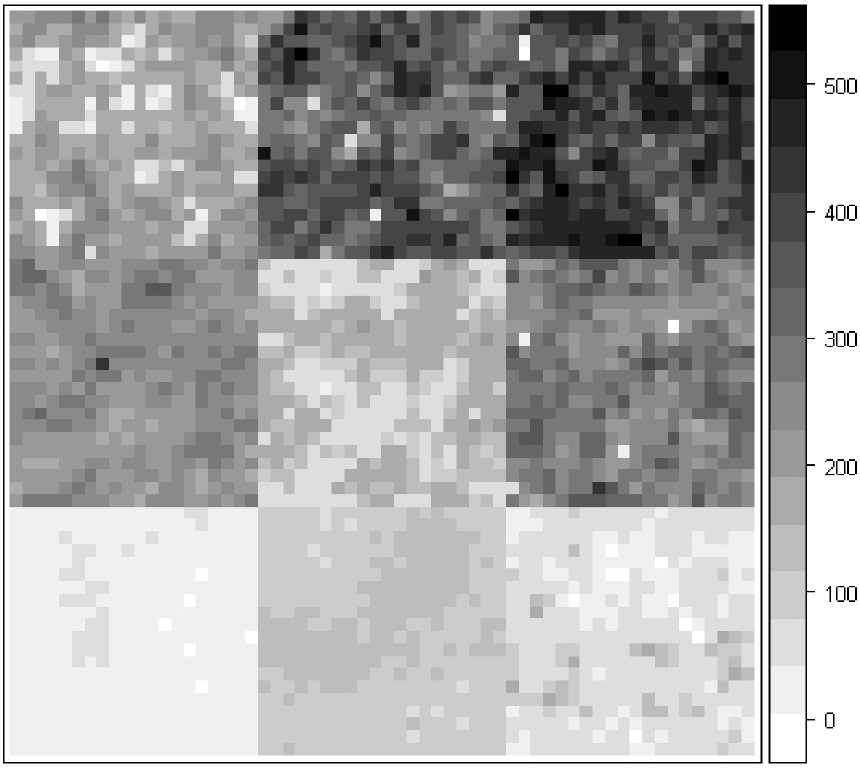

Fig. 20. The final results represented in map form showing the average standing volume in the nine stands of the forest landscape. Forest volume is according to the scale bar at the right side (m^3 ha^{-1}).

In the final solution presented in Fig. 20 stand borders are possible to identify and the within-stand variability is more similar to the simulated data in Fig. 15. One effect of the rearrangement of data is that the positions of individual stands have changed. However, the distribution of stands with different volumes is close to the true state. An overview of the average forest volume is presented in Table 13.

Table 13. *Average volume of the forest stands during the optimisation. Initial values are from the restricted imputation and the final values are from after the optimisations.*

	Volume (m^3 ha^{-1})		
	Initial	Final	True
Stand 1	34	16	10
Stand 2	115	107	52
Stand 3	171	54	102
Stand 4	183	238	146
Stand 5	209	119	201
Stand 6	231	272	247
Stand 7	242	174	282
Stand 8	263	341	368
Stand 9	274	402	388

Research on how spatially comprehensive data for forestry scenario analysis should be assessed is sparse. The framework suggested here is one possible approach, whereby data-introduced errors in the scenario analysis result can be avoided. Overall the presented solution in the case study performs as expected. However, in this thesis the methodological framework is only outlined in coarse terms, and further research and development is required before the framework can be applied in practice.

Discussion

Methods to evaluate data acquisition strategies

Evaluating data acquisition strategies is not straightforward. In the development of new inventory methods, the evaluation of data quality are often limited to the accuracy of single variable estimates as the only quality characteristics. The connection between forest inventory and the usability of data including the consequences of data quality in the decision-making are often weak. However, there are studies and methods that connect forest inventory and decision-making (see Duvemo & Lämås, 2006). With more complex forestry planning problems, and with more advanced decision support systems, these methods are limited in their ability to evaluate data acquisition strategies. New means to evaluate data acquisition strategies in multi-resource forestry are needed, as well as evaluation on the usability of inventory data at the national level.

In Paper I, a framework for evaluating data acquisition strategies was proposed. The method is applicable in practical applications, but would include some subjectivity since it is based on the knowledge and experience of the user. However, based on prior knowledge provided by the research community it can be used as a support when making decisions about forest inventory. The unique possibilities are that multi-resources and national level requirements can be considered, which could be valuable in future national forestry scenario analysis (e.g., Gustafsson, 2000; Lämås & Eriksson, 2003).

In Paper II, a simulation approach was used to study the effect of data quality in connection to decision-making. With this approach the effects of data on decision-making can be evaluated for multi-resource analyses, and also spatially comprehensive data can be considered during the analysis. Using more advanced decision support systems such as Heureka (Lämås & Eriksson, 2003) or Monsu (Pukkala, 2004), the consequences of data quality on different resources could be studied. It would be a relevant approach at the national level and can be used in practice when a data acquisition strategy for a certain forestry scenario analysis is evaluated. However, the analysis is complex due to many correlated results, such as cutting volume and standing volume. An exact comparison on the effect of data between different resources is not a simple task. With a national economic approach, the decision loss could be considered and a comparison between data acquisition strategies could be done. This would still be rather complex and require simplifications.

The simulation approach has been used in many studies, but with the means of cost-plus-loss (Holmström, Kallur & Ståhl, 2003; Eid, Gobakken & Næsset, 2004; Holopainen & Talvitie, 2006). This approach was used in Paper III. The advantages are that it summarises the effect over the complete planning period and different acquisition methods can easily be compared in a numerical approach. A drawback is that only net present value is considered and it is difficult to evaluate resources that cannot be defined in monetary terms. In analysis at the national level there are no direct connections between inventory cost and the loss in net present value at the forest stand level. Thus, this type of analysis is more relevant at the level of the forest owner. From the perspective of national level forestry planning, these analyses are valuable for providing more knowledge on different data acquisition strategies.

An analytical approach of cost-plus-loss was used in Paper IV. The cost of inventory at the national level was linked with decision loss in the forest industry. Here an assumption was made that the decision loss in the forest industry affected the national level with the same strength. The aim was to determine an appropriate accuracy level for NFI data when data are used to determine national sustainable harvesting levels. The results give an indication of the order of magnitude the sample size of a NFI should be. Generally, this is not known. As in most applications of cost-plus-loss analysis, the weakest parts are the loss functions. Here, the loss function was based on a typical scenario. The scenarios can be determined in many different ways, and are dependent on the knowledge and experience of the user. A general limitation of the method is that only sampling intensity is optimised and an evaluation using spatially comprehensive data would be difficult. Another difficulty is in determining the loss where multiple decisions are involved. However, this method is able to numerically determine a minimum accuracy level.

When suggesting an overall method to be used in planning a forest survey, evaluation would include many of the tools listed above (or similar methods). Planning data acquisition strategies for national forestry scenario analysis involves decisions that are dependent upon the experience and knowledge of the decision-maker. There is no complete tool for an evaluation of data acquisition strategies in different decision situations, especially when multiple resources are considered in national forestry scenario analyses. However, the evaluations of different data acquisition strategies are important, especially those where data are linked with decisions. The knowledge gained from numerical evaluations will make a good basis when a final decision is made as suggested in Paper I. A wider evaluation of data acquisition strategies in the early stages in the development of inventory data is valuable. As an example, Wallerman & Holmgren (2007) considered tree species composition and within-stand variability in addition to accuracy in their evaluation of spatially comprehensive data.

Data acquisition strategies for forestry scenario analysis

In Paper I arguments for spatially comprehensive data are given. Detailed data on single trees and information about landscape patterns are required in many applications. Evaluations were done in Papers II and III, mainly in timber oriented forestry scenario analyses based on such data. The accuracy of the different carrier data has an impact on the results: the performance of the laser-based data is considerably better than the SPOT-based imputation. This is reported in both studies. Laser-based data have been evaluated with simulated cost-plus-loss analysis and produced similar results, but using stand based data. Eid, Gobakken & Næsset (2004) compared laser scanner data with photo-interpreted data and concluded that laser data were more efficient. Holopainen & Talvitie (2006) concluded that when inventory costs were also considered, the field based inventory was still more efficient than the laser-based data. However, in larger landscapes the lower cost of remote sensing data would favour these methods.

In the evaluation scenario analysis in Paper II the cutting schedule is slightly delayed and the harvesting is overestimated, probably due to a poor composition of the landscape data. The poor results obtained when using accurate laser scanning data may be due to the methodology. An early evaluation of the laser data concluded that the accuracy level was good enough for forest planning, but that poor tree species composition and poor within-stand variability would be problematic (Wallerman & Holmgren, 2007). Maltamo *et al.* (2006) presents some examples on predicted volume at plot level, estimated with a non-parametric method. In their study, a saturation effect could be noticed in the aerial photograph, while the laser scanner data were considerably better. However, due to the methodology, it was impossible to find suitable neighbours for high values in the laser scanner data, and these were therefore underestimated. Correspondingly, the estimated plots with low values were overestimated.

The decision loss that was used in the cost-plus-loss analysis in Paper IV would have a direct effect on the budgets of the Swedish forest industries. Thus, if these data had not been available from the authorities, it would be economically motivated for the forestry sector to acquire these inventory data. However, the NFI data are required for many other decisions and reporting, and the inventory cost should therefore be well-motivated for the governmental organisations as well.

In Paper IV, the accuracy level of the Swedish NFI is motivated just for the use of one single decision. This does not necessarily mean that more forest data are required, but it is conceivable to conclude that detailed forest data would be beneficial to the assessment. In cost-plus-loss analysis, the decision loss seems to be more critical than inventory cost. In Paper III, the more expensive sample-plot data is still the most efficient data acquisition strategy when inventory cost is considered (cf., Holopainen & Talvitie, 2006). Thus, the value of forest data should not be underestimated.

Enhance data usability

Evaluating spatially comprehensive data in national forestry scenario analysis indicates that decisions are affected by poor quality data. Spatially consistent data are required and improvements on the methodology seem to show high potential. In Paper V, the spatial consistency is improved within a forest stand, with a methodology that considers adjacent units. The method can utilise much of the spatial information in the carrier data and could be useful in large stands or in areas with no distinct stand borders. A positive effect of the method for managed forest areas is the ability to capture the variability within the forest stand. The spatial consistency within a forest stand may not be as critical as between stand consistency at the landscape level. Enhancement of spatial consistency at the landscape level is more complex. Based on the knowledge from Paper V, a concept is suggested in Paper VI as to how this issue can be handled on the landscape level.

A major challenge in Paper V was how to describe spatial consistency. The metric has to determine the variability between the variables in such a way that the accuracy of important spatial characteristics is determined. Within a forest stand, the semi-variogram can be used to describe the variability (Cressie, 1993). Also simpler metrics, used in Paper V, seem to be able to improve the spatial patterns as well. However, which is the most efficient and accurate approach has not been evaluated.

Determining the arrangement of different forest stands at the landscape level is complex. A first attempt was to use a similar approach as in Paper V at the landscape level. There are a large number of different landscape metrics that are used to explain the patterns of a forest landscape (Gustafson, 1998). The aim of such an approach would be to improve the accuracy of the landscape metrics that are of interest for the forestry scenario analysis. However, the NFI data do not provide any information of the spatial configuration in the landscape, at least not in Sweden. Thus, it was asserted that configuration should be adequately enough captured with remote sensing data. On the other hand, overall landscape composition at the scale of plots is available in the NFI data. This was utilised by assuring that the composition in the "wall-to-wall" data should be identical to that in NFI data.

In common for both of these methodologies, but most important at the landscape level, is that the exact positions of different resources are not necessary in a national forest scenario analysis. In strategic planning at national level, sustainability of different resources is analysed. The aim of the decision-maker is to make sure that all resources are available in a landscape over time, but does not necessarily need to provide the geographic location of these resources. This would not be the case in a scenario analysis of a landowner who needs to know where different timber assortments are found and where to allocate conservation areas. As an example, the aim of analysing the clustering of harvesting activities (Öhman & Lämås, 2003) could be different depending on the aim of a scenario analysis. For the landowner, the aim would be to optimise machinery cost and logistics and subsequently to know where and when to cut different stands. In a national forestry scenario analysis, the aim would be to improve the analysis with more details in the prognosis or studying the overall effect in optimising cutting locations. In this case it is important that the initial data accurately describe the spatial distribution between different stand types.

Conclusions

A generic approach of cost-plus-loss analysis can be used to find arguments in the discussion about different data acquisitions strategies. By considering data requirements of the resource indicators and the tools that will be used in the forest scenario analysis, much of the minimum data requirements can be recognized. By linking data requirements to the quality characteristics of data, data acquisition strategies can be ranked. This approach would be useful in practical planning of forest surveys, especially in national forestry scenario analysis where multi-objective resources are considered.

Numerical methods can be used to evaluate data acquisition strategies in applications for specific decisions. These methods provide valuable knowledge in the planning of forest survey and when data acquisition strategies are being ranked. Furthermore, in the development of data acquisition strategies, it is important that the evaluation links data quality with decision-making. These approaches may be best applied in a research environment in order to then pass on knowledge to the users.

In Paper II, it is clear that data quality will influence the decision. In the evaluation of spatially comprehensive data the more accurate laser-based imputation performed better than the SPOT-based imputations. However, the poor composition of the landscape data affected the cutting patterns. Total cutting volumes were delayed in time and affected the decision. The usability of these data in a national forestry analysis is dependent on an improvement of data composition. This was also concluded in Paper III due to the very high decision losses in contrast to the inventory costs.

The accuracy level of the Swedish NFI is appropriate if the inventory data would be used only to determine national sustainable harvesting levels. A conclusion is that the current level of data acquisition is economically motivated and reducing the accuracy level would not reduce the total cost-plus-loss. Considering that these data are also used for making many other decisions, a wider cost-plus-loss analysis may motivate an improvement of the current accuracy level.

Spatial consistency within a forest stand can be improved with new imputation algorithms. This might be especially useful in forestry scenario analyses at national or sub-national level, where reference sample-plot data are available from national forest inventories and carrier data can be cost effectively assessed using remote sensing. This type of data is already available in many countries, including Sweden. However, the methodology is only suitable for national level policy-making and the case studies made in Papers V and VI are simplifications. Thus, for use in practical applications, further development is required.

Future research

There are several possibilities for further development of the methods for evaluating data acquisition strategies. However, it is not likely there will ever be a "complete" method available to evaluate the consequences of using different data sources as a basis for decision-making. Thus, planning forest surveys will always include a good proportion of uncertainty. Formal evaluation techniques in many ways provide good support to the decision-maker when planning a forest inventory. For example, simple methods to determine data usability would be valuable to apply at early stages of the development of new data acquisition strategies. This is important since data requirements continue to increase over time and there is still need for substantial improvement of the data acquisition methodologies. Further development of the methodology for assessing spatially comprehensive data is required, and the methods suggested in this thesis have not yet been tested in practise. A major challenge, which has not been considered in this thesis, is how spatially comprehensive data should be acquired for the tactical and operational planning levels. Thus, further research is needed, and this work has just started!

Literature

Achim, A., Ruel, J.-C., Gardiner, B.A., Laflamme, G. & Meunier, S. 2005. Modelling the vulnerability of basam fir forests to wind damage. *Forest Ecology and Management 204*, 35-50.

Agestam, E. 1985. A growth simulator for mixed stands of pine, spruce and birch in Sweden. *Swedish University of Agriculture Sciences, Department of Forest Yield Research, Report 15.* 150 pp.

Andalo, C., Beaulieu, J. & Bousquet, J. 2005. The Impact of climate change on growth of local white spruce populations in Québec, Canada. *Forest Ecology and Management 205*, 169-182.

Andersson, Å., Kallio, M. & Seppälä, R. 1986. System analysis for the forest sector. In: *System analysis in forestry and forest industry*. (Eds. M. Kallio, Å. Andersson, R. Seppälä & A. Morgan). Elsevier Science Publishers B.V. Amsterdam, The Netherlands.

Andersson, D. & Eriksson, L.-O. 2007. Effects of temporal aggregation in integrated strategic/tactical and strategic forest planning. *Forest Policy and Economics 9*, 965–981.

Andreassen, K. & Tomter, S.M. 2003. Basal area growth models for individual trees of Norway spruce, Scots pine, birch and other broadleaves in Norway. *Forest Ecology and Management 180*, 11-24.

Angelstam, P. & Andersson, L. 2001. Estimates of the needs for forest reserves in Sweden. *Scandinavian Journal of Forest Research 16*, 38-51.

Anon. 2006. Estate and ownership structure. In: *Swedish statistical yearbook of forestry 2006*. (Ed. J.-O. Loman). Swedish Forest Agency. Jönköping, Sweden.

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.

Bailey, R.L. & Dell, T.R. 1973. Quantifying diameter distributions with the Weibull function. *Forest Science 19*, 97-104.

Bettinger, P. & Chung, W. 2004. The key literature of, and trends in, forest-level management planning in North America, 1950-2001. *International Forestry Review 6*, 40-50.

Bettinger, P., Graetz, D., Boston, K., Sessions, J. & Chung, W.D. 2002. Eight heuristic planning techniques applied to three increasingly difficult wildlife planning problems. *Silva Fennica 36*, 561-584.

Bettinger, P., Sessions, J. & Boston, K. 1997. Using Tabu search to schedule timber harvests subject to spatial wildlife goals for big game. *Ecological Modelling 94*, 111-123.

Biging, G.S. & Dobbertin, M. 1992. A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. *Forest Science 38*, 695-720.

Bitterlich, W. 1984. *The relascope idea: Relative measurements in forestry.* Farnham Royal CAB. London, UK. 242 pp.

Blennow, K. & Sallnäs, O. 2004. WINDA - a system of models for assessing the probability of wind damage to forest stands within a landscape. *Ecological Modelling 175*, 87-99.

Bollmann, K., Weibel, P. & Graf, R.F. 2005. An analysis of central alpine capercaillie spring habitat at the forest stand scale. *Forest Ecology and Management 215*, 307-318.

Boman, M., Bostedt, G. & Persson, J. 2003. The bioeconomics of the spatial distribution of an endangered species: The case of the Swedish wolf population. *Journal of Bioeconomics 5*, 55-74.

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

Cashore, B., Auld, G. & Newsom, D. 2003. Forest certification (eco-labeling) programs and their policy-making authority: explaining divergence among North American and European case studies. *Forest Policy and Economics 5*, 225-247.

Congalton, R.G. 1988. Using spatial auto-correlation analysis to explore the errors in maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing 54*, 587-592.

Congalton, R.G. & Green, K. 1999. *Assessing the accuracy of remotly sensed data: Principles and practices*. Lewis Publishers. Boca Raton, U.S. 137 pp.
Cressie, N.A. 1993. *Statistics for spatial data*. Revised edition. John Wiley & Sons, Inc. New York, U.S. 900 pp.
Cubbage, F.W. & Newman, D.H. 2006. Forest policy reformed: A United States perspective. *Forest Policy and Economics 9*, 261-273.
Czaplewski, R.L. & Catts, G.P. 1992. Calibration of remotely sensed proportion or area estimates for misclassification error. *Remote Sensing of Environment 39*, 29-43.
Davis, L.S., Johnson, K.N., Bettinger, P.S. & Howard, T.E. 2000. *Forest management*. 4th edition. McGraw-Hill Science/Engineering/Math. New York, U.S. 790 pp.
Davis, R.G. & Martell, D.L. 1993. A decision support system that links short-term silvicultural operating plans with long-term forest-level strategic plans. *Canadian Journal of Forest Research 23*, 1078-1095.
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.
Dueck, G. & Scheuer, T. 1990. Threshold Accepting – A deneral-purpose optimization algorithm appearing superior to Simulated Annealing. *Journal of Computational Physics 90*, 161-175.
Duvemo, K. & Lämås, T. 2006. The influence of forest data quality on planning processes in forestry. *Scandinavian Journal of Forest Research 21*, 327-339.
Dykstra, D.P. 1984. *Mathematical programming for natural resource management*. McGraw-Hill Book Company. New York, U.S. 318 pp.
Edenius, L. & Mikusinski, G. 2006. Utility of habitat suitability models as biodiversity assessment tools in forest management. *Scandinavian Journal of Forest Research 21*, 62-72.
Efron, B. & Tibshirani, R. 1993. *An introduction to the bootstrap*. Chapman & Hall. New York, U.S. 436 pp.
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. & Næsset, E. 2004. Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analyses. *Scandinavian Journal of Forest Research 19*, 512-523.
Eid, T. & Hobbelstad, K. 2000. AVVIRK-2000: A large-scale forestry scenario model for long-term investment, income and harvest analyses. *Scandinavian Journal of Forest Research 15*, 472-482.
Eid, T. & Næsset, E. 1998. Determination of stand volume in practical forest inventories based on field measurements and photo-interpretation: The Norwegian experience. *Scandinavian Journal of Forest Research 13*, 246-254.
Eid, T. & Tuhus, E. 2001. Models for individual tree mortality in Norway. *Forest Ecology and Management 154*, 69-84.
Ekelund, H. & Hamilton, G. 2001. Skogspolitisk historia. *Swedish Forest Agency, Report 8A*. 1-251. ISSN 1100-0295.
Fleming, K.K., Didier, K.A., Miranda, B.R. & Porter, W.F. 2004. Sensitivity of a white-tailed deer habitat-suitability index model to error in satellite land-cover data: Implications for wildlife habitat-suitability studies. *Wildlife Society Bulletin 32*, 158-168.
Foody, G.M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment 80*, 185-201.
Fransson, J.E.S. & Israelsson, H. 1999. Estimation of stem volume in boreal forests using ERS-1 C- and JERS-1 L-band SAR data. *International Journal of Remote Sensing 20*, 123-137.
Fransson, J.E.S., Walter, F., Blennow, K., Gustavsson, A. & Ulander, L.M.H. 2002. Detection of storm-damaged forested areas using airborne CARABAS-II VHFSAR image data. *IEEE Transactions on Geoscience and Remote Sensing 40*, 2170-2175.
Fransson, J.E.S., Walter, F. & Ulander, L.M.H. 2000. Estimation of forest parameters using CARABAS-II VHFSAR data. *IEEE Transactions on Geoscience and Remote Sensing 38*, 720-727.

Fridman, J. & Ståhl, G. 2001. A three-step approach for modelling tree mortality in Swedish forests. *Scandinavian Journal of Forest Research 14*, 455-466.
von Gadow, K. & Puumalainen, J. 2000. Scenario planning for sustainable forest management. In: *Sustainable Forest Management*. (Eds. K. von Gadow, T. Pukkala & M. Tomé). Kluwer Academic Publishers. Dordrecht, Netherlands.
Gobakken, T., Lexerød, N. & Eid, T. 2004. A growth simulator based on models for individual trees. In: *Forest Inventory and Planning in Nordic Countries, SNS Meeting at Sjusjøen, Norway*, (Ed. K. Hobbelstad). Norweigan Institute of Land Inventory.
Guisan, A. & Zimmermann, N.E. 2000. Predictive habitat distrubition models in ecology. *Ecological Modelling 135*, 147-186.
Gustafson, E.J. 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems 1*, 143-156.
Gustafsson, K. 2000. Skogliga konsekvensanalyser 1999. *Swedish Forest Agency, Report 2*. 1-331. ISSN 1100-0295.
Gustafsson, K. & Hägg, S. 2004. Skogliga konsekvensanalyser 2003. *Swedish Forest Agency, Report 2*. 1-38. ISSN 1100-0295.
Hagner, O. 1990. Computer aided forest stand delineation and inventory based on satellite remote sensing. In: *SNS/IUFRO Workshop in Umeå, 26-28 February 1990: The usability of remote sensing for forest inventory and planning*. Swedish University of Agriculture Sciences.
Hamilton, D.A. 1978. Specifying precision i natural resource inventories. In: Integrated inventories of renewable resources. *USDA Forest Service, General Technical Report RM-55*. 276-281. ISSN 0277-5786.
Hamilton, D.A. 1979. Setting precision for resource inventories: The manager and the mensurationist. *Journal of Forestry 77*, 667-670.
Hauska, H. 1999. Datainsamling. In: *Geografisk informationsbehandling: Metoder och tillämpningar*. (Eds. L. Eklundh & W. Arnberg). Byggforskningsrådet. Stockholm, Sweden.
Heinonen, T., Kurttila, M. & Pukkala, T. 2007. Possibilities to aggregate raster cells through spatial optimization in forest planning. *Silva Fennica 41*, 89-103.
Hirzel, A.H., Helfer, V. & Metral, F. 2001. Assessing habitat-suitability models with a virtual species. *Ecological Modelling 145*, 111-121.
Hoen, H.F., Eid, T. & Økseter, P. 2001. Timber production possibilities and capital yields from the Norwegian forest area. *Silva Fennica 35*, 249-264.
Hof, J.G. & Joyce, L.A. 1993. A mixed-integer linear-programming approach for spatially optimizing wildlife and timber in managed forest ecosystems. *Forest Science 39*, 816-834.
Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scandinavian Journal of Forest Research 19*, 543-553.
Holmgren, J. & Persson, Å. 2004. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment 90*, 415–423.
Holmgren, L. 2002. Internationella konventioner, avtal och processer med koppling till skog. *Swedish Forest Agency, Report 2*. 1-33. ISSN 1100-0295.
Holmgren, P. & Thuresson, T. 1997. Applying objectively estimated and spatially continuous forest parameters in tactical planning to obtain dynamic treatment units. *Forest Science 42*, 317-326.
Holmström, H. 2001. Data aquisition for forestry planning by remote sensing based sample plot imputation. *Acta Universitatis agriculturae Sueciae. Silvestria 201*, 1-41.
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.
Holmström, H., Nilsson, M. & Ståhl, G. 2001. Simultaneous estimations of forest parameters using aerial photograph interpreted data and the k nearest neighbour method. *Scandinavian Journal of Forest Research 16*, 67-78.
Holmström, H., Nilsson, M. & Ståhl, G. 2002. Forecasted reference sample plot data in estimations of stem volume using satellite spectral data and the kNN method. *International Journal of Remote Sensing 23*, 1757-1774.

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.
Huang, S. & Titus, S.J. 1999. An individual tree height increment model for mixed white spruce-aspen stands in Alberta, Canada. *Forest Ecology and Management 123*, 41-53.
Hyyppä, J., Kelle, O., Lehikoinen, M. & Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing 39*, 969-975.
Hyyppä, J., Pulliainen, J., Hallikainen, M. & Saatsi, A. 1997. Radar-derived standwise forest inventory. *IEEE Transactions on Geoscience and Remote Sensing 35*, 392-404.
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.
Johnson, K.N., Duncan, S. & Spies, T.A. 2007. Regional policy models for forest biodiversity analysis: Lessons from Coastal Oregon. *Ecological Applications 17*, 81-90.
Johnson, K.N., Stuart, T.W. & Crim, S.A. 1986. *FORPLAN version 2: An overview*. USDA Forest Service Land Management Planning System Section. Washington DC, U.S.
Jones, J.G., Meneghin, B.J. & Kirby, M.W. 1991. Formulating adjacency constraints in linear optimization models for scheduling projects in tactical planning. *Forest Science 37*, 1283-1297.
Jonsson, B., Jacobsson, J. & Kallur, H. 1993. The Forest Management Planning Package: Theory and application. *Studia Forestalia Suecica 189*, 1-56.
Kallio, M., Moiseyev, A. & Solberg, B. 2004. The global forest sector model EFI-GTM — The model structure. *European Forest Institute, Technical Report 15.* 1-24.
Kangas, A., Heikkinen, E. & Maltamo, M. 2004. Accuracy of partially visually assessed stand characteristics: a case study of Finnish forest inventory by compartments. *Canadian Journal of Forest Research 34*, 916-930.
Kangas, A. & Kangas, J. 2004. Yield tables, forecasting, modelling, and simulation. In: *Encyclopedia of Forest Sciences*. (Eds. J. Burley, J. Evans & J.A. Youngquist). Elsevier Ltd. Amsterdam, Netherlands. pp. 2061.
Kangas, A. & Maltamo, M. 2000. Calibrating predicted diameter distribution with additional information. *Forest Science 46*, 390-396.
Kangas, A.S. 1997. On the prediction bias and variance in long-term growth projections. *Forest Ecology and Management 96*, 207-216.
Kangas, A.S. & Kangas, J. 1999. Optimization bias in forest management planning solutions due to errors in forest variables. *Silva Fennica 33*, 303-315.
Karppinen, H. 1998a. Objectives of non-industrial private forest owners: Differences and future trends in southern and northern Finland. *Journal of Forest Economics 4*, 147-174.
Karppinen, H. 1998b. Values and objectives of non-industrial private forest owners in Finland. *Silva Fennica 32*, 43-59.
Karppinen, H. 2005. Forest owners choice of reforestation method: an application of the theory of planned behavior. *Forest Policy and Economics 7*, 393-409.
Kayitakire, F., Hamel, C. & Defourny, P. 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment 102*, 390-401.
Kleinn, C. 2000. Estimating metrics of forest spatial pattern from large area forest inventory cluster samples. *Forest Science 46*, 548-557.
Kleinn, C. & Vilčko, F. 2006. A new empirical approach for estimation in k-tree sampling. *Forest Ecology and Management 237*, 522-533.
Kolström, M. 1998. Ecological simulation model for studying diversity of stand structure in boreal forests. *Ecological Modelling 111*, 17-36.
Kruys, N., Fries, C., Jonsson, B.G., Lamas, T. & Stal, G. 1999. Wood-inhabiting cryptogams on dead Norway spruce (Picea abies) trees in managed Swedish boreal forests. *Canadian Journal of Forest Research 29*, 178-186.
Kruys, N., Jonsson, B.G. & Ståhl, G. 2002. A stage-based matrix model for decay-class dynamics of woody debris. *Ecological Applications 12*, 773-781.
Lähde, E., Laiho, O., Norokorpi, Y. & Saksa, T. 1999. Stand structures as the basis of a diversity index. *Forest Ecology and Management 115*, 213-220.

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.
Langford, W.T., Gergel, S.E., Dietterich, T.G. & Cohen, W. 2006. Map misclassification can cause large errors in landscape pattern indices: Examples from habitat fragmentation. *Ecosystems 9*, 474-488.
Larson, M.A., Thompson, F.R., Millspaugh, J.J., Dijak, W.D. & Shifley, S.R. 2004. Linking population viability, habitat suitability, and landscape simulation models for conservation planning. *Ecological Modelling 180*, 103-118.
Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A. & Harding, D. 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment 70*, 339-361.
LeMay, V. & Temesgen, H. 2005. Comparison of nearest neighbor methods for estimating basal area and stems per hectare using aerial auxiliary variables. *Forest Science 51*, 109-119.
Lexerød, N. 2005. Recruiment models for different species in Norway. *Forest Ecology and Management 206*, 91-108.
Lillesand, T.M. & Kiefer, R.W. 2000. *Remote sensing and image interpretation*. Fourth edition edition. John Wiley & Sons, Inc. New York, U.S. 724 pp.
Lind, T. 2000. Strategic forestry planning - evaluation of variable spatial aggregations and forest landscapes. *Acta Universitatis agriculturae Sueciae. Silvestria 149*, 1-28.
Lindhagen, A. & Hörnsten, L. 2000. Forest recreation in 1977 and 1997 in Sweden: changes in public preferences and behaviour. *Forestry 73*, 143-153.
Lockwood, C. & Moore, T. 1993. Harvest scheduling with spatial constraints - A Simulated Annealing approach. *Canadian Journal of Forest Research 23*, 468-478.
Lohmander, P. 2000. Optimal sequential forestry decisions under risk. *Annals of Operations Research 95*, 217-228.
Lönnstedt, L. 1986. A dynamic simulation model for the forest sector with an illustration for Sweden. In: *System analysis in forestry and forest industry*. (Eds. M. Kallio, Å. Andersson, R. Seppälä & A. Morgan). Elsevier Science Publishers B.V. Amsterdam, The Netherlands.
Lönnstedt, L. 1998. Calculating non-industrial private forest owners' cuttings. *Scandinavian Journal of Forest Research 13*, 215-223.
Lu, D. & Weng, Q. 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing 28*, 823-870.
Lu, F. & Eriksson, L.O. 2000. Formation of harvest units with genetic algorithms. *Forest Ecology and Management 130*, 57-67.
Lundström, A. & Söderberg, U. 1996. Outline of the HUGIN system for long-term forecasts of timber yields and possible cut. In: *Large-scale forestry scenario models: Experiences and requirements*, Eds. T. Päivinen, L. Roihuvuo & M. Siitonen). EFI, Joensuu, Finland, 15-22 June, 1995, pp. 318.
Magnusson, M. 2006. Evaluation of remote sensing techniques for estimation of forest variables at stand level. *Acta Universitatis agriculturae Sueciae 85*, 1-38.
Magnusson, M. & Fransson, J.E.S. 2005. Estimation of forest stem volume using multispectral optical satellite and tree height data in combination. *Scandinavian Journal of Forest Research 20*, 431-440.
Maltamo, M., Malinen, J., Packaln, P., Suvanto, A. & Kangas, J. 2006. Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand-register data. *Canadian Journal of Forest Research 36*, 426-436.
Martell, D.L., Gunn, E.A. & Weintraub, A. 1998. Forest management challenges for operational researchers. *European Journal of Operational Research 104*, 1-17.
Mattsson, L. & Li, C. 1993. The non-timber value of northern Swedish forests: an economic analysis. *Scandinavian Journal of Forest Research 8*, 426-434.
McRoberts, R.E., Tomppo, E., Finley, A.O. & Heikkinen, J. 2007. Estimating areal means and variances of forest attributes using the k-Nearest technique and satellite imagery. *Remote Sensing of Environment (In Press).*

Mehtätalo, L. & Kangas, A. 2005. An approach to optimizing field data collection in an inventory by compartments. *Canadian Journal of Forest Research 35*, 100-112.

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.

Moberg, L. 2006. Predicting knot properties of Picea abies and Pinus sylvestris from generic tree descriptors. *Scandinavian Journal of Forest Research 21*, 48-61.

Moeur, M. & Stage, A.R. 1995. Most Similar Neighbor - An improved sampling inference procedure for natural-resource planning. *Forest Science 41*, 337-359.

Monserud, R.A. & Sterba, H. 1996. A basal area increment model for individual trees growing in even- and uneven-aged stands in Austria. *Forest Ecology and Management 80*, 57-80.

Murray, A.T. & Church, R.L. 1995. Heuristic solution approaches to operational forest planning problems. *OR Spektrum 17*, 193-203.

Næsset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing 52*, 49-56.

Næsset, E. 2004. Accuracy of forest inventory using airborne laser scanning: Evaluating the first Nordic full-scale operational project. *Scandinavian Journal of Forest Research 19*, 554-557.

Næsset, E., Gobakken, T. & Hoen, H.F. 1997. Economic analysis of timber management practices promoting preservation of biological diversity. *Scandinavian Journal of Forest Research 12*, 264-272.

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

Nelson, R., Valenti, M.A., Short, A. & Keller, C. 2003. A multiple resource inventory of Delaware using airborne laser data. *Bioscience 53*, 981-992.

Nilsson, M. 1996. Estimation of tree weights and stand volume using an airborne lidar system. *Remote Sensing of Environment 56*, 1-7.

Nilsson, M. 1997. Estimation of forest variables using satellite image data and airborne Lidar. *Acta Universitatis Agriculturae Sueciae, Silvestria 17*, 1-31.

Nyström, K. & Kexi, M. 1997. Individual tree basal area growth models for young stands of Norway spruce in Sweden. *Forest Ecology and Management 97*, 173-185.

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

Öhman, K. & Eriksson, L.A. 2002. Allowing for spatial consideration in long-term forest planning by linking linear programming with simulated annealing. *Forest Ecology and Management 161*, 221-230.

Öhman, K. & Eriksson, L.O. 1998. The core area concept in forming contiguous areas for long-term forest planning. *Canadian Journal of Forest Research 28*, 1032-1039.

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

Ohmann, J.L. & Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. *Canadian Journal of Forest Research 32*, 725-741.

Olofsson, K., Wallerman, J., Holmgren, J. & Olsson, H. 2006. Tree species discrimination using Z/I DMC imagery and template matching of single trees. *Scandinavian Journal of Forest Research 21*, 106-110.

Pal, N.R. & Pal, S.K. 1993. A review on image segmentation techniques. *Pattern Recognition 26*, 1277-1294.

Pekkarinen, A. 2002. A method for the segmentation of very high spatial resolution images of forested landscapes. *International Journal of Remote Sensing 23*, 2817-2836.

Peng, C.H. 2000. Growth and yield models for uneven-aged stands: past, present and future. *Forest Ecology and Management 132*, 259–279.

Pesonen, M. 1995. Non-industrial private forest landowners' choices of timber management strategies and potential allowable cut: case of Pohjois-Savo. *Acta Forestalia Fennica 247*, 1-31.

Petersson, H. & Ståhl, G. 2006. Functions for below-ground biomass of Pinus sylvestris, Picea abies, Betula pendula and Betula publescens in Sweden. *Scandinavian Journal of Forest Research 21*, 84-93.

Pettersson, F. & Högbom, L. 2004. Long-term growth effects following forest nitrogen fertilization in *Pinus sylvestris* and *Picea abies* stands in Sweden. *Scandinavian Journal of Forest Research 19*, 339-347.

Pitkanen, J. 2001. Individual tree detection in digital aerial images by combining locally adaptive binarization and local maxima methods. *Canadian Journal of Forest Research 31*, 832-844.

Porté, A. & Bartelink, H.H. 2002. Modelling mixed forest growth: a review of models for forest management. *Ecological Modelling 150*, 141-188.

Proceviat, S.K., Mallory, F.F. & Rettie, W.J. 2003. Estimation of arboreal lichen biomass availible to woodland caribou in Hudson Bay lowland black spruce sites. *Rangifer Special Issue No. 14*, 95-99.

Pukkala, T. 1999. Monsu-metsäsuunnitteluohjelma. Ohjelmiston toiminta ja käyttö (User's guide of the Monsu forest planning program). *University of Joensuu, Finland, Report.* 1-72.

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

Pukkala, T., Ketonen, T. & Pykäläinen, J. 2003. Predicting timber harvests from private forests – a utility maximisation approach. *Forest Policy and Economics 5*, 285-296.

Pukkala, T. & Kurttila, M. 2005. Examining the performance of six heuristic optimisation techniques in different forest planning problems. *Silva Fennica 39*, 67-80.

Pukkala, T., Möykkynen, T., Thor, M., Rönnberg, J. & Stenlid, J. 2005. Modeling infection and spread of *Heterobasidion annsosum* in even-aged Fennoscandian coniferous stands. *Canadian Journal of Forest Research 35*, 74-84.

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

Pussinen, A., Schelhaas, M.J., Verkaik, E., Heikkinen, E., Liski, J., Karjalainen, T., Päivinen, R. & Nabuurs, G.J. 2001. Manual for the European Forest Information Scenario Model (EFISCEN 2.0). *European Forest Institute, Internal Report 5.* 1-49.

Ranneby, B., Cruse, T., Hägglund, B., Jonasson, H. & Swärd, J. 1987. Designing a new national forest survey for Sweden. *Studia Forestalia Suecica 177*, 1-29.

Raulier, F., Pothier, D. & Bernier, P.Y. 2003. Predicting the effect of thinning on growth of dense balsam fir stands using a process-based tree growth model. *Canadian Journal of Forest Research 33*, 509-520.

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.

Riitters, K.H., Oneill, R.V., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timmins, S.P., Jones, K.B. & Jackson, B.L. 1995. A factor-analysis of landscape pattern and structure metrics. *Landscape Ecology 10*, 23-39.

Ringvall, A., Patil, G.P. & Taillie, C. 2000. A field test of surveyors' influence on estimates in line transect sampling. *Forest Ecology and Management 137*, 103-111.

Ringvall, A. & Ståhl, G. 1999. Field aspects of line intersect sampling for assessing coarse woody debris. *Forest Ecology and Management 119*, 163-170.

Rolff, C. & Ågren, G.I. 1999. Predicting effects of different harvesting intensities with a model of nitrogen limited forest growth. *Ecological Modelling 118*, 193-211.

Saaty, T.L. 1985. *Analytical planning. The organization of systems*. Pergamon Press. Oxford, U.K. 208 pp.

Sandström, P., Pahlen, T.G., Edenius, L., Tommervik, H., Hagner, O., Hemberg, L., Olsson, H., Baer, K., Stenlund, T., Brandt, L.G. & Egberth, M. 2003. Conflict resolution by participatory management: Remote sensing and GIS as tools for communicating land-use needs for reindeer herding in northern Sweden. *AMBIO 32*, 557-567.

Scheller, R.M., Domingo, J.B., Sturtevant, B.R., Williams, J.S., Rudy, A., Gustafson, E.J. & Mladenoff, D.J. 2007. Design, development, and application of LANDIS-II, a spatial

landscape simulation model with flexible temporal and spatial resolution. *Ecological Modelling 201*, 409-419.
Schreuder, H.T., Ernst, R. & Ramirez-Maldonado, H. 2004. Statistical techniques for sampling and monitoring natural resources. *Department of Agriculture, Forest Service, Rocky Mountain Research Station, General Technical Report RMRS-GTR-126.* 1-111.
Shifley, S.R., Thompson, F.R., Dijak, W.D., Larson, M.A. & Millspaugh, J.J. 2006. Simulated effects of forest management alternatives on landscape structure and habitat suitability in the Midwestern United States. *Forest Ecology and Management 229*, 361-377.
Siitonen, M. 1995. The MELA system as a forestry modelling framework. *Lesnictvi-Forestry 41*, 173-178.
SIMO – SIMulation and Optimisation for next generation forest planning system. University of Helsinki, Finland. http://www.mm.helsinki.fi/MMVAR/SIMO/ (assessed 16 Aug 2007).
Söderberg, U. 1986. Functions for forecasting timber yield: increment and form height of individual trees of native tree species in Sweden. *Swedish University of Agriculture Sciences, Deparment of Biometry and Forest Management, Report 14.* 1-251. ISSN 0349-2133.
Ståhl, G. 1992. A study on the quality of compartmentwise forest data acquired by subjective inventory methods. *Swedish University of Agriculture Sciences, Deparment of Biometry and Forest Management, Rapport 24.* 1-128. ISSN 0349-2133.
Ståhl, G. 1994. 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-acquisition policies. *Forest Science 40*, 630-649.
Talkkari, A., Peltola, H., Kellomaki, S. & Strandman, H. 2000. Integration of component models from the tree, stand and regional levels to assess the risk of wind damage at forest margins. *Forest Ecology and Management 135*, 303-313.
Tamhane, A.C. & Dunlop, D.D. 2000. *Statistics and data analysis from elementary to intermediate*. Prentice-Hall, Inc. Upper Saddle River, New Jersey. 722 pp.
Temesgen, H., LeMay, V.M., Froese, K.L. & Marshall, P.L. 2003. Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *Forest Ecology and Management 177*, 277-285.
Thies, M., Pfeifer, N., Winterhalder, D. & Gorte, B.G.H. 2004. Three-dimensional reconstruction of stems for assessment of taper, sweep and lean based on laser scanning of standing trees. *Scandinavian Journal of Forest Research 19*, 571-581.
Thompson, S.K. 2002. *Sampling*. John Wiley & Sons, Inc. New York, U.S.
Thor, M., Ståhl, G. & Stenlid, J. 2005. Modelling root rot incidence in Sweden using tree, site and stand variables. *Scandinavian Journal of Forest Research 20*, 165-176.
Tittler, R., Messier, C. & Burton, P.J. 2001. Hierarchical forest management planning and sustainable forest management in the boreal forest. *Forestry Chronicle 77*, 998-1005.
Tokola, T., Pitkanen, J., Partinen, S. & Muinonen, E. 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. *International Journal of Remote Sensing 17*, 2333-2351.
Tomppo, E. 1990. Designing a satellite image-aided national forest survey in Finland. In: *SNS/IUFRO Workshop in Umeå, 26-28 February 1990: The usability of remote sensing for forest inventory and planning*. Swedish University of Agricultural Sciences.
Tomppo, E. & Halme, M. 2004. Using coarse scale forest variables as ancillary information and weighting of variables in k-N-N estimation: A genetic algorithm approach. *Remote Sensing of Environment 92*, 1-20.
Tomppo, E., Nilsson, M., Rosengren, M., Aalto, P. & Kennedy, P. 2002. Simultaneous use of Landsat-TM and IRS-1C WiFS data in estimating large area tree stem volume and aboveground biomass. *Remote Sensing of Environment 82*, 156-171.
Valinger, E. & Fridman, J. 1997. Modelling probability of snow and wind damage in Scots pine stands using tree characteristics. *Forest Ecology and Management 97*, 215-222.

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

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.

Watt, P.J. & Donoghue, D.N.M. 2005. Measuring forest structure with terrestrial laser scanning. *International Journal of Remote Sensing 26*, 1437-1446.

Wehr, A. & Lohr, U. 1999. Airborne laser scanning – an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 68-82.

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

Wickham, J.D., Oneill, R.V., Riitters, K.H., Wade, T.G. & Jones, K.B. 1997. Sensitivity of selected landscape pattern metrics to land-cover misclassification and differences in land-cover composition. *Photogrammetric Engineering and Remote Sensing 63*, 397-402.

Wikström, P. & Eriksson, L.-O. 2000. Solving the stand management problem under biodiversity-related considerations. *Forest Ecology and Management 126*, 361-376.

Wilhelmsson, L. 2006. Two models for predicting the number of annual rings in cross-sections of tree stems. *Scandinavian Journal of Forest Research 21*, 37-47.

Wilhelmsson, L., Arlinger, J., Spangberg, K., Lundqvist, S.O., Grahn, T., Hedenberg, O. & Olsson, L. 2002. Models for predicting wood properties in stems of *Picea abies* and *Pinus sylvestris* in Sweden. *Scandinavian Journal of Forest Research 17*, 330-350.

Zeng, H.C., Pukkala, T. & Peltola, H. 2007. The use of heuristic optimization in risk management of wind damage in forest planning. *Forest Ecology and Management 241*, 189-199.

Zheng, D., Freeman, M., Bergh, J., Røsberg, I. & Nilsen, P. 2002. Production of *Picea abies* in south-east Norway in response to climate change: A case study using process-based model simulation with field validation. *Scandinavian Journal of Forest Research 17(1)*, 35 - 46.

Acknowledgement

This thesis is a result of four and a half years of studies. In my work I have received support and inspiration from a large number of people. When thinking about which people I would like to thank, there are so many moments that come up that it would be impossible to mention all of you by name. However, a couple of persons have been deeply involved and must be mentioned. First of all I thank my supervisor Göran Ståhl for giving me this opportunity and who never failed in being a good supervisor. There has always been time for discussions and instant comments during this work, which has inspired and allowed further improvements. This has, not least, been valuable during the intense work in the finalising of the thesis. I also thank my co-supervisors, who both by their profession and personality have demonstrated their abilities in a good manner: Hans Petersson for inspiring me to do research and for assistance in issues such as sampling, modelling, and international conventions; Torgny Lind for good support in decision-making and decision support systems; and finally Jörgen Wallerman for valuable contributions and knowledge in spatial statistics and remote sensing. Only a few students are privileged with such good supervision! My two colleagues: Karl Duvemo has contributed to parts of this thesis and has also been a good partner for discussions during my studies and Heather Reese who has corrected the English in the thesis.

The subject of the thesis has included a variety of disciplines within forest management and this thesis would not have been possible without all the aggregated knowledge and helpful people at the Department of Forest Management at SLU. Colleagues with knowledge in forest planning, remote sensing, optimisation, programming, statistics, and the NFI have generously supported me whenever needed. Furthermore, I appreciate the good cooperative atmosphere and service at the department. I have enjoyed my stay from the first day and will remember the everyday coffee breaks, and other social activities and joyful moments. A special thought goes to all the other PhD-students I meet during the years, both Swedish and International, both at the Department and at SLU, but also at partner universities through meetings in Finland, Norway, Germany, Czech Republic, and Italy. These have all been joyful and inspiring moments!

Finally, I thank my family and friends. Ines, for believing in me and giving so much inspiration and energy! I look forward to our future adventures in kayaking, skiing, and farming together with our growing family!

Sammanfattning på svenska

Den här avhandlingen behandlar frågor som berör datafångst för planering av skog över stora områden, främst på regional och nationell nivå. Beslut inom skogsbruket har under de senaste decennierna blivit mer komplexa, då såväl ekonomiska, ekologisk och sociala värden väger tungt när samhällets sätter upp mål för att nyttja skogens resurser. Som en följd av detta utvecklas nya, mer avancerade, prognos- och planeringssystem som kan användas för att samlat planera nyttjandet av skogens resurser. Att kombinera analyser av många olika resurser ställer dock höga krav på skogliga data. Som komplement till de data, insamlat i nationella fältinventeringar, som tidigare var tillräcklig för att göra skogliga scenarioanalyser på nationell nivå, krävs i många fall även data på landskapets sammansättning. Dessa rumsligt heltäckande data är nödvändig för att kunna inkludera vissa resurser i en scenarioanalys, som till exempel vissa indikatorer för biodiversitet. Rumsligt heltäckande data möjliggör också mer detaljerade prognoser. Risken för stormskador eller effekter av olika skötselstrategier kopplat till markägare är två exempel på hur analyser kan bli mer detaljerade. För att försörja nationella prognos- och planeringssystem med skogliga data, behövs en kombination av data från flera källor. Heltäckande beskrivningar av landskapet kan erhållas med fjärranalys, exempelvis satellitbilder eller flygburna laser skannrar. Mer detaljerade data kan mätas in i provytor genom nationella fältinventeringar. Flera datakällor kan sedan kombineras genom imputering, och heltäckande data med trädlistor som beskriver skogstillståndet för varje enskilt bestånd i landskapet kan skapas.

Syftet med den här avhandlingen är att utvärdera olika strategier för datafångst och att testa ett antal metoder för att utvärdera dessa strategier. Ett stort fokus ligger också på att förbättra redan gällande metoder för imputering, så att data blir bättre anpassat för skogliga prognos- och planeringssystem. I avhandlingen ingår totalt sex studier. I studie 1 beskrivs vilka komponenter som igår i ett prognos- och planeringssystem, och vilka krav detta har på dataförsörjningen. I kombination med egenskaper hos olika datafångstmetoder kan ett resonemang kring olika datakällors lämplighet avgöras. En slutsats i denna studie är att ingen av våra tre nationella datakällor ensamt kan förse en skoglig scenarioanalys på nationell nivå med skogliga data. I studie 2 och 3 utvärderas två imputeringar av provytor som indata i ett prognos- och planeringssystem där bärardata i de båda fallen baseras på flygburen laser och satellitbilder. I studie 2 studeras utfallet i olika planeringsperioder med syfte att studera effekterna av de olika datakällornas påverkan på beslut som fattas på regional- och nationell nivå. I studie 3 används cost-plus-loss med simulering som utvärderingsmetod. I båda studierna blir slutsatserna att metoderna för imputering bör förbättras för att data skall uppnå en kvalité lämplig för skogliga scenarioanalyser. I studie 4 används en analytisk cost-plus-loss för att utvärdera behovet av nationella data för att bestämma en uthållig avverkningsnivå. Analysen baseras på ett scenario för att beräkna kostnadsförlusten då beslut fattas på felaktiga data. Slutsatsen blir att dagens nivå på den svenska Riksskogstaxeringen är i linje med det beräknade behovet. Beräkningarna är dock förenklade och bortser ifrån att data används som underlag till en rad andra beslut och rapporteringar. I de två avslutade studierna föreslås nya metoder att skapa rumsligt heltäckande data som kan användas i prognos- och planeringssystem. I studie 5 föreslås en utvecklad metod av imputering som tar hänsyn till variation och spatial konfiguration inom ett bestånd. Enklare test visar på lovade resultat. I studie 6 presenteras ett ramverk för att skapa realistiska beskrivningar av ett helt landskap. Ett vanligt problem är ofta att arealen av den yngre och äldre skogen underskattas. För att komma till rätta med det föreslås en modifierad metod av imputering där landskapets komposition först uppskattas baserat på fältinventeringar. Optimering används sedan för att flytta omkring imputerat data för att på så sätt skapa realistiska bestånd. Även här visar enklare tester lovande resultat som kan användas för datainsamling på nationell och regional nivå. Vidare utveckling är dock nödvändig för att kunna använda metoden i praktiska tillämpningar.