Dissertationes Forestales 97

Uncertainty in forest simulators and forest planning systems

Antti Mäkinen
Department of Forest Resource Management
Faculty of Forestry and Agriculture
University of Helsinki

Academic dissertation

To be presented with the permission of the Faculty of Agriculture and Forestry, University of Helsinki, for public critisism in Infokeskus II, Viikinkaari 11, Helsinki, on February 12th 2010 at 12 o'clock.

Title of dissertation: Uncertainty in forest simulators and forest planning systems

Author: Antti Mäkinen

Dissertationes Forestales 97

Thesis Supervisors:
Prof. Annika Kangas
Department of Forest Resource Management, University of Helsinki, Finland
Prof. Timo Tokola
Faculty of Forestry, University of Joensuu, Finland

Pre-examiners:
Prof. Mikko Kurttila
Finnish Forest Research Institute, Finland
Dr. Mathieu Fortin
Ministère des Ressources naturelles et de la Faune du Québec, Canada

Opponent:
Prof. Göran Ståhl
Department of Forest Resource Management, SLU, Umeå, Sweden

ISSN 1795-7389
ISBN 978-951-651-283-2 (PDF)

(2010)

Publishers:
Finnish Society of Forest Science
Finnish Forest Research Institute
Faculty of Agriculture ad Forestry of the University of Helsinki
Faculty of Forest Sciences of the University of Joensuu

Editorial Office:
Finnish Society of Forest Science
P.O. Box 18, FI-01301 Vantaa, Finland
http://www.metla.fi/dissertationes

Mäkinen, A. 2009. Uncertainty in forest simulators and forest planning systems. Dissertationales Forestales 97. 38 p. Available at http://www.metla.fi/dissertationes/df97.htm

The forest simulator is a computerized model for predicting forest growth and future development as well as effects of forest harvests and treatments. The forest planning system is a decision support tool, usually including a forest simulator and an optimisation model, for finding the optimal forest management actions.

The information produced by forest simulators and forest planning systems is used for various analytical purposes and in support of decision making. However, the quality and reliability of this information can often be questioned. Natural variation in forest growth and estimation errors in forest inventory, among other things, cause uncertainty in predictions of forest growth and development. This uncertainty stemming from different sources has various undesirable effects. In many cases outcomes of decisions based on uncertain information are something else than desired.

The objective of this thesis was to study various sources of uncertainty and their effects in forest simulators and forest planning systems. The study focused on three notable sources of uncertainty: errors in forest growth predictions, errors in forest inventory data, and stochastic fluctuation of timber assortment prices. Effects of uncertainty were studied using two types of forest growth models, individual tree-level models and stand-level models, and with various error simulation methods. New method for simulating more realistic forest inventory errors was introduced and tested. Also, three notable sources of uncertainty were combined and their joint effects on stand-level net present value estimates were simulated.

According to the results, the various sources of uncertainty can have distinct effects in different forest growth simulators. The new forest inventory error simulation method proved to produce more realistic errors. The analysis on the joint effects of various sources of uncertainty provided interesting knowledge about uncertainty in forest simulators.

Keywords: forest planning, growth model, uncertainty, inventory error

Mäkinen, A. 2009. Uncertainty in forest simulators and forest planning systems. Dissertationales Forestales 97. 38 p. Available at http://www.metla.fi/dissertationes/df97.htm

Metsäsimulaattori on tietokoneistettu malli, jolla ennustetaan metsän kasvua ja tulevaa kehitystä, sekä hakkuiden ja käsittelyiden vaikutuksia metsiin. Metsäsuunnittelujärjestelmä on päätöstukijärjestelmä, joka yleisesti koostuu metsäsimulaattorista ja optimointimallista ja jonka avulla metsien käsittelyitä voidaan optimoida.

Metsäsimulaattoreilla ja metsäsuunnittelujärjestelmillä tuotettua tietoa käytetään monenlaisiin analyyseihin sekä metsien käyttöä koskevan päätöksenteon perustaksi. Tiedon laatu ja luotettavuus on kuitenkin usein kyseenalaista. Esimerkiksi metsien kasvun luontainen vaihtelu sekä virheet metsien nykytilaa koskevissa tiedoissa aiheuttavat metsien tulevan kehityksen ennusteissa epävarmuutta. Epävarmuudella, tai tiedon puutteella on monia epäedullisia seurauksia. Epävarmalle tiedolle perustuvat päätökset johtavat usein ei-toivottuun lopputulokseen.

Tämän väitöskirjan tarkoituksen oli tarkastella erilaisia epävarmuuden lähteitä sekä niiden vaikutuksia metsäsimulaattoreissa ja metsäsuunnittelujärjestelmissä. Tutkimuksessa tarkasteltiin pääasiassa kolmea merkittävää epävarmuuden lähdettä: metsien kasvuennusteiden virheitä, metsien nykytilaa kuvaavan tiedon virheitä sekä puutavaralajien hintojen satunnaisvaihtelua. Epävarmuuden seurauksia tarkasteltiin sekä yksittäisten puiden että metsikkötason kasvumalleilla ja käyttäen erilaisia virheiden simulointimenetelmiä. Tutkimuksessa kehitettiin uusi menetelmä entistä realistisempien metsien inventointivirheiden simulointiin. Lisäksi työssä tarkasteltiin kolmen merkittävän epävarmuustekijän yhteisvaikutuksia metsikkötason nettonykyarvojen ennustamisessa.

Tutkimuksen tärkeimmät tulokset osoittivat eri epävarmuuden lähteiden vaikuttavan selvästi eri tavoin eri metsäsimulaattoreissa. Työssä esitetyllä inventointivirheiden simulointimenetelmällä voidaan jatkossa tuottaa selvästi realistisempia virhejakaumia. Eri epävarmuustekijöiden yhteisvaikutusten tarkastelu syvensi tietämystä epävarmuuden vaikutuksista metsäsimulaattoreissa.

Asiasanat: metsäsuunnittelu, kasvumalli, epävarmuus, inventointivirhe

ACKNOWLEDGEMENTS

There are many people that have somehow contributed to the process of completing this thesis. They all deserve my most sincerest gratitudes. First and foremost I would like to thank my supervisors Professors Annika Kangas and Timo Tokola. The help, encouragement, and ideas I have gotten from the both of you during these years have been priceless.

During the last four years, I've had the pleasure of working with Doctor Jussi Rasinmäki and Jouni Kalliovirta, first at the Department of Forest Resource Management and later at Simosol Oy. This co-operation has been productive, I have learned a great deal, and what is important, it has been very enjoyable. I have no reason to think that it would not continue that way in the future also.

I have had the opportunity to cooperate with many researchers and co-authors before and during this thesis. My first instructor to scientific work was Doctor Ilkka Korpela, whose enthusiasm and determination has been a great source of inspiration for me. I want to thank Doctor Markus Holopainen for the cooperation in various projects and for all of the support I have had. Doctor Lauri Mehtätalo, who acted as my stand-in supervisor for one year, increased my knowledge in statistics, offered me numerous excellent ideas, and provided detailed comments about this thesis. I would also like to thank all my co-authors: Doctor Kari Hyytiäinen, Ilona Pietilä, Esko Välimäki and Saeed Bayazidi. Professor Mikko Kurttila and Doctor Mathieu Fortin acted as the official pre-examiners of this thesis and provided excellent comments, thus helping me to improve this thesis. I want to warmly thank both of them.

The people at the soon to become Department of Forest Sciences have contributed to the pleasant working environment and made the daily working life enjoyable, of which I am grateful. Luckily, there is also life outside the University walls. I want to thank all of my friends, both at the university and elsewhere, in no particular order for the great times we've had and surely will be having. My whole family has always supported me in all my aspirations without any preconditions, which makes me feel extremely privileged. Finally, I want to offer my warmest thanks to Minttu for making our home a much sunnier place.

LIST OF ORIGINAL ARTICLES

This dissertation consists of a summary and the four following articles, which are referred to by roman numerals I-IV. Articles I and II are reprints of previously published articles, reprinted with the permission of the publisher. Articles III and IV are submitted manuscripts.

I Mäkinen, A., Kangas, A., Kalliovirta, J., Rasinmäki, J. & Välimäki, E. 2008. **Comparison of treewise and standwise forest simulators by means of quantile regression.** Forest Ecology and Management 255: 2709-2717.

II Mäkinen, A., Kangas, A., Holopainen, M. & Rasinmäki, J. 2009. **Propagating the errors of initial forest variables through stand- and tree-level growth simulators.** European Journal of Forest Research. Published Online. DOI: 10.1007/s10342-009-0288-0

III Mäkinen, A., Kangas, A. & Mehtätalo, L. **Correlations, distributions and trends of forest inventory errors and their effects on forest planning.** Manuscript.

IV Holopainen, M., Mäkinen, A., Rasinmäki, J., Hyytiäinen, K., Bayazidi, S. & Pietilä, I. 2009. **Comparison of various sources of uncertainty in stand-level net present value estimates.** Manuscript.

Antti Mäkinen designed and implemented the simulation studies, analyzed the results and was the main author in Papers I, II and III. In Paper IV, Mäkinen was responsible for implementing the simulation system, performing the simulations and analyzing the results. Paper IV was written together by all authors.

CONTENTS

INTRODUCTION

Forest simulators and forest planning systems

The forest simulator is a tool for predicting the future development of forests. Forest simulators typically consist of multiple sub-models, or growth equations, that together form a simplified representation, or model, of the forest ecosystem. This model describes the various processes that take place in a forest, such as growth, mortality and other changes in the forest structure. Common uses for forest simulators include updating previously measured forest data and predicting future forest development. Predictions of future development are used for evaluating silvicultural treatments, management planning and harvest scheduling (Burkhart 1993). Another commonly used term for a forest simulator is a *forest growth model*. In this thesis, the term 'forest growth model' refers to the abstract representation of forest dynamics, and the term 'forest simulator' refers to a computerised growth modeling system (e.g. the implementation of the growth model).

The first, but still widely used, forest growth models, known as yield tables, were already in use in the late 18th century (Vuokila & Väliaho 1980). A simple yield table could be, for example, a tabular representation of forest attribute values, such as the total volume or basal area of the trees, at given ages. Nowadays, the term 'forest growth model' covers a vast number of models that vary in complexity as well as in the theoretical framework on which they are based. Thus forest growth models can be divided into various categories. One common convention is to divide growth models into empirical models and mechanical process models. Typically, empirical models are estimated from measured data by using statistical methods and are usually based on some common growth equations (Zeide 1993). Process models aim at modeling the detailed eco-physiological processes of individual trees and are based on ecological theory (Mäkelä et al. 2000, Kokkila et al. 2006). Traditionally, empirical models have served mostly for prediction, whereas process models have been used to understand the various natural processes in forest ecosystems. This division is not necessarily so strict, as most of the models lie somewhere between purely data-driven models and theory-based models. In fact, good and usable models should combine a solid theoretical background with parameters estimated from data.

Another commonly used separation is categorising forest simulators by their level of organisation. According to Munro (1974), forest growth models can be divided into stand-level models, distance-independent tree-level models and distance-dependent tree-level models. Tree-level models predict the growth of individual trees, and stand-level models predict the growth of some aggregate variable. Distance-dependent, or spatial models, incorporate information about neighbouring trees and other spatial indices describing, for example, within-forest competition (Tomé & Burkhart 1989, Vettenranta 1999). The distinction between stand-level and tree-level models is not so strict as there is rather a continuum of different model types between the two levels, including, for example, diameter distribution, or size class, models (Vanclay 1994). Besides the aforementioned forest growth model types, there is a number of other types, such as succession models (Shugart & West 1980) and transition matrix models (Buongiorno & Michie 1980, Kolström 1993).

This thesis focuses strictly on empirical forest growth models, as they are more commonly used in practical forest planning computations. A notable trend in the development of empirical forest growth models has been a shift from stand-level models to tree-level models. Stand-level growth models can have sufficiently good predictive qualities, especially in even-aged forests, and they are reportedly more accurate than tree-level models in some cases (Burkhart 2003). Stand-level growth models are usually less complex than tree-level growth models, which makes them computationally more efficient (Vanclay 1994, Atta-Boateng & Moser 2000). On the other hand, stand-level models cannot necessarily describe the complex dynamics of forests, such as between-tree competition (Porté & Bartelink 2002). Also, a *forest stand* is not as easily conceivable a biological entity as a single tree is (Garcia 2001). Huston et al. (1988) stated that biological processes should be modeled using individual-based models, such as tree-level models, rather than aggregated models, such as stand-level models.

The set of trees in a tree-level growth model is either empirical (i.e. trees measured in the field) or theoretical (i.e. set of trees generated using a distribution model) (Kangas & Maltamo 2003). Forest dynamics are easier to capture with tree-level models, and for mixed forest or forests with an otherwise complex structure, tree-level models may be the only feasible option. Although complex interactions and dynamics between trees can be incorporated into tree-level growth models, the actual interactions are sometimes so complex that even tree-level models cannot take them into account (Zeide 1993, Sterba et al. 2002).

Despite the trend of favouring tree-level growth models, stand-level growth models can provide a good alternative for tree-level models in certain situations (Garcia 2001). And more important than the level of aggregation is that the model is biologically and logically sound, its statistical properties are satisfactory, and that the reliability of the predictions is sufficient (Vanclay & Skovsgaard 1997). The goodness of a model is always dependent on the use of the model also; if high accuracy and precision are unnecessary, then a less accurate and precise model will suffice.

In general, forest planning aims to identify the optimal way to utilise forest resources. This usually means maximising the forest owner's utility, limited by a given set of constraints (Pukkala 2002). Forest planning is generally carried out with the aid of a *forest planning system* (FPS). In a broad sense, the term forest planning system can refer to the whole forest planning framework, beginning with the collection of field data and covering each step in the planning process. These steps can include, for example, preprocessing, validation and storage of inventoried data, interaction between forest planners, forest owners and other stakeholders, assessment of alternative activities using simulation and optimisation models, and reporting of the resulting forest plan. In addition, many steps of this process can be iterative and controlled by various laws and regulations, making the whole forest planning system very complex. On the other hand, the term forest planning system can also refer to a simple computerized decision support tool that is often used for small or medium scale forest planning tasks. The repertory of existing FPSs is vast and diverse, as systems are available in different countries and regions for different types of forest planning tasks and on various scales. Examples of FPSs in Finland include MELA (Hynynen et al. 2005), MOTTI (Salminen et al. 2005), MONSU (Pukkala 2004), and SIMO (Tokola et al. 2006, Rasinmäki et al. 2009). Commonly used FPSs in the other Nordic countries include Gaya-JLP (Hoen & Gobakken 1990), Forest Management Planning Package FMPP (Jonsson et al. 1993), Hugin (Lundström & Söderberg 1996), Avvirk-2000 (Eid & Hobbelstad 2000), Heureka (Lämås & Eriksson 2003) and T (Gobakken 2008). The core of an FPS is its forest simulator, which is used for simulating possible future developments. Many FPSs also include an optimisation model that is used for choosing the optimal management activities from a group of alternative simulated treatment schedules. The optimisation models in various FPSs have been either exact, such as linear programming algorithms, or heuristic algorithms (Pukkala & Kangas 1993, Pukkala & Kurttila 2005, Heinonen 2007). FPSs commonly have different types of user interfaces and interfaces with GIS. In this thesis, FPS refers to a computer *decision support system* (DSS) that consists of a forest simulator and an optimisation model for selecting optimal management activities from a set of simulated alternatives. In this thesis, the term 'forest planning system' refers to a forest planning software containing simulation and optimisation models.

Sources of uncertainty in forest planning systems

Forest simulators are often quite complicated systems, as are forest planning systems, which, in addition to the simulator, include the optimisation model as well. When such a complex system is used to predict natural processes, such as forest development, the outputs contain a lot of uncertainty, which stems from many sources. The term *uncertainty* can be understood as a lack knowledge. This can mean, for example, unknown probabilities for various events or unknown distributions of some variables of interest (Pukkala & Kangas 1995). For example, in the forest planning context we may not know the exact current state of a forest, as the forest inventory data contain errors. Alternatively, we cannot know the exact states of a forest stand in the future, as growth predictions contain errors. In forest planning systems, uncertainty also stems from random variation in timber assortment prices, known market risk, and the risk of natural hazards. Even if uncertainty precludes the knowledge of exact values, such uncertainty can be quantified and analysed in many ways. A common way to quantify uncertainty is to estimate the distributions that are associated with and influenced by uncertainty. Uncertainty in forest growth predictions, for example, could be described with bias and variance in the predicted forest attribute values at different points in time. Different approaches to defining, classifying and managing uncertainty in decision making have been discussed comprehensively in Kangas & Kangas (2004).

The concept of *risk* is closely related to uncertainty and, in most cases, bears a negative connotation to it. In the forest planning context, risk can mean, for example, the risk of a forest to incur damage in a forest fire or windstorm. The probability of risk can be quantified if the distributions affiliated with the risk are known. Risk is typically described by the severity of realised risk, the *potency*, and the likelihood that the risk will be realized, or the *exposure* (Mowrer 2000).

Predictions of forest growth and yield inherently contain some measure of uncertainty, which stems from: (1) *model misspecification*, (2) *random estimation errors of the model coefficients*, (3) *residual variation in the model*, and (4) *errors in the independent variables of the models*. The errors in the independent variables may result from sampling errors, measurement errors, grouping errors and prediction errors (Kangas 1999). These sources of uncertainty apply especially to statistical models. The prediction errors can be either systematic, which is described with *bias*, or random, which is described with *variance*. One of the difficult things about predicting future developments in natural systems, such as forest growth, is that natural processes tend to involve a lot of random variation. A large proportion of the uncertainty related to growth prediction error stems directly from this natural variation. Even though the predictive properties of a forest growth model can be good, the actual growth in individual forests is not exactly the same as the prediction because the predicted growths are based on averages (Mowrer 2000).

Forest inventory provides the input data for forest planning calculations and describes the current state of the forests. Because forest areas are very large, it is practically impossible to make accurate and extensive measurements that would provide an exact description of the forest properties. Instead, the properties of only a small sample are normally measured or estimated. In many cases the measurements or estimates are somewhat inaccurate, and contain both bias and variance, which lead to uncertainty in the forest inventory data. Traditionally, forest planning data in Finland are collected with a stand-level field inventory in which stand-level aggregate attributes are partly visually estimated and partly measured from subjectively located sample plots (Poso 1983, Laasasenaho & Päivinen 1986, Haara 2003, Kangas et al. 2004, Saari & Kangas 2005). The subjective nature of this method makes it prone to assessment errors, and the quality of the data collected with this method is considered low. Assessment errors in a forest inventory can result from (1) *measurement errors*, (2) *sampling errors*, (3) *prediction errors* and (4) *classification errors* (Gertner 1986, 1991, Haara 2005).

Remote sensing methods have been applied in forest inventory in recent decades, but have only recently become a feasible alternative to stand-wise field inventory. Estimation methods utilising airborne laser scanning (ALS) data and digital photogrammetry can now provide data of comparable reliability to traditional stand-level field inventory data and at competitive cost (Naesset 1997, Korpela 2004, Naesset 2004, Packalen 2009). The ALS estimation and digital photogrammetry-based estimation methods can provide accurate estimates of some stand-level attributes, such as mean height and number of stems, but the accuracies of some other attributes, such as mean diameter and tree species, are poor. The aforementioned remote sensing methods can also provide tree-level data, such as the locations and dimensions of individual trees instead of aggregate attributes (Kaartinen & Hyyppä 2008). Even though these novel remote sensing methods are quite accurate, their estimates contain both systematic and random variation, and thus are uncertain.

Besides the uncertainty resulting from errors in forest growth models and forest inventory data, economic aspects also lead to uncertainty in the forest planning context. One such source of uncertainty is the random fluctuation in timber assortment prices, also known as market risk or price risk. Market risk means that, because future prices are unknown, the forest owner cannot know when is the optimal time to sell timber in order to earn maximal profits. Thus the exact net present values (NPV) of future income from timber sales cannot be known. Consequently, the fluctuating timber assortment prices cause uncertainty about when to harvest and what is the theoretical value of forest when the value estimate is based on future timber sales. Newman (2002) has extensively reviewed the literature on economic uncertainty in forestry.

Natural hazards also constitute a source of uncertainty in forest planning. The hazards are usually events that damage forests and decrease their value. Potential hazards to forests include wildfires, high winds, heavy snow, flooding, insects and pathogens. Risk and hazards in forest planning, especially in harvest planning and scenario modelling, have been studied by, for example, von Gadow (2000).

Additional important source of uncertainty in forest planning process is the preferences of the forest owner. In many cases, the forest owner may have various objectives, both quantitative and qualitative, which can also be conflicting. For forest owners, as well as for forest planning experts, defining the objectives unambiguously in multi-objective forest planning can be challenging. This is even more difficult in participatory forest planning, in which the number of stakeholders with different objectives can be substantive. A number tools for aiding the decision making process have been developed, such as Analytical Hierarchical Process (AHP) (Saaty 1980), which has been utilized, along with other similar tools, for some time in forest planning context (Kangas 1992, Kangas et al. 2002).

Analysing the uncertainty

Uncertainty and its consequences in forest simulators and forest planning systems have been studied using different approaches. One typical approach has been to study the variance of a certain variable of interest, such as stand volume, and to see how a given source of uncertainty affects it over a period of time. This type of analysis can be carried out using different methods and has especially been used in studying uncertainty due to growth prediction errors. One possibility is to conduct a straightforward empirical simulation study using the MC method. Kangas (1997, 1999) used this technique for examining growth prediction errors and (Gertner & Dzialowy 1984) the effects of forest inventory errors. Another option is to examine the distribution of the variable of interest analytically using the Taylor series approximation or similar error propagation method. Gertner (1987) and Mowrer (1991) adopted this approach for examining growth prediction errors. Gertner (1987) reported that using an error propagation method instead of crude the MC simulation reduced computational cost by a factor of 2000. However, the MC method is straightforward and in many cases easier to implement than, for example, the Taylor series approximation for complex model systems. Moreover, according to Gertner (1987), the analytical method underestimated the uncertainty in the growth predictions. Another type of analytical approach was adopted by Ståhl (1994) for determining the optimal timings for harvest and inventory actions at stand-level. This approach used Bayesian theory and the decision variables were probability distributions instead of point estimates. Yet another approach for evaluating uncertainty in growth predictions is simply to compare predicted growths to observed growths (Välimäki & Kangas 2009, Haara & Leskinen 2009). This type of approach is suitable for the validation and comparison of alternative forest simulators. Studying the variances of various attributes in a simulation system can provide confidence intervals for the attributes and thus help to estimate the reliability of the whole system. Typical measures of uncertainty in these studies include standard error (SE), root mean square error (RMSE), and coefficient of variation (CV).

Though the variance of the various attributes of a forest simulator tells something about the uncertainty, it does not necessarily provide an accurate impression of the practical implications of the uncertainty. One approach for elucidating these practical implications has been calculating so-called *inoptimality losses*. The idea of inoptimality losses is based on the assumption that if we can determine the optimal way to manage a forest, then every other management option that differs from the optimal one is either equal or inferior to it. When the decision-making process, such as the scheduling of harvests, involves sources of uncertainty, the decision may end up being something other than optimal. If an inoptimal management option is chosen, the maximum utility, in many cases quantified as the NPV of the forest stand or holding, cannot be attained. Thus, inoptimality loss is simply the difference between the utilities of inoptimal and optimal management options. Inoptimality loss can be used to estimate the *value of information* (VOI) or the *expected value of perfect information* (EVPI), the theoretical value that a decision-maker would be willing to pay for such information before making a decision. Such an approach has especially been used in studying the effects of forest inventory errors in forest planning. Alternative forest inventory methods have been compared using *cost-plus-loss* (CPL) analysis, where both the inoptimality losses and the costs of the alternative inventory methods are considered (Burkhart et al. 1978, Hamilton 1978). The use of CPL analysis in the forest planning context has been covered extensively by Duvemo & Lämås (2006). In management science context, uncertainty and it's effects on decision analysis have been studied using an approach called Robust Portfolio Modeling (RPM) (Liesiö et al. 2007). In RPM approach, robust portfolios are formed from subsets of decision alternatives, which in forest planning context could be alternative treatment schedules. The aim is to find such combinations, or portfolios, that are not likely to change even if the uncertainty would decrease, i.e. more accurate forest inventory data would be available.

Forest inventory errors used for studying inoptimality losses and in CPL analysis have been either real errors observed in an actual forest inventory in case-study fashion or simulated random errors. Real observed errors have been used, for example, by Eid et al. (2004), Duvemo et al. (2007), Juntunen (2006), Väisänen (2008) and simulated errors by Eid (2000), Holmström et al. (2003), Holopainen & Talvitie (2006), Borders et al. (2008), Islam et al. (2009). One of the strengths of the case-study approach is that the observed errors are realistic and can be used as such. However, using observed errors requires a lot of data and is restricted to existing inventory methods. Using simulated errors gives more flexibility as the errors are not restricted to existing inventory methods, and real inventory data is not needed. On the other hand, in order to simulate realistic errors, the distributions and correlation structures of the errors must be known and taken into account. Most of these studies assume the errors are normally distributed. This, however may, according to Canavan & Hann (2004) and Westfall & Patterson (2007), be an invalid assumption. Moreover, many previous studies have disregarded the possibility that the errors can be correlated, although Sprängare (1978) and Eid (1993), for example, suggest that the errors are in fact correlated.

Existing studies on economic uncertainty have concentrated mostly on optimal harvest scheduling. Traditionally optimal harvest schedules have been determined using Faustmann's rule, which maximises the NPVs of future harvest income. This rule assumes constant timber prices, but in reality prices fluctuate, which affects optimal harvest scheduling. Random fluctuation in timber prices reportedly increases rotation ages and forest NPVs (Brazee & Mendelsohn 1988, Haight & Holmes 1991, Thomson 1992, Tahvonen & Kallio 2006). Various models for stochastic timber assortment prices appear in the existing literature (Yoshimoto & Shoji 1998, 2002, Insley & Rollins 2005). In most cases, economic risks in the forest planning context have been studied separately from other sources of uncertainty, but exceptions do exist. The effects of economic uncertainty have been studied concurrently with growth prediction errors (Pukkala & Kangas 1995, Pukkala & Miina 1997) and natural hazards (Valsta 1992). Reed & Haight (1996) took both growth prediction uncertainty and natural hazards into account when determining NPV distributions for forest stands.

Due to the multifaceted and complex nature of the uncertainty in forest simulators and FPS, various simplifications have been made in existing studies. Most existing studies, such as those on growth prediction errors and CPL analysis of forest inventory errors, consider only one source of uncertainty at a time. One of the reasons behind CPL analysis is that it produces a more tangible and realistic measure of uncertainty. However, including only a single source of uncertainty inevitably decreases the realism of the CPL analysis. The effects of growth prediction errors and economic uncertainty have been studied concurrently, but no studies to date have studied uncertainty due to forest inventory errors at the same time with growth prediction errors and economic uncertainty.

OBJECTIVES OF THE THESIS

The general objective of this thesis was to study different sources of uncertainty and their effects in forest simulators and forest planning systems. The sources of uncertainty considered in this thesis were: 1. uncertainty due to forest inventory errors, 2. uncertainty stemming from growth prediction errors, and 3. economic uncertainty resulting from stochastic timber assortment prices. Different forest growth models presumably respond to uncertainty in different ways, thus two alternative forest growth models were used in the various simulations of this thesis. In addition, different forest inventory methods were considered when studying the effects of forest inventory errors.

One of the main objectives was to study whether notable differences related to uncertainty exist between tree-level and stand-level growth models (Papers I and II). Other, more specific questions in this study were: "Can quantile regression provide more information about the growth prediction error distributions than standard regression analysis?" (Paper I) and "How are forest inventory errors propagated in different forest simulators?" (Paper II).

In order to understand the practical implications of uncertainty, Paper III explored economic inoptimality losses resulting from forest inventory errors. One objective was to examine the detailed properties of forest inventory errors and to determine how the different properties of simulated errors affect inoptimality losses. A related objective was to outline and to test a method for simulating realistic forest inventory errors (Paper III).

One aim was to acquire an understanding of the proportions and joint effects of various sources of uncertainty in a fairly complex forest simulator by examining the distributions of predicted stand-level NPVs. Another objective was to obtain an understanding of the level of accuracy in stand-level NPV predictions (Paper IV).

MATERIALS

Tree increment sample plot measurements (I)

The dataset used as the reference for forest growth predictions in Paper I consisted of 60 sample plots located in 30 forest stands. The stands were located in central Finland, and the field measurements were conducted in summer 2005. Because one of the aims was to evaluate the performance of growth models in extreme conditions, the selected stands had been unmanaged over the past 20 years and were denser than typical commercially managed forests in Finland. A wide repertoire of different site and age classes was represented in stands dominated by both Norway spruce (*Picea abies* L. Karst.) and Scots pine (*Pinus sylvestris* L.).

The sample trees were measured from two types of circular sample plots: *sample tree plots* and *tally tree plots*, both of which had a radius relative to the stand density. The sample tree plots and tally tree plots overlapped with a joint centre point, but the radius of a sample tree plot was half that of a tally tree plot. In all, 1580 tally trees and 490 sample trees were measured from the sample plots.

The diameter at breast height (dbh) of the tally trees was measured with a caliber, and the species was recorded. For the sample trees, tree height (h), height of crown, thickness of bark and the health of the tree were also recorded. Additionally, each pine and spruce sample tree was also cored at breast height in order to determine the age and dbh growth over the past 20 years by analysing the annual rings.

The sample trees were used to construct linear mixed effects models for predicting dbh, h and age of the tally trees at the present moment and over the past 20 years. This way the reference data included a total of 2070 trees with either measured or predicted dbh, h and age both at present and 5, 10, 15 and 20 years ago. Also, stand level aggregate attributes were calculated from the tree-level values for basal area (G), mean diameter (D_{gM}), mean height (H_{gM}), total volume (V), number of stems per hectare (N) and mean age (Age).

Stand-level field inventory estimates, ALS inventory estimates and reference sample plot measurements (III, IV)

One of the aims of Papers III and IV was to study the detailed properties of assessment errors in different forest inventory methods. The two inventory methods were traditional stand-level field inventory and area-level ALS estimation. The stand-level field inventory dataset (hereafter, FIELD) included estimates by experienced forest planners for 1158 stands and accurate reference measurements from sample plots located within these stands. The aggregate attributes estimated for the stands and measured from the sample plots, stratified by tree species in both cases, were: D_{gM}, H_{gM}, G, V, and N. A more detailed description of the field data collection appears in Haara & Korhonen (2004).

The area-level ALS estimation reference data were collected from two areas in north-eastern Finland, with the first dataset (hereafter ALS1) comprising 89 stands and the second (hereafter, ALS2) 57 stands. Both datasets contain the ALS estimates and accurate reference measurements from the sample plots for aggregate attributes D_{gM}, H_{gM}, G, V, and N, again at tree species stratum-level. The ALS estimation process used in the data collection is described extensively in Packalen & Maltamo (2007).

In both the traditional stand-level field inventory and area-level ALS estimation, the estimates $\hat{x}$ represent the typical accuracy of the corresponding inventory technique. The accurate reference measurements x from the sample plots represent the true values of the attributes. Thus, the estimation error vector $\mathbf{e}_k$ for attribute k in a given dataset was

$$\mathbf{e}_k = \hat{\mathbf{x}}_k - \mathbf{x}_k \tag{1}$$

where $\hat{\mathbf{x}}_k$ and $\mathbf{x}_k$ are the estimate and true value vectors for attribute k, and element e_{ijk} is the value for attribute k in tree species stratum j of stand i, where $i = 1, 2, ...,$number of stands; $j = 1, 2, ...,$number of strata in stand i; $k = D_{gM}, H_{gM}, G, N, V$. The reference measurements x also contain some measure of error due to sampling errors and assessment errors. However, the amount of error in the reference measurements was assumed insignificant.

METHODS

The forest planning framework

All of the Papers I to IV involve some type of forest growth simulation. The simulators used in these studies were all implemented on the SIMO (SIMulation and Optimisation) framework, except for the MOTTI simulator Used in Paper I. SIMO is simulation and optimisation software for various types of forest planning computations originally developed at the Department of Forest Resource Management, University of Helsinki (Tokola et al. 2006, Rasinmäki et al. 2009). SIMO provides a flexible and easily adaptable and extendable framework for implementing different types of forest simulators. The MOTTI simulator, used in Paper I, is a well-established stand-level forest DSS for Finland which uses distance-independent tree-level models for growth prediction (Hynynen et al. 2002, 2005, Salminen et al. 2005).

Two growth and yield simulators, one using tree-level models and the other using stand-level models for growth prediction, were implemented on the SIMO framework. Both simulators were designed to be used for stand-level growth and yield simulations, and both simulators shared some common models for simulating different types of harvests in detail and for calculating various ecological and economic indices. The key difference between these two simulators was the level of organisation in the growth prediction.

The distance-independent tree-level growth models used in the tree-level simulator were largely the same as those used in the MOTTI and MELA systems (Hynynen et al. 2002). These models covered all of Finland's main tree species and forest types with separate model groups for mineral and peat soils. The dependent variables in the tree-level growth equations were the increment of tree height h and the increment of tree basal area g. The independent variables included various attributes describing tree dimensions, site and location indices, within-stand competition measures, and others. The tree-sets used in the simulations were either measured in the field, as in Papers I and II, or generated with a diameter or height distribution model, as in Paper III.

The stand-level simulator included growth models that predicted the increments of stand-level aggregate attributes stratified by tree species. The growth models used for Scots pine and Norway spruce were adopted from Vuokila & Väliaho (1980). The growth models used for birches and other deciduous species were adopted from Saramäki (1977) and Oikarinen (1983). The stand-level simulator included separate submodels for different geographical regions and site types in Finland. The dependent variables in these models were the increments of G, V, basal area under bark G_u and dominant height H_{dom}, depending on the tree species and soil type. The independent variables in the stand-level growth models were, for example, G, G_u, Age, H_{dom} and site class. Although the growth was predicted strictly at stand-level, theoretical tree sets were generated in order to predict N. In addition, the harvest models used in both tree-level and stand-level simulators required tree-level inputs, as each stem of a stand was tapered, and stem cutting was optimised (Papers III and IV).

In Papers I, II and IV, the SIMO system was used only for simulations, but in Paper III, the optimisation module was used for harvest scheduling. Optimisation required multiple alternative harvest schedules for each stand. In the SIMO simulator, this could be achieved so that each simulated harvest created a new "branch" to the set of alternative future developments. By varying the timings of thinnings and clearcuts, and the thinning intensities, the simulator module generated multiple alternative harvest scenarios for each stand. Thus, the alternative developments of a single stand formed a type of tree structure, where the current state is the root and possible future states are the branches.

The optimisation task in Paper III simply maximised the total estate-level NPV with no constraints. In practice, this was equal to maximising separately the NPV_i of each stand i. The optimisation task can be defined formally as

$$NPV = \sum_{i=1}^{100} \left[\sum_{t=1}^{30} \left(\frac{H_{it}}{(1+r)^t} \right) + \frac{PV_{i30} \times AREA_i}{(1+r)^{30}} + PVLAND_i \right] \tag{2}$$

where H_{it} was the discounted net income from harvests for stand i at year t, and PV_{i30} and $PVLAND_i$ were the productive value of the stand and the productive value of land at the end of the 30-year planning period. The total number of stands was 100 and the discount rate was 3% ($r = 0.03$). The PV_{i30} and $PVLAND_i$ values were predicted using models for stand-level productive values by Pukkala (2005). As the optimisation task was very simple, a straightforward heuristic search algorithm was used to find the optimal harvest schedule for each stand. In principle, the algorithm was an adaptation of the HERO algorithm, which was developed specifically for tactical-level forest planning (Kangas & Pukkala 1998, Kangas et al. 2001). In this case, the HERO algorithm searched through all possible schedules of each stand and was therefore able to determine the optimal solution after a single iteration.

Stochastic simulation methods

The simulation studies described in Papers II, III and IV utilised the MC method to estimate the distributions of the different responses of the SIMO system. In Paper II, the responses were various stand-level aggregate attributes. In Paper III, the response was the stand-level inoptimality loss, or EVPI, due to forest inventory errors. And in Paper IV, the response was stand-level NPV. Another option would have been to use some analytical error propagation method, such as the Taylor series approximation. The MC method was preferred over the analytical methods as it is more straightforward and easier to implement for complex model systems (Kangas 1999).

Simulation of forest inventory errors (II, III, IV)

In Papers II, III and IV, different types of forest inventory errors were generated into the input datasets. These errors were either classification errors of discrete attribute values, omission errors of individual trees or random estimation errors for continuous attributes. In Paper II, the classification errors were introduced to the tree species labels of individual trees. A tree with true species p was classified as species q with a conditional probability P_{pq} using a symmetrical classification error matrix. The classification error probabilities were based on the EuroSDR report by Kaartinen & Hyyppä (2008).

Missing trees in individual tree detection with remote sensing, known as omission errors, were simulated in Paper II. Relatively smaller trees have a lower probability of being detected, as the relatively larger trees occlude the smaller ones. The detection probability $P(detected_{tree})$ was modelled as a sigmoid function

$$P(detected_{tree}) = 1 - e^{-60h_{rel}^5} \tag{3}$$

where h_{rel} is the tree's relative height in the sample plot-level height distribution. The function shape and parameters were adopted from Korpela (2004). The simulation study in Paper II used a tree set as input data, and a proportion of the trees were removed according to probability $P(detected_{tree})$. The trees were removed so that for each tree in a sample plot, a uniformly distributed random number between 0 and 1 was drawn, and if the number was lower than the detection probability, the tree was removed.

In Papers II and III, random errors were generated in the continuous attribute values in the input data. A simple method for generating random estimation errors, which had also been used in many previous studies, was to generate a normally distributed error term ε with mean *mean* and standard deviation *SD* for each input variable x. Thus the erroneous value for attribute x was obtained by

$$\hat{x} = x + \varepsilon \tag{4}$$

One of the aims was to examine the properties of forest inventory errors and to evaluate methods for generating realistic errors (Paper III). The estimation error vector $\mathbf{e}_k$ for variable k in the studied datasets showed trends such that large values were underestimated and small values overestimated. To remove this trend, linear regression models of shape

$$\mathbf{e}_k = \beta_{k0} + \beta_{k1}\mathbf{x}_k + \varepsilon_k \tag{5}$$

were fitted to datasets ALS1 and ALS2, and polynomial regression model of shape

$$\mathbf{e}_k = \beta_{k0} + \beta_{k1}\mathbf{x}_k + \beta_{k2}\mathbf{x}_k^2 + \varepsilon_k \tag{6}$$

was fitted to dataset FIELD. The residual vectors $\hat{\varepsilon}_k$ of these models were used in the following analysis as they were the estimation errors without the error trend. Computing a 5×5 correlation matrix $Corr(\hat{\varepsilon}_{.ij})$, where $\hat{\varepsilon}_{.ij} = (\hat{\varepsilon}_{ij,D_{gM}}, \hat{\varepsilon}_{ij,H_{gM}}, \hat{\varepsilon}_{ij,G}, \hat{\varepsilon}_{ij,N}, \hat{\varepsilon}_{ij,V})^T$, indicated that there were significant correlations between the various attributes. In addition, according to the Shapiro-Wilk test of normality, the errors were unlikely to have been produced by a normal distribution.

In order to simulate such non-normal correlated random errors with strong trends, we first fitted various distributions into the error trend model residuals $\hat{\varepsilon}_k$ using the Maximum Likelihood (ML) method. The most suitable distribution shape was the logit-logistic distribution, which is a flexible distribution with minimum and maximum values ψ and λ, as well as shape and scale parameters ϕ and σ (Tadikamalla & Johnson 1982, Wang & Rennolls 2005).

As the distributions of $\hat{\varepsilon}_k$ in the various datasets were something other than normal, a multinormal distribution could not be used to simulate correlated random vectors. Instead, we had to adopt what is known as the Copula approach (Kolev et al. 2006, Mehtätalo et al. 2008). First, the $\hat{\varepsilon}_k$ values were transformed into normally distributed random variables $\mathbf{y}_k$ through

$$y_{ijk} = \Phi^{-1}(F(\hat{\varepsilon}_{ijk}|\hat{\psi}_k, \hat{\lambda}_k, \hat{\phi}_k, \hat{\sigma}_k)) \tag{7}$$

using the ML fitted logit-logistic distribution parameters. Next, a variance-covariance matrix $Cov(\mathbf{y}_{.ij})$, where $\mathbf{y}_{.ij} = (y_{ij,D_{gM}}, y_{ij,H_{gM}}, y_{ij,G}, y_{ij,N}, y_{ij,V})^T$, was constructed from the transformed residuals. With $\mathbf{y}_{.ij}$ known, a required number of correlated multinormal random vectors $\tilde{\mathbf{y}}_{.ij}$ were generated for each stratum and stand j and i using the Cholesky decomposition (Rubinstein 1981). After that, the simulated vectors $\tilde{\mathbf{y}}_k$ were transformed into logit-logistic random variable vectors $\tilde{\varepsilon}_k$ through

$$\tilde{\varepsilon}_{ijk} = F^{-1}(\Phi(\tilde{y}_{ijk}|\hat{\psi}_k, \hat{\lambda}_k, \hat{\phi}_k, \hat{\sigma}_k)) \tag{8}$$

To obtain the simulated errors $\tilde{\mathbf{e}}_k$, the error trend was added to the $\tilde{\varepsilon}_k$ vector. The simulated estimates $\tilde{x}_{est}$ were then obtained with

$$\tilde{\mathbf{x}}_{k\,est} = \mathbf{x}_k + \tilde{\mathbf{e}}_k \tag{9}$$

In addition to the correlated logit-logistic random errors, errors were also generated from logit-logistic distributions without maintaining the correlation structures between the various attributes. In addition, random errors were generated from multinormal distributions such that the correlations were maintained, but the random values were distributed normally (Paper III).

The aim in Paper IV was to study simultaneously the effects of multiple sources of uncertainty. The uncertainty resulting from forest inventory errors was taken into account, not by generating random errors, but by generating true values. First, the difference vector

$$\delta_k = \mathbf{x}_k - \hat{\mathbf{x}}_k \tag{10}$$

between the true values and the estimates of attribute k, separately for datasets FIELD and ALS1, was generated using exactly the same procedure as for the errors $\mathbf{e}_k$ in Paper III. After generating the differences $\tilde{\delta}_k$, the simulated true $\tilde{\mathbf{x}}_{k\,true}$ vectors were obtained by adding the simulated differences to the estimate vectors as in

$$\tilde{\mathbf{x}}_{k\,true} = \tilde{\delta}_k + \hat{\mathbf{x}}_k \tag{11}$$

In this way, a distribution of simulated true values, instead of simulated estimates, was obtained, as in Papers II and III. The former values were *classical* variables where $Var(true) < Var(estimate)$, and the latter were *Berkson* variables where $Var(true) > Var(estimate)$.

Random variation in stand-level growth projections (IV)

The uncertainty in growth prediction was simulated by introducing a random error component u into the growth models. Stand-level growth models were used instead of tree-level models as the error models were easier to introduce into the simpler stand-level growth simulator. The random error u consisted of two individual error components: between-stand error u_B and within-stand error u_W. The error was divided into two components to keep stand-level biases at the same level throughout the simulation. The total random error at time t was then

$$u_t = u_B + u_{Wt} \tag{12}$$

The between-stand error u_B was generated for each tree species stratum of each stand, once at the beginning of the simulation and again if the stand was regenerated. The within-stand error was simulated as an autoregressive process where u_{Wt+1} depended on u_{Wt} such that

$$u_{Wt+1} = \alpha \times u_{Wt} + b_t \tag{13}$$

where α was the autocorrelation coefficient and b_t was a random coefficient at time t. The value for b_t was generated from a normal distribution at each simulation time step. The simulated random error components were added to the increments of H_{dom} and G. This type of growth prediction error simulation was used for all non-seedling stands and for Scots pine-dominated seedling stands. The total variation of u, based on results by Haara & Leskinen (2009), was divided into u_B and u_W by applying the results of Kangas (1999). The value for the correlation coefficient α was calculated also from the models by Haara & Leskinen (2009).

The growth in spruce and birch seedling stands was based on a simple model that predicts the number of years for the stand H_{gM} to reach 1.3 meters. The prediction error was introduced into these stands by adding a normally distributed random error term u_B to variables Age_{spruce} and Age_{birch}. The between -tand errors u_B of the variables H_{dom}, G, Age_{spruce} and Age_{birch} were presumed to be correlated. Thus, the values for these variables were generated from a multinormal distribution (Rubinstein 1981).

When the growth prediction error models were active in the SIMO stand simulator, the root mean square errors (RMSE) and biases in the predicted values of various attributes were of the same magnitude as those in the results of Haara & Leskinen (2009). If the distribution errors in Haara & Leskinen (2009) are normal, or close to normal, the simulated random variance in the growth predictions can be considered realistic.

Stochastic timber assortment price models (IV)

The economic uncertainty in stand-level NPV estimates was simulated using a stochastic timber assortment price model. The model was based on real monthly stumpage prices between January 1986 and August 2008 (Finnish Statistical Yearbook of Forestry, 2008). Separate prices were given for saw logs and pulpwood for Finland's three main commercial tree species: Scots pine, Norway spruce, and birch. The fluctuation of the prices was modelled as a *geometric mean-reverting* (GMR) process, which is not analytically solvable, but has been used in various studies. Another widely used option would have been *geometric brownian motion* (GBM), which can be solved analytically. The stochastic model for the timber assortment prices was

$$dp = \eta(\overline{p} - p)dt + \sigma p dz \tag{14}$$

with $\overline{p}$ being the long-term mean price, parameters η and σ denoting the speed of reversion and the level of annual variation, and dz representing the increment of the Wiener Process (Dixit & Pindyck 1994). For a detailed description of the estimation of the parameters, refer to Paper IV. Because strong correlations were observed in previous timber assortment prices, future predictions were generated from multinormal distribution by multiplying the Cholesky decomposition of the timber assortment price variation covariance matrix with a matrix of normally distributed N(0,1) random variables (Rubinstein 1981).

Description of the simulation setups

Paper I

The objective in Paper I was to evaluate the growth prediction accuracies of tree-level and stand-level simulators implemented in the SIMO system, as well as the MOTTI tree-level simulator by comparing the predicted developments to observed developments at sample-plot level. The reference sample plot data of Paper I served as the input data for the simulations. The simulators used in the study took the measured sample tree and predicted tally tree values as input. The growth of each sample plot was predicted for a 20-year period using 1-year time steps in the SIMO simulators and 5-year time steps in the MOTTI simulator. Harvests or any kind of treatments were not simulated as the sample plots had been unmanaged during the past 20 years. The simulations were completely deterministic with no randomness whatsoever, and the three simulators used are referred to as

TREE SIMO tree-level simulator,

STAND SIMO stand-level simulator,

MOTTI MOTTI tree-level simulator.

Paper II

In Paper II, the propagation of two types of forest inventory errors was studied in tree-level and stand-level forest simulators. The two types of errors were those observed in stand-level field inventory and those observed in single tree-level ALS estimation. Input data consisted of 240 tree-wise sample plots measured in 2006 from the Evo area in southern Finland. Sample tree d and h were measured, and tree species recorded on circular sample plots with a 9.77 metre radius. A number of sample plot-level attributes describing plot location, site and soil quality were obtained from an existing forest planning database available for the study area. The simulation period was 30 years in 5-year time steps. The harvest models of SIMO were disabled. With two forest inventory error types and two types of forest simulators, there were four separate simulation setups:

FIELD-TREE stand-level field estimation errors, simulation with tree-level simulator (Simulation setup I in Paper II);

FIELD-STAND stand-level field estimation errors, simulation with stand-level simulator (Simulation setup II in Paper II);

ALS-TREE tree-level ALS estimation errors, simulations with tree-level simulator (Simulation setup III in Paper II);

ALS-STAND tree-level ALS estimation errors, simulation with stand-level simulator (Simulation setup IV in Paper II).

The MC method was applied in the simulations so that each stand was simulated 100 times for each simulation setup, and at the beginning of each iteration, random errors were introduced into the input data. In simulation setups FIELD-TREE and FIELD-STAND, the inventory errors were generated from univariate normal distributions for tree species stratum-level attributes G, D_{gM}, H_{gM}, N and Age, both separately for each attribute and simultaneously for all attributes. The inventory error means and SDs for the stand-level aggregate attributes were based on the existing literature (Poso 1983, Laasasenaho & Päivinen 1986, Kangas et al. 2004).

In simulation setups ALS-TREE and ALS-STAND, the inventory errors were introduced into the single tree-level attributes dbh and h. A portion of the trees were removed with probability $P(detected_{tree})$ in order to simulate omission errors. In addition, a tree species classification error was simulated for some of the trees. All these errors were also generated separately and simultaneously. The tree-level estimation error properties were based on the EuroSDR report by Kaartinen & Hyyppä (2008).

In addition to the MC simulations, a reference simulation with no input data errors was carried out for each stand. The reference simulation was considered to represent the forest development with no uncertainty.

Paper III

Also in Paper III, different types of forest inventory errors were introduced into the input data values using the MC method. However, the whole simulation setup was a bit more complex than that in Paper II. Only the tree-level simulator was used, but instead of simulating a single future development for each stand, multiple alternative developments, or harvest schedules, were simulated. Also, an optimisation module was used for selecting the optimal harvest schedule for each stand. The simulation time period was again 30 years in 5-year time steps. Input data consisted of 100 randomly selected stands from a forest estate located in southern Finland. The data had been collected previously with stand-level field inventory. The values describing stand location and site quality were at stand-level, and the data describing the standing stock with aggregate attributes were stratified by tree species.

The reference simulation was carried out once for each stand as in Paper II, but this time with the harvests and the optimisation of harvest schedules. The generation of random errors in the MC simulations also took place at the beginning of each iteration, but the error generation methods were different from those in Paper II. Random errors were generated for attributes G, D_{gM}, H_{gM}, V and N using four alternative simulation methods:

MLLTr errors simulated from multivariate logit-logistic (MLL) distributions with both the correlations and error trends (Tr);

UN errors simulated from univariate normal (UN) distributions without the correlations or error trends;

ULLTr errors simulated from univariate logit-logistic (ULL) without the correlations, but with error trends;

MNTr errors simulated from multinormal (MN) distributions with the correlations and error trends.

The above four methods were used to simulate errors similar to those observed in datasets FIELD and ALS2. The errors observed in dataset ALS1 were simulated only with methods MLLTr and UN. Thus, altogether 10 different simulation setups were repeated 100 times for each stand in the input data using the MC method.

Paper IV

The simulation in Paper IV incorporated three sources of uncertainty: forest inventory errors, growth prediction errors and stochastic timber prices. The sources of uncertainty are hereafter referred to as U_{PRICE} for price uncertainty, U_{GROWTH} for growth uncertainty and U_{FIELD} and U_{ALS} for inventory uncertainties in stand-level field inventory and area-level ALS estimation, respectively. The various combinations of these sources of uncertainty are listed in table 5. A stand-level simulator was used to predict growth, and different harvests, such as thinnings and clearcuts, were activated. The simulation time period depended on the maturity of each stand, as the stands were simulated until the next regeneration harvest. Harvest scheduling was based on silvicultural recommendations in Finland (Hyvän Metsänhoidon Suositukset, 2006) and regeneration harvests were done after the relative value growth dropped below the chosen interest rate. Input data of the simulations was a synthetic dataset consisting of 40 stands. The data contents were similar to those in the input data of Paper III. The sample plots and stands used as input data in Papers II to IV represented a wide variety of different forest types, tree species compositions and site and age classes typical in Finland. The three sources of uncertainty were simulated separately, in pairs and simultaneously, and all these combinations were simulated at interest rates of 3%, 4% and 5%. Simulations were performed using the MC method with 100 iterations for each combination, yielding a distribution of NPV values for each stand. The NPV for all realisations of each stand i were calculated with Equation 15, where $PVLAND_{iT}$ is the productive value of land after the next regeneration. The timing T_i of the next regeneration harvest of stand i depended on the maturity of the stand.

$$NPV_i = \sum_{t=1}^{T_i} \frac{H_{it}}{(1+r)^t} + \frac{PVLAND_{iT} \times AREA_i}{(1+r)^{T_i}} \tag{15}$$

Statistical analysis of the effects of uncertainty

The effects of various sources of uncertainty were analysed using basic statistics, such as arithmetic mean (16), SD (17) and the RMSE (18) of the variable of interest X, which in many cases was the assessment error of some attribute of stand i. The RMSE is defined here as the square root of the sum of the variance and squared bias B^2 of variable X.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{16}$$

$$SD(X) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2} \tag{17}$$

$$RMSE(X) = \sqrt{SD(X)^2 + B^2} \tag{18}$$

The distributions of the growth prediction errors in Paper I were examined using quantile regression instead of traditional regression analysis. Quantile regression is an extension of linear regression and is used for estimating the conditional quantiles of the dependent variable distribution. It is used in situations in which the variance is heterogenous as it often suggests that no predictive relationships exist between the dependent and independent variables. Even if a relationship did exist in certain parts of the dependent variable distribution, basic regression provides only models for the conditional mean function of the dependent variable distribution. Quantile regression can provide models for the whole range of conditional quantile functions (Koenker & Bassett 1978, Koenker & Hallock 2001). Conditional quantile regression functions for the variable Y can be written as

$$Q_Y(\tau|X) = \beta_0(\tau)X_0 + \beta_1(\tau)X_1 + \beta_2(\tau)X_2 + ... + \beta_p(\tau)X_p \tag{19}$$

where τ refers to the τ'th percentile of the distribution of Y, conditional to X (Cade & Noon 2003).

Analysis of the inoptimality losses

In Paper III, the effects of forest inventory errors were studied by calculating the inoptimality loss, also known as the EVPI, for each realisation of each optimised harvest schedule. First, a number of alternative harvest schedules were simulated for each stand using input data devoid of errors. Then the optimal schedule was selected for each stand using the HERO heuristic search and optimisation task definition (Eq. 2). Based on the income from harvests and stand-level productive value estimates at the end of the planning period, optimal NPV_i was calculated for each stand i applying the same principles as in Equation 2, but without the aggregation to the estate-level.

After determining the true optimal NPV_i values, each stand was simulated using an iterative MC approach. The number of realisations for each stand was 100, and at the beginning of each iteration, random errors were generated into simulation input data. After repeatedly simulating and optimising each stand, 100 optimised harvest schedules were obtained for each stand.

The next step was to simulate each stand and realisation again, but this time without generating errors in the input data values and using the harvest schedules determined in the previous step. Discounted net present value npv_{il} was then calculated for each stand i and realisation l. If the harvest schedule of realisation l of stand i was equal to the optimal schedule of the stand, then NPV_i and npv_{il} were also equal. After this, the inoptimality loss, or $EVPI$, and relative loss $EVPI\%$ could be calculated respectively as

$$EVPI_{il} = NPV_i - npv_{il} \tag{20}$$

$$EVPI\%_{il} = ((NPV_i - npv_{il})/NPV_i) \times 100 \tag{21}$$

RESULTS

In Paper I, the growth prediction accuracies of different forest simulators were compared against real observed growths. The compared simulators were two tree-level growth simulators and one stand-level growth simulator, and the growth period was 20 years.

Table 1. The mean of relative prediction error in 5-year intervals (%) for attributes H_{gM}, D_{gM}, G, and N, from simulators TREE, STAND and MOTTI (Paper I).

simulator	year	H_{gM}	D_{gM}	G	N
TREE	5	-4.5	-2.3	-4.6	-6.1
	10	-6.2	-2.3	-6.3	-9.6
	15	-6.7	-1.6	-4.3	-8.7
	20	-6.6	-0.6	-0.3	-6.1
STAND	5	-4.0	-0.2	0.4	-3.0
	10	-4.9	-0.3	-3.8	-3.9
	15	-5.2	0.0	-3.4	-0.3
	20	-5.2	0.3	-0.2	-6.3
MOTTI	5	-3.2	0.1	-4.8	-9.0
	10	-5.3	-0.8	-9.5	-12.6
	15	-6.1	-0.1	-8.7	-12.2
	20	-6.4	1.2	-5.3	-10.3

The mean of the prediction error at different years for four sample plot-level aggregate attributes appears in Table 1. During the 20-year simulation, the values of all four attributes were underestimated. The predictions of D_{gM} were the most accurate as the means were close to zero. Interestingly, for attributes H_{gM}, G, and N, the underestimation grew initially, but decreased toward the end of the simulation period. The mean prediction errors of the stand-level simulator were on average more accurate than those of the two tree-level simulators.

The relative SDs of the prediction errors appears in Table 2. The variation in the prediction errors grew steadily during the simulation period for all attributes.Predictions of the development of H_{gM} were more precise with the tree-level simulators, but the stand-level simulator predicted the development of D_{gM} and G with greater precision. During the first 15-year period, the precision of predicted N was similar in both tree- and stand-level simulators, and thereafter the variation in the stand-level simulator prediction increased notably.

The prediction errors were also modelled using quantile regression as the variance of the errors was heterogenous. This was the case when the errors, for example, were modelled as a function of simulation time. In Figure , the relative error in predicted G after a 20-year simulation was, using quantile regression, modelled as a function of initial Age. The figure clearly depicts the differences between the SIMO tree and stand simulators in the relationship of predicted G and initial Age. In the SIMO tree simulator, the variance of the prediction error was quite homogenous, whereas in the stand simulator, the variance was high for young stands and notably lower for older stands. Also, the SIMO tree simulator had quite a strong negative bias in the predicted G of younger stands but the SIMO stand simulator did not.

Uncertainty resulting from forest inventory errors was examined first by analysing the

Table 2. The SD of relative prediction error in 5-year intervals (%) for attributes H_{gM}, D_{gM}, G, and N, from simulators TREE, STAND and MOTTI (Paper I).

simulator	year	H_{gM}	D_{gM}	G	N
TREE	5	6.0	4.4	11.1	7.5
	10	8.6	5.2	14.0	10.4
	15	10.3	5.4	16.4	13.7
	20	11.6	5.5	17.9	14.3
STAND	5	8.9	3.4	19.4	7.1
	10	11.2	4.4	13.3	9.6
	15	12.9	4.9	12.0	12.8
	20	14.3	5.2	12.8	17.8
MOTTI	5	5.9	4.6	14.1	11.8
	10	8.5	5.3	15.5	13.6
	15	10.1	6.5	17.4	15.9
	20	11.4	9.6	19.4	16.7

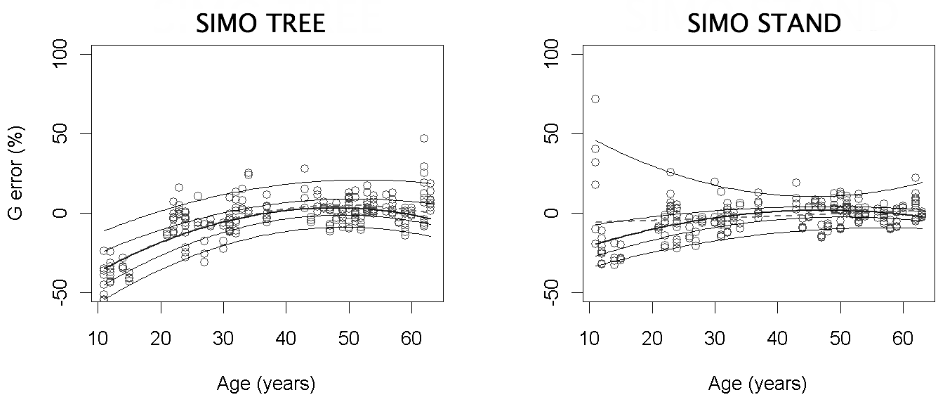

Figure 1. Quantile regression models of the relative error of predicted G for SIMO tree and stand simulators. The independent variable was the initial Age at the beginning of the simulation.

distributions of errors in predicted stand-level aggregate attribute values (Paper II). Table 3 shows the relative SDs of the errors of various stand-level aggregate attributes after a 30-year simulation. The errors in the predictions stem from simultaneously generated random errors in the values of various input attributes. In general, the single tree-level ALS estimation inventory errors (ALS-TREE and ALS-STAND) caused less variation in the values of attributes Age, G, V and N, but more in the values of D_{gM} and H_{gM}. Also, in most cases using a tree-level simulator resulted in lower variation in the stand-level aggregate attributes. Systematic errors, or biases in the predicted values were insubstantial.

Table 3. The SDs of the relative errors (%) in stand-level aggregate attributes due to forest inventory errors after a 30-year growth simulation, using four different forest inventory method and growth simulation combinations (Paper II).

Simulation setup	Age	G	V	N	D_{gM}	H_{gM}
FIELD-TREE	20.8	20.7	21.2	38.5	8.6	11.7
FIELD-STAND	16.2	26.3	43.3	53.4	17.7	16.9
ALS-TREE	15.6	19.7	17.9	19.2	12.4	15.9
ALS-STAND	7.6	26.4	29.1	19.3	14.8	12.9

Table 4. Absolute (euros) and relative (%) inoptimality losses due to forest inventory errors using four alternative error simulation methods (Paper III).

			MLLTr	UN	ULLTr	MNTr
FIELD	Abs.	mean	375	450	353	352
		SD	459	542	492	481
	Rel.	mean	5.6	6.1	5.4	5.0
		SD	6.1	7.0	6.8	6.5
ALS1	Abs.	mean	505	308	-	-
		SD	743	400	-	-
	Rel.	mean	6.2	4.5	-	-
		SD	8.2	5.7	-	-
ALS2	Abs.	mean	515	309	505	510
		SD	782	466	769	767
	Rel.	mean	6.4	4.6	6.2	6.4
		SD	8.6	6.2	8.4	8.6

The effects of the uncertainty due to forest inventory errors were also analysed by calculating the inoptimality losses caused by inoptimal harvest scheduling (Paper III). The input data errors were generated with four alternative error simulation methods. The mean of the absolute inoptimality losses varied between 308 and 515 euros, and the relative losses between 4.5% and 6.4% of the stands NPV, depending on the error simulation method. Uncorrelated normally distributed errors resulted in notably larger losses than correlated logit-logistic errors with trends, when the generated errors were based on the errors observed in dataset FIELD, but were the other way around when the errors were based on datasets ALS1 and ALS2. Uncorrelated logit-logistic errors with trends and correlated normal errors with trends caused similar inoptimality losses to correlated logit-logistic errors with trends, when simulating ALS2 errors. Uncorrelated logit-logistic and correlated normal FIELD errors, both with trends, resulted in a bit smaller losses than correlated logit-logistic errors with trends.

A strong trend in the NPV losses was that the highest losses occurred in mature stands close to their optimal rotation age. The average inoptimality losses in seedling stands and in young stands were smaller than in the mature stands. This trend was especially strong when the simulated errors were based on datasets ALS1 and ALS2. In seedling stands, the average inoptimality losses due to input data errors ranged from 1.3% to 3.1%, whereas in mature stands, the average losses were between 5.4% and 9.8%. These figures excluded seedtree stands in which the relative inoptimality losses were notably higher. The absolute inoptimality losses in the seedtree stands were also quite high, even though the NPV of a seedtree stand is normally small, as the only valuable timber is in the few seedtrees.

Table 5. Average relative biases and SDs (%) of discounted stand NPVs due to three individual sources of uncertainty, calculated separately, paired and simultaneously, and for three separate interest rates (Paper IV).

				3%		4%		5%	
U_{PRICE}	U_{FIELD}	U_{ALS}	U_{GROWTH}	bias	SD	bias	SD	bias	SD
o				-6.1	8.2	-1.5	7.3	-0.9	6.9
	o			-6.8	28.8	-5.4	29.2	-5.7	32.6
		o		1.7	26.5	4.8	26.4	7.3	28.7
			o	-9.5	33.2	-6.7	33.4	-5.8	33.2
o	o			-9.1	29.0	-3.8	32.1	-0.8	33.8
o		o		-1.0	27.4	7.1	28.6	10.2	30.9
o			o	-5.7	34.9	-1.5	35.3	-2.9	34.9
	o		o	-12.5	46.9	-7.9	48.2	-6.4	50.0
		o	o	-2.1	46.5	4.3	46.6	7.0	47.1
o	o		o	-9.2	47.4	-3.6	48.3	-1.8	51.3
o		o	o	0.1	46.5	7.5	46.8	10.0	47.6

To understand the practical implications of different sources of uncertainty, the effects of economic uncertainty, growth prediction errors, and forest inventory errors on stand-level NPV predictions were examined (Paper IV). Average relative biases (%) SDs in NPV predictions were calculated relative to reference NPV simulations, i.e. similar simulations, but without any random variation or uncertainty. The averages of these biases and SDs due to different source of uncertainty and interest rate combinations appear in Table 5. From the individual sources of uncertainty, the economic uncertainty U_{PRICE} caused markedly less variation in the NPV predictions than did the growth prediction errors U_{GROWTH} or forest inventory errors U_{FIELD} and U_{ALS}. The variation caused by U_{PRICE} decreased from 8.3% to 6.9% as the interest rate grew from 3% to 5%. The relationship between the variation caused by forest inventory uncertainty, U_{FIELD} and U_{ALS}, on the contrary, increased with the interest rate. Variation in NPV due to U_{GROWTH} was essentially unaffected by the interest rate, but the average NPV values naturally were markedly affected by the interest rate. Increasing interest rate led to lower NPV values. As expected, the greatest variation in NPV predictions, from 46.5% to 51.3%, was observed when all sources of uncertainty were active simultaneously. In addition, when all sources of uncertainty were considered simultaneously, the variation increased along with the interest rate. The average relative biases were mainly negative, except when U_{ALS} affected the NPVs, in which case the bias was positive. The biases ranged from -12.5% to 10.2%.

DISCUSSION AND CONCLUSIONS

Forest simulators are often complex systems of models used to predict future developments in forest ecosystems in which natural variation can be prominent. The time periods for which the forest growth and yield are predicted can be long, and the input data for the computations are often inaccurate. In addition, a number of other future developments that cannot easily be foreseen may diminish the reliability of forest simulator outputs. These sources of *uncertainty* should be understood and accounted for when forest simulators, or FPSs, serve as decision support tools. The main purpose of this thesis was to examine the various sources of uncertainty in forest simulators and FPSs and to increase our understanding of the implications of uncertainty.

First, the precision of growth prediction and the accuracy of tree-level and stand-level forest simulators were compared (Paper I). Both simulators underestimated the four sample plot-level aggregate attributes of interest: D_{gM}, H_{gM}, BA and N. This relative negative bias seemed not to increase during the simulation time period, but rather to decrease. A relative bias can, however, decrease even though the absolute does not because the absolute reference values increase as the forest grows. These results suggest that the rate of growth in dense forests is actually higher than the growth projections which might affect the forest management guidelines that are based on growth models, as was noted by Välimäki & Kangas (2009). The accuracy of the growth predictions decreased, or the variation of the predictions increased as a function of the simulation time period, which is in line with the results of Kangas (1999). The biases and variances of the predicted values were only slightly higher than the RMSEs reported by Haara & Leskinen (2009), even though the stands were denser than typical commercially managed forests in Finland. In general, both tree-level and stand-level forest simulators can provide fairly reliable projections about forest development. According to Paper I, the stand-level simulator is a bit more robust in dense forests. On the other hand, the tree-level simulator provides more detailed description of forest properties, which is important in, for example, evaluating the effects of management activities, such as thinnings.

The variances of the growth prediction errors were heterogenous and consequently, quantile regression was used for analysing the errors. Quantile regression better described the dependencies between the prediction errors and the various independent variables, such as initial stand Age. The variance of the prediction errors of G was very homogenous in stands of different ages when the tree simulator was used. When the stand simulator was used, the variance was notable in young stands, but not in older stands. Stand Age was an independent variable in stand-level growth models, but not in tree-level models, which was the main reason for this difference. Another option instead of quantile regression would have been to use variance functions.

Next, two types of simulated forest inventory errors were propagated through the tree-level and stand-level forest simulators using the MC method (Paper II). The responses to the input data errors were studied by analysing the variances of the errors in the predicted values of the stand-level aggregate attribute. The variances of the absolute errors decreased, but the variances of the relative errors increased during the simulation time period. The greatest variances were observed in the errors of N, G and V, and the smallest in D_{gM} and H_{gM}. Simulated tree-level ALS estimation errors led to notably higher accuracy in the predicted values of N. In general, the tree-level simulator was more accurate than the stand-level simulator, and especially when the errors were simulated stand-level field inventory errors. Many of the differences between the two simulators resulted from the various interactions between the sub-models of the growth simulators. Errors in stand age, for example, caused substantial error in the stand-level simulator growth projections, but almost none in the tree-level simulator. The reason for this difference is that the tree-level growth models are age-independent, whereas the stand-level growth models are not. In general, interactions in the tree-level simulator are much more difficult to analyse as the tree-level simulator is much more complex than the stand-level simulator and contains many more interactions and different feedback-loops.

One of the objectives in this thesis was to examine the properties of forest inventory errors in detail. The errors observed in stand-level field inventory and area-level ALS estimation were characterised using three error properties: error trend, error distribution shape and correlations between the errors. In most previous studies, the errors have been assumed to be normally distributed and uncorrelated (Eid 2000, Holopainen & Talvitie 2006). According to Paper III, this assumption may be invalid. The error distributions in both stand-level field inventory and area-level ALS estimation were not gaussian, and the correlations between the errors of various attributes were in many cases notable. Moreover, the error trends were strong, especially in datasets ALS1 and ALS2. In these two datasets, the values of the various attributes were underestimated in stands with large trees and overestimated in stands with smaller trees. Similar *averaging effect* could be observed also in dataset FIELD, but it was not as distinct. Paper III described an error simulation method for generating correlated non-normal random errors, which can also serve in further simulation studies for generating more realistic forest inventory errors.

In addition to studying the errors, the effects of the different error properties when simulating inoptimality losses were also examined. The average inoptimality losses relative to stand-level NPV ranged from 4.5% to 6.4%, depending on the error simulation method. Holopainen & Talvitie (2006) reported inoptimality losses ranging from 4.2% to 11.4%, depending on the simulated forest inventory method. The inoptimality losses reported by Eid (2000) were considerably smaller than in this study (on average 0.92%) but the distributions of random errors were also narrower. Islam et al. (2009) reported estate-level inoptimality losses ranging from 0.3% to 3.6% due to inventory errors in G. A notable trend in the inoptimality losses, which Eid (2000) also observed, was that the losses were larger in mature stands close to their optimal rotation ages. The error trend had by far the strongest impact on the inoptimality losses. The shape of the error distribution and the correlation between the various attributes only slightly affect the inoptimality losses. Based on this observation, using non-correlated gaussian random values as simulated inventory errors could be enough in practice, if the error trends are taken into account. Still, it is theoretically more justifiable to take into account the distribution shape and correlation when simulating errors. According to these results, the error trend significantly affects the simulated inoptimality losses, but has largely been overlooked in previous studies.

The differences in the inoptimality losses reported in this thesis and in previous studies stem from many factors. The input data are one factor affecting the outcome, as are the errors generated into the input data. Eid (2000), for example, used fixed error levels, and in Paper III, error properties were derived from real forest inventory results. In addition, the forest growth model, or the simulation implementation, affects the outcome, even if the models are intended for use in the same forests, as is the case with tree-level and stand-level growth models in Finland. The simulation time step affects the resolution for timing the harvests as well as other forestry operations. Besides the simulator, the optimisation method and optimisation task definition also affects such simulation studies. Moreover, the length of the planning period, or in other words, the number of decisions based on the current data, does affect the inoptimality losses.

The joint effects of three considerable sources of uncertainty were studied by simulating stand-level NPV distributions using the MC method (Paper IV). To date, no other studies have aimed to take into account these three sources of uncertainty in a complex forest simulation system. The uncertainty in the NPV estimates was quantified as the average bias and SD of NPV, relative to reference NPV estimated without any uncertainty. The simulated NPVs were compared to NPVs predicted with the NPV models of Pukkala (2005) and were found out to be very close, if not exactly the same. The total SD, when all three sources of uncertainty were taken into account simultaneously, was between 46.5% and 51.3%, depending on the simulated forest inventory method and interest rate. When all three sources of uncertainty were considered simultaneously, the SD of the NPV was only slightly higher than the SD resulting from input data errors or growth model errors separately. The NPV distributions affected by U_{ALS} were positively biased, which was caused by the averaging effect in the inventory method, observed also in Paper III (i.e. large values are underestimated and small ones overestimated). In this case, the simulated true forest attribute values were larger in forests with big trees than the estimated values. The other sources of uncertainty led to negative biases. The observations in Paper IV imply that the cumulation of the variation is not straightforward, as the sum of variances due to individual sources of uncertainty does not equal the variance of the same source of uncertainty combination. Moreover, these results suggest that improving the accuracy of forest inventory data can improve the accuracy of forest NPV predictions only to a certain extent. In other words, uncertainty in forest value predictions cannot be totally avoided with more precise and accurate data. Previous studies on the VOI of forest inventory data, many of them applying CPL analysis, have not accounted for uncertainty caused by growth model errors and natural variation in forest growth. As different sources of uncertainty cumulate in a non-linear manner, the decrease of uncertainty due to the use of more accurate data is probably not linear. The results of Papers I and IV suggest that the uncertainty attributed to growth prediction errors is a major source of uncertainty in forest simulators.

The results of this thesis illuminated some of the various sources of uncertainty in forest simulators and FPSs. In addition, the magnitude and the effects of various sources of uncertainty were examined and reported, and a new methodology was applied. Although the individual Papers of this thesis included a number of advances in studying sources of uncertainty, some shortcuts were taken to simplify the simulations and the analysis. The dataset in Paper I was rather small and included only stands that were distinctly denser than typical commercially managed forests. Because of this, some of the results, such as the average growth prediction accuracies, cannot be generalised to all finnish forests.

The generation of forest inventory errors in Paper II failed to take into account error correlations, trends or distribution shapes. Even though the omission errors, i.e. undetected trees, in single tree-level ALS estimation were modeled, the commission errors, i.e. the possibility of detecting non-existing trees, were not considered as there was no decent model available. In Paper III, error correlations, trends and distribution shapes were considered, and the error trend in particular was found to have a strong effect. If the study presented in Paper III would have been done before Paper II, the simulation of forest inventory errors would probably have been similar to that in Paper III. The forest planning simulation in Paper III was simplified, so the results may not be applicable to practical forestry as such. However, the aim of Paper III was to gain insight about error properties and to outline an error simulation method that can be used to increase realism in simulation studies. Duvemo (2009) has proposed more realistic methods for simulating the whole forest planning process.

The simulation of various sources of uncertainty in Paper IV included the uncertainty due to growth prediction, forest inventory errors and stochastic timber prices. Uncertainty resulting from natural hazards and other risks were omitted from the analysis. In Papers I, III and IV, the reference measurements that were considered to represent the true values inevitably contained some amount of errors, but were assumed to be so close to the true values that the errors could be disregarded. Errors in reference measurements and their effects have been discussed throughly in Välimäki & Kangas (2009).

Uncertainty originating from various sources is an inseparable part of forest simulators and FPSs. Some parts of uncertainty can be decreased; for example, the uncertainty due to errors in the input data can be reduced by acquiring more precise and accurate data. This, however, is often an unfeasible option as the costs of data acquisition increase rapidly as the quality of data improves. Some parts of uncertainty, such as natural variation in forest growth and natural hazards, cannot be eliminated so easily. However, uncertainty can be quantified and taken into account when using forest simulators and FPSs for decision support. For example, the NPV distributions in Paper IV can be used to estimate confidence intervals for stand-level NPV estimates. A forest NPV estimate with confidence intervals is much more valuable information than an NPV point estimate when forest investments are in question.

In Papers I and II, uncertainty in the predicted values of some of the interesting stand-level aggregate attributes was reported with figures such as bias, SD, variance and RMSE. These figures may be informative for researchers, but not so much for practical users. The same applies to the SD of the NPV distributions of Paper IV. The inoptimality losses reported in Paper III are simpler to interpretate, as they are presented in monetary units, and thus provide a more practical measure of uncertainty. There is a pitfall in this quantification also. The simulated forest planning processes, the simulated sources of uncertainty and the models used in this and in various related studies are in the end simple abstractions of real world processes. Consequently, reported inoptimality losses do not necessarily represent the actual losses and should be used with caution. For some uses, however, such as comparing alternative inventory methods, this type of approach may be suitable.

A large part of growth prediction errors result from natural variation in forest growth, which can be difficult to decrease by improving the growth models (Mowrer 2000). Natural variation in forest growth is a sum of many factors. Some of this variation is due to stand- or site-specific factors, which could be taken into account to reduce uncertainty. Stand-specific past growth information, were it available, could perhaps be used to calibrate growth predictions. In theory, such information could be available in Finland, as large parts of finnish forests have been inventoried repeatedly. In practice, however, such information may not be easily available (Rasinmäki et al. 2009). The past growth could be estimated also by coring sample trees and measuring past growth from annual rings. This, however, can be quite expensive as it requires a lot of manual work (Holm 1980).

According to the results of Papers I and II, there are differences in how different simulators response to uncertainty, i.e. how forest inventory errors propagate in the simulator, or what is the accuracy or growth projections in different types of forests. These type of differences could be taken into account, for example, by using certain growth models or simulators only for certain types of simulations.

A number of studies have dealt with various sources of uncertainty in forest planning and management. Yet many questions remain to be answered. Paper IV described a step forward in taking uncertainty in forest growth and yield simulations into account comprehensively. However, the simulation task was only a simplified harvest scheduling, and some sources of uncertainty were still not taken into consideration. Most of the CPL analysis to date have considered only uncertainty resulting from forest inventory errors, so the next obvious step would be to incorporate more sources of uncertainty, namely random variation in growth and economic risk, into these analysis. In addition, to make the results more applicable into practice, the forest planning process should be modeled in greater detail, such as suggested in Duvemo (2009). Taking into account other sources of uncertainty besides inventory errors in the CPL analysis could provide information about how much the inoptimality losses could really be reduced with more accurate and precise data. If only the inventory errors are considered, no errors in the data means no inoptimality losses, which is rather unrealistic. One example of a novel CPL application in forest planning is the calibration of forest growth models using data on past growth. As there are methods for increasing growth prediction accuracy using data on past growth, CPL approach could be used to estimate the profitability of acquiring such data and calibrating the growth predictions.

As a conclusion, it can be said that a forest planning system as a whole is influenced by multiple sources of uncertainty. This thesis dealt with some of the obvious sources of uncertainty, but many were left outside the scope of this study. As forest planning systems are used for decision support, this uncertainty should be accounted for in the systems and in decision making process in order to make good decisions, or at least to avoid bad ones.

References

Atta-Boateng, J. & Moser, J.W. 2000. A compatible growth and yield model for the management of mixed tropical rain forest. Canadian Journal of Forest Research 30: 311–323.

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

Brazee, R. & Mendelsohn, R. 1988. Timber harvesting with fluctuation prices. Forest Science 34: 359–372.

Buongiorno, J. & Michie, B.R. 1980. A matrix model of uneven-aged forest management. Forest Science 26(4): 609–625.

Burkhart, H.E. 1993. Tree and stand models in forest inventory. In: A. Nyyssönen, S. Poso & J. Rautala (eds.), Proceeding of the Ilvessalo Symposium on National Forest Inventories. IUFRO S4.02, volume 444 of Metsäntutkimuksen tiedonantoja, pp. 164–170.

— 2003. Suggestions for choosing an appropriate level for modeling forest stands. In: A. Amaro, D. Reed & P. Soares (eds.), Modeling forest systems, p. 398. CABI Publishing, Oxfordshire.

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

Cade, B. & Noon, B. 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1: 412–420.

Canavan, S.J. & Hann, D.W. 2004. The Two-Stage Method for The Two-Stage Method for Measurement Error Characterization. Forest Science 50(6): 743–756.

Dixit, A.K. & Pindyck, R.S. 1994. Investment Under Uncertainty. Princeton University Press.

Duvemo, K. 2009. The Influence of Data Uncertainty onPlanning and Decision Processes inForest Management. Ph.D. thesis, Department of Forest Resource Management, Faculty of Forest Sciences, SLU, Umeå.

Duvemo, K., Barth, A. & Wallerman, J. 2007. Evaluating sample plot imputation techniques as Evaluating sample plot imputation techniques as input in forest management planning. Canadian Journal of Forest Research 37: 2069–2079.

Duvemo, K. & Lämås, T. 2006. The influence of forest data quality on planning processes in forestry. Scandinavian Journal of Forest Research 21: 327–339.

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

— 2000. Use of Uncertain Inventory Data in Forestry Scenario Models and Consequential Incorrect Harvest Decisions. Silva Fennica 34(2): 89–100.

Eid, T., Gobakken, T. & Naesset, E. 2004. Comparing Stand Inventories for Large Areas

Based on Photo-interpretation and Laser Scanning by Means of Cost-plus-loss Analyses. Scandinavian Journal of Forest Research 19: 512–523.

Eid, T. & Hobbelstad, K. 2000. AVVIRK-2000: A Large-scale Forestry Scenario Model for Long-term Investment, Income and Harvest Analyses. Scandinavian Journal of Forest Research 15: 472–482.

Finnish Statistical Yearbook of Forestry 2008. Finnish Statistical Yearbook of Forestry. Finnish Forest Research Institute, Finland.

Garcia, O. 2001. On Bridging the Gap between Tree-Level and Stand-Level Models. In: Proceedings of the IUFRO4.11 Conference held at the University of Greenwich, Forest biometry, modelling and information science.

Gertner, G.Z. 1986. Post calibration sensitivity procedure for regressor variable error. Canadian Journal of Forest Research 16(1120-1123).

— 1987. Approximating precision in simulation projections: an efficient alternative to Monte Carlo methods. Forest Science 33(1): 230–239.

— 1991. Prediction bias and response surface curvature. Forest Science 37: 755–765.

Gertner, G.Z. & Dzialowy, P.J. 1984. Effects of measurement errors on an individual tree-based growth projection system. Canadian Journal of Forest Research 14(3): 311–316.

Gobakken, T. 2008. T: A forest simulator for bioeconomic analyses based on models for individual trees. Scandinavian Journal of Forest Research 23: 250–265.

Haara, A. 2003. Comparing Simulation Methods for Modelling the Errors of Stand Inventory Data. Silva Fennica 37(4): 477–491.

— 2005. The uncertainty of forest management planning data in Finnish non-industrial private forestry, volume 8 of Dissertationes Forestales. Faculty of Forest Sciences, University of Joensuu.

Haara, A. & Korhonen, K. 2004. Kuvioittaisen arvioinnin luotettavuus. Metsätieteen aikakauskirja 4: 489–508.

Haara, A. & Leskinen, P. 2009. The Assessment of the Uncertainty of Updated Stand-Level Inventory Data. Silva Fennica 43(1): 87–112.

Haight, R.G. & Holmes, T.P. 1991. Stochastic price models and optimal tree cutting: results for loblolly pine. Natural Resource Modeling 5(4): 423–443.

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

Heinonen, T. 2007. Developing spatial optimization in forest planning, volume 34 of Dissertationes Forestales. The Finnish Society of Forest Research.

Hoen, H.F. & Gobakken, T. 1990. A model for analysis of treatment strategies for a forest applying stand-vice simulations and linear programming. Technical Report 9/90, Norwegian Forest Research Institute.

Holm, S. 1980. Att utnyttja historisk tillväxt i tillväxtprognoser. Hugin rapport no 19 p. 88.

Holmström, H., Kallur, H. & Ståhl, G. 2003. Cost-Plus-Loss Analyses of Forest Inventory Strategies Based on kNN-Assigned Reference Sample Plot Data. Silva Fennica 37(3): 381–298.

Holopainen, M. & Talvitie, M. 2006. Effect of Data Acquisition Accuracy on Timing of Stand Harvests and Expected Net Present Value. Silva Fennica 40(3): 531–543.

Huston, M., DeAngelis, D. & Post, W. 1988. New Computer Models Unify Ecological Theory. BioScience 38(10): 682–691.

Hynynen, J., Ahtikoski, A., Siitonen, J., Sievänen, R. & Liski, J. 2005. Applying the MOTTI simulator to analyse the effects of alternative management schedules on timber and nontimber production. Forest Ecology and Management 207: 5–18.

Hynynen, J., Ojansuu, R., Hökkä, H., Siipilehto, J., Salminen, H. & Haapala, P. 2002. Models for predicting stand development in MELA system. Finnish Forest Research Institute Research Papers p. 835.

Hyvän metsänhoidon suositukset 2006. Hyvän metsänhoidon suositukset. Metsätalouden kehittämiskeskus Tapio, Helsinki, Finland.

Insley, M. & Rollins, K. 2005. On solving the multirotational timber harvesting with stochastic prices: a linear complementarity formulation. Journal of Agricultural Economics 87: 735–755.

Islam, N., Kurttila, M., Mehtätalo, L. & Haara, A. 2009. Analyzing the Effects of Inventory Errors on Holding-Level Forest Plans: the Case of Measurement Error in the Basal Area of the Dominated Tree Species. Silva Fennica 43(1): 71–85.

Jonsson, B., Jacobsson, J. & Kallur, H. 1993. The forest management planning package. Theory and application. Studia Forestalia Suecacia 189.

Juntunen, R. 2006. Puustotiedon laadun vaikutus metsän käsittelyn optimoinnin tuloksiin – UPM Metsän laserkeilausaineiston ja kuviotiedon vertailu. M.sc. thesis, Helsinki University Department of Forest Economics.

Kaartinen, H. & Hyyppä, J. 2008. EuroSDR-Project Commission 2, "tree extraction", final report. EuroSDR - European Spatial Data Research, Dublin, Ireland.

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

— 1999. Methods for assessing uncertainty in growth and yield predictions. Canadian Journal of Forest Research 29: 1357–1364.

Kangas, A., Heikkilä, E. & Maltamo, M. 2004. Accuracy of partially visually assessed stand characteristics: a case study of Finnish forest inventory by compartments. Canadian Journal of Forest Research 34: 916–930.

Kangas, A. & Kangas, J. 2004. Probability, possibility and evidence: approaches to consider risk and uncertainty in forestry decision analysis. Forest Policy and Economics 6: 169–188.

Kangas, A. & Maltamo, M. 2003. Calibrating predicted diameter distribution with additional information in growth and yield predictions. Canadian Journal of Forest Research 33: 430–434.

Kangas, J. 1992. Multiple-use planning of forest resources by using the Analytic Hierarchy Process. Scandinavian Journal of Forest Research 7: 259–268.

Kangas, J., Kangas, A., Leskinen, P. & Pykäläinen, J. 2002. MCDM methods in strategic planning of forestry on state-owned lands in Finland: applications and experiences. Journal of Multi-Criteria Decision Analysis 10(5): 257–271.

Kangas, J. & Pukkala, T. 1998. HERO heuristic optimization for forest management planning: applications and experiences. In: J. Saarikko (ed.), Resource Technology '98 Nordic. International Symposium on Advanced Technology Environmental and Natural Resources, p. 13. Rovaniemi, Finland.

Kangas, J., Pukkala, T. & Kangas, A. 2001. HERO: heuristic optimisation for multi-criteria forestry decision analysis. In: D. Schmoldt, J. Kangas, G. Mendoza & M. Pesonen (eds.), The Analytic Hierarchy Process in Natural Resource and Environmental Decision Making, pp. 51–65. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Koenker, R. & Bassett, G. 1978. Regression quantiles. Econometrica 46(1): 33–50.

Koenker, R. & Hallock, K.F. 2001. Quantile regression. Journal of Economic Perspectives 15(4): 143–156.

Kokkila, T., Mäkelä, A. & Franc, A. 2006. Comparison of Distance-Dependent and Distance-Independent Stand Growth Models—Is Perfect Aggregation Possible? Forest Science 52(6): 623–635.

Kolev, N., dos Anjos, U. & Mendes, B.V.M. 2006. Copulas: a review and recent developments. Stochastic Models 22: 617–660.

Kolström, T. 1993. Modelling the development of an uneven-aged stand of Picea abies. Scandinavian Journal of Forest Research 8(1): 373–383.

Korpela, I. 2004. Individual Tree Measurements by Means of Digital Aerial Photogrammetry. Silva Fennica Monographs 3: 93 p.

Laasasenaho, J. & Päivinen, R. 1986. On the checking of inventory by compartments. Folia Forestalia 664: 19.

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

Liesiö, J., Mild, P. & Salo, A. 2007. Preference programming for robust portfolio modeling and project selection. European Journal of Operational Research 181: 1488–1505.

Lundström, A. & Söderberg, U. 1996. Outline of the Hugin system for long-term forecasts of timber yields and possible cuts. In: T. Päivinen, L. Roihuvuo & M. Siitonen (eds.), Large-scale forestry scenario models: experiences and requirements. EFI, Joensuu, Finland.

Mäkelä, A., Landsberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Ågren, G., Chadwick, D.O. & Puttonen, P. 2000. Process-based models of forest ecosystem management: current state-of-art and challenges for practical implementation. Tree Physiology 20: 289–298.

Mehtätalo, L., Gregoire, T.G. & Burkhart, H.E. 2008. Comparing strategies for modeling tree diameter percentiles from remeasured plots. Environmetrics 19: 529–548.

Mowrer, H.T. 1991. Estimating components of propagated variance in growth simulation model projections. Canadian Journal of Forest Research 21(3): 379–386.

— 2000. Uncertainty in natural resource decision support systems: sources, interpretation, and importance. Computers and Electronics in Agriculture 27: 139–154.

Munro, D.D. 1974. Forest growth models—a prognosis., pp. 7–21. Number 30 in Research Notes. Royal College of Forestry, Stockholm.

Naesset, E. 1997. Estimating Timber Volume of Forest Stands Using Airborne Laser Scanner Data. Remote Sensing of Environment 61: 246–253.

— 2004. Practical Large-scale Forest Stand Inventory Using a Small-footprint Airborne Scanning Laser. Scandinavian Journal of Forest Research 19: 164–179.

Newman, D.H. 2002. Forestry's golden rule and the development of the optimal forest rotation literature. Journal of Forest Economics 8: 5–27.

Oikarinen, M. 1983. Etelä-Suomen viljeltyjen rauduskoivikoiden kasvumallit. Commun Inst. For. Fenn. 113.

Packalen, P. 2009. Using airborne laser scanning data and digital aerial photographs to estimate growing stock by tree species, volume 77 of Dissertationes Forestales. Faculty of Forest Sciences, University of Joensuu.

Packalen, P. & Maltamo, M. 2007. The k-MSN method in the prediction of species specific stand attributes using airborne laser scanning and aerial photographs. Remote Sensing of Environment 109: 328–341.

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

Poso, S. 1983. Basic features of forest inventory by compartments. Silva Fennica 17: 313–343.

Pukkala, T. 2002. Multi-objective forest planning, chapter Introduction to multi-objective forest planning, pp. 11–21. Kluwer Academic Publishers.

— 2004. Dealing with Ecological Objectives in the Monsu Planning System. Silva Lusitana 12 Especial: 1–15.

— 2005. Metsikön tuottoarvon ennustemallit kivennäismaan männiköille, kuusikoille ja rauduskoivikoille. Metsätieteen aikakauskirja 3: 311–322.

Pukkala, T. & Kangas, J. 1993. A heuristic optimization method for forest planning and decision making. Scandinavian Journal of Forest Research 8(1): 560–570.

— 1995. A Method for Integrating Risk and Attitude Toward Risk into Forest Planning. Forest Science 42(2): 198–205.

Pukkala, T. & Kurttila, M. 2005. Examining the Performance of Six Heuristic Optimisation Techniques in Different Forest Planning Problems. Silva Fennica 39(1): 67–80.

Pukkala, T. & Miina, J. 1997. A method for stochastic multiobjective optimization of stand management. Forest Ecology and Management 98: 189–203.

Rasinmäki, J., Mäkinen, A. & Kalliovirta, J. 2009. SIMO: An adaptable simulation framework for multiscale forest resource data. Computers and Electronics in Agriculture 66: 76–84.

Reed, W.J. & Haight, R.G. 1996. Predicting the present value distribution of a forest plantation investment. Forest Science 42: 378–388.

Rubinstein, R. 1981. Simulation and Monte Carlo method. John Wiley & Sons, Inc.

Saari, A. & Kangas, A. 2005. Kuvioittaisen arvioinnin harhan muodostuminen. Metsätieteen aikakauskirja (1): 5–18.

Saaty, T.L. 1980. The analytical hierarchy process. Planning, priority setting, resource allocation. McGraw-Hill, New York.

Salminen, H., Lehtonen, M. & Hynynen, J. 2005. Reusing legacy FORTRAN in the MOTTI growth and yield simulator. Computers and Electronics in Agriculture 49: 103–113.

Saramäki, J. 1977. Ojitettujen turvemaiden hieskoivikoiden kehitys Kainuussa ja Pohjanmaalla. Metsäntutkimuslaitoksen julkaisuja 91(2).

Shugart, H.H. & West, D.C. 1980. Forest succession models. BioScience 30(5): 308–313.

Sprängare, B. 1978. A method for analysing the sensitivity of the long-range planning which depends on errors in stand data. Number 87 in Research Notes. Royal College of Forestry, Stockholm.

Ståhl, G. 1994. Optimizing the utility of forest inventory activities. Ph.D. thesis, Swedish University of Agricultural Sciences, Department of Biometry and Forest Management, Umeå.

Sterba, H., Blab, A. & Katzensteiner, K. 2002. Adapting an individual tree growth model for Norway spruce Picea abies L. Karst.) in pure and mixed species stands. Forest Ecology and Management 159: 101–110.

Tadikamalla, P. & Johnson, N. 1982. Systems of frequency curves generated by transformations of logistic variables. Biometrika 69(2): 461–465.

Tahvonen, O. & Kallio, M. 2006. Optimal harvesting of forest age classes under price uncertainty and risk aversion. Natural Resource Modeling 19(4): 557–585.

Thomson, T.A. 1992. Optimal forest rotation when stumpage prices follow a diffusion process. Land Economics 68: 329–342.

Tokola, T., Kangas, A., Kalliovirta, J., Mäkinen, A. & Rasinmäki, J. 2006. SIMO—SIMulointi ja optimointi uuteen metsäsuunnitteluun. Metsätieteen aikakauskirja 1: 60–65.

Tomé, M. & Burkhart, H.E. 1989. Distance-dependent competition measures for predicting growth of individual trees. Forest Science 35: 816–831.

Väisänen, M. 2008. Kuvioittaisen arvioinnin ja laserkeilauksen kustannustehokkuuden arviointi mäntyvaltaisten metsiköiden inventoinnissa. M.sc. thesis, Faculty of Forestry, University of Joensuu.

Välimäki, E. & Kangas, A. 2009. Kasvumallien toiminnan validointi ylitiheissä metsiköissä. Metsätieteen aikakauskirja 2: 97–112.

Valsta, L. 1992. A Scenario approach to stochastic anticipatory optimization in stand management. Forest Science 38: 430–447.

Vanclay, J.K. 1994. Modeling forest growth and yield. CAB International, Wallingford UK.

Vanclay, J. & Skovsgaard, J. 1997. Evaluating Forest Growth Models. Ecological Modelling 98: 1–12.

Vettenranta, J. 1999. Distance-dependent models for predicting the development of mixed coniferous forests in Finland. Silva Fennica 33: 51–72.

von Gadow, K. 2000. Evaluating risk in forest planning models. Silva Fennica 34(2): 181–191.

Vuokila, Y. & Väliaho, H. 1980. Viljeltyjen havumetsiköiden kasvumalllit. Metsäntutkimuslaitoksen julkaisuja 99(2).

Wang, M. & Rennolls, K. 2005. Tree diameter distribution modeling: introducing a logit-logistic distribution. Canadian Journal of Forest Research 35: 1305–1313.

Westfall, J.A. & Patterson, P.L. 2007. Measurement variability error for estimates of volume change. Canadian Journal of Forest Research 37: 2201–2210.

Yoshimoto, A. & Shoji, I. 1998. Searching for an optimal rotation age for forest stand management under stochastic log price. European Journal of Operations Research 105: 100–112.

— 2002. Comparative analysis of stochastic Models for Financial uncertainty in forest management. Forest Science 48(4): 755–766.

Zeide, B. 1993. Analysis of growth equations. Forest Science 39: 594–616.