Data Acquisition for Forestry Planning
by Remote Sensing Based Sample Plot Imputation

Hampus Holmström

Department of Forest Resource Management and Geomatics

Acta Universitatis Agriculturae Sueciae
Silvestria 201

Abstract

In forestry planning, accurate description of the state of the forests is essential. Advanced planning models often require input data with high resolution, *i.e.* data at the single-tree level. Field inventory procedures based on sample plot measurements are usually employed. However, such methods are expensive, so cost-efficient alternatives would be attractive.

In the work described in this thesis, inventory methods based on imputation of reference sample plot data were evaluated. The reference material consisted of data from previously performed field inventories. The *k* nearest neighbour (*k*NN) method was used, in which all variables available at the reference plots were simultaneously estimated for the target areas. The imputations were based on information derived from interpretations of aerial photographs, optical satellite data, radar data (airborne sensor), and existing stand records. To account for differences in the qualities of the different information sources combined in the *k*NN estimations, distance metrics were defined and applied using regression functions.

The utility of various types of forecasted reference material was evaluated. Increasing the length of forecasts of reference sample plot data increased the mean square error (*MSE*) in stem volume estimates. However, by excluding disturbed plots (due to thinnings) from the reference material, plot data forecasted for up to 25 years could be used without severely decreasing the accuracy of the estimations.

Using aerial photo interpretations together with stand records, *k*NN estimates of stem volume with relative root *MSE*s (*RMSE*s) of 14-20% at the stand level were obtained. More accurate estimates were obtained for a northern test site, in comparison with results from southern Sweden. Combining optical satellite data and radar data significantly improved results, giving a *RMSE* in standwise stem volume estimates of 22%, compared to 30% for the best single-sensor case.

Consequences of using *k*NN estimations in a planning context were evaluated by a cost-plus-loss approach. The total cost of undertaking an inventory was obtained by summing the actual inventory cost and the net present value of expected future losses due to non-optimal decisions caused by erroneous data. Input data obtained by imputation of reference sample plots were compared with traditional field sample plot inventories. Results indicated that the total cost of an inventory could be reduced by 15-50% by integrating different methods; imputation could be applied for some types of stands while more accurate field inventories should be carried out in others. It is not necessarily the most valuable stands that should be inventoried by careful field measurements, but many of those with a treatment impending.

Keywords: aerial photography, cost-plus-loss analysis, forest assessment, individual tree data, multisource data, nonparametric estimation.

Author's address: Hampus Holmström, Department of Forest Resource Management and Geomatics, SLU, S-901 83 Umeå, Sweden, *e-mail*: hampus.holmstrom@resgeom.slu.se

Contents

Appendix

Papers I-V

This thesis is based on the following papers, which are referred to by the corresponding Roman numerals in the text:

I.   Holmström, H., Nilsson, M., and Ståhl, G. 2001. Forecasted reference sample plot data in estimations of stem volume using satellite spectral data and the kNN method. *International Journal of Remote Sensing* (in press).

II.   Holmström, H., Nilsson, M., and 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.

III.   Holmström, H. and Fransson, J.E.S. 2001. Combining remotely sensed optical and radar data in kNN estimation of forest variables. *Forest Science* (in press).

IV.   Holmström, H. 2001. Estimation of single-tree characteristics using the kNN method and plotwise aerial photograph interpretations. *Forest Ecology and Management* (in press).

V.   Holmström, H., Kallur, H., and Ståhl, G. 2001. Cost-plus-loss analyses of forest inventory strategies based on kNN-assigned reference sample plot data. (Manuscript).

Papers I-IV are reproduced with the kind permission of the publishers: Taylor and Francis (Papers I and II), the Society of American Foresters (Paper III), and Elsevier Science (Paper IV).

Introduction

Forests produce goods that are useful for many purposes. In Sweden, the focus today is on wood products, especially timber, pulpwood, and fuelwood (Anon. 2001a). Further valued aspects of forests, such as their wildlife, recreational, aesthetic, and existence amenities, have been given greater prominence in recent years (*e.g.*, Franklin 1989, Angelstam 1992, Rülcker *et al.* 1994, Lämås and Fries 1995, Kohm and Franklin 1996, Törnquist 1996, Fries *et al.* 1997). Wood products are now extracted in a manner that pays consideration to these aspects, as well as to general environmental concerns. Preservation of biological diversity is regulated by international agreements (Anon. 1992; 1993) and is, in Sweden, incorporated in legislation governing the use of natural resources (*e.g.*, Skogsstyrelsen 1994). It is, however, a fair assumption that forestry, in the context of promoting the growth and deriving products of commercial value from the trees, will continue to be a very important activity.

Sustained utilization of the forests requires planning. Inevitably, man will be exposed to the consequences of actions taken in the past. The proportions of forest resources that can be utilized today and in the future need careful consideration. Thus, a planning problem occurs (*e.g.*, Aplet *et al.* 1993).

The starting point in a planning process is the current state. Any major errors present at this stage can ruin the entire idea of planning, making proposals for future activities inadequate or even impossible to follow. No matter how accurately future developments are modelled and suitable treatment schemes are selected, an erroneous description of the initial state of the forest will lead to incorrect conclusions (*e.g.*, Bell 1982; 1985, Pickens *et al.* 1991, Weintraub and Abramovich 1995). On the other hand, too detailed information can seldom be used in planning models, which are by definition simplifications of the real world.

Forest inventories should deliver information that is accurate enough to enable sufficiently adequate decisions to be made (*e.g.*, Cochran 1977, Burkhart *et al.* 1978, Hamilton 1978, Jacobsson 1986, Ståhl 1994). A plethora of forest inventory methods have been developed, and suggested to be cost-efficient for certain planning situations (*e.g.*, Loetsch and Haller 1973, Schreuder *et al.* 1993, Avery and Burkhart 1994). A challenge for research in forest inventory is to determine which methods are best suited for a certain purpose, especially since new techniques and information sources are continuously being proposed.

Forest management planning

Planning can be done in numerous ways, involving anything from pure intuition to complex scenario analyses and optimisation. Theoretically, a distinction

between formal and incremental planning can be made (*e.g.*, Saaty 1985). In formal planning, mathematical models are assumed to depict aspects of reality that influence the activity to be planned. These models, which collectively comprise the planning system, will then define the necessary input data. From all the possible treatment options generated by the system, the one that comes closest to maximising the stated goals is chosen and implemented. In contrast, instead of relying on models, incremental planning describes attempts to achieve set goals by decisions based on experience and intuition. In these cases, information is added in an informal and arbitrary manner, relying on past and present observations.

In forestry, both of the two outlined approaches to planning have advantages and drawbacks. Thus, a combination of formal and incremental planning is usually desirable. Forest planning systems should be regarded, and used, as decision support systems (*e.g.*, Kilkki 1985, Jonsson *et al.* 1993, Lämås 1996). In this respect, use of models (defining a set of input data necessary) is balanced with intuition and logical reasoning, adding additional input data to the planning process in a continuous and flexible manner.

Another way to approach the complexity of forestry planning is to separate the process into hierarchical levels: strategic, tactical, and operational (*e.g.*, Weintraub and Cholaky 1991, Davis and Martell 1993, Davis *et al.* 2001). At different levels, different objectives can then be treated. Strategic planning extends over long time-periods, from decades to infinity, and usually deals with sustainable harvest levels and timber flows in the largest areas being considered (forest estates, regions, and nations). The tactical and operational planning phases, which have respectively 5 to 10-years and short-term time horizons, involve translating results derived at higher levels, and connecting them to specific objects, usually the forest stands. Implementation of superior goals can be complex (*e.g.*, Ozbekhan 1969). Moreover, before an action is carried out, the randomness of many events can often lead to intuitive decisions that deviate from the formal plan.

Despite the deviations from planning outcomes that occur during implementation, the usefulness of forestry planning is normally indisputable. Once entering the planning process, forestry planners face several intricate challenges in the search for appropriate management strategies to deal with possible and uncertain future scenarios. Usually, planning outcomes are based on several non-trivial assumptions. Argued by Jonsson (1982), complete planning models should include components for goal formulation, forecasting, optimisation (the search for actions that lead to desired results), and acquisition of forest data.

The goals of forestry may vary, though maximising the net present value (*NPV*) is a common desire (*e.g.*, Faustmann 1849, Dykstra 1984, Johansson and Löfgren 1985). The fact that *NPV*s are expressed in monetary terms, simplifies measures

of goal fulfilment. Traditionally, the revenues arise from wood products. Forestry may also strive to maximise other additional values, and these objectives are sometimes treated as restrictions in the planning process (*e.g.*, Hansen *et al.* 1991, Dahlin and Sallnäs 1993, Wikström 2000).

Several planning models have been developed to derive guidelines for use in forestry. The Finnish MELA system (Siitonen 1995, Siitonen and Nuutinen 1996) was primarily used for national and regional-level analyses, based on national forest inventory (NFI) data. The system has been further refined since then, to enable it to be used at lower scales (*i.e.* communities and forest holdings of relatively small areas). The system includes optimisation routines based on linear programming (Lappi *et al.* 1996). The Swedish HUGIN system (Lundström and Söderberg 1996) also utilizes high resolution NFI data but, in contrast to MELA, it is used to analyse management programs set in advance, *i.e.* for simulation instead of optimisation. The corresponding Norwegian system is AVVIRK (Hobbelstad 1988, Eid and Hobbelstad 2000), which is also based on simulation and designed for analyses at the national and regional levels. Both in Scandinavia and in the rest of the world, still many other additional forestry planning models are available (*e.g.*, Nabuurs and Päivinen 1996).

Common to the above listed planning models is the use of forest input data at the single (individual) tree level. Such a format is usually essential for accurate prognoses. Furthermore, it enables tests of different management alternatives and more precise cost and revenue calculations when economic issues are considered (*e.g.*, Jacobsson 1986, Getz and Haight 1989). Modeling the forest by single trees conserve valuable structural information that can be used in the analyses (*e.g.*, Eriksson 1994).

A Swedish system for aiding decision-making in forestry is the Forest Management Planning Package (FMPP; Jonsson *et al.* 1993). The FMPP, developed in the 1970s, has a basically strategic nature, focusing on planning for timber production. Several sub-models are coordinated within the system. Typical outcomes, maximising the *NPV*, with options to smooth the net revenue profile over time, are accommodated by the model. Suggestions for scheduling harvest activities (with several different types of thinning and clearcutting) are generated in different forest stands in 5-years periods. The single tree is the basic unit in most of the FMPP sub-models. The forest inventory method in FMPP prescribes plotwise field measurements within a sample of stands in the total forest holding. Within the sampled stands, approx. 10 plots per stand are inventoried, normally using a square lattice (randomly positioned). Since only a sample of stands is inventoried, the transfer of results to all stands of the holding is elaborate (but enabled by using the stand records available). To obtain results representative for the entire forest holding, the initial sampling of stands needs to be statistically sound. The FMPP was primarily intended for major forest holdings (covering thousands of hectares) and the system has been used by several forest companies

in Sweden. For small forest holdings the system is assumed to be less suitable, mainly because of the expensive assessment of forest data involved.

In Sweden, the non-industrial private forest (NIPF) owners hold about 50% of the total forest land, mostly in southern Sweden in areas with a relatively high productivity (Anon. 2000). The NIPF holdings typically are small forest estates, with a median area of about 50 ha (Anon. 2001a). Consequently, the planning process here partly differs from the normal practice at the forest companies. As a group, the NIPF owners are rather heterogeneous with regard to forest values and sometimes have vaguely formulated goals of forestry. Further, harvest decisions often tend to be effected by factors that are difficult to control in a planning process (*e.g.*, Lönnstedt 1985, Carlén 1990). However, forestry plans are often used at such holdings, based on data from subjective field inventories, with treatments suggested more or less intuitively.

Data acquisition

Forests are multi-dimensional and can, therefore, be described from several different perspectives (*e.g.*, Schreuder *et al.* 1993, Dahlin *et al.* 1997). In addition to data about the forests, other information is also required for planning, such as interest rates, timber prices, and costs for silvicultural and cutting activities. However, a major problem is associated with describing the forests, which are usually dispersed over large geographical areas.

Field methods

In field-based forest inventories, both subjective and objective methods are frequently used. Methods where data are collected in a subjective manner (*i.e.* purposive or non-statistical sampling) will generate estimates with reliance on the skill of the surveyor. Objective methods rely on statistical sampling theory, meaning that representative data are ideally assessed independent of the person taking the measurements. In general, objectively derived estimates are unbiased and the precision of these can be determined (*e.g.*, Thompson 1992, Schreuder *et al.* 1993).

Nearly all forest land in Sweden is described in some sort of stand record, which is the basis of the forestry plan traditionally used by NIPF owners. For each stand, the records include a vector of mean values for a set of variables, usually describing the trees (stem volume, age, *etc.*) and the site (vegetation, ground, *etc.*). These data are normally collected by subjective field methods based on ocular judgements. Despite often supported by relascope measurements, the quality of stand record data depends on the surveyors' skill and experience, and hence varies from case to case (*e.g.*, Eriksson 1990, Ståhl 1992).

The need for unbiased data in long-term prognoses has been noted in previous sections. Further, objective inventories are motivated in environmental monitoring, to enable analyses of changes over time and compare conditions in different regions (*e.g.*, Berg *et al.* 1994, Ringvall 2000).

The Swedish national forest inventory (NFI) is conducted annually and collects, according to present design, data at both temporary and permanent plots, with radii of respectively 7 m and 10 m (Ranneby *et al.* 1987). The Swedish NFI is based on a stratified sample with clustered plots. At the plots, intensive measurements of several forest characteristics are made. Results from the NFI present the national and regional timber production capacity as well as estimates of other forest environment states (Anon. 2000). The annual sample is relatively sparse, meaning that results at lower geographical levels (*i.e.* communities, forest holdings) need data aggregated from several years, intensified measurements or other additional information. NFIs are performed in many other countries (*e.g.*, Köhl and Pelz 1991, EFICS 1997a; 1997b), in each case designed for efficiency for the specific objectives and forest conditions.

Plotwise inventories are widely used to derive estimates for stand-level characteristics. Plots are then systematically sampled within stands, to account for spatial autocorrelation and thus decrease the sampling errors (*e.g.*, Lindgren 1984, Schreuder *et al.* 1993). In general, within a fixed radius all trees at the plots are callipered (at breast height) and additional measurements are made for a subset of the trees of characters such as age and tree height. Such sample tree data are later used to improve the estimates for the trees that have only been callipered (*e.g.*, Holm *et al.* 1979, Jonsson 1995, Korhonen and Kangas 1997).

Line-based, as opposed to plot-based, inventories can be more efficient in certain situations (*e.g.*, Schreuder *et al.* 1993, Ringvall 2000). Other ways to enhance precision (and cost-efficiency) in field inventories involve the use of some auxiliary information, for example remotely sensed data, in stratification approaches, *etc*. At the same time, many inventory methods based on remotely sensed data need access to a field measured reference material, for purposes of translation and levelling to present conditions. Usually, these remote sensing applications highlight the importance of accurately geo-referenced field data (*e.g.*, Fazakas and Nilsson 1996, Nilsson 1997).

Remote sensing
An attractive property of remotely sensed information is its coverage; information about all parts of large areas is usually obtained within an image. This leads to low costs of data per unit area. Covering an equal area with a field-based method would be very time consuming and, hence, remote sensing is often involved in cost-efficient solutions for forest inventory.

In forest inventory, the use of aerial photography as a source of information goes back to the 1920s (*e.g.*, Loetsch and Haller 1973). Aerial photographs of forest areas are used for delineating stands and for deriving various stand features. By analysing stereo-images, certain forest characteristics can be measured and estimated with relatively high accuracy (*e.g.*, Emanuelsson *et al.* 1983, Ericson 1984, Næsset 1996, Eid and Næsset 1998). In the analogue images, scale and film type effect the results. For example, the separation of tree species is usually improved by using colour or colour-infrared photos instead of panchromatic photos (*e.g.*, Stellingwerf and Hussin 1999). Visual interpretations, where tree height is measured and stocking ('crown closure') estimated, are used in a forest inventory system with long practical experience (Åge 1985). Since it is well established in forestry, aerial photograph interpretation is employed in many different situations. The transition from analogue to digital aerial photographs has further improved the usefulness, making automatic image-processing possible (*e.g.*, Dralle 1997, Brandtberg 1999, Hill and Leckie 1999).

Optical satellite data, *i.e.* measurements of reflected light within the visible and infrared electromagnetic spectrum, have been used in forestry applications since the launch of the first Landsat satellite in 1972 (*e.g.*, Kuusela and Poso 1975, Ahmad *et al.* 1992, Kramer 1994, Trotter *et al.* 1997). As well as giving extensive geographical coverage, the often high temporal resolution is used to obtain updated information and thus identify changes of different kinds caused by diverse factors, such as natural growth or harvest activities (*e.g.*, Olsson 1994, Coppin and Bauer 1996). Several forest characteristics show a relationship with optical satellite data and methods for estimating features like total biomass, stem volume, basal area, and tree species composition, have been developed (*e.g.*, Poso *et al.* 1987, Hagner 1989, Ripple *et al.* 1991, Cohen and Spies 1992, Bauer *et al.* 1994, Tokola and Heikkilä 1997, Holmgren *et al.* 2000, Katila and Tomppo 2001). However, problems usually arise in dense forests since the reflectance changes very little once the canopy is closed (*e.g.*, Guyot *et al.* 1989). This leads to under-estimation in high-volume classes, which is unfortunate since the highest forest values are usually found in the mature and dense forests (*e.g.*, Ardö 1992). The spatial resolution ('pixel size') varies between different sensors and this effects the suitability of any given information source for specific purposes. Pixel-level estimates generally have low accuracy, although aggregation up to stand or landscape levels improve the results (*e.g.*, Nilsson 1997, Päivinen and Anttila 2001). New satellite sensors offer resolutions with 'several pixels per tree' instead of 'several trees per pixel' as before, assumingly generating more intricate estimation problems.

Active sensors, with internal illumination sources, have been used in estimations of forest characteristics (*e.g.*, Lillesand and Kiefer 1994, Leckie and Ranson 1998). Both spaceborne and airborne sensors are available, while the airborne types (mounted on aeroplanes or helicopters) have proven more useful in forestry applications. With radar techniques, using wavelengths in the range of

12

centimeters up to 15 m, relatively high accuracy in estimating variables such as stem volume has been reported (*e.g.*, Fransson 1999, Walter 1999, Smith 2000). In comparison with optical image data, the signal saturation limit seems to be very high when using relatively long wavelengths, *i.e.* >3 m (*e.g.*, Ulander *et al.* 2000). Also, weather independence is an attractive property of the radar sensors, capable of acquiring data despite presence of a cloud cover.

With sensors based on laser techniques, forest characters like tree height have been successfully estimated (*e.g.*, Næsset 1997, Nelson *et al.* 1997, Nilsson 1997, Magnussen and Boudewyn 1998, Means *et al.* 1999). With profiles or surfaces of the forest canopy, similar forest characteristics as those derived from manual aerial photograph interpretation (mean tree height and stocking) can be estimated. Here, the automatic processing should improve cost-efficiency even further.

The development of the Global Positioning System (GPS) has played a key role in advances in forest inventory based on remotely sensed information (*e.g.*, Leckie 1990, Deckert and Bolstad 1996). The relationship between remotely sensed information and ground states can be estimated with accurately geo-referenced field measurements as reference data (*e.g.*, Næsset 1999, Sigrist *et al.* 1999, Holmström 2001). Remote sensing images are usually geometrically precision corrected to fit a selected co-ordinate system, using sensor parameters and ground reference points or internal sensor-navigation systems (*e.g.*, Ackermann 1996). Validation of the image data is especially important, since this information often is influenced by factors other than the ground state, *e.g.*, atmospheric conditions, which may vary between image acquisition occasions.

Combining information sources

In forestry planning, multi-variate descriptions of the forest are usually demanded. Data from a certain remote sensor may supply accurate information about certain forest characteristics while being less informative for others. Thus, combining information from different sources is valuable. In large-scale forest inventories at forest holdings or landscapes, field sampling methods can be supported by information from auxiliary sources (maps, aerial photos, satellite images, *etc*.) in methods such as multi-phase sampling or stratification (*e.g.*, Thompson 1992).

Stand level forest data, acquired by using ocular methods, are sometimes calibrated based on objective assessments in a sample of the stands (*e.g.*, Li 1988). Further, several different information sources might be available for all stands. The information can stem from different remote sensors, or be accessible from maps, stand records, *etc*. The quality of the sources may differ greatly. The main problem is then how to make use of each information source in the best possible way, to combine the different sources in composite estimators. The estimates can be obtained as weighted averages; weights should be set inversely proportional to the variance (or *MSE*) of the estimate from a certain information

source (*e.g.*, Raj 1968, Ståhl 1992). Many forestry applications have been based upon such procedures (*e.g.*, Burk *et al.* 1982, Hagner 1990, Poso *et al.* 1999, Tuominen and Poso 2001).

In several applications, multiple regression analysis has been used to estimate forest variables from remotely sensed data (*e.g.*, Tomppo 1988, Hagner 1990, Thuresson 1995). If several forest variables are to be estimated, multivariate regression has been suggested (*e.g.*, Holm 1980). Multivariate regression analysis accounts for any correlation between the dependent variables; nevertheless, the estimates are obtained as interpolated (or extrapolated) values that may show unnatural relationships between the variables. Another result of regression is the reduction of variation among the estimated variables, a weakness in the context of supplying planning models with input data (*e.g.*, Holm 1980).

Estimation by imputation

Imputation involves techniques where known characteristics for a certain unit are imputed ('assigned', 'allotted') to another unit, for which data on some of the characteristics are not available. In the statistical literature, the use of such a 'stand-in estimator' is sometimes referred to as synthetic estimation (*e.g.*, Rao 1998). Imputation has been applied in several different situations, for example dealing with non-respondents in social surveys (*e.g.*, Rubin 1987), and for updating purposes (*e.g.*, McRoberts and Hansen 1999). An example of imputation in forest inventory is the grid method (Holm *et al.* 1979, Hägglund 1981); data obtained from measurements of sample trees are assigned to trees that have only been callipered.

The *k* nearest neighbour (*k*NN) method is regarded as non-parametric regression. The method is data-driven, *i.e.* no assumptions about the distribution of the variables involved are made (*e.g.*, Linton and Härdle 1998, Efromovich 1999). Nearest neighbour methods have been thoroughly examined, triggered by applications for discrimination, classification, and pattern recognition (*e.g.*, Hardin 1994, Yu 1994, Ripley 1996).

In estimating forest characteristics, the *k*NN method has been extensively used. As a forestry application, the *k*NN method has been employed in the Finnish multi-source (MS) NFI since 1990 (*e.g.*, Tomppo 1990, Tomppo 1991, Tomppo and Katila 1992, Tomppo 1993, Katila and Tomppo 2001), based on earlier ideas pointed out by Poso (1972), Kilkki and Päivinen (1987), and Muinonen and Tokola (1990).

In the Finnish MS-NFI, field and satellite data are integrated. Areas known only by their spectral signatures in satellite images are assigned field data values as weighted means of the *k* nearest field plots; nearness is measured in a feature space defined by the different spectral wavelength bands of the satellite image. Sample plot data from previous field inventories are used as references. Thus,

14

field measured variable values are assigned to non-inventoried plot locations (represented by pixels in the satellite image) based on the similarity of 'target' and reference plots. In general, after defining a distance measure (often Euclidian), k = 5-10 reference plots are assigned to each target plot; the reference plot values weighted according to the distances.

Further applications using the kNN method have been developed. In some of these, the satellite image data are used together with auxiliary information derived from maps (Tokola and Heikkilä 1997), stand records (Tomppo *et al.* 1999), and remotely sensed tree heights (Nilsson 1997). Combining different information sources will add dimensions to the feature space in which nearness is measured and, hence, strengthen the association between the target plots and the reference plots (*e.g.*, Tokola *et al.* 1996, Tomppo *et al.* 1999).

Several variables, often many hundred (*e.g.*, Anon. 2001b), are measured at each field plot used in a reference material. By assigning entire (weighted) reference plots, all of the associated variables are estimated simultaneously. The natural and often complex relationships between the variables are hence retained. The importance of this, if the estimates are to be used as input in a planning model, is stressed by Moeur and Stage (1995). The compromise between preserving the natural correlation structures and deriving more accurate estimates (with lower *MSE*s) is discussed by Franco-Lopez *et al.* (2001). However, instead of calculating weighted mean values of the k assigned reference plots, these k plots may enter a planning system in their original format (with a weight assigned to each plot), *i.e.* imputation of reference sample plot data. Although the information will be imputed rather than field-measured for the certain plot, it will appear in a format that may readily be used in a traditional planning system such as the FMPP (with formal norms regarding the input data). The high-resolution, *i.e.* the single-tree characteristics of the reference data, is preserved by the imputation of entire (weighted) plots.

Data acquisition for forestry planning

All forest descriptions are accompanied with certain errors. However, the value of the data mainly depends on their usefulness in supporting decisions to be made (*e.g.*, Jacobsson 1986, Jonsson *et al.* 1993). Based on information about the present state of forests, planning models are used to obtain treatment schedules, *i.e.* timing of certain activities in certain stands, that maximises a defined goal. In the search for suitable forest inventory strategies in a planning context, cost-plus-loss analysis has been suggested for comparing different methods (*e.g.*, Cochran 1977, Hamilton 1978). The problem is approached by considering both the cost of data acquisition and the expected losses due to erroneous decisions caused by errors in the data.

Inventory costs can generally be determined rather easily. To determine the expected loss due to non-optimal decisions is considered the difficult part in this type of analyses (*e.g.*, Burkhart *et al*. 1978, Ståhl 1994). To estimate the expected loss, simulation of the consequences of using data with certain errors in a planning system has often been employed (*e.g.*, Sprängare 1975, Larsson 1994, Kangas and Kangas 1999, Eid 2000). The cost of the inventory plus the expected loss is the inventory method's total cost, *TC*. Minimising *TC* should then lead to the optimal choice of inventory method in a certain planning situation, as illustrated in Figure 1.
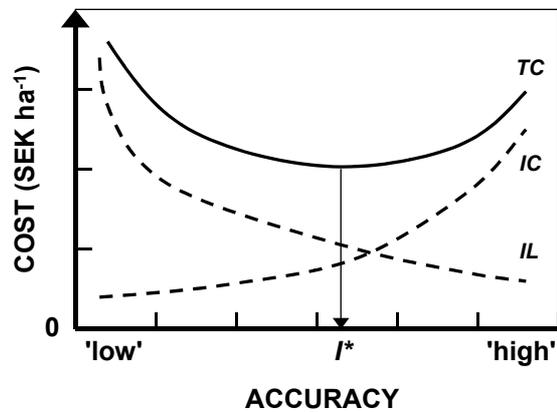

Figure 1. General view of cost-plus-loss analysis showing the inventory cost (*IC*), the inoptimality loss (*IL*; Jacobsson 1986), and the sum of these representing the total cost for an inventory (*TC*). *I** indicates the optimal accuracy, corresponding to the use of a certain inventory method.

16

Objectives

The main objective of the work presented in this thesis was to develop and evaluate forest data acquisition methods based on imputation. Imputation of reference sample plot data was made using aerial photograph interpretations, optical satellite data, radar data, and existing stand record information. The evaluations focused on the accuracy of stand-level estimates of forest characteristics and the usefulness of the imputed data in forestry planning.

The specific objectives of the studies described in Papers I-V were:

I. To evaluate the utility of forecasted reference sample plot data in kNN estimations. Plotwise stem volumes were estimated using simulated Landsat TM satellite data and reference sample plot data forecasted for up to 25 years were evaluated.

II. To evaluate the kNN estimation method when aerial photograph interpretations, in combination with stand record information, were used as carrier data. In addition, the accuracy of the plotwise aerial photograph interpretation was evaluated.

III. To evaluate the combination of two different remote sensing data sources: optical satellite data (SPOT-4 XS) and radar data (airborne CARABAS-II VHF SAR), used in kNN estimations of forest characteristics.

IV. To further evaluate the kNN method based on plotwise aerial photograph interpretations described in Paper II. In this case, carrier data were only available for the target areas and, hence, adaptations of the distance measures were made and evaluated.

V. To study the planning consequences of using imputation as a data acquisition strategy. In cost-plus-loss analyses, imputations were compared with traditional field inventories to assess each method's suitability for providing input data for forest management planning.

The studies were carried out within the framework of a project aiming at further development of the FMPP. The particular scope of the studies on imputation techniques was to make the planning system economically available to NIPF-owners, by reducing the costs for data acquisition.

Material and Methods

Study areas and field data

The characteristics of Swedish forests vary between different parts of the country. Changes along the north-south gradient are most distinct; the productivity of wood in southern Sweden can be many times higher than in the northernmost parts (Anon. 2000). The density of the forests, the tree species composition, the site conditions, *etc.* will most likely effect the results obtained from studies of inventory methods. Conclusions from different studies depend on the specific study area and its representativity. Further, the properties of the utilized reference material will effect the accuracy of the estimations, especially when the estimates are obtained from imputation of reference data. In the following sections, the study areas and corresponding field data used in Papers I-V are described.
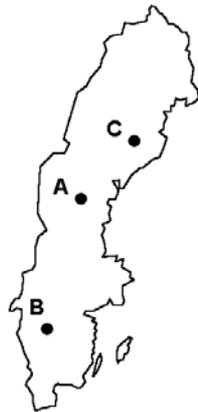
Figure 2. Map of Sweden showing the location of the study areas: Bräcke (A), Remningstorp (B), and Brattåker (C).

Bräcke (Paper I)

The Bräcke study area is located in central Sweden (62°30′ N; Figure 2) and covers about 300 000 ha. It is part of the forest company SCA's holdings. A stand record from 1997 with data from 32 000 stands within the area was used for simulations of plot-level forest characteristics. Plot data (one plot per stand) were simulated from stands sampled by probability proportional to size (PPS), with respect to area, and restricted to ages greater than 40 years. In a cross-validation approach, these plots were used in *k*NN estimations as both target and reference material (Table 1). In the simulations of forecasted plot data, the reference plots were coupled with assumed states; plot states forecasted for a certain number of years were used as reference values in the estimations.

To enable errors in forecasts of plot data to be derived, permanent plots from the Swedish NFI were used (Table 1). The permanent plot data were collected between 1983 and 1995, each plot being measured at least twice with an interval of 5 years. Prognoses using the HUGIN system (Lundström and Söderberg 1996) were compared with field measured plot-state differences, estimating the error variances in the forecasts. Derived error components were used in the simulations of the forecasted states of reference plots.

Remningstorp (Papers II, III, and V)

The Remningstorp study area is located in southern Sweden (58°30′ N, 13°40′ E; Figure 2). In total, the estate comprises 1 200 ha of forest land divided into 340 stands. The estate was considered a suitable study area while representing a large proportion of southern Swedish forests, being dominated by Norway spruce (*Picea abies*) on relatively fertile sites (average potential productivity approx. 10 $m^3 ha^{-1} yr^{-1}$).

During 1997-99, a sample of 49 stands was carefully inventoried in the field according to the FMPP guidelines. Within each stand approx. 12 plots with a 10 m radius were systematically sampled. These plots and stands were used as evaluation objects, representing true states (Table 1). In Paper III, only 47 of the stands were considered, since forestry activities had taken place between the field inventory and the acquisition of remote sensing data. In Paper V, only 35 of the stands were analysed since the study focused on established forests, *i.e.* >25 years.

In addition to the standwise inventories, 251 plots were systematically sampled over the estate and inventoried according to the same procedures. These plots were used as reference material in the estimations reported in Papers II and III (Table 1). All plots within the Remningstorp estate were geo-referenced by high accuracy differential GPS, with horizontal mean errors of approx. 1 m. In Paper V, the analyses were made using a reference material of 3 937 plots, inventoried during 1985-2000 within an area approx. corresponding to the counties of south-western Sweden (Table 1).

Brattåker (Papers IV and V)

The Brattåker study area is located in northern Sweden (63°35′ N, 20°15′ E; Figure 2), with altitudes varying between 150-400 m. Most of the area, covering approx. 10 000 ha, is owned by Holmen Skog AB or, to a lesser extent, NIPF owners. The forest in the area is dominated by Scots pine (*Pinus sylvestris*) followed by Norway spruce, in both pure and mixed-conifer stands. 35 stands were used as evaluation objects, inventoried in 1993 according to the FMPP, with approx. 7 plots per stand (Table 1). In Paper V, 33 of these stands were used. Plot centre co-ordinates were obtained with differential GPS.

Reference data were obtained from an inventory conducted in 1996 (Wallerman 1998), where 2 383 plots were systematically sampled within the Brattåker area (Table 1).

Table 1. Study areas and corresponding field data, *n* = number of units (plots or stands).

| Study area | *n* | Stem volume ($m^3 ha^{-1}$) | | |
		Mean	Min.	Max.
Bräcke				
Evaluation plots	4752	182.2	10	656
Permanent NFI plots	1748	163.6	2	561
Remningstorp				
Evaluation stands	49	179.2	0	426
Evaluation plots	563	173.5	0	743
Reference plots[a]	814	178.2	0	743
Reference plots[b]	3937	184.6	0	955
Brattåker				
Evaluation stands	35	139.3	0	253
Evaluation plots	255	139.5	0	413
Reference plots[c]	2383	128.3	0	479

[a] Within the Remningstorp study area (Papers II and III).
[b] Within the corresponding region (Paper V).
[c] Within the Brattåker study area (Papers IV and V).

Remotely sensed and auxiliary data

The term carrier data refers to the information used in the imputation; to 'carry' data from a reference plot to a target area. Final estimations are obtained from imputed reference sample plot data. In the following sections, the carrier data employed in Papers I-V are described.

Aerial photograph interpretation (Papers II, IV, and V)

Aerial photographs were acquired by the Swedish National Land Survey (NLS) for the Remningstorp study area in 1996 and for Brattåker in 1993. The images were panchromatic and taken from normal flight height, 4 600 m. From the photographs, plotwise mean tree height, stocking (Jonsson 1914), and tree species composition were measured and estimated. The interpretations were done by professionals using advanced stereo-instruments according to principles suggested by Åge (1985). The plot centre co-ordinates (with a plot radius of 10 m) were superimposed onto the aerial photographs and interpretations were made for both the target and the reference plots used in Paper II. In the study described in Paper IV, aerial photo interpretations were made only for the target plots.

The interpretations simulated in Paper V were generated using the error components (and their correlations) derived in previous studies, where the interpretations had been made for plots with field-measured states.

Optical data and radar data (Papers I and III)
The simulated optical data employed in Paper I corresponded to a Landsat-5 TM satellite scene, derived according to a previous study by Nilsson (1997). The simulations of pixelwise values in the wavelength bands 1-5 and 7 were generated from the corresponding plotwise stem volume using regression functions including stochastic error components.

As reported in Paper III, the Remningstorp study area was covered by a SPOT-4 XS satellite image, acquired in 1999. The image was geometrically precision corrected to the Swedish national grid, RT90. The pixel-size was 20×20 m and values in the four available bands, approximately corresponding to green (G), red (R), near-infrared (NIR), and mid-infrared (MIR) reflected light, were used in the estimations.

In the same study (Paper III), the optical satellite data were combined with radar data in the estimations of forest characteristics. The radar data were obtained from CARABAS-II VHF SAR, an airborne sensor operating within wavelengths between 3-15 m (Hellsten *et al.* 1996). The flight campaign over Remningstorp took place in 1999, at an altitude of 3 600 m. Since it is an active sensor, signals were transmitted and received after ground interference. The recorded backscattering ('radar echo') was processed to yield an image with a ground resolution of approx. 3×3 m (Walter *et al.* 1999).

Stand record information (Papers II, IV, and V)
Estimates of stand record variables are usually based on quick, visual field inventories (Ståhl 1992). Moreover, stand record information can be updated for several years by relatively coarse methods before any new field inventories are conducted. Thus, the quality of these data varies considerably from case to case (Eriksson 1990). In Papers II and IV, the quality of the stand records was first estimated, then the data were used in combination with the aerial photograph interpretations to estimate forest characteristics by means of imputation. For Brattåker and Remningstorp, the stand records showed *RMSE*s in stem volume estimates of about 25%, while age was somewhat more accurately estimated. These figures are slightly higher than what has been found in previous studies (*e.g.*, Eriksson 1990, Ståhl 1992).

The stand record information in Paper V was simulated in such a way as to be consistent with the errors obtained in the previous studies. Variables used from the stand records were: stem volume ($m^3 ha^{-1}$), age (yrs), site index (m) (Hägglund and Lundmark 1981), and tree species composition (%).

The *k*NN estimation method

Results presented in Papers I-V are based on estimates derived using the *k*NN method (Kilkki and Päivinen 1987, Muinonen and Tokola 1990, Tomppo 1990), sometimes referred to as the reference sample plot (RSP) method. The carrier data were used to define distances between reference sample plots and target plots, and the estimates were obtained as weighted means of the values at the *k* nearest reference plots.

From the field-measured variable *X* at the reference plots, a *k*NN estimator $\hat{Y}_t$ of the corresponding variable was obtained as:

$$\hat{Y}_t = \sum_{i=1}^{k} w_{t,i} X_i$$

The weight *w* was set with regard to the distance *a* between target plot *t* and the *j*th of the *k* nearest reference plots according to (*cf.* Isaaks and Srivastava 1989):

$$w_{t,j} = \frac{\dfrac{1}{a_{t,j}^p}}{\displaystyle\sum_{i=1}^{k} \frac{1}{a_{t,i}^p}}$$

In Paper I, the weights were inversely proportional to the squared distance, *i.e.* $p = 2$. In Papers II-IV, the weights were defined in a straightforward manner by using $p = 1$. By using $p = 0$, as in Paper V, the *k* nearest reference plots were assigned with equal weights, independently of differences in distance for these *k* plots. The number of reference plots imputed, *k*, varied in the different studies from 1 to 10. The distance *a* between target plot *t* and reference plot *r* was defined by an Euclidian metric (*e.g.*, Manly 1986) in Paper I. In the following papers, regression transform distances were used (*cf.* Tokola *et al.* 1996) according to:

$$a_{t,r} = \left| \hat{X}_t - \hat{X}_r \right|$$

The difference, in absolute terms, between estimates of *X* at target plot *t* and reference plot *r* here defined the distance. By regression analysis, functions were derived to estimate a variable of interest, *X*, using the *m* carrier data variables, *c*, as independent variables, *i.e.* $\hat{X} = f(c_1, c_2, ..., c_m)$. *X* was set to stem volume (m^3 ha^{-1}), age (yrs), or tree species composition (%). The carrier data set constituted of different information sources combined, for example from photo interpretations and from stand records, of different qualities and with different

abilities to explain a certain variable. Thus, to weight each source according to the additional information supplied, the use of multiple regression was motivated.

In Paper IV, where carrier data were only available at the target areas, the estimates of X for the reference plots were replaced by the corresponding field measured variable values according to:

$$a_{t,r} = \left| \hat{X}_t - X_r \right|$$

Both of these similar distance definitions were evaluated. However, in the latter case, estimates focus solely upon one variable, and the values for all other variables at the reference plots are disregarded, whereas in a planning context, accurate and consistent descriptions of several input variables are required.

In Paper V, where carrier data were simulated for the target areas and related to field measured variable values at the reference plots, the distance measure used was a linear combination of several different regression transform distances (with X as stem volume, age, *etc.*), each standardized and weighted. By this, reference plots were imputed based on nearness in several variables.

In Paper III, both the optical data and the radar data offered a total coverage of the target stands. Estimates were then made using all information pertaining to the stands and not just information describing the target plots. In Papers II-IV, stand-level estimates were obtained as mean values of the estimates for the target plots (and all 'target pixels' in Paper III) within the specific stands.

Evaluations

In the evaluations, the kNN-estimates of stem volume, age, and tree species composition were compared with the objective field measurements of corresponding variables, which were regarded as true values. For plot or stand i, the difference between the estimated and true variable value, $\Delta_i = \hat{Y}_i - X_i$, was used to derive the standard deviation, *Std*, and the average error (*i.e.* empirical bias), $\overline{\Delta}$:

$$Std = \sqrt{\frac{\sum_{i=1}^{n}\left(\Delta_i - \overline{\Delta}\right)^2}{n-1}}$$

$$\overline{\Delta} = \frac{\sum_{i=1}^{n}\Delta_i}{n}$$

The mean square error, *MSE*, was derived from:

$$MSE = Std^2 + \overline{\Delta}^2$$

The root mean square error, *RMSE*, was derived as the square root of *MSE*. The number of units (plots or stands) observed in different evaluations is denoted, in the above equations, by *n*. Relative errors, expressed in percent (%), were obtained in relation to the mean true value of field-measured variables.

Forecasting reference sample plot data (Paper I)

Simulations were employed to evaluate the impact of forecasted reference sample plot data in *k*NN estimations. Reference data forecasted for 5 to 25 years were evaluated. Single-tree growth models (Söderberg 1986) were used to forecast stem volumes of plots. The state of the reference material was kept constant throughout the estimations, to avoid differences in the composition of the reference material influencing the results. Instead, by performing forecasts from 5-25 years back in time, a certain reference plot obtained a set of forecasted values. These values were obtained from pairwise forecasts made with and without stochastic error components (Figure 3).
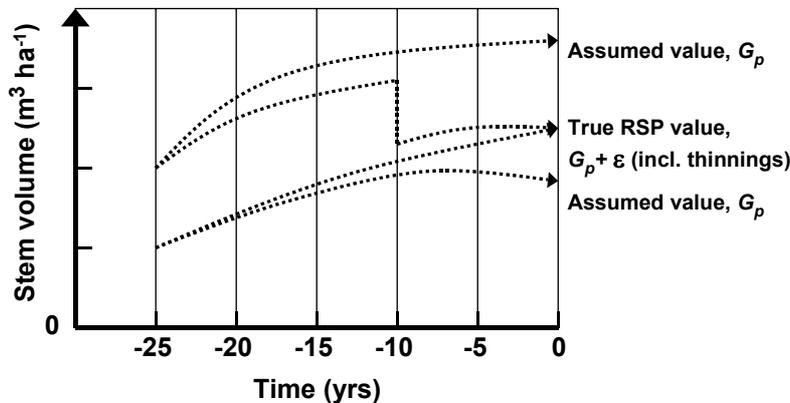
Figure 3. Assumed values forecast by growth functions (G_p) and coupled with trajectories ending at the true reference sample plot (RSP) value when using growth functions with random components, ε, including simulated thinnings. The figure shows two different cases, simulated from different starting points, both ending at the same RSP value. The assumed reference values are, however, different for the two cases.

The size of the errors (in each 5-year forecasting period) was empirically derived. When the forecasts including errors predicted the state would be the same as that of a certain reference plot, the corresponding forecasts made without errors were used as reference data for the plot, representing what was assumed to be its true

state. In the following *k*NN estimations, assumed states of the reference plots were randomly selected from the set of forecasted values, in accordance with the length of the forecasting period being studied.

The importance of excluding plots exposed to thinnings (or similar disturbances causing deviations from normal growth) was evaluated by simulating thinnings in the trajectories representing 'true developments'. Moreover, situations with improved carrier data were evaluated by increasing the correlation between ground state and satellite data, the simulated carrier data in this study.

Cost-plus-loss analyses (Paper V)

The consequences of using imputed reference sample plot data in planning were analysed by a cost-plus-loss approach (Cochran 1977, Hamilton 1978). Here, the cost part corresponds to the actual cost of employing a certain method to acquire forest data. Losses occur when non-optimal decisions are made due to erroneous descriptions of the forest state. Hence, the total cost of a certain method becomes the sum of the inventory cost and the expected losses due to treatments deviating from those leading to a maximum net present value (*NPV*).

Two different *k*NN applications, one based solely on stand record data and the other on aerial photograph interpretations combined with stand records as carrier data, were compared with traditional field sample plot inventories. Each inventory method was used to estimate initial states for two sets of forest stands (two fictitious estates). These stands had been carefully inventoried and were assumed to represent true states. Using the FMPP, a treatment schedule maximising the *NPV* was derived for each stand. Two different interest rates, 2% and 4%, were evaluated while other relevant factors (timber price lists, harvest costs, *etc.*) were kept constant throughout the analyses.

By simulation, the stands were first provided with stand record information and plotwise aerial photograph interpretations. From the field measured variables, carrier data were simulated 50 times per stand using correlated random errors (Ripley 1987). The size of the error components (and their correlations) was based on results from previous studies (Papers II and IV). Reference plot data were then assigned to the stands based on the available carrier data. Plotwise field inventories were simulated by resampling of the stand's original plots, *i.e.* by bootstrapping (Efron and Tibshirani 1993, Hjort 1994), with respectively 5 and 10 plots per stand. With the assigned or resampled data as input in the planning model, optimal treatment schedules for the evaluation stands were derived.

Derivations of losses due to non-optimal decisions, *i.e.* inoptimality losses (*IL*s; Jacobsson 1986) were made with the FMPP. Based on the simulated data, the treatment schedule believed to maximise the *NPV* was selected. When this

treatment schedule was applied to true data, the difference in *NPV* between this schedule and the truly optimal schedule was calculated, and taken as the inoptimality loss. Only decisions (concerning thinning, clearcutting, and 'no treatment') in the first two 5-year periods were considered. Thus, an average inoptimality loss, $\overline{IL}$, was estimated for each stand and inventory method according to:

$$\overline{IL} = \frac{1}{50} \sum_{i=1}^{50} \left(NPV^{opt} - NPV_i^* \right)$$

where NPV^{opt} denotes the maximum net present value and NPV_i^* indicates the net present value for simulation *i*, obtained when a treatment schedule based on non-perfect data was selected. By assuming a certain stand area (ranging from 1-10 ha in the analyses), a total cost, *TC*, was estimated by adding the methods' inventory cost to the corresponding $\overline{IL}$. Results were obtained at the stand level and at the forest estate level.

Results and Discussion

Summary of Paper I

The use of remotely sensed data in estimations of forest characteristics generally needs access to a proper reference material. In Sweden, plot data from the NFI have been used for this purpose, since each plot is geo-referenced and can thus be related to image data (Nilsson 1997, Reese and Nilsson 2000). However, the annual sample of NFI plots is relatively sparse. To obtain a sufficiently large reference material, data from earlier years' inventories can be updated to the time of image acquisition. Plot states are then forecasted using single-tree growth simulators (Söderberg 1986). The risk of introducing errors in the reference data will increase with increasing length of the forecasting period. Plots that have been disturbed by management activities, or damaged, may be especially prone to cause problems.

Results presented in Paper I showed that the *MSE* in plot-level *k*NN estimates of stem volume increased with longer forecasts of reference plot data (Table 2). After 25 years, the *MSE* had increased by 102%, compared to the errors obtained when using non-forecast reference material. However, after excluding reference plots exposed to thinnings during the forecasting period, only minor increases in *MSE*s were obtained. These results indicate the importance of identifying disturbed plots (*i.e.* plots with developments diverging from natural growth) and excluding them from the reference material. When using a reference material where thinned plots had been excluded, the systematical errors considering the entire data set were modest. However, classwise results revealed over-estimations in young forests and under-estimations in old and denser forests.

26

Although the errors introduced by forecasting were found to be relatively substantial in some cases, they were still quite modest in comparison to the errors due to the carrier data applied, *i.e.* the satellite image information. Situations with improved carrier data were also simulated, corresponding to, for instance, combinations of information from different remote sensing sources. In these cases, the *MSE* in the *k*NN estimates increased with increasing length of forecasts. This was also true when disturbed plots were excluded from the reference material (Table 2).

Table 2. Mean square error, *MSE*, and average error, $\overline{\Delta}$, in *k*NN estimates of plotwise stem volume ($m^3\ ha^{-1}$) using optical satellite data and a forecasted reference material (with and without thinned plots).

	MSE ($\overline{\Delta}$ below), Length of forecast					
Carrier data / Reference data	0 yrs	5 yrs	10 yrs	15 yrs	20 yrs	25 yrs
Optical data[a]	11449	12030	13745	16809	20014	23092
	(-2.7)	(8.7)	(23.0)	(39.7)	(53.9)	(67.5)
-"-, thinned plots excluded	11449	11885	11820	11768	11655	11627
	(-2.7)	(-2.5)	(-1.9)	(-0.5)	(1.6)	(4.0)
Improved optical data[b]	1584	1737	2422	3766	5202	6991
	(-1.7)	(6.3)	(16.1)	(26.5)	(36.8)	(47.1)
-"-, thinned plots excluded	1584	1653	1652	1691	1801	1900
	(-1.7)	(-1.1)	(-0.3)	(0.0)	(1.4)	(3.9)
Optical data, accumulated[c]	11470	10526	9923	10081	10772	12372
	(2.8)	(0.3)	(5.7)	(11.0)	(18.5)	(25.6)
-"-, -"-, thinned plots excl.	11470	10644	9430	8577	7815	7224
	(2.8)	(-5.8)	(-3.2)	(-5.0)	(-4.5)	(-3.8)

[a] Simulated Landsat-5 TM information.
[b] Obtained using 0.1 × the error component in the simulations.
[c] Obtained adding 500 forecasted reference plots in each 5-yr period.

Examples were given for situations where the reference data set consisted of plots with differing lengths of forecast. Practical options might be to use a limited but accurate reference material (*i.e.* no forecasted plots) or a larger reference set of material by also including old, forecasted reference plots. In the examples given (when thinned plots were still included in the reference material), the *MSE*s decreased when accepting reference plots forecasted for up to 10 years, and then began to increase when older plots were added (Table 2). An alternative approach that could be used to allow plots forecasted for a certain time period to be included in the reference material, could be to assign weights to the reference plots. Lesser weights should then be given to plots with longer forecasting periods.

In several studies employing the *k*NN method with satellite image data, the importance of a sufficiently large reference material has been stressed (*e.g.*,

Nilsson 1997, Atkinson *et al.* 2000, Katila and Tomppo 2001). Further, increased estimation accuracies have been obtained by using reference plots that are geographically near (*e.g.*, Tokola 2000, Nilsson and Sandström 2001). The present study points out the possibility to use plot data from earlier years' inventories. If accurate growth models are used, it is possible to forecast plot states for relatively long times and still obtain decent estimates. In addition to the growth of measured trees, forecasts of plots should handle in-growth and tree mortality for reliable updates (Fridman 2000, Nyström 2001). Detection of disturbed plots is crucial, and thus reliable methods for identification of changes (other than natural growth) should precede the choice of reference data (Olsson 1994).

Summary of Papers II-IV

The *k*NN estimations presented in Papers II and IV were obtained using the same type of carrier data: aerial photograph interpretations in combination with stand record information. Hence, the accuracy of the following *k*NN estimations depended on the quality of these information sources. The errors in the aerial photograph interpretations and stand records obtained at the two study areas were derived regarding the objective field measurements as true values (Table 3).

The accuracy of the photograph interpretations was lower for Remningstorp than for Brattåker, valid for all interpreted variables: mean tree height, stocking, and tree species composition. At Remningstorp, the more dense forests complicated ground visibility and made the plotwise height measurements in the photos less accurate. In comparison with other studies, stand level tree heights have been estimated with standard deviations usually between 1-2 m (*e.g.*, Ståhl 1992, Eid and Næsset 1998). Moreover, systematical errors in height measurements have been reported, usually in the range of –1.5 m to 2.5 m, and often assumed to depend on forest type and interpreter (*e.g.*, Ståhl 1992, Eid and Næsset 1998).

The stand records available at Remningstorp and Brattåker showed quite large errors in stem volume estimates, with relative *RMSE*s of respectively 26% and 24%. However, the systematical errors were relatively small. In general, this is not the case; Eriksson (1990) found systematic under-estimation of volume in stand record data of about 20%. Age in stand records are usually more accurately determined than volume (*e.g.*, Eriksson 1990, Ståhl 1992), and this was also valid for Remningstorp and Brattåker. Age estimates derived from aerial photo interpreted data are usually less accurate (*e.g.*, Åge 1985).

Table 3. Root mean square error, *RMSE*, and average error, $\overline{\Delta}$, in aerial photo interpretations and stand records at Remningstorp and Brattåker study areas (Papers II and IV, respectively).

Carrier data	Remningstorp		Brattåker	
	RMSE	$\overline{\Delta}$	*RMSE*	$\overline{\Delta}$
Mean tree height (dm)[a]	20.9 (14.9%)	-10.1	18.2 (13.5%)	-7.9
Stocking (%)[a]	17.0 (27.0%)	0.1	14.4 (24.0%)	-1.1
Proportion of spruce (%)[a]	39.6 (63.7%)	-19.6	34.7 (85.3%)	-12.9
Proportion of conifers (%)[a]	32.6 (40.8%)	6.0	28.1 (34.4%)	12.5
Stem volume (m³ ha⁻¹)[b]	46.1 (25.8%)	4.9	33.0 (23.7%)	-2.7
Age (yrs)[b]	6.9 (15.0%)	-1.4	18.5 (22.3%)	5.6
SI, spruce (m)[b]	2.7 (9.6%)	0.6	3.3 (17.1%)	0.9

[a] Aerial photograph interpretation.
[b] Stand record information.

Employing the aerial photograph interpretations as carrier data in *k*NN estimations was evaluated in Papers II and IV. In the first case, interpretations were made for both target and reference plots, in the latter case only for the target plots. In Paper II (Remningstorp), a relative *RMSE* of 21% was obtained in stem volume estimates at the stand level. Improved results were obtained in Paper IV (Brattåker), showing a corresponding *RMSE* of 14% (Table 4). The use of a larger set of reference data and improved interpretations was assumed to be the major reason for this increase in accuracy. The additional information obtained from the stand records was of importance especially in the age estimates; when using the photo interpretations alone, age was estimated with a *RMSE* of 21%. Combined with stand record information, a *RMSE* of 16% was obtained. Despite using carrier data with, in some cases, significant systematic errors, the final *k*NN estimates showed only minor bias.

Many *k*NN applications estimating forest characteristics are based on optical satellite image data. In these cases, the *RMSE*s are usually in the range of 25% to 50% for stand level estimates of stem volume (*e.g.*, Nilsson 1997, Poso *et al.* 1999, Tomppo *et al.* 1999). For practical use in forestry planning, this is normally regarded as too high (*e.g.*, Holmgren and Thuresson 1998, Hyyppä *et al.* 2000). Standwise stem volume estimates from aerial photo interpretations are, however, more reliable, usually with mean errors of 15% to 30% (*e.g.*, Emanuelsson *et al.* 1983, Ericson 1984, Ståhl 1992, Eid and Næsset 1998). Traditional field inventories with approx. 10 plots per stand usually show standard errors of approx. 10% in stand-level estimates of volume (*e.g.*, Lindgren 1984). Clearly, this level of accuracy could not be met by the inventory methods presented here. However, in comparison with other remote sensing based methods, the accuracy in the estimates was quite good, and assumingly acceptable for use in forest management planning. The long experience of aerial photographs in forestry (*e.g.*, Poso 1972, Talts 1977), and the generally high accuracy in estimates of variables such as tree height and stocking, form a reliable basis for using this method.

29

Table 4. Root mean square error, *RMSE*, and average error, $\overline{\Delta}$, of the *k*NN estimations of stand-level stem volume and age using different carrier data.

Carrier data	Stem volume (m^3 ha^{-1})		Age (yrs)	
	RMSE	$\overline{\Delta}$	*RMSE*	$\overline{\Delta}$
Aerial photo interpretations[a]	37.7 (21.4%)	0.1	9.7 (21.2%)	-0.9
-"- and stand records[a]	35.5 (20.1%)	3.0	7.5 (16.3%)	0.2
-"-[b]	19.1 (13.7%)	1.7	13.7 (16.6%)	-0.3
Optical data[c]	50.5 (29.9%)	6.7	13.3 (29.6%)	-0.7
Radar data[d]	64.6 (38.2%)	2.2	16.9 (37.6%)	-0.9
Optical data and radar data	37.0 (21.9%)	2.8	11.2 (24.9%)	-0.5

[a] Remningstorp study area (Paper II).
[b] Brattåker study area (Paper IV).
[c] SPOT-4 XS (Paper III).
[d] CARABAS-II VHF SAR (Paper III).

In Paper III, results from *k*NN estimations of forest stand characteristics using optical satellite data from SPOT-4 and radar data from the airborne CARABAS-II system are presented. A relative *RMSE* of 22% was obtained in stem volume estimates using the combination of optical and radar data (Table 4). This was a significant improvement compared to the best single-sensor case, when using multi-spectral optical data, which gave a *RMSE* of 30%. For the current average stand area, approx. 5 ha, results were comparatively good (*cf*. Hyyppä *et al*. 2000). While data are becoming commercially available from a considerable number of new remote sensors, the presented *k*NN application can state as an example of possible integration of information of different kinds. The assumption of complementary information in the two sources was verified.

Summary of Paper V

Paper V presents the consequences of using *k*NN imputed reference sample plot data in a planning context. Cost-plus-loss analyses were performed for two sets of stands (Remningstorp and Brattåker), regarded as fictitious forest estates. The different inventory methods evaluated were: *k*NN imputation using stand record information (M1), *k*NN imputation using plotwise aerial photograph interpretations in combination with stand record information (M2), field sample plot inventory with 5 plots per stand (M3), and field sample plot inventory with 10 plots per stand (M4).

The average inoptimality losses for each method ranged between 78 SEK ha^{-1} (M4 at the Remningstorp estate using an interest rate of 4%) and 1345 SEK ha^{-1} (M1 in the same case). For Brattåker, the difference in average *IL*s was less pronounced between the different inventory methods, approx. four times higher for M1 compared to M4.

Total costs at the estate level for each method were derived, assuming that all stands had equal size: 2, 5, or 10 ha (Table 5). In the case with small stands, kNN imputations were the most appropriate methods. With larger stand areas, field inventory with 5 or 10 plots per stand resulted in the lowest total cost. However, the average estate level cost-plus-losses could be reduced by in each stand applying the method with lowest total cost, in Table 5 noted as method integration. At the stand level, the cost-plus-losses varied considerably between the different inventory methods, this being especially pronounced for the two kNN methods.

In the presented cases, the optimal choice of method did not depend very much on the interest rate used in the *NPV* calculations. The effect of varying interest rate can be assumed to differ depending on the composition of forest stands.

Table 5. Total cost, *TC*, for different stand areas, *A*, for two different interest rates *r* used in the *NPV* calculations. **Bold** indicates minimum *TC*, for the method performing best on average.

| | TC (SEK ha^{-1}) | | | | | |
| | Brattåker, $n = 33$ | | | Remningstorp, $n = 35$ | | |
Inventory method	$A = 2$ ha	$A = 5$ ha	$A = 10$ ha	$A = 2$ ha	$A = 5$ ha	$A = 10$ ha
$r = 2\%$						
M1	**427**	**415**	411	702	690	686
M2	504	425	399	**663**	584	558
M3	888	420	**264**	955	**487**	331
M4	1105	481	273	1157	533	**325**
M, integration[a]	*290*	*189*	*140*	*507*	*362*	*268*
$r = 4\%$						
M1	**420**	408	404	1367	1355	1351
M2	463	**384**	358	1289	1210	1184
M3	931	463	307	**924**	**456**	300
M4	1131	507	**299**	1118	494	**286**
M, integration[a]	*329*	*246*	*188*	*656*	*372*	*245*

[a] In each stand selecting the inventory method with the lowest cost-plus-loss (indicated by values in *italics*).

Results indicated that the total cost for an inventory could be decreased by 15-50%, in comparison to the method performing best on average, if the optimal method was employed in each stand. However, it was not always easily decided which stands should be inventoried with a certain method, especially if this decision should be made from some general stand description. All the same, the study pointed out that large valuable stands as well as stands with a treatment impending should generally be selected for careful field inventories. In other stands, imputation could be employed.

Conclusions

Imputation of reference sample plot data appears to be an interesting option for cost-efficient data acquisition in forestry planning. The presented kNN applications based on carrier data from stand records, aerial photograph interpretations, and satellite and radar data, resulted in estimates with reasonable quality at relatively low costs. Especially the applications based on stand record and photo interpretation data should be straightforward to introduce in practice, since they are based on well-known standard methods.

By imputation, data at the single-tree level can be obtained. Thus, planning models that require these kind of data can be applied, although the carrier data may regard stand-level estimates only. Further, planning models usually encompasses several different sub-models, each of which is generally designed to use a multivariate description of the state of the forest as input. Obtaining accurate outcomes from planning requires both accurate estimates of the many variables and accurate estimations of the relationships between them. When entire plots are imputed, the natural relationships between the variables will be preserved.

The simplicity by which different kinds of data can be combined and support the imputation of reference plots is appealing. This makes it possible to operationally utilize the many different new information sources that are currently becoming available (*e.g.*, radar and lidar), and combine them with existing sources of information. In this context, it is important to weigh each data source in relation to its quality.

The access to a sound reference material is important. However, objective sample plot inventories are performed in many different situations, for example in NFIs, and existing old data may be updated without seriously decreasing the precision in the final estimates.

Supported by the imputation technique presented, the FMPP and similar planning systems can be used for relatively small forest holdings without involving very high costs for the data acquisition. To provide input for such applications, field inventory methods should be employed in certain stands while data for other stands can be obtained by imputation. Cost-plus-loss analyses indicate that in large valuable stands, especially in stands with a treatment impending, quite intensive sample plot inventories should be used in order to obtain a low cost-plus-loss. However, in many other stands imputation based on stand record information or stand record information in combination with aerial photograph interpretation is an appropriate data acquisition method.

References

Ackermann, F. 1996. Georeferencing requirements for airborne photogrammetry and remote sensing. *International Archives of Photogrammetry and Remote Sensing Volume XXXI*, Part B6, 7-9.

Åge, P.-J. 1985. Forest inventory – photo interpretation. *National Land Survey*. Report 13. 21 pp.

Ahmad, W., Jupp, L.B., and Munez, M. 1992. Land cover mapping in a rugged terrain area using Landsat MSS data. *International Journal of Remote Sensing* 13, 673-683.

Angelstam, P. 1992. Conservation of communities – the importance of edges, surroundings and landscape mosaic structure. In: *Ecological Principles of Nature Conservation*, 7-70. Elsevier Applied Science. London, UK.

Anon. 1992. Agenda 21 & UNCED Proceedings.

Anon. 1993. Om godkännande av konventionen om biologisk mångfald. Regeringens proposition 92/93: 227. Stockholm, Sverige. 51 pp. (In Swedish).

Anon. 2000. Skogsdata – aktuella uppgifter om de svenska skogarna från Riksskogstaxeringen. *Swedish University of Agricultural Sciences*, *Department of Forest Resource Management and Geomatics*. 110 pp. (In Swedish).

Anon. 2001a. Statistical yearbook of forestry. *Official Statistics of Sweden*, *National Board of Forestry*. Jönköping, Sweden. 337 pp. (In Swedish with English summary).

Anon. 2001b. Instruktion för fältarbetet vid Riksskogstaxeringen år 2001. *Swedish University of Agricultural Sciences*, *Department of Forest Resource Management and Geomatics*. (In Swedish).

Aplet, G.H., Olson, J.T., and Alaric Sample, V. 1993. *Defining sustainable forestry.* Island Press. Covelo, CA. 328 pp.

Ardö, J. 1992. Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. *International Journal of Remote Sensing* 13, 1779-1786.

Atkinson, P.M., Foody, G.M., Curran, P.J., and Boyd, D.S. 2000. Assessing the ground data requirements for regional scale remote sensing of tropical forest biophysical properties. *International Journal of Remote Sensing* 21, 2571-2587.

Avery, T.E. and Burkhart, H.E. 1994. *Forest measurements*. 4th edition. McGraw-Hill. Boston, Mass. 331 pp.

Bauer, M.E., Burk, T.E., Ek, A.R., Coppin, P.R., Lime, S.D., Walsh, T.A., Walters, D.K., Befort, W., and Heinzen, D.F. 1994. Satellite inventory of Minnesota forest resources. *Photogrammetric Engineering and Remote Sensing* 60, 287-298.

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

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

Berg, Å., Ehnström, B., Gustafsson, L., Hallingbäck, T., Jonsell, M., and Weslien, J. 1994. Threatened plant, animal, and fungus species in Swedish forests – distributions and habitat associations. *Conservation Biology* 8, 718-731.

Brandtberg, T. 1999. Automatic individual tree-based analysis of high spatial resolution remotely sensed data. *Swedish University of Agricultural Sciences*, *Silvestria* 118.

Burk, T.E., Hansen, M.H., and Ek, A.R. 1982. Combining sources of information for improved in-place inventory statistics. In: Brann, T.B. (ed.). *In-place resource inventories – principles and practices*. National workshop proceedings. University of Maine, Orono.

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

Carlén, O. 1990. Private nonindustrial forest owners' management behaviour – an economic analysis based on empirical data. *Swedish University of Agricultural Sciences, Department of Forest Economics*. Report 92. 104 pp.

Cochran, W.G. 1977. *Sampling techniques*. 3rd edition. John Wiley & Sons. New York. 428 pp.

Cohen, W.B. and Spies, T.A. 1992. Estimating structural attributes of Douglas Fir/Western Hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment* 41, 1-17.

Coppin, P.R. and Bauer, M.E. 1996. Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sensing Reviews* 13, 207-234.

Dahlin, B., Ekö, P.-M., Holmgren, P., Lämås, T., and Thuresson, T. 1997. Heureka – a model for forest resource management. *Swedish University of Agricultural Sciences, Faculty of Forestry*. Report 17. 115 pp.

Dahlin, B. and Sallnäs, O. 1993. Harvest scheduling under adjacency constraints – a case study from Swedish sub-alpine region. *Scandinavian Journal of Forest Research* 8, 281-290.

Davis, L.S., Johnson, K.N., Bettinger, P.S., and Howard, T.E. 2001. *Forest Management*. 4th edition. McGraw-Hill. New York. 804 pp.

Davis, R.G. and 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.

Deckert, C. and Bolstad, P.V. 1996. Forest canopy, terrain, and distance effects on Global Positioning System point accuracy. *Photogrammetric Engineering and Remote Sensing* 62, 317-321.

Dralle, K. 1997. Locating trees by digital image processing of aerial photos. *Royal Veterinary and Agricultural University*, Denmark. Dina Research Report 58.

Dykstra, D.P. 1984. *Mathematical programming for natural resource management*. McGraw-Hill. New York. 318 pp.

EFICS. 1997a. Study on European Forestry Information and Communication System: reports on forestry inventory and survey systems. Volume 1. Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, and Liechtenstein. *Office for Official Publications of the European Communities*, 1-673. Luxembourg.

EFICS. 1997b. Study on European Forestry Information and Communication System: reports on forestry inventory and survey systems. Volume 2. Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, Czech Republic, Hungary, and Poland. *Office for Official Publications of the European Communities*, 675-1328. Luxembourg.

Efromovich, S. 1999. *Nonparametric curve estimation – methods, theory, and applications*. Springer-Verlag. New York. 411 pp.

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

Emanuelsson, M., Engsjö, T., Linderson, B., and Moberg, J. 1983. Skogsvärdering med flygbildsteknik. Ett praktikfall från Hammar-Komsta, Ånge kommun, Västernorrlands län. *National Land Survey*. Report 4. 95 pp. (In Swedish).

Ericson, O. 1984. Beståndsinventering med flygbild. *Forskningsstiftelsen Skogsarbeten*. Redogörelse 8. 19 pp. (In Swedish).

Eriksson, L. 1990. Quality of data and proposals of activities in our forest plans. *Swedish University of Agricultural Sciences*, *Department of Forest-Industry-Market Studies*. Report 11. 74 pp. (In Swedish with English summary).

Eriksson, L.O. 1994. Two methods for solving stand management problems based on a single tree model. *Forest Science* 40, 732-758.

Faustmann, M. 1849. On the determination of the value which forest land and immature stands posses for forestry. In: *Martin Faustmann and the evolution of discounted cash flow*. Oxford Institute. Paper 42 (1968), 27-55.

Fazakas, Z. and Nilsson, M. 1996. Volume and forest cover estimation over southern Sweden using AVHRR data calibrated with TM data. *International Journal of Remote Sensing* 17, 1701-1709.

Franco-Lopez, H., Ek, A.R., and Bauer, M.E. 2001. Estimation and mapping of forest density, volume, and cover type using the *k*-nearest neighbors method. *Remote Sensing of Environment* 77, 251-274.

Franklin, J.F. 1989. Toward a new forestry. *American Forests* 95(11/12), 37-44.

Fransson, J. 1999. Analysis of synthetic aperture radar images for forestry applications. *Swedish University of Agricultural Sciences*, *Silvestria* 100.

Fridman, J. 2000. Damaged and dead trees in Swedish forests – assessment and prediction based on data from the National Forest Inventory. *Swedish University of Agricultural Sciences*, *Silvestria* 146.

Fries, C., Johansson, O., Pettersson, B., and Simonsson, P. 1997. Silvicultural models to maintain and restore natural stand structures in Swedish boreal forests. *Forest Ecology and Management* 94, 89-103.

Getz, W.M. and Haight, R.G. 1989. *Population harvesting – demographic models of fish, forest, and animal resources*. Princetown University Press. New Jersey. 391 pp.

Guyot, G., Guyon, D., and Riom, J. 1989. Factors affecting the spectral response of forest canopies: a review. *Geocarto International* 3, 3-18.

Hägglund, B. 1981. Forecasting growth and yield in established forests – an outline and analysis of the outcome of a subprogram within the HUGIN project. *Swedish University of Agricultural Sciences*, *Department of Forest Survey*. Report 31. 145 pp.

Hägglund, B. and Lundmark, J.-E. 1981. Handledning i bonitering. *National Board of Forestry*. Jönköping, Sweden. 124 pp. (In Swedish).

Hagner, O. 1989. Computer aided forest mapping and estimation of stand characteristics using satellite remote sensing. *Swedish University of Agricultural Sciences*, *Department of Biometry and Forest Management*. 11 pp.

Hagner, O. 1990. Computer aided stand delineation and inventory based on satellite remote sensing. In Proceedings: *The usability of remote sensing for forest inventory and planning*. Swedish University of Agricultural Sciences, Remote Sensing Laboratory. Report 4, 94-105.

Hamilton, D.A. 1978. Specifying precision in natural resource inventories. In: *Integrated Inventories of Renewable Resources*. USDA Forest Service. General Technical Report RM-55, 276-281.

Hansen, A.J., Spies, T.A., Swanson, F.J., and Ohmann, J.L. 1991. Conserving biodiversity in managed forests. *BioScience* 41, 382-392.

Hardin, P.J. 1994. Parametric and nearest neighbor methods for hybrid classification – a comparison of pixel assignment accuracy. *Photogrammetric Engineering and Remote Sensing* 60, 1439-1448.

Hellsten, H., Ulander, L.M.H., Gustavsson, A., and Larsson, B. 1996. Development of VHF CARABAS II SAR. *SPIE Proceeding of Radar Sensor Technology* 2747, 48-60.

Hill, D.A. and Leckie, D.G. 1999. Automated interpretation of high spatial resolution digital imagery for forestry. In: *International Forum Proceeding*. Pacific Forestry Centre, Victoria, BC.

Hjort, J.S.U. 1994. *Computer intensive statistical methods. Validation*, *model selection and bootstrap*. Chapman & Hall. New York.

Hobbelstad, K. 1988. Planleggningsprogrammet AVVIRK3. *The Agricultural University of Norway*, Ås. Communications of the Norwegian Forest Research Institute 42. 38 pp. (In Norwegian).

Holm, S. 1980. An introduction to multivariate regression analysis – ideas, applications and results. *Swedish University of Agricultural Sciences, Department of Biometry and Forest Management*. Report 20. 64 pp.

Holm, S., Hägglund, B., and Mårtensson, A. 1979. En metod för generalisering av riksskogstaxeringens provträdsdata. *Swedish University of Agricultural Sciences, Department of Forest Survey*. Report 25. 94 pp. (In Swedish with English summary).

Holmgren, J., Joyce, S., Nilsson, M., and Olsson, H. 2000. Estimating stem volume and basal area in forest compartments by combining satellite image data with field data. *Scandinavian Journal of Forest Research* 15, 103-111.

Holmgren, P. and Thuresson, T. 1998. Satellite remote sensing for forestry planning – a review. *Scandinavian Journal of Forest Research* 13, 90-110.

Holmström, H. 2001. Averaging absolute GPS positionings made underneath different forest canopies – a splendid example of bad timing in research. *Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics*. Working Paper 82. 13 pp.

Hyyppä, J., Hyyppä, H., Inkinen, M., Engdahl, M., Linko, S., and Zhu, Y.-H. 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management* 128, 109-120.

Isaaks, E.H. and Srivastava, R.M. 1989. *An introduction to applied geostatistics*. Oxford University Press. New York. 561 pp.

Jacobsson, J. 1986. Optimization and data requirements – a forest management planning problem. *Swedish University of Agricultural Sciences, Section of Forest Mensuration and Management.* 183 pp.

Johansson, P.O. and Löfgren, K.G. 1985. *The economics of forestry and natural resources.* Oxford. Blackwell. 292 pp.

Jonsson, B. 1982. A forest management planning package. In: *Statistics in theory and practice. Essays in honour of Bertil Matérn*. Swedish University of Agricultural Sciences, Section of Forest Biometry, 159-171.

Jonsson, B. 1995. Age imputation to single trees. *Swedish University of Agricultural Sciences, Section of Forest Mensuration and Management*. Report 30. 35 pp.

Jonsson, B., Jacobsson, J., and Kallur, H. 1993. The Forest Management Planning Package. Theory and application. *Studia Forestalia Suecica* 189. 56 pp.

Jonsson, T. 1914. Om boniering av skogsmark. *Svenska Skogsvardsfoereningens Tidskrift*. Volume 5, 369-392. (In Swedish).

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

Katila, M. and Tomppo, E. 2001. Selecting estimation parameters for the Finnish multisource national forest inventory. *Remote Sensing of Environment* 76, 16-32.

Kilkki, P. 1985. Timber management planning. *Silva Carelica* 5. 160 pp.

Kilkki, P. and Päivinen, R. 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. In: *Remote sensing-aided forest inventory, SNS and Forest Mensurationist Club*, Hyytiälä. University of Helsinki, Department of Forest Mensuration and Management. Research Notes 19, 209-215.

Köhl, M. and Pelz, D.R. 1991. *Forest inventories in Europe with special reference to statistical methods*. IUFRO Symposium S.4.02. Birmensdort, Switzerland. 239 pp.

Kohm, K.A. and Franklin, J.F. 1996. *Creating a forestry for the 21st century – the science of ecosystem management.* Island Press. Washington, DC. 475 pp.

Korhonen, K.T. and Kangas, A. 1997. Application of nearest-neighbour regression for generalizing sample tree information. *Scandinavian Journal of Forest Research* 12, 97-101.

Kramer, H.J. 1994. *Observation of the earth and its environment, survey of missions and sensors.* Springer-Verlag. Berlin, Germany. 580 pp.

Kuusela, K. and Poso, S. 1975. Demonstration of the applicability of satellite data to forestry. *Communicationes Instituti Forestalis Fenniae* 83, 1-31.

Lämås, T. and Fries, C. 1995. Emergence of a biodiversity concept in Swedish forest policy. *Water, Air, and Soil Pollution* 82, 57-66.

Lämås, T. 1996. Forest management planning for biodiversity and timber planning. *Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics.* Report 3. 56 pp.

Lappi, J., Nuutinen, T., and Siitonen, M. 1996. A linear programming software for multi-level forest management planning. In: *Management systems for a global economy with global resource concerns*, 470-482. Proceedings of the symposium on system analysis in forest resources. Pacific Grove, CA.

Larsson, M. 1994. The significance of data quality in compartmental forest registers in estimating volume growth and non-optimal losses. A study of final felling compartments in northern Sweden. *Swedish University of Agricultural Sciences, Department of Biometry and Forest Management.* Report 26. 73 pp. (In Swedish with English summary).

Leckie, D.G. 1990. Advances in remote sensing technologies for forest surveys and management. *Canadian Journal of Remote Sensing* 20, 464-483.

Leckie, D.G. and Ranson, K.J. 1998. Forestry applications using imaging radar. In: *Principles & applications of imaging radar, Manual of remote sensing.* 3rd edition, 435-510. John Wiley & Sons. New York.

Li, C. 1988. Mathematical models in forest resource management planning – an integrated study of calibration, prediction and optimal decision models. *Swedish University of Agricultural Sciences, Department of Biometry and Forest Management.* Report 18. 197 pp.

Lillesand, T.M. and Kiefer, R.W. 1994. *Remote sensing and image interpretation.* 3rd edition. John Wiley & Sons. New York. 750 pp.

Lindgren, O. 1984. A study on circular plot sampling of Swedish forest compartments. *Swedish University of Agricultural Sciences, Department of Biometry and Forest Management.* Report 11. 153 pp.

Linton, O. and Härdle, W. 1998. Nonparametric regression. In: Kotz, S., Read, C.B., and Banks, D.L. (eds.). *Encyclopedia of statistical sciences.* Update vol. 2, 470-485. Wiley. New York.

Loetsch, F. and Haller, K.E. 1973. *Forest inventory – volume 1.* 2nd edition. BLV. München, Germany. 469 pp.

Lönnstedt, L. 1985. De enskilda skogsägarnas avverkningsintensitet – analys av en intervjuundersökning. *Swedish University of Agricultural Sciences, Department of Operational Efficiency.* Report 165. (In Swedish with English summary).

Lundström, A. and Söderberg, U. 1996. Outline of the HUGIN system for long-term forecasts of timber yields and possible cut. *EFI Proceeding* 5, 63-78.

Magnussen, S. and Boudewyn P. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research* 28, 1016-1031.

Manly, B.F.J. 1986. *Multivariate statistical methods – a primer.* Chapman and Hall. New York. 159 pp.

McRoberts, R.E. and Hansen, M.H. 1999. Annual forest inventories for the north central region of the United States. *Journal of Agricultural, Biological, and Environmental Statistics* 4, 361-371.

Means, J.E., Acker, S.A., Harding, D.J., Blair, J.B., Lefsky, M.A., Cohen, W.B., Harmon, M.E., and McKee, W.A. 1999. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sensing of Environment* 67, 298-308.

Moeur, M. and Stage, A.R. 1995. Most similar neighbor - an improved sampling inference procedure for natural resource planning. *Forest Science* 41, 337-359.

Muinonen, E. and Tokola, T. 1990. An application of remote sensing for communal forest inventory. In Proceedings: *The usability of remote sensing for forest inventory and planning*. Swedish University of Agricultural Sciences, Remote Sensing Laboratory. Report 4, 35-42.

Nabuurs, G.J. and Päivinen, R. 1996. Large scale forestry scenario models – a compilation and review. *European Forest Institute*. Working Paper 10. 174 pp.

Næsset, E. 1996. Determination of number of stems in coniferous forest stands by means of aerial photo interpretation. *Scandinavian Journal of Forest Research* 11, 76-84.

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. 1999. Point accuracy of combined pseudorange and carrier phase differential GPS under forest canopy. *Canadian Journal of Forest Research* 29, 547-553.

Nelson, R., Oderwald, R., and Gregoire, T.G. 1997. Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass. *Remote Sensing of Environment* 60, 311-326.

Nilsson, M. 1997. Estimation of forest variables using satellite image data and airborne lidar. *Swedish University of Agricultural Sciences*, *Silvestria* 17.

Nilsson, M. and Sandström, P. 2001. Evaluation of a non-parametric method for estimation of wood volume and biomass in boreal forests by combining remote sensing and field data. *International Journal of Remote Sensing* (in press).

Nyström, K. 2001. Growth models for young stands – development and evaluation of growth models for commercial forests in Sweden. *Swedish University of Agricultural Sciences*, *Silvestria* 180.

Olsson, H. 1994. Monitoring of local reflectance changes in boreal forests using satellite data. *Swedish University of Agricultural Sciences*, *Remote Sensing Laboratory*. Report 7.

Ozbekhan, H. 1969. Toward a general theory of planning. In: *Perspectives of Planning*, OECD. Paris, France.

Päivinen, R. and Anttila, P. 2001. How reliable is a satellite forest inventory? *Silva Fennica* 35, 125-127.

Pickens, J.B., Hof, J.G., and Kent, B.M. 1991. Use of chance-constrained programming to account for stochastic variation in the A-matrix of large-scale linear programs. A forestry application. *Ann. Operations Research* 31, 511-526.

Poso, S. 1972. A method of combining photo and field samples in forest inventory. *Commun. Inst. For. Fenn.* 76.1, 1-133.

Poso, S., Paananen, R., and Simila, M. 1987. Forest inventory by compartments using satellite imagery. *Silva Fennica* 21, 69-94.

Poso, S., Wang, G., and Tuominen, S. 1999. Weighting alternative estimates when using multi-source auxiliary data for forest inventory. *Silva Fennica* 33, 41-50.

Raj, D. 1968. *Sampling theory*. McGraw-Hill. New York. 302 pp.

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

Rao, J.N.K. 1998. Small area estimation. In: Kotz, S., Read, C.B., and Banks, D.L. (eds.). *Encyclopedia of statistical sciences*. Update vol. 2, 621-628. Wiley. New York.

Reese, H. and Nilsson, M. 2000. Wood volume estimations för Älvsbyn kommun using SPOT satellite data and NFI plots. *Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics*. Working Paper 79. 21 pp.

Ringvall, A. 2000. Assessment of sparse populations in forest inventory – development and evaluation of probability sampling methods. *Swedish University of Agricultural Sciences, Silvestria* 151.

Ripley, B.D. 1987. *Stochastic simulation*. John Wiley & Sons. New York. 237 pp.

Ripley, B.D. 1996. *Pattern recognition and neural networks*. University Press. Cambridge, UK. 403 pp.

Ripple, W.J., Wang, S., Isaacson, D.L., and Paine, D.P. 1991. A preliminary comparison of Landsat Thematic Mapper and SPOT-1 HVR multispectral data for estimating coniferous forest volume. *International Journal of Remote Sensing* 12, 1971-1977.

Rubin, D.B. 1987. *Multiple imputation for nonresponse in surveys*. Wiley. New York. 258 pp.

Rülcker, C., Angelstam, P., and Rosenberg, P. 1994. Ecological forestry planning – a proposed planning model based on the natural landscape. *SkogForsk*. Redogörelse 8. 44 pp. (In Swedish with English summary).

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

Schreuder, H.T., Gregoire, T.G., and Wood, G.B. 1993. *Sampling methods for multisource forest inventory*. John Wiley & Sons. New York. 446 pp.

Sigrist, P., Coppin, P., and Hermy, M. 1999. Impact of forest canopy on quality and accuracy of GPS measurements. *International Journal of Remote Sensing* 18, 3595-3610.

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

Siitonen, M. and Nuutinen, T. 1996. Timber production analyses in Finland and the MELA system. *EFI Proceedings* 5, 89-98.

Skogsstyrelsen. 1994. Skogsvårdslagen – handbok. Jönköping, Sweden. 66 pp. (In Swedish).

Smith, G. 2000. Radar remote sensing of forests using CARABAS and ERS. *Chalmers University of Technology, Department of Radio and Space Science*. Technical Report No. 393.

Söderberg, U. 1986. Functions for forecasting of timber yields – increment and form height for individual trees of native species in Sweden. *Swedish University of Agricultural Sciences, Section of Forest Mensuration and Management*. Report 14. 251 pp. (In Swedish with English summary).

Sprängare, B. 1975. A method for analysing the sensitivity of the long-range planning which depends on errors in stand data. *Swedish University of Agricultural Sciences, Department of Forest Engineering*. Report 87. 63 pp. (In Swedish with English summary).

Ståhl, G. 1992. A study on the quality of compartmentwise forest data acquired by subjective inventory methods. *Swedish University of Agricultural Sciences, Department of Biometry and Forest Management*. Report 24. 127 pp. (In Swedish with English summary).

Ståhl, G. 1994. Optimizing the utility of forest inventory activities. *Swedish University of Agricultural Sciences, Section of Forest Mensuration and Management*. Report 27.

Stellingwerf, D.A. and Hussin, Y.A. 1999. *Measurements and estimations of forest stand parameters using remote sensing*. VSP. Utrecht, The Netherlands. 273 pp.

Talts, J. 1977. Photogrammetric measurements for stand cruising. *Royal College of For., Dept. of Forest Mensuration and Management*. Research Notes 6. 102 pp. (In Swedish with English summary).

Thompson, S.K. 1992. *Sampling*. John Wiley & Sons. New York. 343 pp.

Thuresson, T. 1995. Forest valuation using a two phase inventory method and satellite data. *Scandinavian Journal of Forest Research* 10, 65-73.

Tokola, T. 2000. The influence of field sample data location on growing stock volume estimation in Landsat TM-based forest inventory in eastern Finland. *Remote Sensing of Environment* 74, 422-431.

Tokola, T. and Heikkilä, J. 1997. Improving satellite image based forest inventory by using a priori site quality information. *Silva Fennica* 31, 67-78.

Tokola, T., Pitkänen, J., Partinen, S., and 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. 1988. Standwise forest variate estimation by means of satellite images. In: *IUFRO Proceedings* S4.02.05, Forest Station Hyytiälä. University of Helsinki, Department of Forest Mensuration and Management. Research Notes 21, 103-111.

Tomppo, E. 1990. Designing a satellite image-aided national forest survey in Finland. In Proceedings: *The usability of remote sensing for forest inventory and planning*. Swedish University of Agricultural Sciences, Remote Sensing Laboratory. Report 4, 43-47.

Tomppo, E. 1991. Satellite image-based national forest inventory of Finland. In Proceedings: Symposium on global and environmental monitoring, techniques and impact. Victoria, BC. *International Archives of Photogrammetry and Remote Sensing* 28, 419-424.

Tomppo, E. 1993. Multi-source national forest inventory of Finland. In: *Symposium on National Forest Inventories*, 17-21 August 1993, Ilvessalo, 52-59.

Tomppo, E., Goulding, C., and Katila, M. 1999. Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. *Scandinavian Journal of Forest Research* 14, 182-192.

Tomppo, E. and Katila, M. 1992. Multi-source national forest inventory of Finland. *International Archives of Photogrammetry and Remote Sensing* 29, 671-674.

Törnquist, K. 1996. Ecological landscape planning in Swedish forestry. In Proceedings: *Integrating environmental values into forest planning.* EFI No. 13.

Trotter, C.M., Dymond, J.R., and Goulding, C.J. 1997. Estimation of timber volume in coniferous plantation forest using Landsat TM. *International Journal of Remote Sensing* 18, 2209-2223.

Tuominen, S. and Poso, S. 2001. Improving multi-source inventory by weighting auxiliary data sources. *Silva Fennica* 35, 203-214.

Ulander, L.M.H., Askne, J., Fransson, J.E.S., Gustavsson, A., Le Toan, T., Manninen, T., Martinez, J.-M., Melon, P., Smith, G., and Walter, F. 2000. Retrieval of stem volume in coniferous forest from low VHF-band SAR. In: *Proceedings of IGARSS 2000 Symposium*, *IEEE*, 441-443. Piscataway, NJ.

Wallerman, J. 1998. Brattåkerinventeringen – dokumentation av databasen och avdelningsskattningarna. *Sveriges lantbruksuniversitet, Institutionen för skoglig resurshushållning och geomatik*. Arbetsrapport 43. 37 pp. (In Swedish).

Walter, F. 1999. Extraction of forest stand parameters from CARABAS VHF SAR images. *Swedish University of Agricultural Sciences, Silvestria* 115.

Walter, F., Fransson, J.E.S., and Frölind, P.-O. 1999. Fully automatic geo-coding of CARABAS-II VHF SAR images. In: *Proceeding of IGARSS Symposium*, 569-573.

Weintraub, A. and Abramovich, A. 1995. analysis of uncertainty of future timber yields in forest management. *Forest Science* 41, 217-234.

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

Wikström, P. 2000. Solving stand-level planning problems that involve multiple criteria and single-tree growth model. *Swedish University of Agricultural Sciences, Silvestria* 167.

Yu, J. 1994. Nearest neighbor probability density estimators. Doctoral dissertation. *University of Umeå, Institute of Mathematical Statistics*.

Acknowledgements

There are, I have been told, two ways of expressing one's gratitude to the people whose support helped one complete a commitment like a doctoral thesis. One can mention them all, with a sentence explaining how each and every one has contributed. In my case, this would truly be a very long list. The other way is to name no one individually, and just say 'Thank you all'. Here, I will try something in between. I must name and acknowledge a sample of outstanding people who gave inestimable help.

Prof. Göran Ståhl, who never failed, even for a single moment, in his duty as my supervisor. Careful guidance was always guaranteed. I extend my greatest thanks to Göran and his family, since I am fully aware of all the precious time I have occupied.

I also thank:
Prof. Håkan Olsson, Dr. Mats Nilsson, and Dr. Johan Fransson, who collectively formed the best possible group of co-supervisors.
Prof. Ljusk Ola Eriksson, the leader of the pack, who was also involved in initiating the 'Housekeeping' research project.
Torleif Carlsson, the project leader, who was always willing to share his deep and pragmatic knowledge of forestry planning.
Hans Kallur, an excellent programmer with the ability to both pinpoint and solve almost every problem. Furthermore, with experience from several decades of forest research, a true mentor.
My beautiful daughters, Felise and Frances. You brought light into my life. Really, nothing else matters to me. This one goes to the two I love.

And to all the others:
sponsors (especially SkogForsk, the Forestry Research Institute of Sweden, for supplying half of the funding for the project), colleagues (*e.g.*, Dr. Ulf Söderberg and Dr. Karin Öhman, for extraordinary care), teachers, students, my friends and my family, thanks.