

# ACTA FORESTALIA FENNICA



Jori Uusitalo

Pre-harvest Measurement of Pine Stands  
for Sawing Production Planning

259 · 1997

**Publishers** The Finnish Society of Forest Science  
The Finnish Forest Research Institute

**Editors** Editor-in-chief Eeva Korpilahti  
Production editors Seppo Oja, Tommi Salonen

**Editorial Office** Unioninkatu 40 A, FIN-00170 Helsinki, Finland  
Phone +358 9 857 051, Fax +358 9 625 308, E-mail silva.fennica@metla.fi,  
WWW <http://www.metla.fi/publish/acta/>

**Managing Board** Erkki Annala (The Finnish Forest Research Institute), Jyrki Kangas (The Finnish Forest Research Institute), Esko Mikkonen (University of Helsinki), Lauri Valsta (The Finnish Forest Research Institute), Harri Vasander (University of Helsinki), and Seppo Vehkamäki (University of Helsinki)

**Editorial Board** Per Angelstam (Grimsö Wildlife Research Station, Sweden)  
Julius Boutelje (Swedish University of Agricultural Sciences, Sweden)  
Finn H. Brække (Swedish University of Agricultural Sciences, Sweden)  
J. Douglas Brodie (Oregon State University, USA)  
Raymond L. Czaplewski (USDA Forest Service, USA)  
George Gertner (University of Illinois, USA)  
Martin Hubbes (University of Toronto, Canada)  
William F. Hyde (Virginia Polytechnic Institute and State University, USA)  
Jochen Kleinschmit (Lower Saxony Forest Research Institute, Germany)  
Michael Köhl (Dresden University of Technology, Germany)  
Noel Lust (University of Gent, Belgium)  
Bo Långström (Swedish University of Agricultural Sciences, Sweden)  
William J. Mattson (USDA Forest Service, USA)  
Robert Mendelsohn (Yale University, USA)  
Hugh G. Miller (University of Aberdeen, United Kingdom)  
John Pastor (University of Minnesota, USA)  
John Sessions (Oregon State University, USA)  
Jadwiga Sienkiewicz (Environment Protection Institute, Poland)  
Richard Stephan (Federal Research Centre for Forestry and Forest Products, Germany)  
Elon S. Verry (USDA Forest Service, USA)  
A. Graham D. Whyte (University of Canterbury, New Zealand)  
Claire G. Williams (Texas A&M University, USA)

**Aim and Scope** Acta Forestalia Fennica publishes monographs in forest science. These may be original research articles or comprehensive reviews aiming at a synthesis of knowledge in a particular topic. The series covers all aspects of forest science, both basic and applied subjects. Manuscripts are subject to peer review.

---

# ACTA FORESTALIA FENNICA

259 · 1997

---

Jori Uusitalo

## Pre-harvest Measurement of Pine Stands for Sawing Production Planning

---

The Finnish Society of Forest Science — The Finnish Forest Research Institute

Uusitalo, J. 1997. Pre-harvest measurement of pine stands for sawing production planning. Acta Forestalia Fennica 259. 56 p.

To enhance the utilization of the wood, the sawmills are forced to place more emphasis on planning to master the whole production chain from the forest to the end product. One significant obstacle to integrating the forest-sawmill-market production chain is the lack of appropriate information about forest stands. Since the wood procurement point of view in forest planning systems has been almost totally disregarded there has been a great need to develop an easy and efficient pre-harvest measurement method, allowing separate measurement of stands prior to harvesting. The main purpose of this study was to develop a measurement method for pine stands which forest managers could use in describing the properties of the standing trees for sawing production planning.

Study materials were collected from ten Scots pine stands (*Pinus sylvestris*) located in North Häme and South Pohjanmaa, in southern Finland. The data comprise test sawing data on 314 pine stems, dbh and height measures of all trees and measures of the quality parameters of pine sawlog stems in all ten study stands as well as the locations of all trees in six stands. The study was divided into four sub-studies which deal with pine quality prediction, construction of diameter and dead branch height distributions, sampling designs and applying height and crown height models. The final proposal for the pre-harvest measurement method is a synthesis of the individual sub-studies.

Quality analysis resulted in choosing dbh, distance from stump height to the first dead branch (dead branch height), crown height and tree height as the most appropriate quality characteristics of Scots pine. Dbh and dead branch height are measured from each pine sample tree while height and crown height are derived from dbh measures by aid of mixed height and crown height models. Pine and spruce diameter distribution as well as dead branch height distribution are most effectively predicted by the kernel function. Roughly 25 sample trees seems to be appropriate in pure pine stands. In mixed stands the number of sample trees needs to be increased in proportion to the intensity of pines in order to attain the same level of accuracy.

**Keywords** crown height estimation, diameter distribution, forest sampling, height models, lumber quality prediction

**Author's address** University of Helsinki, Department of Forest Resource Management, PL 24, 00014 University of Helsinki, Finland. **Fax** 358-0-1917755

**Email** jori.uusitalo@helsinki.fi

**Accepted** 12 November 1997

ISBN 951-40-1599-1

ISSN 0001-5636

Tampere 1997, Tammer-Paino Oy

## Contents

1	INTRODUCTION .....	6
2	LITERATURE REVIEW .....	8
2.1	Predicting the quality of pine stems .....	8
2.2	Diameter distribution .....	10
2.3	Forest sampling methods .....	11
2.4	Tree height prediction .....	14
3	PURPOSE OF THE STUDY .....	15
4	EXPERIMENTAL STUDIES .....	16
4.1	Predicting the quality of pine stems .....	16
4.1.1	Material and methods .....	16
4.1.1.1	Field material and test sawings .....	16
4.1.1.2	Applying logistic regression to modeling lumber quality ...	17
4.1.2	Results .....	19
4.1.2.1	Yield of sawn timber .....	19
4.1.2.2	General lumber quality models .....	19
4.1.2.3	Between-stand variation .....	25
4.1.3	Appropriate stem characteristics for practical purposes .....	27
4.2	Predicting diameter and dead branch height distribution .....	28
4.2.1	Material and methods .....	28
4.2.1.1	Study material .....	28
4.2.1.2	Methods for predicting distributions .....	28
4.2.1.3	Random repetition technique .....	30
4.2.2	Results .....	30
4.3	Sampling design studies .....	32
4.3.1	Material and methods .....	32
4.3.1.1	Study material .....	32
4.3.1.2	Sampling simulations .....	32
4.3.2	Results .....	34
4.3.2.1	Precision of the sampling designs .....	34
4.3.2.2	Evaluation of the optimal sample size .....	38
4.4	Tree height and crown height model studies .....	40
4.4.1	Material and methods .....	40
4.4.2	Results .....	42
5	PRE-HARVEST MEASUREMENT APPLICATION .....	45
6	GENERAL DISCUSSION .....	47
6.1	Materials and methods .....	47
6.2	Results .....	47
6.3	Evaluation of the proposed pre-harvest measurement method .....	49
6.4	Future perspectives and final remarks .....	50
	REFERENCES .....	51

## Abbreviations

Variable	Unit	Explanation	Variable	Unit	Explanation
A	years	Age	$H_{2, \text{cut}}$	m	Second cut height, distance from stump height to the point of second cut of the stem
Dbh	cm	Diameter at breast height (1.3 m)	$I_{2-4}$		Early growth rate at stump height, number of annual rings between the distance 2...4 cm from the pith
H	m	Tree height	$N_{\text{ib}}$		Number of inner boards in a log. A board is an inner board if at least one other board is sawn from outer face at the same side of the log.
$H_c$	m	Crown height, distance from stump height to the base of live crown	$N_{\text{ob}}$		Number of outer boards in a log. A board is an outer board if there is no other board sawn from outer face at the same side of the log.
$H_{\text{dbr}}$	m	Dead branch height, distance from stump height to the first dead branch (min 10 mm in diameter for trees under 20 cm in dbh, otherwise min 15 mm)			
$H_{\text{lbr}}$	m	Live branch height, distance from stump height to the first live branch			
$H_{\text{l-gr}}$	m	First grade height, distance from stump to the end of I-grade log			
$H_{\text{ls}}$	m	Height of log section, distance from stump height to the end of log section			

## Acknowledgements

This study was carried out at the Department of Forest Resource Management, University of Helsinki during 1992–1995. I would like to express my deepest gratitude to Prof. Esko Mikkonen for providing me with good research facilities and for his support and encouragement especially during the early stages of my work. I am greatly obliged to Prof. Arto Usenius from the Technical Research Center of Finland (VTT) for his marked guidance, criticism and encouragement, and for all the support provided by his laboratory during my research. I would also like to express my thanks to Prof. Rihko Haarlaa for his valuable advice and comments during my doctoral studies. I really appreciate his attitude towards young researchers; especially for giving his time to all the questions and queries that a young researcher has to face.

I am very grateful to Mr. Veli-Pekka Kivinen, M.Sc. and Mr. Timo Säarentaus, M.Sc. for all their help I have received both in field work and data processing. This extremely laborious task would not have been possible to carry out without exceptional, conscientious and assiduous research assistants. I would also like to thank Dr. Herman Hakala for his invaluable comments and support during my doctoral studies.

I am obliged to the staff of Aureskoski Oy for providing our group significant technical backing during this process. I would especially like to thank Mr. Juha Hakala, general director; Mr. Toivo Paloneva, head of the wood procurement department and his staff for providing study stands and for technical support in logging and haulage operations, and Mr. Juha Yli-Koski for his support in grading arrangements. I am also indebted to the Hietaniemi family for providing us with excel-

lent test sawing facilities and Mr. Timo Vainio for his beneficial advice in test sawing arrangements.

I thank Dr. Juha Lappi and Mr. Esa Ärölä for their help with the secrets of mixed models and Mr. Hannu Rita, Phil.lic. for his guidance in logistic regression analysis. I am also grateful to Prof. Mats Nylinder, Prof. Pentti Hyttinen, Prof. Rolf Birkeland and the two anonymous reviewers for their comments and criticism, to Ass.Prof. Jouko Laasasenaho for his comments on the sampling simulations and to Prof. Simo Poso for his comments on this manuscript.

I am deeply indebted to the Academy of Finland for enabling me to concentrate on this work three solid years. This study has also been supported by the Technology Development Center, Finland (TEKES) and by a grant from the Emil Aaltonen Foundation. I also wish to thank Dr. Roderick McConchie and Ms. Sara Verville for proofreading my English.

I would also like to extend my thanks to the personnel of our department for their helpful attitude as well as to my friends at the University of Tampere and the Tampere Technology Centre Hermia, where I have partly written this thesis.

Many thanks to my parents for their encouragement and support during my doctoral studies.

Lastly, I wish to express my warmest gratitude to my wife, Kaarina, for her love and support and to our precious children Antti, Heikki, Hanna and Annika who really have made our lives worth living.

Champaign, Illinois in June 1997

*Jori Uusitalo*

# 1 Introduction

The competitiveness of the Finnish sawmill industry is largely based on the high quality of the raw material. Due to short, cool summer periods the forests in Finland grow slowly with the consequence that the most commercially valuable conifers, Scots pines (*Pinus sylvestris* L.) and Norway spruces (*Picea abies* Karst.) are mostly less branched and hard in wood structure. The highest quality grades of sawn goods of the Finnish sawmills are mainly exported to joinery, furniture and panel board factories in Europe while the remainder, poorer quality grades, are sold mainly to the Finnish and European construction industry.

In spite of high efficiency in all branches of sawmill industry (i.e. wood procurement, sawing, marketing, transportation, etc.) the sawmills cannot completely exploit the great variety and the high quality of Finnish forests. Poor selection of stands to use, inappropriate cross-cutting of stems and abandonment of the best selling and secondary processing opportunities for each quality grade mean that sawmills lose a great amount of money yearly. Accordingly, to enhance the utilization of the wood the sawmills are forced to place more emphasis on planning to master the whole production chain from the forest to the end product.

For a decade already, sawmill managers have emphasized the importance of incorporating information about the quality and dimensions of the forests and markets into production planning (Lönnner 1985). Rapid changes in the international timber trade are making the need for this new approach to traditional production planning even more important. Today timber from Finnish sawmills is more often traded direct to customers and sawn to the customer's special dimensions (Niemelä 1993). Commercial activities as well as transportation are also running at increasingly high tempo, placing more demands on production planning. Consequently, beside the new market-oriented strategy, more advanced decision support

systems (DSS) are required at the sawmills.

A consequence of this market orientation is that in wood procurement of sawmills the management actions are increasingly focusing on tree-by-tree management. When logging is carried out by the log-length method the decision to direct crosscutting according to customers' orders has to be done in the forest. The information about the length and dimension combinations of the logs needed in production has to be sent into the forest to the harvester operators. The technology for optimizing the crosscutting of the stems is already available to modern harvesters. The optimization of crosscutting according to customers' needs may, however, lead us to totally inappropriate results if the crosscutting optimization is carried out in unsuitable stands. The optimal utilizing of the wood may be achieved only by taking the whole forest reserve to be harvested during the planning period into account and by conducting the cross-cutting at a central location.

One significant obstacle to integrating the forest-sawmill-market production chain is the lack of appropriate information about the forest stands. In order to develop a more custom-oriented and more integrated production chain, we need more specific and accurate information about the dimensional and quality distributions of the stands. Such information about the forests enables the sawmills to allocate the raw material available in the best way possible. With the help of better integration the sawmills may reduce trimming losses, avoid unmarketable wood raw material and steer the flow of wood into the best secondary processing destinations.

Unfortunately, the information provided by traditional forest planning systems is insufficient to fulfil the needs of sawmill production planning. Despite the great demand, the need to incorporate the view point of wood procurement and sawing production planning into forest planning systems has until recently been disregarded. This is, how-

ever, partly due to the lack of skill of sawmill and wood procurement managers in determining the changes required for these systems.

Shortcomings in the traditional forest planning systems have forced sawmill and wood procurement managers to find other solutions to provide more information about the forests. The Finnish forest industry has shown a particular interest in developing some kind of pre-harvest measurement method that could provide more accurate planning information about forests (e.g. Lemmetty and Mäkelä 1992). The experiences from these separate measurement trials are, however, not encouraging. The methods applied have been inefficient and too laborious to be feasible. Since an interface with the production planning systems has been lacking the measurement objectives have remained unclear. This being the case, it is no wonder that this kind of pre-harvest measurement, where separate measurement of stands prior to harvesting is done by wood procurement managers, has not become general.

Some scepticism about the importance of the information provided by pre-harvest measurement has also been expressed. The sceptics have pointed out that cross-cutting by modern harvester optimization systems is a sufficient tool to extend the custom-oriented philosophy to wood procure-

ment. The sawmill managers, especially those who are in charge of production planning have, however, emphasized that generous information about the quality and size of the timber coming to the sawmill during the planning period would facilitate the execution of production planning, and what is more important, would enhance production and increase income. Moreover, the significance of prior information in production planning is thought to be more important if it could be automatically interfaced with operative production planning systems.

Altogether, it has been proved that new kind of measurement method would be needed to provide information about the forests for sawing production planning. Since the wood procurement managers are in every case daily visiting in the stands they are supposed to buy, efficient information collection would not greatly inconvenience their work but would provide significant additional benefits to the sawmill. A prerequisite in developing a new measurement method is that it should be linked with the operational production planning systems in order to achieve as many advantages as possible. The method should also be easy, efficient and sufficiently accurate to dovetail with the manager's daily program.

## 2 Literature Review

### 2.1 Predicting the Quality of Pine Stems

The skill of predicting the quality of standing trees prior to harvesting has awakened interest among foresters for various reasons. Earlier studies concerning the quality of stems or logs have mostly dealt with the value relations of sawlogs of different grades. Since the poorer quality logs inevitably lead to the production of large quantities of lower quality grades and reduced profit, the sawmill managers have fought for log pricing systems that ensure appropriate payment for quality raw material. Study of the log or stem valuation systems has been most active in the United States (e.g. Brundage 1936, Kellogg 1941, Gregory and Pearson 1949) but similar studies have also been carried out in Finland (Vuoristo 1936, Heiskanen 1951, 1954a, Kärkkäinen 1980, Uusitalo 1989) and in Sweden (Orver 1970, Weslien 1983).

More recently, it has been noticed that information about the value relations of the sawlogs does not fully satisfy the needs of modern production scheduling. A normal way to improve the accuracy of information required is to construct a linear regression model for predicting yields of individual grades of lumber from sawlogs. The models have been derived both for logs (e.g. Yaussy 1986, Howard and Yaussy 1986, Howard and Gasson 1989, Howard 1991) when independent variables are formed from log characteristics and for standing trees (e.g. Stayton et al. 1971, Hanks 1976, Kärkkäinen 1980) when independent variables are formed from stem characteristics.

The traditional linear regressions have some serious shortcomings in quality prediction however, since grading makes the quality of lumber a categorical variable. A characteristic of sawmill yield data is that yields of individual grades from single logs are often zero, especially for the higher grades. The situation is more problematic if the percentage of grade yields only are predicted, since the percentage of different grade may yield only

a few values. When regression analysis is applied to the data, where the dependent variable gets only few values, the use of least square methods (LS) creates ineffective estimators and may result in biased models. Furthermore, regression models derived from the grade yield data may give predictions including negative values for some grades.

To avoid the most serious problems of the LS method Howard and Gasson (1989) applied so-called TOBIT procedure to derive the estimators for regression models of lumber grade yields. The TOBIT model is based on the assumption that the same parameters determine both the value of the dependent variable when it is greater than the bound and the probability that it is at the bound. Uusitalo (1994a) analyzed different kinds of statistical methods for predicting the quality of the stems and chose logistic regression analysis as the method to create models that predict the appearance of lumber grades in the stems. Unlike the linear regression models logistic models can deal with the relative changes in the comparison between the dependent variable and predictors. This is extremely valuable in the prediction of stems, since the relation between the main quality predictors such as distance from ground to the lowest dead branch and the quality is not linear (Uusitalo 1994a). As distinct from other distribution functions proposed for use in the analysis of a dichotomous outcome variable, logistic distribution is from a mathematical point of view an extremely flexible and easily used function producing a biologically meaningful interpretation (Hosmer and Lemeshow 1989).

In Uusitalo's studies (1994a, 1994b) binary variables only were used as dependent variables in the regression models created for stems. Although most of the studies, especially in medical science, have applied only binary dependent variables in logistic regression analysis, there are, however, no reasons to restrict the use of logistic

regression to binary dependent variables in the case of predicting the appearance of lumber grades. Polytomous data may be modeled for example by using proportional odds model (McCullah and Nelder 1989), where the cumulative probabilities are transformed into the logistic scale and modelled using parallel linear regression. In forestry, proportional odds model has been applied by Leinonen and Rita (1995) and Schabenberger (1995).

Since the sophisticated lumber grade models have to be created separately for the battens and for different groups of boards, the number of alternatives among dependent variables is not expanding too much. The most typical set-ups in cant-sawing are based on separating two battens from the pith when the appearance of a single grade in the battens gets the values of 0, 1 or 2. The remaining boards may also be separated into different groups that are most meaningful from the production planning point of view.

There are, however, other efforts to predict quality from the logs or from the stems. Some efforts have been made to quantify the relationships between the outer quality parameters of logs (Samson 1993a) or stems (Kärkkäinen 1986) and knots in lumber. This approach may be very interesting in the future if methods such as that proposed by Samson (1993b) are able to assess the effect of knots and convert the logs into structural lumber as a result of various sawing patterns. The rapid development of scanning technology also presents an alternative opportunity to assess the internal log defects (e.g. Chang 1990).

Unlike many parts of the world, the timber sawn in Finnish conditions from Scots pine does not suffer greatly from insects, cankers and other pests. Instead, the number of knots, their size and whether they are alive or dead, is by far the most important property for classifying boards into separate quality grades (Vientisahatavaran ... 1979). Therefore, in predicting the lumber quality of pine stems all the stem characteristics associated biologically in these lumber knots properties have to be considered as possible predictors.

Historically, the sawmill managers in Finland have been very interested in the amount of u/s grade sawn from the pine logs; quite understandably because the value of u/s quality in the market may be more than twice that of the lower quality

grades. The u/s grade is a combination of the first four (I...IV) quality grades, which are very seldom sold separately in the European market. The u/s lumber is obtained almost entirely from the butt logs, where the knots are small and as a consequence of natural pruning the outer boards may be entirely knotless. Recently, there have been some changes in the appreciation of other quality grades. The respect for sound knots as a result of their considerable usefulness in furniture manufacturing processes has clearly increased. This has resulted in the development of a new quality grade, knotty pine, in which the appearance of dead knots is totally forbidden. Knotty pine may be obtained only from the live crown section. The great desire for logs including only live branches as well as the development of new sawing technology have changed the traditional grouping of timber assortments in Finland.

For many years there has been activity and interest in studying the effect of different stand characteristics including stand density and site fertility (Kellomäki and Tuimala 1981, Kärkkäinen and Uusvaara 1982, Kellomäki and Väisänen 1986, Turkia and Kellomäki 1987) as well as the effect of silvicultural actions like regeneration method, thinning, pruning and fertilization (Heiskanen 1965, Uusvaara 1974, Kärkkäinen and Uusvaara 1982, Varmola 1982, Mäkinen and Uusvaara 1993) on the quality of Scots pine. As a result of this work we know that stem branchiness decreases within limits with increasing stand density and the high growth rate affected greatly by the site fertility increases stem branchiness. Despite the clear correlation between site fertility and the quality of lumber, the forest type classified according to Cajander's (1926) system has not been found to be a good predictor of quality (Heiskanen 1965, Kärkkäinen 1980, Turkia and Kellomäki 1987).

Since the stand characteristics are unable to predict the quality of the stand, the appropriate predictors of the stem quality have to be found among the stem characteristics. The previous studies show that three important stem characteristics seem to play a significant role in predicting the quality of Scots pine butt logs; namely, diameter at breast height (dbh), the distance from stump height to the lowest dead branch (dead branch height) and early growth rate.

From these three stem characteristics dead branch height is considered to be the best individual predictor of the quality of Scots pine (e.g. Kärkkäinen 1980, Uusitalo 1994a). The correlation between the dead branch height and the quality of the butt logs can be explained by the diameter growth of branches at the early age. If the diameters of branches in the butt become large, natural pruning takes a long time with the consequence of low dead branch height and poor stem quality. Early growth rate, usually expressed in the thickness of the annual rings around the pith in the butt of the tree, is found to be another quite reliable predictor of quality (Heiskanen 1954b, 1965, Orver 1970, Weslien 1983). The branches of the butt logs are usually considered to be thicker the faster the tree has developed when young. Although the growth rate is also considered to be associated with the dbh, it does not mean that wide diameter indicates poor quality. On the contrary, large trees are commonly associated with good quality, largely because large amounts of u/s boards may only be obtained from large butt logs.

## 2.2 Diameter Distribution

The simplest and most widely used density estimator for diameter distributions is the histogram. Since the diameters are in practice classified into 2 cm or 1 inch classes the diameter distribution may be depicted simply by summarizing the observations in each diameter class. Despite its simplicity, the histogram has many shortcomings both in practical applications as well as in scientific calculations. Firstly, sampling or complete enumeration of the diameters in the forest has in many cases proved to be a too time-consuming task. Secondly, the discontinuity of histograms causes extreme difficulties in many mathematical calculations, e.g. in derivations (Silverman 1986). Altogether, the appropriateness of histogram in depicting the diameter distribution in increasingly demanding decision-making systems has to be questioned.

For almost a century there has been activity and interest in describing the frequency distribution of diameters in the forest by using probability density functions. Since the work of de Licourt (1898) applying exponential distribution to fre-

quency data from all-aged forest, diameter distributions in forest stands have been modeled by numerous probability density functions. Although it has been generally known that the normal distribution is inappropriate to describe the shape of diameter distributions, there have been many investigations trying to apply normal distribution for forestry purposes by adequate transformations. As early as 1914 a well-known Finnish forest scientist Werner Cajanus applied Gram-Charlier expansion to describing the diameter distributions in even-aged stands. Later, Petterson (1955) used truncated normal distribution and Bliss and Reinker (1964) lognormal distribution for construction of diameter distribution.

The Weibull function is no doubt the most extensively used model for depicting the diameter distribution of trees. Since the landmark paper of Bailey and Dell (1973) presenting the application of the Weibull function for quantifying the diameter distribution of trees, it has been applied in numerous investigations. The reason for its popularity is not based on biological explanations, but purely on its mathematical simplicity and flexibility in depicting the full range of unimodal, continuous shapes. Although the Weibull function has proved to be easily used and sufficiently flexible to fit a relative broad spectrum of shapes, the estimation of its parameters can be very laborious. Bailey and Dell (1973) suggested that with access to an adequate computer, maximum likelihood estimation is the best method while in calculations restricted to hand computations easier methods such as percentile estimation should be used. Zarnoch and Dell (1985) and later Shiver (1988) strengthened the opinion that the maximum likelihood method is superior in accuracy compared to the percentile method in evaluating the Weibull parameters for diameter distribution. Hyink and Moser (1983) proposed an alternative method, parameter recovery, for same purpose. Rather than directly predicting the future values of the parameters, parameters are recovered from estimates of stand attributes which can be expressed in terms of the dbh distribution. Burk and Newberry (1984) elaborated on this recovery technique and developed a simple algorithm for moment-based recovery of the Weibull distribution parameters.

It seems that the academic discussion about the

method providing the most accurate parameters for the Weibull distribution has in many cases forgotten the restrictions in forestry in applying these methods in practice. One interesting practical approach has been taken by a Finnish group (Mykkänen 1986, Kilkki et al. 1989) who tried to generalize the shape of the Weibull distribution. They created regression models for spruce (*Picea abies*) and pine (*Pinus sylvestris*) from the Finnish National Forest Inventory data to derive the Weibull parameters when stand characteristics like basal area, median basal area diameter and age only are known. Swindel et al. (1987) have offered another practical approach by developing functions and diameter distribution algorithms for hand-held calculator.

Beside the Weibull distribution many similar distributions have been applied for forestry purposes. Zöhrer (1969) and Päivinen (1980) have proposed beta, Nelson (1964) gamma and Hafley and Screuder (1977) Johnson's  $S_B$  distribution to describe diameter distributions. Recently, there has been much use of nonparametric methods to determine the diameter distribution. It has been noticed that the parametric probability functions are not always adequate for diameter distributions, especially in uneven-aged stands; hence greater flexibility than is possible with unimodal distribution functions is desirable. The percentile-based method developed by Borders et al. (1987) does not necessitate the assumption of a particular probability distribution. Instead of using percentiles for determining the parameters, they utilized the information from observed stand tables more completely and defined an empirical probability density function with 12 percentiles. Droessler and Burk (1989) compared nonparametric methods, the kernel estimator and frequency polygon-averaged sifted histogram (FP-ASH) estimator to the Weibull function in their pioneering study. Since the smoothed nonparametric methods can describe every detail in any diameter distribution they are superior in bimodal or thinned populations to the Weibull function.

The kernel estimator is actually an aggregate of symmetric probability functions, kernels, centred at the observations. The choice of kernel function has been shown not to be critical and is usually based on computational reasons. The most widely used kernel function is the Epanechnikov (1969)

function. The problem of defining the appropriate smoothing parameter value is of crucial importance in density estimation. Although numerous methods, more or less automatic, have been developed for choosing the smoothing parameter (see Silverman 1986, p. 43–61), it has to be stressed that its users should not violate the nature of the object. As to the diameter distribution of trees, we have to consider how the nearest diameter classes centred round the observation can be weighted. If the smoothing parameter value is too large, the ability to depict the details in the distribution disappears. On the other hand, shrinkage of the smoothing parameter reduces the effect of smoothing and forces us to increase the sample size. The smoothing parameter needs not to be the same at every point. In adaptive kernel method we may use a broader kernel in region of low densities (Silverman 1986, p. 100–110.).

## 2.3 Forest Sampling Methods

Systematic sampling with various modifications is, no doubt, the most extensively used forest sampling method for its convenience and efficiency in practice. Systematic sampling compared to random techniques are thought to produce a more representative picture of the forest. Furthermore, random sampling has proved to be very complicated in the field. In addition to that systematic sampling has in most cases proved to be superior in accuracy to random sampling. The earlier studies of systematic sampling concentrated mainly on assessing the most efficient size and shape of the sample plots for large areas (e.g. Ilvessalo 1935, Meyer 1949, Johnson and Hixon 1952, Mesavage and Grosenbaugh 1956). Very soon however it was realized that the accuracy and above all the efficiency of systematic sampling may be increased by stratification the managed area. It amounts to dividing an area into a series of distinct classes (strata) and then sampling within each class. The greater the difference among classes the more advantageous the stratification is. Its effective use reduces the variation among individuals of the same class by increasing the differences among the class means. There are several criteria which may be used to define the classes. Delineation of classes and determination of stra-

tum areas may be done most easily with aerial photographs (Bickford 1952, 1961). The stratification may also be based on so-called treatment classes (Nyyssönen and Kilkki 1966) or forest site. Nyyssönen and Vuokila (1963) applied stratification based on the dominant height calculated as the mean height of the 100 trees with the largest diameter per hectare.

Beside systematic sampling with fixed plots, angle count sampling (ACS) better known as the relascope method or point sampling developed by Austrian forest scientist Bitterlich (1949) has become a very popular forest sampling method. The advantage of this method is that the cruiser can avoid the measurement of dbh and obtains reliable estimates of basal area with a relatively small amount of work. In predicting the mean volume or basal area of the stand, the relascope method has proved to be more efficient than systematic or random sampling with fixed plots in most of the studies (e.g. Avery and Newton 1965, O'Regan and Arvanitis 1966, Wensel and John 1969). Despite its advantages both theoretically and in operational experiments, the relascope method has not received unreserved approval. First of all, it is recognized in practise as quite vulnerable to various sources of error (e.g. instrument error, slopover and edge bias, hidden trees). Another unpleasant feature of the relascope method is the evaluation of so-called borderline trees. Regardless of the chosen angle, some portion of the trees is always so close to the edge of the tree's theoretical plot that it is hard to judge whether the tree is "in" or "out". A common procedure is to control the correctness of the borderline trees using diameter caliper and measurement tape, which is a time-consuming task. According to Iles and Fall (1988) the number of trees checked is usually from 10 to 18 % of the total cruise. Considering the relation in time taken, it is obvious that control of borderline trees is in many cases more time-consuming than the sampling itself. Control of borderline trees is thus left out in some easier samplings.

Since Bitterlich's paper (1949), the interest in applying the relascope in the forest sampling has been overwhelming in many parts of the world. Many practical applications based on the idea of ACS have been developed to estimate mean volumes of the stand (Grosenbaugh 1952, Nyyssönen

1954, Strand 1964), tree height (Hirata 1955), thinning needs (Nyyssönen 1963), stand density (Unterdörfer 1955), just to mention a few. Kuusela (1966) even proposed that almost every essential stand characteristic may be calculated from the mean tree weighted by the basal area. ACS has also stimulated researchers to construct other tree sampling methods with unequal selection probabilities. Instead of sampling with a probability proportional to tree size (p.p.s.) like in ACS, Grosenbaugh (1965, 1967) developed a sampling method where trees are selected with probability proportional to prediction (the 3P method). In this method the selection probabilities were made proportional to any variables of interest which are ocularly estimated for the setting of selection probabilities. Zöhrer (1978, 1979) developed a point sampling called sampling with programmed probability (SPP sampling) where sampling tree selection may be chosen on a large scale from fixed plots to ACS according the characteristics of the forest. The weakness of these methods is that they need prior knowledge of the forest which means that in order to choose the optimal parameters the forest needs to be cruised twice.

Attempts to increase the accuracy of sampling have led to development of various types of simultaneous or successive sampling designs. A common feature of subsampling or multi-phase sampling is that we use auxiliary variables to derive ratio or regression estimates of the variable of interest. At its simplest, deriving the tree height from the breast height diameter by utilizing the relationship established from the first-phase sample trees may be considered as an example of multi-phase sampling. However, many multi-phase applications suggested for forest sampling require prior knowledge of the forest area in the form of aerial photographs, ocular estimates or even measured estimates. In most intensive methods, sampling the area twice or even more is required. The rapid development in computer techniques enabling large sets of simulations have led to development of even more complicated sampling procedures. These are in many cases combinations of the methods already mentioned above including multi-phase features and regression models between the variables (e.g. Screuder et al 1987, Screuder and Ouang 1992). The usefulness of these methods in common practice is unclear be-

cause their benefit has been demonstrated only in sampling simulations.

The superiority of ACS compared to systematic sampling with fixed plots seems to be clear where mean volume or basal area of the forest is chosen as the predicted parameter. Due to the evident need to estimate the parameters describing the features of the forest, the popularity of ACS has markedly decreased during the past decade. On the other hand, the popularity of systematic sampling with fixed plots seems to be increasing since it enables the estimation of mean volume, stand density and diameter distribution at the same time. The Nordic Countries especially seem more interested in applying circular plot sampling in smaller forest areas (Lindgren 1984, Lemmetty and Mäkelä 1992). One unpleasant feature in non-intensive circular plot sampling is that because of the clustering of the stand, the number of trees in an individual plot may be very small or even zero. To avoid this problem Jonsen et al. (1992) developed a density-adapted method where the size of sample plots can be fixed according to the density of the stand. The advantage of this method is that the work per sample plot is proportional to the amount of useful information gathered per sample plot. These methods were found to be little biased and since the researchers have also developed practical electronic instrumentation, it may have some advantage in practice.

A desirable feature of any sampling method is the knowledge of its precision. Since no valid theory exists for estimating the standard error of the systematic sampling, foresters have often applied standard random sampling variance formulae to estimate the precision of systematic samples. It is generally recognized that treating systematic samples as simple random samples results in an overestimate of sampling error. To avoid this, Nyyssönen et al. (1967) established a regression equation to estimate the standard error of systematic sampling. Payandeh and Paine (1971) suggested the solution of the so-called nonrandomness index as an aid in predicting the relative precision of systematic sampling. Shiue (1960) and Shiue and John (1962) suggested the use of multiple random starts for same purpose. This approach may be however considered almost entirely impossible in practice.

Optimization of plot size and shape is one of the basic interests in the literature on forest sampling. It is generally known that an increase in sample size results in a decrease in the coefficient of variation. However, the increase in sample size inevitably leads to a reduced number of plots if the cruising time factor is limited. Hence, a common way to analyze the optimum size of the plot is to compare the coefficient of variation with the cruising time. The most efficient sampling method is then that which provides the desired precision at the lowest cost or the greatest precision for a limited amount of money. Numerous empirical studies concerning optimum plot size and shape have been conducted. Some of the more outstanding papers are: Johnson and Hixon (1952), Mesavage and Grosenbaugh (1956), Strand (1957), Prodan (1958), Kangas (1959–60), Nyyssönen and Vuokila (1963), Nyyssönen and Kilkki (1965), Avery and Newton (1965) and Kulow (1966). Techniques that optimize the right plot size have developed by Freese (1961), Wiant and Yandle (1980) and Zeide (1980).

The results of the studies cited deviate considerably from each other, both in the size and shape of the plot. Due to the variations in diversity, purpose of the inventories or the time calculations used, the optimum plot size has been estimated as from 100 (Kangas 1959–60) to 1349 m<sup>2</sup> (Johnson and Hixon 1952). Correspondingly, both long and narrow rectangular, square as well as circular shapes has been proposed as best plot shape. A general conclusion has been, however, that plot shape does not usually affect the size of sampling error and the optimal plot shape may be determined with the aid of time calculations.

Efforts to improve the quality of the information needed the forest management planning has recently inspired researchers to incorporate spatial within-stand variation into management planning procedures (Tomppo 1986, Pukkala 1990, Penttinen et al. 1992, Fuldner 1995). The disadvantage in feeding spatial information into forest management is that it inevitably increases sampling time. Such being the case, forest managers need in future to assess whether it is reasonable to meet additional costs in order to manage the forest by spatial data.



## 2.4 Tree Height Prediction

A common procedure to obtain a height estimate for each tree in the stand is to relate dbhs to tree heights measured from the sample trees and then derive the height estimates from dbh measures. The height-diameter curve is best expressed as a single equation whose parameters are usually fitted by linear regression methods. For practical reasons, it is very important that the number of tallied trees is minimized since the measurement of tree height is relatively laborious. A wide variety of equations have been used world-wide for describing the height-diameter relationship (See, e.g. Curtis 1967). The most widely used tree height model used in Finland is Näslund's (1936) which has the form

$$h = d^2 / (a + bd)^2 + 1.3 \quad (1)$$

The reason for its popularity is its flexibility and simplicity in mathematical calculations. Since no common estimates for the parameters  $a$  and  $b$  has been derived, the parameter estimates have to be calculated for each stand separately. As a consequence, formulation of a reliable height curve in a stand necessitates a great number of tallied trees.

Alternatively, the use of Näslund's formula or some other similar formula with general fixed parameters calculated from the population inevitably provides biased predictions in a stand that deviates from the population average (Lappi 1991a).

Instead of using general parameters or stand specific parameters, the height curve for a stand can also be formulated by combining prior knowledge about the population and the information obtained from the trees measured. The less the number of sample trees is the greater is the benefit of using prior knowledge in the model. The simplest way to utilize prior knowledge is to apply a general height model and formulate a calibration coefficient by relating the height estimates of this general model to the calibration trees. This means that the height curve is transferred linearly from the basic level to the calibrated level of the stand. A more sophisticated procedure is to apply mixed models containing both fixed and random parameters. In this approach random effects both for the constant as well as for slope are calculated from calibration trees. Theoretically, a mixed linear height model like Lappi's (1991a) should be more explicit compared to use of a calibrated general model because in this approach the stand-specific relationship between height and diameter is also taken into account.

## 3 Purpose of the Study

The purpose of this study is to develop a measurement method for pine stands which foresters can use in describing the properties of the standing trees for sawing production planning. The method will be applied in providing data for a comprehensive stem-optimization sawing model. Therefore, the whole stands will be described through a total enumeration of trees or alternatively with 30...100 "element trees" characterized by the variables chosen.

A comprehensive sampling method comprises several elements. Before the method can be approved one must be able to demonstrate its practicality in various details including measured variables, sample size, time consumption, types and extent of error and methodological restrictions. To be able to resolve as many difficulties in this list as possible I have had to divide the study into several sub-studies. In the first sub-study (section 4.1) I shall present a large set of models that predict the appearances of the quality grades. The basic aim of the quality analysis is to evaluate the effect of the stem characteristics in predicting the quality of pine stems and to choose the most appropriate characteristics for the method. Once the stem characteristics most suitable for the method have been chosen, the techniques for estimating the values and the distributions of these characteristics within the stand are discussed in the following sub-studies.

In section 4.2 I shall test the various ways to construct the diameter at breast height and the dead branch height distribution of a stand. The aim of this part is to determine the most efficient ways of predicting relative diameter and dead branch height distributions and to evaluate the effect of sample size on the precision of these distributions. In the following section (4.3) I shall compare different sampling designs with regard

to the real frequencies of trees per hectare and to dead branch height distributions. This analysis uses a sampling simulator specially programmed for this study. This section also includes analysis of the optimal sample size. In the last sub-study (section 4.4) I shall test the efficiency of different approaches in constructing the height and crown height model for each stand.

After the individual sub-studies I shall describe the method as a synthesis of the sub-studies. This is followed by an analysis of how the variables estimated by the distributions and models can be linked together in order to construct a realistic description of the trees in a stand.

The method is supposed to be used operatively in conjunction with the manager's daily timber buying that impose some restrictions on it. Firstly, the time that can be spent for this measurement has to be short, a maximum one hour per stand. If the measurement work is too time-consuming, the operative utility of the method becomes questionable. Secondly, only ordinary cheap and light measuring instruments can be used in measurement work. In the ideal case, however, the measurement results are stored on a hand-held field computer which provides an interface with PCs. Due to the strict restriction of the measurement method, all the traditional guidelines of forest inventory or forest sampling have to be questioned. Additionally, the research in the modeling and estimating of diameters, heights and the qualities of the trees has to be utilized in order to attain more accurate estimates with the minimum number of sample trees. The method to be developed does not have to be valid in all kinds of forests. In the first phase, the pre-harvest measurement is only meant to be used in sawmill wood procurement which means that measurement work is restricted mainly to the final cut stands.

## 4 Experimental Studies

### 4.1 Predicting the Quality of Pine Stems

#### 4.1.1 Material and Methods

##### 4.1.1.1 Field Material and Test Sawings

The data used in this study were collected in close collaboration with Aureskoski Oy in 1993. The Aureskoski sawmill is located in Parkano, in southern Finland, 80 km Northwest of Tampere, producing approximately 140 000 m<sup>3</sup> sawn goods annually and about 60 % of which is marketed in foreign countries (e.g. Germany, Denmark, the Netherlands, the United Kingdom, France, etc.). Wood comes almost entirely from local private forest owners and is procured by its local wood procurement department. The remainder is supplied by exchange deals with other wood procurement companies.

The study material consists of 10 Scots pine stands (*Pinus sylvestris*) located in North Häme and South Pohjanmaa, in southern Finland (see Table 1). Due to the wood selling strike during the summer in 1993 in Finland the choice of study stands was restricted. Because more than half of the pine stands in that area are regenerated naturally it was desirable that the study stands represent all cutting methods used in final cutting areas. It was also wanted that size and quality variation between the study stands is as large as possible. The number of pines in the stand shows that half were pure pine and the other half mixed pine-spruce and pine-spruce-birch stands.

In each stand 30 pine sawlog stems (dbh>18 cm) and 5 small pine sawlog stems (dbh>15 cm) from stands 2, 5 and 7 were selected as sample trees for test sawing. The sample trees were chosen from 4 to 7 sample plots by choosing 5 to 8 trees nearest to the sampling point as sample trees. The sampling points were laid out in a square pattern. The sample trees were numbered and for each tree height (H), diameter at breast height (dbh) and the height of log section (H<sub>ls</sub>) were measured. The distances from stump height to the lowest

dead branch (dead branch height, H<sub>abr</sub>), to the lowest live branch (live branch height, H<sub>lbr</sub>) and to the live crown (crown height, H<sub>c</sub>) were also measured. The minimum diameter for the lowest dead branch was determined as 15 mm for stems exceeding 20 cm in dbh and 10 mm for stems being lower than 20 cm in dbh. The crown height and the live branch height had different values only when at least two dead branch whorls separated the lowest live branch and continuous live crown. Moreover, the distance from stump height to the first branch or defect that lowers the pine log grade from the first grade to the second grade (first grade height, H<sub>1-gr</sub>) was determined applying Heiskanen and Siimes (1959) log grading rules. The thicknesses of the bark were measured from heights of 0.2, 1.00, and 1.80 m above the stump height with a bark gauge. The age of the tree (A) and the growth rate in early stage of life was measured with age auger. The drillings were made at the stump height from the direction of the sampling point. The number of annual rings between 2 to 4 cm from the pith was selected as a suitable variable (I<sub>2-4</sub>) to describe early growth rate.

The stems that were selected for test sawing were felled manually by an experienced feller applying general instructions for bucking. After felling and bucking, the logs of the sample trees were marked carefully and hauled to the sawmill. The code number marked on the logs indicated the location of the log (i.e. stand, stem and location in stem). Numerous measurements of taper, sweep, diameter, knottiness, bark and age were made before sawing.

The log properties measured were diameters from the top and from the butt, diameters at 0.5 m, 1.5 m, 2.5 m, etc. from the butt of the log, length, maximum sweep, the distance of possible defective section cut away from the log, the thickness of the bark from the top and from the butt of the log. Occurrences of possible defects were also registered properly. Furthermore, small slices

Table 1. Study stands.

Stand No.	Location	Forest type	Cutting method	Mean age <sup>1)</sup> years	Area ha	Stand density stems/ha	No of pines %	Mean dbh cm	Mean H m	Mean H <sub>abr</sub> m
1	Kuru	MT	clear cut	100	0.8	450	77	26.5	20.8	5.3
2	Ikaalinen	MT/VT	clear cut	95	2.1	600	63	22.4	15.5	4.2
3	Kuru	MT/VT	rem.of seed trees	105	2.4	90	100	27.3	18.4	5.3
4	Ruovesi	MT	seeding felling	130	0.6	1130	36	28.1	22.0	4.3
5	Virrat	VT	seeding felling	120	0.9	680	96	23.1	19.7	6.6
6	Virrat	MT	clear cut	120	1.2	510	33	26.2	17.6	6.2
7	Kauhajoki	CT	seeding felling	115	1.6	330	100	21.3	13.9	3.3
8	Kauhajoki	VT	rem.of seed trees	115	1.3	140	100	26.1	16.1	3.1
9	Virrat	MT	rem.of seed trees	90	1.1	130	98	28.2	22.0	4.5
10	Virrat	MT/VT	seeding felling	100	1.2	620	69	24.3	17.9	4.5

<sup>1)</sup> Average age of sample trees. Annual rings were calculated from small slice cut from the butt. No other years were added.

(disks) were cut from the butts of every butt log to enable the recalculation of annual rings. Special drawings on these slices ensured that such comparable age calculations were done in the same direction as the calculations done with an age auger in the forest.

Test sawing was carried out in a small circular sawmill Hietaniemen Saha Ky, located in Viljakkala, a city near Parkano. During the sawing the locations of sawn goods were registered and marked with special code numbers. The sawing pattern is described in Appendix 1. After the sawing the pieces were visually edged and trimmed if necessary by the grader and their dimensions measured. Grading was done by an experienced grader applying the Aureskoski Oy grading rules. These follow the old Finnish export rules (Vientisahatavaran ... 1979) but have stricter knottiness rules. The strictness of the grading rules applied in the study agrees almost entirely with the new Scandinavian export rules (Pohjoismainen ... 1994).

#### 4.1.1.2 Applying Logistic Regression to Modeling Lumber Quality

Logistic regression analysis was applied to develop the quality grade prediction models. As distinct from my earlier studies (Uusitalo 1994a, 1994b) the quality grades were handled as polytomous dependent variable instead of a dichotomous dependent variable. It is naturally possi-

ble to replace a model with a polytomous dependent variable by several models using a dichotomous dependent variable. In many cases this is however not a good alternative because there is no guarantee that the probabilities of each class can be summed up.

Since the lumber quality varies greatly from the surface to the pith, especially in the lowest part of the stem, in modeling the quality of the logs the location of the lumber has to be taken into account. The battens sawn from the pith of the log were fitted independently. In addition, the outermost board from each of the four faces of the log were classified as an outer board and the boards beneath were classified as inner boards. In modeling phase these two board groups were also fitted independently which in turn reduced the number of response levels and improved the goodness of fit of the models.

Regarding the boards sawn from the logs, a single saw log was divided into five response classes (*j*) according to the number of the boards that falls into the certain grade *y*:

1. all boards are grade *y*
2. three of the boards are grade *y*
3. two of the boards are grade *y*
4. one of the boards is grade *y*
5. none of the boards is grade *y*.

In battens, that include only two sawn pieces, the number of response classes (*j*) is only three:

1. both of the battens are grade  $y$
2. one of the battens is grade  $y$
3. none of the battens is grade  $y$ .

Let the response category probabilities be  $\pi_j$ , where  $j = 1, \dots, m$  is number of the class  $j$ . Given the vector at explaining variables  $x = (x_1, \dots, x_k)$ ,

$$\gamma_j(x) = \pi_1(x) + \dots + \pi_j(x) \quad (2)$$

is the probability that a single board falls into the class  $j$ ,  $j = 1, \dots, m$ , given  $x$ .

As the dependent variable classes are naturally ordered, the effects of the factors were modeled using the proportional odds model where the cumulative probabilities  $\gamma_j(x)$  are transformed into the logistic scale and modeled using parallel linear regression as follows (McCullagh and Nelder 1989):

$$\log \frac{\gamma_j(x)}{1 - \gamma_j(x)} = \theta_j - \beta' x, \quad j = 1, 2, \dots, m \quad (3)$$

The parameters  $\theta_j$  are analogous to constants in ordinary regression models. They have no obvious interpretation in this study and are regarded as nuisance parameters. In the proportional odds model, there is only regression coefficient  $\beta$  for each factor, which is independent of the choice of the category ( $j$ ). According to structure of the model the different probabilities  $\pi_j$ ,  $j = 1, 2, \dots, m$  are:

$$\begin{aligned} \pi_1 &= \frac{e^{\theta_1 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\theta_1 + \beta_1 x_1 + \dots + \beta_k x_k}}, \quad x = (x_1, \dots, x_k) \\ \pi_2 &= \frac{e^{\theta_2 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\theta_2 + \beta_1 x_1 + \dots + \beta_k x_k}} - \frac{e^{\theta_1 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\theta_1 + \beta_1 x_1 + \dots + \beta_k x_k}} \\ \pi_j &= \frac{e^{\theta_j + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\theta_j + \beta_1 x_1 + \dots + \beta_k x_k}} - \frac{e^{\theta_{j-1} + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\theta_{j-1} + \beta_1 x_1 + \dots + \beta_k x_k}} \\ \pi_m &= 1 - \frac{e^{\theta_{m-1} + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\theta_{m-1} + \beta_1 x_1 + \dots + \beta_k x_k}} \end{aligned} \quad (4)$$

The logistic regression models were fitted by using SAS PROBIT procedure and SAS LOGISTIC procedure (SAS 1989). Note, that PROBIT and LOGISTIC have different link functions which provide slightly different parameter estimates. Basically they give however same result.

The subjects of interest in predicting the quality of pine stems as well as the main external stem

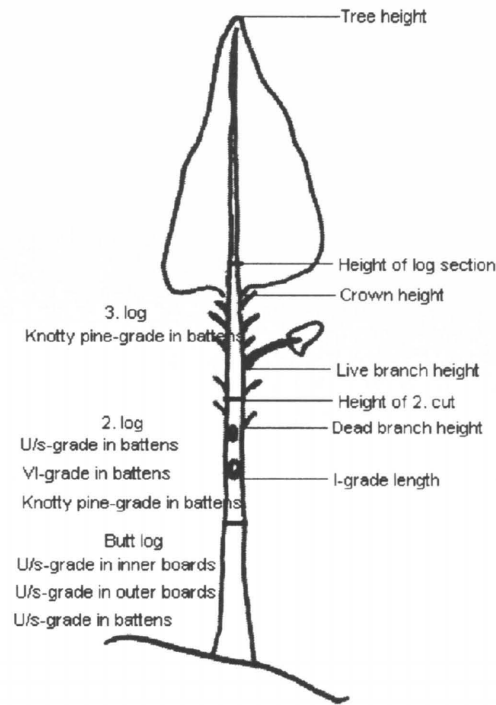


Fig. 1. The subjects of interests in predicting the quality of pine stems (lefthand side) and clarification of some external characteristics used in the study (righthand side).

characteristics used as independent variables are shown in Figure 1. The subjects of interest totalled seven: u/s quality in butt logs battens, u/s quality in the butt log inner boards and butt log outer boards, u/s quality, knotty pine and VI grade in second log battens and knotty pine in third log battens. The Aureskoski sawmill managers consider that these satisfy the needs of production planning.

The logistic regression analyses were done in two phases. First, selecting the most appropriate external stem variables was begun with a careful univariate analysis of each variable. Once the statistics of the univariate logistic models had been analyzed, the most promising combinations of the variables were modeled in the multivariate context. Assessing the models was based on the value of log likelihood and on the log likelihood ratio test, in which test the statistic deviance resulting from comparing the predicted values to values resulting from the saturated model is tested against

Table 2. Yield of sawn goods in percentage log volume including bark by the location of the log and top diameter. Log volume has been calculated by Huber<sup>1)</sup> equation using exact diameters measured every one metre.

Top diameter cm	1			2			3			4			All		
	$\bar{x}$	s	n	$\bar{x}$	s	n	$\bar{x}$	s	n	$\bar{x}$	s	n	$\bar{x}$	s	n
13	39.2	6.5	13	45.8	4.6	19	48.6	4.9	11	—	—	—	44.5	6.4	43
15	48.7	5.6	21	51.6	5.4	60	54.7	4.9	43	47.2	9.8	4	52.0	5.8	128
17	48.6	5.4	42	52.9	5.1	64	55.8	4.0	44	53.9	1.7	2	52.6	5.6	152
19	50.7	5.9	51	54.1	4.9	30	56.2	2.5	11	—	—	—	52.5	5.6	92
21	51.3	4.6	59	54.2	3.0	39	62.6	2.5	4	—	—	—	52.8	4.6	102
23	50.2	5.8	38	54.2	3.9	28	52.2	10.8	2	—	—	—	51.9	5.5	68
25	50.5	5.0	36	56.0	3.2	11	—	—	—	—	—	—	51.8	5.1	47
27	50.6	4.8	31	55.1	4.1	3	—	—	—	—	—	—	51.0	4.8	34
29	51.3	7.2	10	57.1	—	1	—	—	—	—	—	—	51.8	7.1	11
31	46.0	—	1	—	—	—	—	—	—	—	—	—	46.0	—	1
33	51.6	8.3	3	—	—	—	—	—	—	—	—	—	51.6	8.3	3
All	49.8	5.9	305	52.8	5.1	255	54.9	5.0	115	49.4	8.3	6	51.8	—	681

<sup>1)</sup> Huber equation is  $v = \sum_{i=1}^n g_i l_i$ , where  $g_i$  is the cross-sectional area measured in the middle of the piece of wood and  $l_i$  is the length of the piece.

$\chi^2$ -distribution. The greater the p-value is, the better is the fit of the model. The significance of the independent variables was tested by the Wald test. Since the lumber quality varies greatly from the surface to the pith, especially in the lowest part of the stem, in modeling the quality of the butt logs the location of the lumber has to be taken into account. The outermost board from each of the four faces of the log were classified as an outer board and the boards beneath were classified as inner boards.

#### 4.1.2 Results

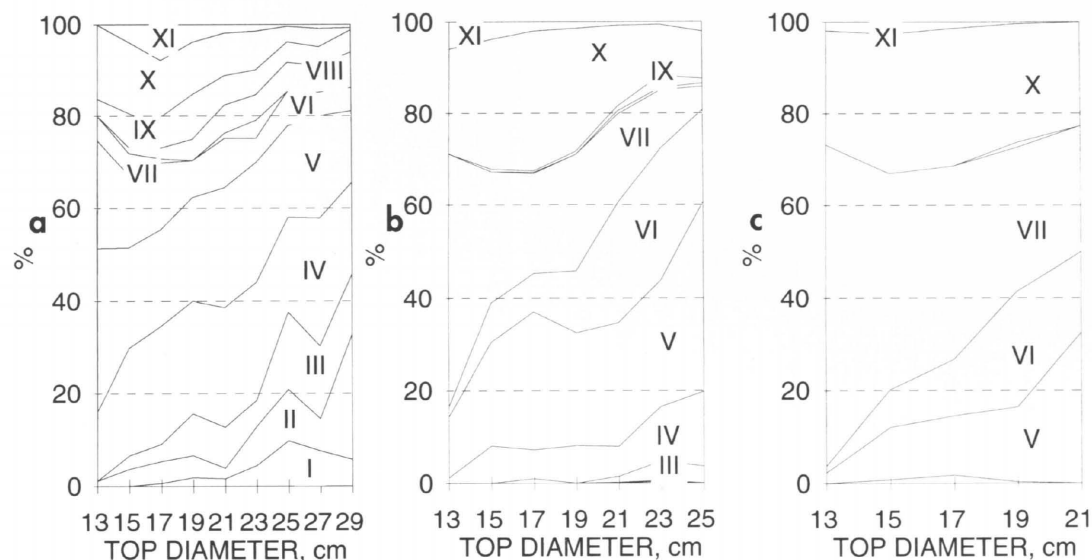
##### 4.1.2.1 Yield of Sawn Timber

The yields of sawn goods are in accordance with the earlier Finnish studies by Kärkkäinen (1980), Hakala (1992) and Uusitalo (1994a) even though the present yield is a little bit higher than the yield obtained from those studies (Table 2). This is the consequence of the exceptional precision of the sawyer used in this study. The yield increases a little with increasing log diameter. The location of the log has a clear effect on the yield. Due to the differences in taper the yield obtained from the upper logs is clearly higher than the yield from the butt logs.

The proportion of u/s grade increases with the increasing diameter of the log (Fig. 2). There are two basic reasons for this. Firstly, the grading rules are to some extent dimension-dependent in the sense that the grading is more strict with the smaller pieces. Secondly, the larger the pines are in diameter, the thicker the amount of slightly knotty or knottless zone in the butt is when other quality characters are assumed to remain stable. Knotty pine (VII) plays a significant role in the second and in the third logs. It also seemed to be very diameter-dependent since its proportion clearly decreases with increasing diameter.

##### 4.1.2.2 General Lumber Quality Models

A correlation analysis of the independent variables was carried out before modeling (Table 3). The correlation matrix gives a good overview of the relationships between the main predictors of quality. Perhaps most interesting part of the matrix is the correlations between early growth rate ( $I_{2-4}$ ) and the other quality parameters. There is no correlation between early growth rate ( $I_{2-4}$ ) and dead branch height ( $H_{dbr}$ ), even though both of them are generally considered good predictors of the quality of butt logs. From the modelling point



**Fig. 2.** Proportion of grades I...XI according to the dbh of stem in a) butt logs ( $n = 301$ ) b) second logs ( $n = 251$ ) and c) third logs ( $n = 113$ ). Grades I...VI are the traditional Nordic grading classes (Vientisahasatavaran... 1979) and grades VII...XI the traditional Finnish wane grades (*oksaton vajaasärmä, höylävajaasärmä, pl/vl* and *pl/kl*). Grade VII is the sound-knot grade (knotty pine) determined only from battens.

**Table 3.** Correlation coefficients between the most appropriate stem characteristics.  $N = 314$ .

	Dbh	H	H <sub>dbr</sub>	H <sub>c</sub>	H <sub>I-gr</sub>	A	I <sub>2-4</sub>
Dbh	1.000						
H	0.626	1.000					
H <sub>dbr</sub>	0.125	0.343	1.000				
H <sub>c</sub>	0.289	0.720	0.374	1.000			
H <sub>I-gr</sub>	0.085	0.313	0.374	0.267	1.000		
A	0.168	0.007	0.078	0.068	0.171	1.000	
I <sub>2-4</sub>	-0.393	-0.274	0.018	-0.213	0.190	0.482	1.000

of view this may be regarded only as a positive matter because there is no fear of the presence of multicollinearity if both of these variables are entered in the model. It is, however, clear that growth rate does not affect the absolute value of the dead branch height. Slow early growth rate implies in many cases a weak growth site with a consequent slow growth rate later in the life, a small tree size both in diameter and in height and further, a slow natural pruning rate.

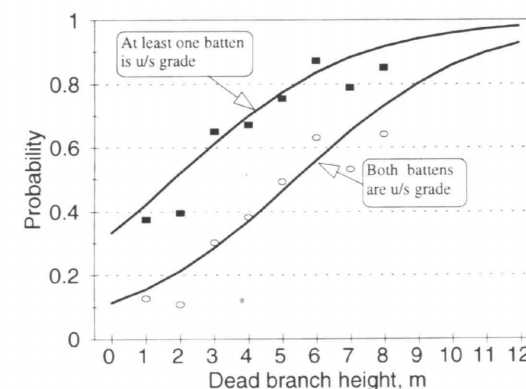
First grade height ( $H_{I-gr}$ ) and dead branch height ( $H_{dbr}$ ) are apparently the best single predictors of

the u/s grade in butt log battens (Table 4). Of these variables, the first grade height ( $H_{I-gr}$ ) has the greatest log likelihood while the dead branch height has the greatest p-value. Early growth rate ( $I_{2-4}$ ) as a single variable is a poor predictor of the quality of the butt logs.

The relationship between dead branch height and the probability of the u/s battens in butt logs is illustrated in Figure 3. The lines describe Model 2 from Table 4 and the square and circle marks are the corresponding probabilities calculated from the data. The upper line describes the probability

**Table 4.** Univariate logistic regression models for predicting the probability of u/s grade in butt log battens. Dependent variable: both (class 1), one (class 2) or none (class 3) of butt log battens are u/s-grade.  $N = 302$ .

Model	Variable	Const. 1	Const. 2	Coeff.	Std error	Wald test p-value	Log likelihood	Deviance p-value
1	$H_{I-gr}$	-1.94	-0.52	0.439	0.063	0.0001	-293.9	0.8980
2	$H_{dbr}$	-2.08	-0.70	0.386	0.059	0.0001	-299.1	0.9671
3	H	-2.95	-1.67	0.148	0.045	0.0011	-314.0	0.0131
4	$H_s$	-1.56	-0.29	0.122	0.028	0.0001	-314.9	0.2482
5	Dbh	-2.22	-0.98	0.078	0.023	0.0007	-318.9	0.7523
6	$I_{2-4}$	-0.94	0.29	0.048	0.017	0.0039	-320.4	0.0776
7	$H_c$	-1.13	0.04	0.086	0.034	0.0109	-321.5	0.7021
8	A	-1.37	-0.15	0.012	0.005	0.0205	-322.0	0.1018
9	$H_{lbr}$	-1.01	0.19	0.077	0.034	0.0225	-322.1	0.5979



**Fig. 3.** The probability of the battens in butt logs being u/s grade by the dead branch height. The lines describe Model 2 from Table 4. Square marks (level 1) and circular marks (level 2) describe the relative frequencies as calculated from the data. Because of the rather small number of observations occurring relative frequencies for dead branch height values 9...12 were left out of the figure.  $N = 292$ .

that at least one batten falls into the u/s grade. Surprisingly, this line is almost identical to the corresponding model formed in my earlier study (Uusitalo 1994a), although lumber qualities were graded by different grader and the sample trees were collected about 200 km away from each other. This implies that the grading rules as well as the quality of the pines do not necessarily differ significantly in southern Finland.

It was apparent that either dead branch height or first grade height should be excluded in multivariate models, because they are substitute variables. There were two primary reasons why dead branch height was preferred in the multivariate models for predicting the butt log quality. Firstly, dead branch height was judged to be the more objective variable because identifying the knobs on the surface of the tree is generally considered to be a difficult task. Secondly, preliminary calculations dealing with upper logs indicated that dead branch height is superior to first grade height in predicting the quality of upper logs and is naturally a consistent choice for butt logs as well. Consequently, a central task in multivariate model-building strategy is to choose the most appropriate variables already containing dead branch height for the model.

A carefully considered collection of the multivariate logistic models for predicting the probability of u/s grade in butt log battens appears in Table 5. In selecting the variables, a p-value of 0.02 was used as the significance level for rejection.

Entering the early growth rate (Model 1 from Table 5) as well as dbh (Model 2 from Table 5) in the model already containing dead branch height gives relatively little advantage. However, entering both the early growth rate and dbh (Model 4) enhances the fit of the model markedly. These tree variables seem to have a clear interaction in predicting the quality of butt logs. The best fit is attained by including dead branch height, early

**Table 5.** Multivariate logistic regression models for predicting the probability of u/s grade in butt log battens. Dependent variable: both (class 1), one (class 2) or none (class 3) of butt log battens are u/s grade. N = 302. Standard errors are given in parentheses.

Model	Const. 1	Const. 2	H <sub>dbr</sub>	I <sub>2-4</sub>	Dbh	H	H <sub>c</sub>	H <sub>ls</sub>	Log likelihood	Deviance p-value
1	-2.81 (0.396)	-1.39 (0.138)	0.386 (0.059)	0.050 (0.017)					-294.6	0.5971
2	-3.68 (0.664)	-2.27 (0.138)	0.375 (0.060)		0.066 (0.024)				-295.3	0.5165
3	-3.39 (0.619)	-1.98 (0.137)	0.345 (0.061)			0.083 (0.034)			-296.1	0.5347
4	-6.88 (0.993)	-5.37 (0.148)	0.371 (0.060)	0.100 (0.022)	0.137 (0.029)				-282.9	0.8156
5	-5.42 (0.832)	-3.95 (0.144)	0.323 (0.062)	0.077 (0.020)		0.139 (0.038)			-287.4	0.7542
6	-6.05 (0.961)	-4.55 (0.155)	0.352 (0.068)	0.083 (0.022)		0.251 (0.062)	-0.159 (0.060)		-283.4	0.8164
7	-6.09 (0.911)	-4.55 (0.157)	0.310 (0.069)	0.109 (0.024)		0.186 (0.055)	-0.206 (0.060)	0.144 (0.051)	-279.3	0.8676

**Table 6.** Univariate logistic regression models for predicting the probability of u/s grade in butt log inner boards. Dependent variable: three (class 1), two (class 2), one (class 3) or none (class 4) of butt log inner boards are u/s-grade. N = 302.

Model	Variable	Const. 1	Const. 2	Const. 3	Coeff.	Std error	Wald test	p	Log likelihood	Deviance p-value
Stem variables										
1	Dbh	-12.2	-10.0	-8.53	0.304	0.035	75.0	0.0001	-231.4	0.2966
2	H <sub>ls</sub>	-6.13	-4.15	-2.88	0.197	0.042	21.5	0.0001	-262.0	0.2155
3	H	-8.27	-6.27	-5.00	0.232	0.090	6.7	0.0098	-262.2	0.3583
4	H <sub>I-g</sub>	-5.18	-3.18	-1.90	0.309	0.054	32.3	0.0001	-264.0	0.8909
5	H <sub>dbr</sub>	-5.16	-3.19	-1.95	0.262	0.053	24.2	0.0001	-268.6	0.3648
6	H <sub>c</sub>	-4.91	-2.97	-1.79	0.107	0.037	8.3	0.0039	-277.0	0.4729
7	H <sub>lbr</sub>	-4.82	-2.88	-1.71	0.103	0.038	7.6	0.0059	-277.3	0.7453
8	I <sub>2-4</sub>	-3.69	-1.77	-0.62	-0.00391	0.021	0.035	0.8507	-281.1	0.1760
Set-up variables										
9	N <sub>ib</sub>	-7.48	-5.00	-3.19	2.19	0.584	14.1	0.0002	-196.1	0.0000

growth rate, tree height, crown height and height of the log section in the model. However, in order to avoid overfitting, Model 4 may be regarded as superior to these models.

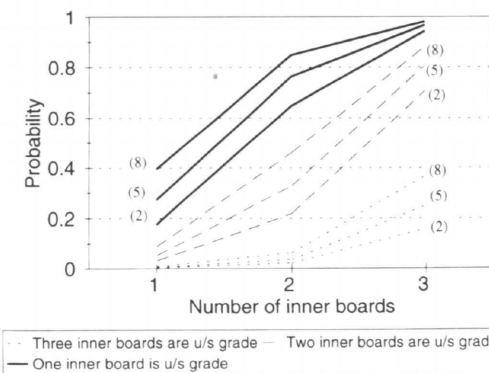
In the case of inner boards size of the tree (Dbh, H, H<sub>ls</sub>) seem to predict the probability of u/s grade quite well (Table 6). This may, however, fundamentally be considered a matter of course because inner boards are generally sawn only from large

logs. To avoid this self-evident truth, the number of inner boards sawn from the log (N<sub>ib</sub>) was chosen as an auxiliary variable to assist us in clarifying our understanding of this phenomenon.

As may be seen in Table 6, the number of inner boards is clearly the best predictor of u/s boards. We may even say that if the number of inner boards cannot be predicted, there is no sense in predicting the number of u/s boards. In any case multi-

**Table 7.** Multivariate logistic regression models for predicting the probability of u/s grade in butt log inner boards. Dependent variable: three (class 1), two (class 2), one (class 3) or none (class 4) of butt log inner boards are u/s-grade. N = 302. Standard errors are given in parentheses.

Model	Const. 1	Const. 2	Const. 3	N <sub>ib</sub>	H <sub>dbr</sub>	I <sub>2-4</sub>	Dbh	Log likelihood	Deviance p-value
Models excluding set-up var.									
1	-13.9 (1.24)	-11.6 (0.389)	-10.0 (0.422)		0.259 (0.057)		0.310 (0.037)	-220.8	1.0000
2	-17.3 (1.54)	-11.8 (0.418)	-13.1 (0.458)		0.260 (0.059)	0.100 (0.023)	0.376 (0.040)	-210.8	1.0000
Models including set-up var.									
3	-10.5 (0.921)	-7.63 (0.469)	-5.60 (0.518)	0.270 (0.254)		0.130 (0.024)		-180.2	1.0000
4	-8.48 (0.721)	-5.95 (0.416)	-4.06 (0.462)	2.15 (0.218)	0.186 (0.062)			-191.6	1.0000
5	-11.3 (1.00)	-8.40 (0.474)	-6.30 (0.527)	2.65 (0.257)	0.162 (0.064)	0.127 (0.024)		-177.0	1.0000
6	-11.1 (1.41)	-8.54 (0.419)	-6.63 (0.465)	1.82 (0.257)	0.197 (0.062)		0.110 (0.048)	-189.0	1.0000
7	-15.9 (1.78)	-12.8 (0.495)	-10.6 (0.555)	2.17 (0.283)	0.182 (0.064)	0.147 (0.026)	0.176 (0.050)	-170.9	1.0000



**Fig. 4.** The probability of butt log inner boards being u/s grade by the number of inner boards and dead branch height. The values of dead branch height are shown in parentheses (in meters). The figure describes Model 4 from Table 7.

variate models both including and excluding the number of inner boards were constructed (Table 7). Due to the uneven response profile of the models the p-values tend to fall at 1 which implies that there is some lack of large logs in the data. The evaluation of the models has thus to be

based only on log likelihood. The models that include the number of inner boards are clearly superior in predicting the number of u/s boards. In addition to the variables related to size of the tree (i.e. dbh, N<sub>ib</sub>), both dead branch height and early growth rate may be used separately or together as an auxiliary variable. The influence of the number of inner boards on the number of u/s boards is apparently greater than the influence of dead branch height (Fig. 4).

Owing to the large amount of wane in outer boards the best wane grades, *oksaton vajasärmä* (VIII, knotless wane grade) and *höylävajasärmä* (IX, plane quality wane grade), were included in the u/s grade. Likewise with inner boards, the size of the tree influences the probability of u/s grade in butt log outer boards most (Tables 8 and 9). In the case of outer boards the significance of the set-up variables is not so important. In fact, dbh has an ability to predict the probability of u/s grade in outer boards equal to the number of outer boards. The same interaction of dbh, dead branch height and early growth rate as appeared with battens and inner boards may also be discovered in the case of outer boards. If one of these variables is rejected, the fit of the model is clearly worse.

**Table 8.** Univariate logistic regression models for predicting the probability of u/s grade in butt log outer boards. Dependent variable: all (class 1), three (class 2), two (class 3), one (class 4) or none (class 5) of butt log outer boards are u/s grade. N = 302.

Model	Variable	Const. 1	Const. 2	Const. 3	Const. 4	Coeff.	Std error	Wald test	p	Log likelihood	Deviance p-value
Stem var.											
1	Dbh	-6.34	-5.37	-4.58	-3.45	0.195	0.024	64.4	0.0001	-447.5	0.4087
2	H <sub>1-gr</sub>	-2.86	-1.94	-1.19	-0.07	0.370	0.051	51.6	0.0001	-453.1	0.1740
3	H <sub>dbr</sub>	-3.12	-2.20	-1.45	-0.34	0.358	0.051	50.2	0.0001	-454.7	0.3300
4	H <sub>1s</sub>	-3.11	-2.23	-1.50	-0.43	0.163	0.028	34.6	0.0001	-463.8	0.3539
5	H	-4.84	-3.96	-3.23	-2.16	0.191	0.053	12.9	0.0003	-464.0	0.3090
6	H <sub>c</sub>	-2.36	-1.53	-0.86	0.15	0.105	0.032	10.8	0.0010	-477.0	0.6110
7	H <sub>lbr</sub>	-2.29	-1.46	-0.79	0.21	0.102	0.032	10.2	0.0014	-477.4	0.5752
8	I <sub>2-4</sub>	-1.72	-0.90	-0.26	0.73	0.033	0.016	4.47	0.0344	-480.2	0.1740
Set-up var.											
9	N <sub>ib</sub>	-2.41	-1.42	-0.62	0.50	0.965	0.170	32.3	0.0001	-447.0	0.0140
10	N <sub>ob</sub>	-5.16	-4.25	-3.48	-2.31	1.09	0.192	31.1	0.0001	-447.9	0.0012

**Table 9.** Multivariate logistic regression models for predicting the probability of u/s grade in butt log outer boards. Dependent variable: all (class 1), three (class 2), two (class 3), one (class 4) or none (class 5) of butt log outer boards are u/s grade. N = 302. Standard errors are given in parentheses.

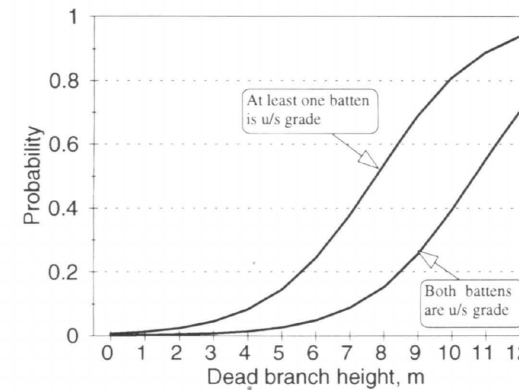
Model	Const. 1	Const. 2	Const. 3	Const. 4	N <sub>ob</sub>	N <sub>ib</sub>	H <sub>dbr</sub>	I <sub>2-4</sub>	Dbh	Log likelihood	Deviance p-value
Models excluding set-up var.											
1	-11.1 (1.01)	-10.0 (0.146)	-9.10 (0.179)	-7.88 (0.218)				0.137 (0.020)	0.304 (0.031)	-420.8	1.0000
2	-7.81 (0.725)	-6.74 (0.138)	-5.84 (0.176)	-4.60 (0.217)			0.335 (0.052)		0.184 (0.025)	-424.8	1.0000
3	-12.8 (1.10)	-11.6 (0.159)	-10.6 (0.198)	-9.23 (0.242)			0.338 (0.053)	0.137 (0.021)	0.297 (0.032)	-399.0	1.0000
Models including set-up var.											
4	-6.44 (0.597)	-5.49 (0.130)	-4.57 (0.166)	-3.29 (0.212)	0.967 (0.142)		0.312 (0.051)			-427.9	1.0000
5	-9.14 (0.822)	-8.05 (0.140)	-7.15 (0.177)	-5.80 (0.226)	1.32 (0.164)		0.305 (0.052)	0.0981 (0.018)		-412.1	1.0000
6	-5.96 (0.594)	-4.87 (0.141)	-3.94 (0.181)	-2.62 (0.226)	0.651 (0.152)	0.620 (0.133)	0.297 (0.052)			-416.8	1.0000
7	-9.50 (0.844)	-8.25 (0.162)	-7.22 (0.202)	-5.79 (0.252)	0.988 (0.168)	0.923 (0.146)	0.287 (0.052)	0.138 (0.021)		-390.8	1.000

As with the butt log battens, dead branch height and first grade height proved to be the best single predictors of the probability of u/s battens in the second logs (Table 10). In fact, they were the only single significant predictors. It was also found that the distance from stump height to the end of the second log which may also be called the second

cut height (H<sub>2.cut</sub>) has a certain influence on the probability of u/s grade. Although it was not significant as a single variable it was found to be a good auxiliary variable in the model already containing dead branch height. The negative value of the coefficient shows that the probability of u/s grade increases with decreasing second cut height.

**Table 10.** Logistic regression models for predicting the probability of u/s grade in second log battens. Dependent variable: both (class 1), one (class 2), none (class 3) of second log battens are u/s grade. N = 211. Standard errors are given in parentheses.

Model	Const. 1	Const. 2	H <sub>dbr</sub>	H <sub>1-gr</sub>	H <sub>2.cut</sub>	Log likelihood	Deviance p-value
1	-6.83 (0.742)	-4.97 (0.316)	0.640 (0.089)			-106.7	0.7363
2	-5.61 (0.602)	-3.81 (0.309)		0.555 (0.081)		-111.8	0.7340
3	-1.37 (2.08)	0.60 (0.334)	0.724 (0.099)		-0.617 (0.189)	-100.6	1.0000



**Fig. 5.** The probability of second log battens being u/s grade by dead branch height. The figure describes Model 1 from Table 10.

As may be noticed, the probability of u/s battens is very small with low dead branch height values (Fig. 5). It is quite natural that noticeable probabilities of u/s grade in second log battens may be obtained only from those stems where dead branch height exceeds the second cut height.

The same variables as proved significant in predicting the probability of u/s grade in second logs were also significant in predicting the probability of VI grade and knotty pine in second log battens (Table 11 and Table 12). The probability of VI grade increases with decreasing dead branch height (Model 1, Table 12). The fit of that model may be improved a little if the second cut height is entered in the model (Model 2, Table 12). The negative value of the coefficient tells us that the lower the second cut is made the greater is the probability of VI grade. Low second cut height

increases the probability that the second log will be cut from the dead-branch section of the stem which leads to greater probability of VI grade. Besides dead branch height and second cut height, crown height seemed to be significant in predicting the probability of knotty pine. Second cut height and crown height entered separately in the model already containing dead branch height (Models 2 and 3, Table 12) do not improve the fit of the model noticeable. Nevertheless, all these tree variables together have a strong interaction in predicting knotty pine in second log battens.

Crown height as a predictor of sound knot section is quite insufficient without a knowledge of the location of the second log. The higher the second cut is made and the lower the crown height is, the higher is the probability that the log will be cut from the sound knot section, and vice versa. The interaction of these two variables may also be found in the models created for predicting the probability of knotty pine in third log battens (Model 3, Table 13). Low dead branch height (Model 1, Table 13) and low crown height (Model 2, Table 13) as a single predictor also seems to indicate a high probability of knotty pine.

#### 4.1.2.3 Between-Stand Variation

A crucial point in assessing the validity of the proposed quality models is the significance of the different variance components. If the between-stand variation tends to be large, the general models will provide biased predictions in a stand which deviates from the average stand of the data used in this study. The between-stand variation was

**Table 11.** Logistic regression models for predicting the probability of knotty pine grade in second log battens. Dependent variable: both (class 1), one (class 2) or none (class 3) of the second log battens are knotty pine grade. N = 211. The models are available for logs greater than 139 mm in top diameter. Standard errors are given in parentheses.

Model	Const. 1	Const. 2	H <sub>dbr</sub>	H <sub>c</sub>	H <sub>2, cut</sub>	Log likelihood	Deviance p-value
1	0.0397 (0.373)	1.88 (0.210)	-0.416 (0.074)			-185.4	0.7188
2	1.51 (0.619)	3.50 (0.219)	-0.396 (0.076)	-0.147 (0.049)		-180.7	0.9652
3	-3.00 (1.16)	-1.02 (0.217)	-0.432 (0.075)		0.330 (0.120)	-181.4	0.9691
4	-2.53 (1.18)	-0.41 (0.235)	-0.420 (0.079)	-0.218 (0.053)	0.523 (0.133)	-172.4	0.9949

**Table 12.** Logistic regression models for predicting the probability of VI-grade in second log battens. Dependent variable: Both (class 1), one (class 2) or none (class 3) of the second log battens are VI-grade. N = 211. The models are available for the logs that are greater than 139 mm in top diameter. Standard error of the estimates are given in parentheses.

Model	Const. 1	Const. 2	H <sub>dbr</sub>	H <sub>2, cut</sub>	Log likelihood	Deviance p-value
1	-1.88 (0.442)	-0.22 (0.262)	-0.178 (0.075)		-145.4	0.6472
2	1.21 (1.38)	2.91 (0.266)	-0.167 (0.076)	-0.347 (0.149)	-142.4	1.0000

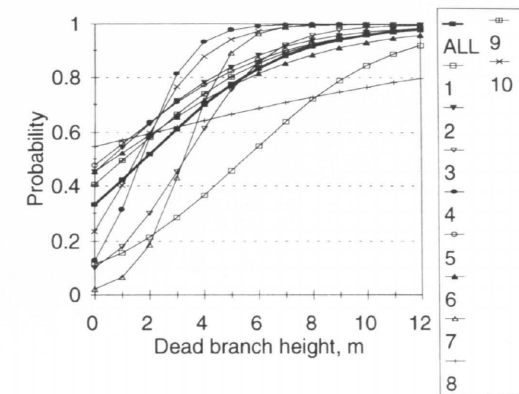
**Table 13.** Logistic regression models for predicting the probability of knotty pine grade in third log battens. Dependent variable: both (class 1), one (class 2) or none (class 3) of the third log battens are knotty pine grade. N = 81. The models are available for logs greater than 139 mm in top diameter. Standard errors are given in parentheses.

Model	Const. 1	Const. 2	H <sub>dbr</sub>	H <sub>c</sub>	H <sub>2, cut</sub>	Log likelihood	Deviance p-value
1	0.986 (0.518)	2.38 (0.256)	-0.197 (0.082)			-82.5	0.5424
2	2.12 (1.00)	3.50 (0.255)		-0.181 (0.078)		-82.7	0.4991
3	-1.94 (1.79)	-0.45 (0.276)		-0.282 (0.092)	0.592 (0.227)	-78.9	0.4911

considered here only in respect of u/s grade in butt log battens. Instead of presenting the results of residual analysis, the between-stand variation is illustrated by depicting the first level of the stand-specific quality models created for u/s

grades of butt logs using dead branch height as an independent variable (Fig. 6).

As may be noticed, four of the ten stands deviate little from the general model (ALL). In addition, in five other stands the differences near the



**Fig. 6.** The probability of at least one of the butt log battens is u/s grade by dead branch height and stand number (1–10). N = 302.

average of the dead branch height in the whole population (about 4.7 m) are not particularly big. Since the differences in the shapes of the curves of these stands compared to that of the general model are largely due to the small number of observations, the deviation may not be regarded as great. The deviation of the curve created for stand 1 may however be regarded significant, since the level of that curve is clearly beneath the population average. However, it is worthy of consideration that stand 1 had grown fastest, and the probability level would be nearer the population average if early growth rate ( $I_{2-4}$ ) had been included in the models. Altogether, the appropriateness of the general models in all kinds of final cuttings may be considered fairly good. It has to be emphasized that the stand specific curves depicted in Figure 6 describe only the variation of true population is considerably smaller since the variation of stand estimates includes also estimation error. However, the final acceptance of the validity of these models may be judged only by independent data.

#### 4.1.3 Appropriate Stem Characteristics for Practical Purposes

From the stem characteristics dbh, dead branch height, early growth rate and crown height may be regarded as the most important quality predic-

tors in Scots pine. Moreover, since tree height is crucial to assessing the shape of trees, it needs to be added to this list although it seems to predict quality poorly. Despite their ability in prediction of quality, all of these characteristics may not be considered appropriate for practical purposes. From the measurement point of view dbh and dead branch height may be done quite simply in the forest. Due to their importance dbh and dead branch height may be called primary variables of the method. The distributions of these variables are formed on the grounds of single observations within the stands.

As a result of its laboriousness the height of each tree is traditionally derived from dbh measures once the dbhs have been related to the heights measured from sample trees. This kind of approach may also be applied in the case of crown height since in most cases it needs to be measured with a hypsometer, like tree height. The usefulness of crown height as a predictor is largely dependent on whether the estimate of crown height can be derived from a crown height model or not as in the case of tree height. Owing to the principle that height and crown height are derived from dbh measures or dbh/height measures they may be called secondary variables of the method.

Unfortunately, determination of the early growth rate is even more difficult to carry out than measurement of height and crown height, since it has to be done with an age auger. Drilling the tree is an inconvenient task in many ways. Firstly, if drilling is done during summer, many weeks before felling, it may impair the lumber quality. In any case there is always a small risk that the drill hole does not fit in that 0...10 cm area of butt supposed to be cut away in trimming, which may lead to trimming losses. Secondly, drilling and calculating annual rings takes quite a long time – from experiences obtained in this study, about 2...5 minutes per tree. Although more advanced augers for determination of growth rate are available (see Rinn 1994) they are now too expensive and too heavy to be used daily by forest managers. Consequently, until a more suitable tool for determining the growth rate becomes available we are forced to leave variable early growth rate out of our method.

## 4.2 Predicting Diameter and Dead Branch Height Distribution

### 4.2.1 Material and Methods

#### 4.2.1.1 Study Material

The data used in the analysis dealing with pine diameter distributions and dead branch height distributions were collected from the same stands used in quality analysis (Table 1, p. 17). Once the measurements of the sample trees selected for test sawing were carried out, every tree in the stand was measured for dbh and height. Every pine saw log stem was also measured for dead branch height, I-grade height and crown height. Ten stands supplied the entire data. The size of the study stands varied from 0.6 to 2.4 hectares and the number of trees measured varied from 143 to 1262.

#### 4.2.1.2 Methods for Predicting Distributions

Diameter distribution of pines and spruces were constructed in five different ways. The methods were:

1. Sample with no modifications.
2. Application of the Weibull function where the parameters were estimated by regression models developed for Finnish conditions both for pine (Mykkänen 1986) and for spruce (Kilikki et al. 1989).
3. Application of the kernel function with a 1 cm smoothing parameter.
4. Application of the kernel function with a 2 cm smoothing parameter.
5. Application of the kernel function with a 3 cm smoothing parameter.

Although the last four methods provide continuous functions, the relative proportion of trees in each diameter class (2 cm) were formulated as one value for each diameter class. The range of the diameter class distribution was limited from 9 cm to 47 cm which means that the error terms were calculated from 20 diameter classes. The left endpoint of the Weibull function is determined by the location parameter and was restricted to being greater or equal to 7 cm. In the use of kernel functions the smoothing effect was restricted in diameter classes 9, 11, 13 and 15. If the smooth-

ing gave a weighting for diameter lower than 9 cm it was rejected as well as the same weight for the diameter as far from the right size of the observation. The right side of the distribution was left without truncation, since the highest diameter class clearly exceeds the largest diameter in these stands. In applying the Weibull function the proportion of trees in each diameter class was calculated by subtracting the cumulative value of the lower boundary of diameter class from the cumulative value of the upper boundary. In Kernel-function applications the value of diameter class center was used as the value of the diameter class.

The cumulative distribution function of the Weibull is

$$F(x) = 1 - \exp(-((x - a) / b)^c), \quad (5)$$

when  $a \leq x < \infty$ ; otherwise 0

where parameter  $a$  determines the location, parameter  $b$  the scale and parameter  $c$  the shape. In calculations the minimum diameters of both tree species (parameter  $a$ ) were adjusted to be one diameter class smaller than the minimum observed diameter class. The estimate of parameter  $c$  for diameter distribution of pines was derived from regression model:

$$c = \exp(0.622877 - 0.000405 Alt - 0.004854 G - 0.053183 a + 0.052858 d_{gM}) \quad (6)$$

where  $Alt$  is altitude above the sea level (m),  $G$  is stand basal area ( $m^2/ha$ ),  $d_{gM}$  is basal area median of the sample trees (cm) and  $a$  is the location parameter (Mykkänen 1986). Parameter  $b$  was then solved from the formula

$$b = (d_{gM} - a) / (-\ln(0.5))^{1/c} \quad (7)$$

The estimate of parameter  $b$  for diameter distribution of spruces was derived from regression model:

$$b = 0.629537 + 1.050618 d_{gM} - 1.020776 a + 0.014405 G - 0.001986 T \quad (8)$$

where  $T$  is age in years and other symbols as mentioned above (Kilikki et al. 1989). In this case parameter  $c$  may be derived from formula

$$c = (\ln(-\ln(0.5))) / \ln((d_{gM} - a) / b) \quad (9)$$

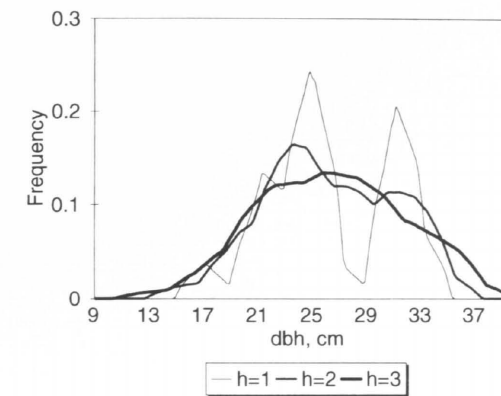


Fig. 7. The effect of smoothing on the shape of the diameter distribution by the different values of smoothing parameters ( $h$ ) incorporated with the kernel function.

Basal area was estimated by presuming that trees were located systematically in the stand. The theoretical size of the plot was derived from the sample size (i.e. each tree equals a certain size of the area) and the basal area of the stand was calculated by dividing the basal area of the sample trees by the theoretical size of the plot. The altitude of 80 m above the sea level where applied as constant.

The kernel estimator with kernel  $K$  is defined as:

$$f(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (10)$$

where  $(X_i; i=1, \dots, n)$  is a random variable,  $x$  is a sample from an unknown density function  $f(x)$ ,  $h$  is a smoothing parameter and  $K$  is a kernel function (Silverman 1986). The widely used Epanechnikov (1969) function was selected as an adequate kernel function. This function is defined as:

$$K(t) = \frac{3}{4} \left(1 - \frac{1}{5} t^2\right) / \sqrt{5} \quad (11)$$

for  $|t| < \sqrt{5}$ , 0 otherwise

The effect of the smoothing parameter is illustrated in figure 7 by a sample including 20 trees. It shows in detail how the value of the smoothing parameter reflects the shape of the distribution. Figure 8 presents the differences among the five

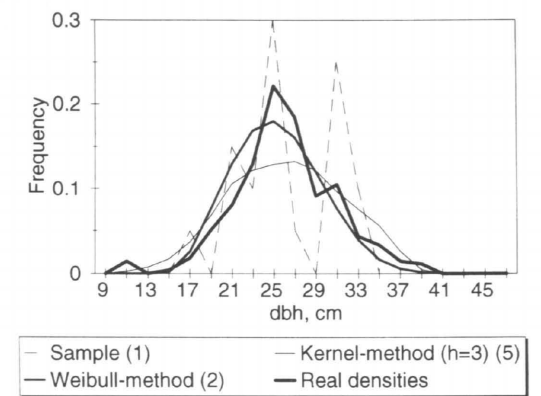


Fig. 8. Different principles in constructing the diameter distribution of the trees used in this study.

methods compared to the real population after the real proportions of diameter classes have been constructed by distribution functions.

The dead branch height distributions of pines were constructed following the same principles applied in diameter distributions. Apart from the methods applying diameter distributions, the Weibull function (method 2) was replaced with normal distribution in the case of dead branch height. The parameters of normal distribution (i.e. mean and standard deviation) were calculated in each replication from a sample. In addition, values of 0.5 m, 1 m and 1.5 m were chosen as smoothing parameters for the kernel function. Thus the complete list of alternative methods applied to construction of dead branch height distribution were as follows:

1. Sample with no modifications.
2. The standard normal distribution where the parameters mean and standard deviation are calculated from a sample.
3. Application of the kernel function with a 0.5 m smoothing parameter.
4. Application of the kernel function with a 1 m smoothing parameter.
5. Application of the kernel function with a 1.5 m smoothing parameter.



#### 4.2.1.3 Random Repetition Technique

The ability of the methods chosen to construct the diameter distributions of trees with different sample sizes were studied by repeating a random selection of the tree species in each stand a hundred times. In order to facilitate the comparison between the constructed distribution and real densities, the distributions were converted to relative distributions. From each repetition the error terms were calculated independently after which the means of the chosen error term were calculated.

A traditional approach in the process of validating a diameter distribution model is to use a goodness-of-fit test. In predicting the real diameter distributions the information resulting from these tests is however, insufficient, since the size of the error is very hard to determine. Among the numerous candidates, the statistics sum of square errors (SSE) and Kolmogorov-Smirnov (KS) test value were selected as the most appropriate error terms to describe the error between the real and the estimated diameter distributions. The KS test value is determined as

$$KS = \max|S(x) - F(x)| \quad (12)$$

where  $x$  is a random variable,  $F(x)$  is the real cumulative distribution function and  $S(x)$  is the cumulative distribution function derived from the sample.

The KS measure detects skewness errors better while the SSE can describe the ability of the model to predict jagged distributions better. The SSE recalls the error index suggested by Reynolds et al. (1988) because the only difference between these statistics is that the error index summarizes errors and the SSE summarizes square errors in diameter densities.

#### 4.2.2 Results

As expected, strongly smoothed distributions like the Weibull distribution (Method 2) and the kernel function with smoothing values 2 (Method 4) and 3 (Method 5) are superior with small sample sizes compared to the less smoothed (Method 3) or unsmoothed (Method 1) distributions (Fig. 9).

The methods based on the application of the kernel function with smoothing values 2 and 3 prove to be most precise almost regardless of the sampling intensity. The mean SSE of these distributions seems to decrease little after 20...25 trees. The results are parallel in all stands. Neither the area of the stand, density nor tree species proportions seem to affect the error level of the distributions. Moreover, the error term used seems to have no influence on the relationships between the methods chosen or on the shape of the distributions. Calculations based on KS test value gave similar results to the mean SSE.

The results from all ten study stands favour the use of the kernel function in construction of diameter distribution. Among the smoothing values, 3 cm seems to suit best when the sample size is less than 20, while 2 cm and 3 cm are equally good when the sample size varies between 20–30 and 2 cm is most suitable when sample size exceeds 30. Naturally, smaller smoothing values are more precise when the sample size exceeds the upper value of the scale.

The results of the diameter distribution of spruce in mixed pine-spruce stands are illustrated in Figure 10 with the examples drawn from stands 1 and 4. Pine is clearly the dominant tree species in stand 1 while spruce is clearly dominant in stand 4. Figure 10b is almost equal to the earlier figures drawn from the error levels of pine distributions (Fig. 9). When spruce becomes dominant in the forest its diameter distribution seems to form into regular unimodal shapes. When spruce remains clearly in the minority however, the diameter distribution often forms bimodal or multimodal shape as in our example stand 1 (Fig. 10a). This explains the weakness of methods 2, 4 and 5 with greater sample sizes. However, the use of methods 1 and 3 tend to be inappropriate in construction of the diameter distribution for spruce. In order to increase the number of spruce sample trees in pine-dominated stands we have to double or even triple the total size of the sample. Hence methods 4 and 5 seems to suit small sample sizes best in predicting the diameter distribution of spruce as well. Just as in the case of pine, the use of the KS test value as a measure did not affect the relative superiority of the methods.

The differences between the methods in constructing the dead branch height distribution of

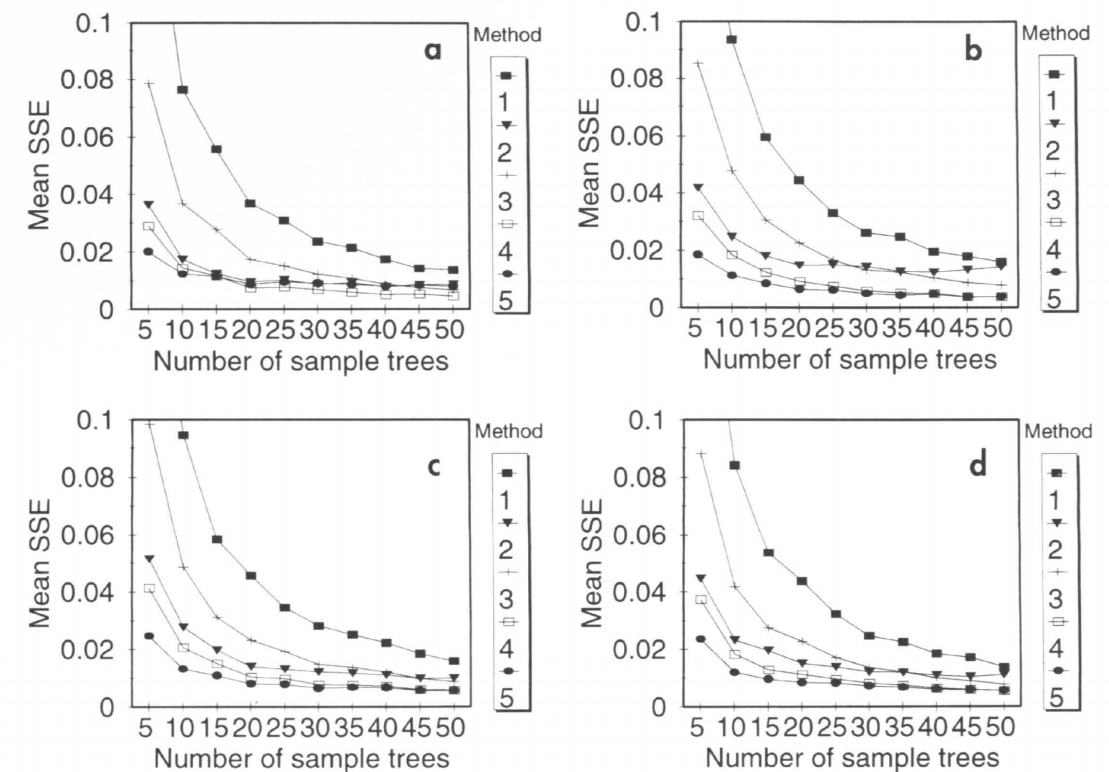


Fig. 9. The mean SSE on 100 random replications for the proportion of pines per diameter class in stands 3 (a), 5 (b), 2 (c) and 6 (d).

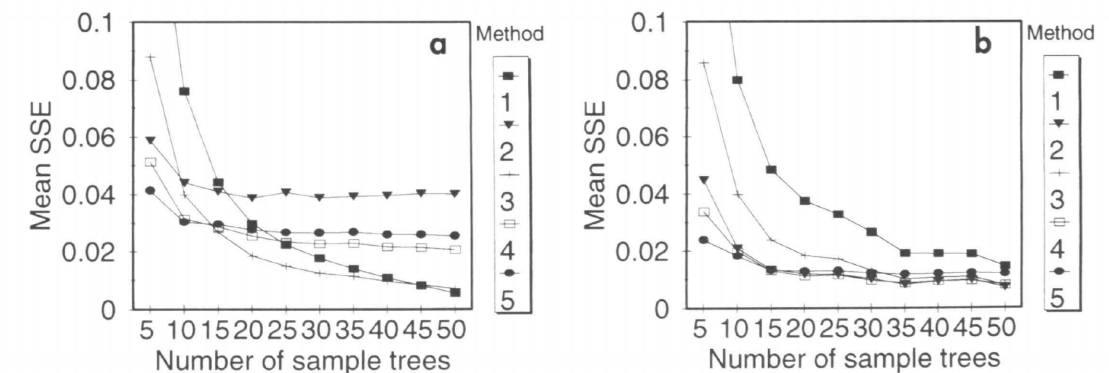


Fig. 10. The mean SSE on 100 random replications for the proportions of spruces per diameter class in stands 1 (a) and 4 (b).

pinus are illustrated by two example stands (Fig. 11). Parallel to the results regarding diameter distribution, the most strongly smoothed kernel distributions (methods 4 and 5) proved to be most

precise at least with small sample sizes. The smoothing value of 3 m was best with sample size 5 but as a whole 2 m seems to suit better because in six out of ten stands with sample size 10 and in

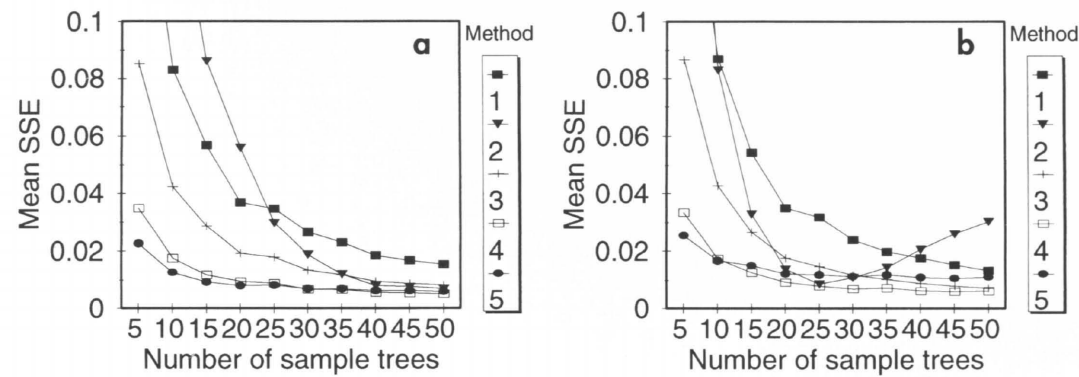


Fig. 11. The mean SSE on 100 replications for the proportion of pines per dead branch height class in stands 5 (a) and 6 (b).

seven out of ten with sample size 15, method 4 was more precise than method 5. Normal distribution did not fit well for prediction of dead branch height distribution. As may be noticed from Figure 11b the precision of the normal distribution (method 2) decreases markedly first but begins to increase after 25 sample trees. This phenomenon appeared in almost every stand. However, the range of distribution affects the precision of the method greatly because, since dead branch height values were measured only with an accuracy of one metre it yields only a few values in some stands. The broader the range of distribution was the better the fit of normal distribution was. For example, in stand 5 (Fig. 11a) the range of distribution was very broad (2...12 m) with the consequence that precision of the normal distribution improves logically with increasing sample size. The appropriateness of normal distribution might have improved if dead branch heights were measured more accurately.

### 4.3 Sampling Design Studies

#### 4.3.1 Material and Methods

##### 4.3.1.1 Study Material

Study stands 1, 2, 3, 5, 6 and 7 used with the analyses of quality prediction and diameter and dead branch height distributions were used to analyze sampling designs (Table 1, p. 17). In the field work, besides the measures explained in section 4.2.1.1, the locations of every tree were fixed in these six stands with a special surveying instrument, the

tacometer. The lens was placed at a certain height on the tree (e.g. breast height) and tacometer stored the coordinates (X, Y and Z) of the lens in the memory. If there were obstacles (e.g. other trees or branches) between the tacometer and the tree, the lens was moved to the nearest point from the tree where it could be seen and the location of the tree was registered there. The distance and the point of the compass of this point from the tree were also registered so that the lens coordinates could be moved on the tree afterwards. This problem was encountered in about 10 % of all trees. Later, when the coordinates were combined into the tree characteristic data, surface coordinates were transformed to polar coordinates by the location of the tacometer, the location of the lens and dbh.

##### 4.3.1.2 Sampling Simulations

A computer program that simulates sampling was developed for the present study to investigate the effectiveness of different sampling techniques as well as the effect of changing the plot size. The simulator was planned to simulate two kinds of cruising techniques, circular plot sampling and combined nearest trees/BA sampling (Table 14). The sampling points were laid out systematically in a square grid pattern. Circular plot sampling was simulated by selecting the trees within a certain radius of the sample points. Both the diameter distribution of each tree species as well as the basal area of the stand were calculated from the

Table 14. Sampling designs used for systematic sampling in with the number of sample plots and and plot size. In relascope sampling only BAF = 1 m<sup>2</sup>/ha was used. Number of sample plots  $N_s$  = Number of sample trees (n) / (Plot size (ha) × Stand density (n/ha)).

Design no	Removal of seed trees stands		Seeding felling pine stands		Mixed pine-spruce (-birch) stands	
	No of plots (relascope plots)	Plot size or No of nearest trees	No of plots (relascope plots)	Plot size or No of nearest trees	No of plots (relascope plots)	Plot size or No of nearest trees
<b>Level 1</b>						
1.1	10 (3)	1 tree	5 (3)	3 trees	7 (4)	3 trees
1.2	4 (3)	3 trees	3 (3)	5 trees	4 (4)	5 trees
1.3	$N_s$	314 m <sup>2</sup>	$N_s$	50 m <sup>2</sup>	$N_s$	100 m <sup>2</sup>
1.4	$N_s$	706.5 m <sup>2</sup>	$N_s$	100 m <sup>2</sup>		
<b>Level 2</b>						
2.1	20 (5)	1 tree	7 (4)	3 trees	10 (5)	3 trees
2.2	7 (4)	3 trees	4 (4)	5 trees	6 (6)	5 trees
2.3	$N_s$	314 m <sup>2</sup>	$N_s$	50 m <sup>2</sup>	$N_s$	100 m <sup>2</sup>
2.4	$N_s$	706.5 m <sup>2</sup>	$N_s$	100 m <sup>2</sup>		
<b>Level 3</b>						
3.1	30 (6)	1 tree	10 (5)	3 trees	15 (8)	3 trees
3.2	10 (5)	3 trees	6 (6)	5 trees	9 (9)	5 trees
3.3	$N_s$	314 m <sup>2</sup>	$N_s$	50 m <sup>2</sup>	$N_s$	100 m <sup>2</sup>
3.4	$N_s$	706.5 m <sup>2</sup>	$N_s$	100 m <sup>2</sup>		
<b>Level 4</b>						
4.1					20 (10)	3 trees
4.2					12 (12)	5 trees
4.3					$N_s$	100 m <sup>2</sup>

circular plots. In nearest trees/BA sampling a certain number of the nearest trees (e.g. nearest one, nearest three, nearest five) from the sample point were selected as sample trees. Beside the nearest trees' selection, the basal area of the stand was determined by the relascope method alternatively in all, in every third or in every fourth sample point. Correspondingly, the diameter distributions of the stand were calculated from the nearest tree's selections and the basal area of the stand was calculated from the relascope measures. Finally, the diameter distributions of each tree species were proportioned to the real scale according to the mean of the relascope measures and proportions of the basal area of the tree species.

The boundaries of each stand were defined with straight lines  $ax + by + c = 0$ , where  $x$  and  $y$  are corresponding values of coordinates  $x$  and  $y$ ,  $c$  is a constant and  $a$  and  $b$  are the slope parameters. The boundary lines formed a basis for determination of boundary points, which were chosen at intervals of five metres along the boundary lines. X and Y coordinates as well as slope of the line

from which the point was selected were listed in separate files. In each replication the starting point for sampling was chosen randomly among the boundary points. The first sample point was laid out a half line distance apart from the starting point right-angled to the boundary line and the rest were fixed in square grid pattern through the stand. Before accepting the final locations of each sample point, the program tested whether the sample point was inside the stand or not. In the first stage, only complete sample plots were acceptable. A sample point was acceptable if the radius of the sample plot did not exceed the distance from sample point to the boundary line. In nearest trees/BA sampling the radius of the sample plots were determined by relating the proportion of the sampled area to the relation between the number of sample trees and stand density. In the second stage, the group of incomplete sample plots were fixed in descending order and the size of the area inside the stand was calculated. If the size of the area summarized from the incomplete sample plots exceeded the area of one complete sample plot, the largest sample plot was moved towards stand

until it was as far from the boundary line as its radius. Similarly, as many sample plots were moved as often as the size of the summarized area exceeded a complete sample plot. Both methods introduce some bias because their edge is not represented with the same probability as the rest of the stand.

The choice of the sampling designs was based on experience obtained from diameter distribution analysis. The leading principle in selection of the sampling designs was its appropriateness in practical use. The number of basal area measures in nearest trees/BA sampling was increased with increasing sampling intensity so that the relation between basal area measures and number of sampled trees was about 1/5 regardless of the sampling intensity. A hundred replications of each sampling design were done in each stand. Apart from the earlier analysis the comparisons of real distributions and estimated distributions were based on real frequencies of trees per hectare. Consequently, the KS measure was the obvious choice for validating the differences between the methods, since it is important to detect the skewness errors of the distributions.

The development of a time consumption function related to sampling field work requires that one specifies the mode of carrying out the work. An extensive time study of the measurement work was carried out during the summer 1993 in the first seven stands in this study (Table 1). The work of two workers using two different methods was videotaped and the duration of each work element was later measured in studio conditions. The detailed results of the time study have already been published by Uusitalo and Kivinen (1994). Those results, which form the basis for the time calculations used in this study, are shown in Appendix 2. The measurement work was divided into seven work elements: moving, establishment of sample plots, measurement of dbh with a calliper, measurement of quality parameters (chiefly dead branch height), measurement of tree height, measurement of basal area with relascope (without checking of borderline trees) and recording of results.

The measurement work was done by two alternative methods: circular plot sampling and three nearest trees/BA sampling. In circular plot sampling the sample points were selected by moving along the square grid by compass and thread meas-

ure. Quality parameters were measured by telescope pole and tree heights were measured by Suunto-hypsometer. In three nearest trees/BA sampling points were defined simply by walking along a subjectively chosen route and defining the sample point, e.g., pacing out 20 meters. Quality parameters as well as tree heights were determined by ocular estimation. Naturally, the work technique for establishing the sample plot also deviated greatly since in circular plot sampling the selection of sampling trees was done by circulating the telescope pole around and in three nearest trees/BA sampling simply by ocularly identifying the three nearest trees.

The moving time comprises two elements: walking speed ( $S$ ) and walking distance ( $L$ ). In the time study, the average number of sample plots was six when the sampling points were fixed in the square grid and ten when the walking route was chosen subjectively. Keeping these values as standards, walking distances for the cases where the number of sampling plots differs from the average values were corrected by two equations. Correction  $L = \sqrt{nA} + 2\sqrt{A}$  ( $L$  = walking distance in hundreds of metres,  $n$  = number of sampling plots and  $A$  = size of area in hectares) suggested by Nyssönen et al. (1971) was applied to systematic sampling using a square grid pattern and correction  $L = \sqrt{nA} + \sqrt{A}$ , suggested by Mesavage and Grosenbaugh (1956), to a subjectively chosen route. Since the time study was carried out in excellent conditions during the summer, it is no wonder that the time consumption values are considerably smaller than those in many other studies (Johnson and Hixon 1952, Mesavage and Grosenbaugh 1956, O'Regan and Arvanitis 1966, Nyssönen et al. 1971). Furthermore, cruising may be performed at a brisk pace since the work is not supposed to take longer than one hour.

#### 4.3.2 Results

##### 4.3.2.1 Precision of the Sampling Designs

The superiority of the methods in predicting the diameter distribution continues to be the same as in section 4.2.2 even though we are now dealing with real frequencies of trees per hectare. Therefore, in constructing the pine diameter distribution the kernel function with the 3 cm smoothing

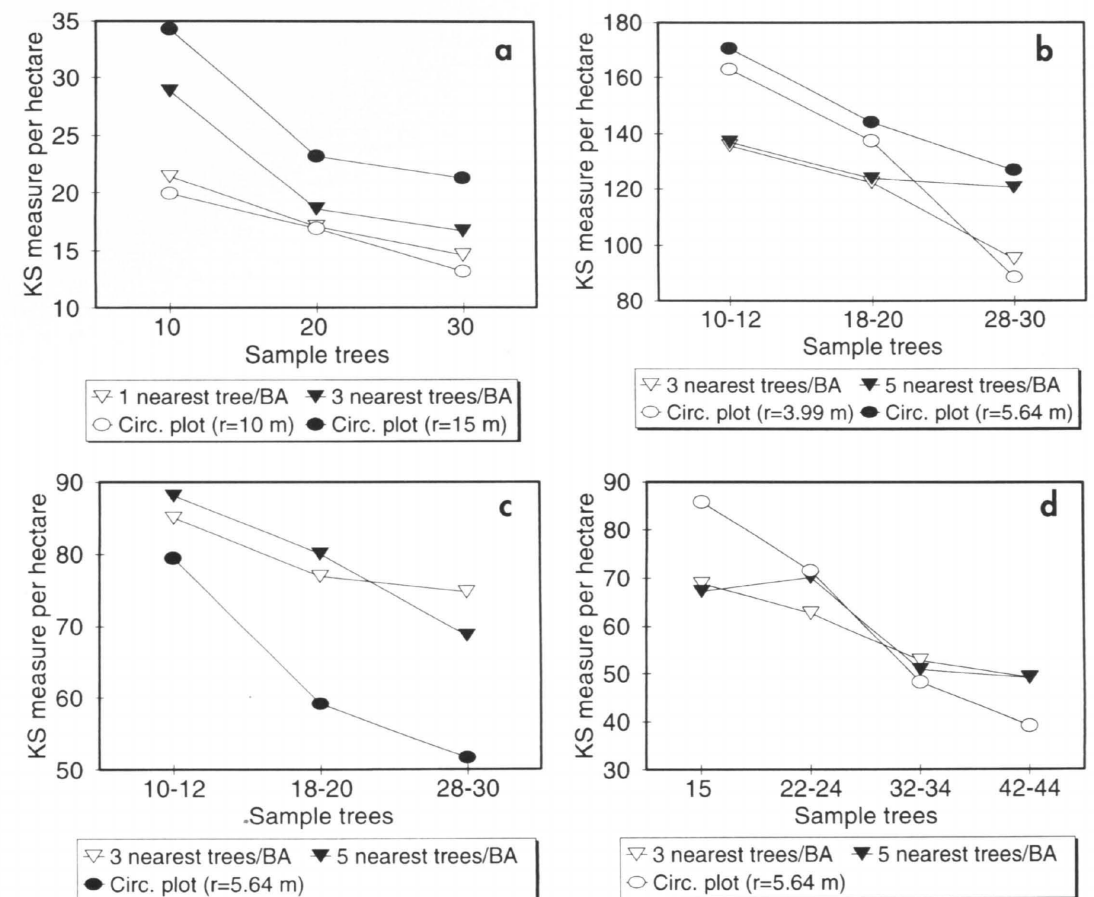


Fig. 12. The precision of the chosen sampling designs on 100 replications in terms of KS measure per hectare for the diameter distribution of pines in stands 3 (a), 5 (b), 7 (c) and 1 (d).

value was used when sample sizes were less than 30 trees and the kernel function with the 2 cm smoothing value when the sample size exceeded 30 trees.

The comparison of the sampling methods implies that there are no big differences in precision between circular plot sampling and nearest trees/BA sampling within the sampling intensity chosen (Fig. 12). Only in stand 7 did circular plot sampling prove to be superior to combined nearest trees/BA sampling (Fig. 12c). One common feature may be found almost in every stand; namely that combined nearest trees/BA sampling is a little better with small sample sizes, the methods are equally good with sample sizes of 20...30 and circular plot sampling is little better with sample

sizes greater than 30. It is quite logical to presume that circular plot sampling is also more precise with the sample sizes that exceed this range. Theoretically, nearest trees/BA area sampling is an inaccurate method since the measurement of basal area and the construction of relative diameter distribution is done partly from different trees. However, this method may be regarded at least as precise as circular plot sampling with small sample sizes. The efficiency of relascope in estimating the basal area of the stand increases the reliability of the real diameter distribution estimates.

The nearest trees/BA method and circular plot sampling were simulated in some stands with two different plot size/number of plot ratios within the same sample size. This enables us to assess

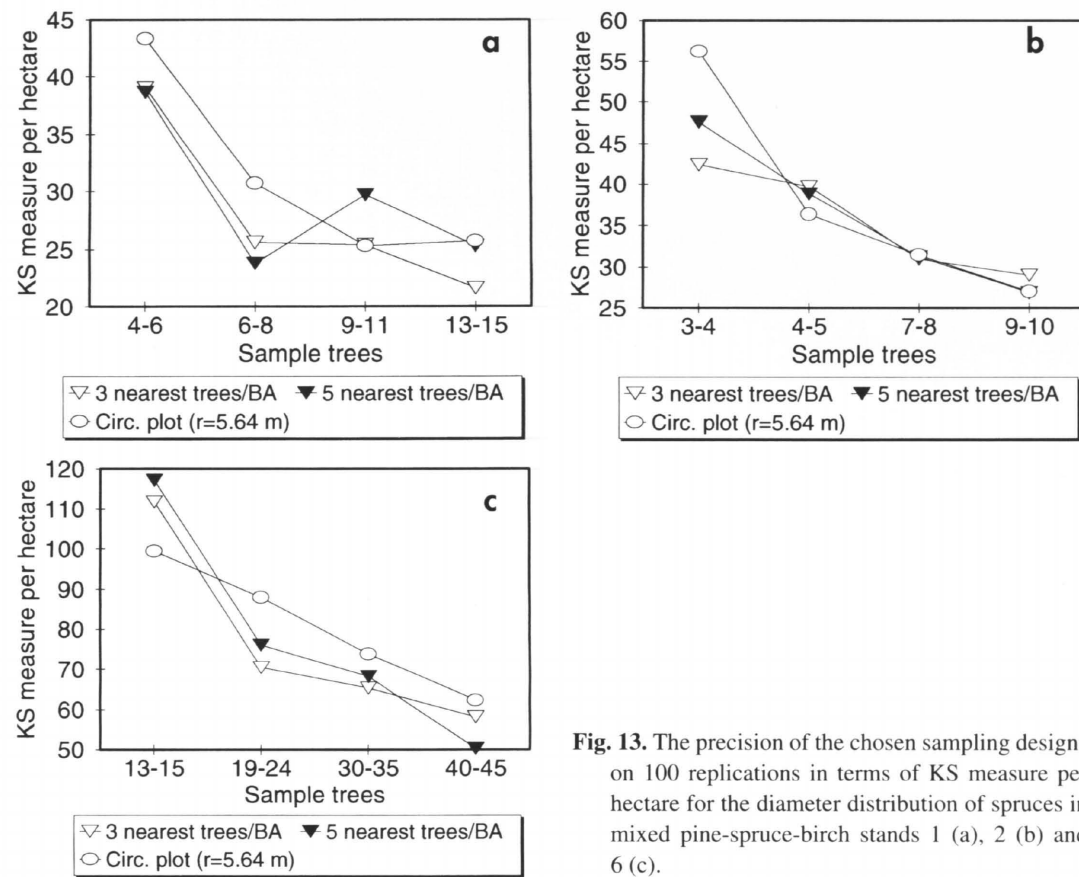


Fig. 13. The precision of the chosen sampling designs on 100 replications in terms of KS measure per hectare for the diameter distribution of spruces in mixed pine-spruce-birch stands 1 (a), 2 (b) and 6 (c).

the optimal relationship between the sample size and number of sample plots. It seems that the right number of sample plots in the stands is more often nearer 10 than 5 almost regardless of the sample size. This observation agrees with the experiences obtained from the USA where so-called ten-point sampling has been widely applied (see e.g. Dilworth and Bell 1973). The name of the method refers to use of ten single points in order to get a good representation of the whole area. Even though the use of ten points has some unquestionable advantages including the ability to provide simultaneous information on area conditions, it may not be the most efficient of the methods considered. When time factor is taken into account the use of ten points may prove to be less efficient than e.g. the use of five or eight points because a greater number of sampling points inevitably leads to longer walking distance. As a matter of fact, Nyysönen and Kilkki (1965) compared the Amer-

ican ten point sampling with 13 other designs in Finnish conditions and found ten-point sampling not to be among the most efficient sampling designs.

Just as with pines there are no big differences in precision between the methods in mixed pine-spruce stands (Fig. 13). The nearest trees/BA sampling seems to be a little better in stands 1 and 6 which had a quite clustered structure. The precision of sampling methods seems to increase quite smoothly with increasing sample size.

Parallel to the results on diameter distributions there seem to be no big differences in precision between the sampling designs in estimating the dead branch height distributions of stands (Fig. 14). Since the comparison of the dead branch height distribution was based on the relative distributions, mean SSE was selected as a suitable measure. However, the results are not as consistent as with diameter distributions. Some incon-

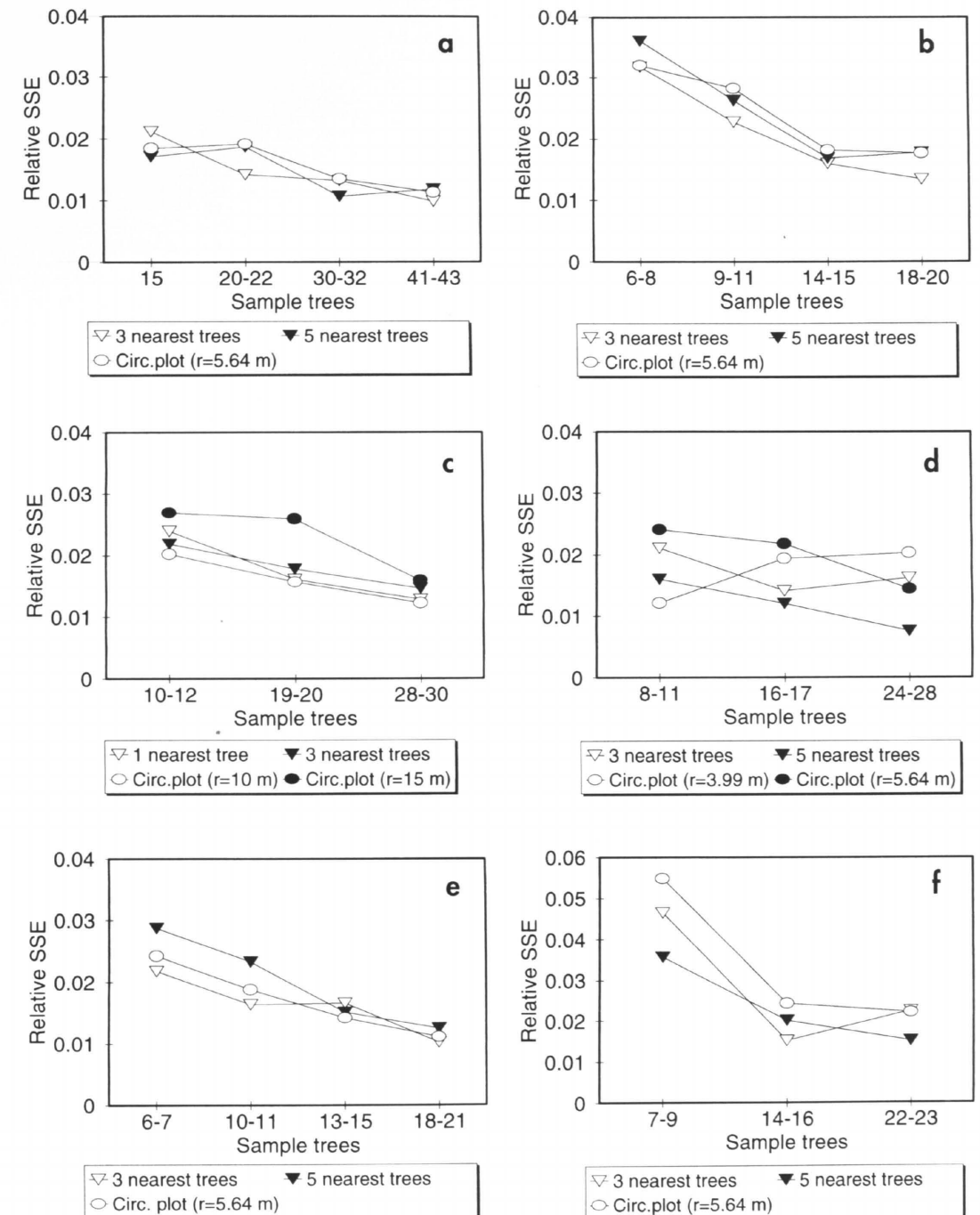


Fig. 14. The precision of the chosen sampling designs on 100 replications in terms of relative SSE measure for the dead branch height distribution of pines in stands 1 (a), 2 (b), 3 (c), 5 (d), 6 (e) and 7 (f).

sistent results, i.e., precision decreases with increasing sampling intensity, implies that the dead branch height varies more than diameter by area within the stand. Therefore, it seems that the placement of sample points all around the stand is perhaps more significant for estimating the dead branch height distribution than for estimating the diameter distribution.

With one exception, a smoothing value of 3 m was applied where the sample size was less than 10 trees and the smoothing value of 2 m was applied when the sample size exceeded 10 trees. Stand 7 differs greatly from the others. Since these pines were growing on very well drained soil [Calluna type (CT) according to Cajander (1926) classification], 95 % of the measures of the dead branch height yield values of 2...4 m. In this case use of 1.5 m or 1 m as the smoothing value of the kernel function gives poor estimates of the dead branch height distribution. The use of 0.5 m as smoothing value however proved to be suitable in stand 7 regardless of the sample size. This divergent observation shows that the deviation of the dead branch height values should be taken into account when the optimal smoothing value is selected. Apart from the diameter distribution, the dead branch height distribution deviates in a relatively narrower range if it is measured with an accuracy of one metre.

Owing to the error in the dead branch height distribution we may conclude that there is no need to increase the total sample size for the sake of this measure. As in the case of diameter distribution the error seems to diminish less markedly after 20–30 trees, in some stands even with a clearly smaller sample size. This being the case, the optimal sample size of the pre-harvest method for pine stands may be derived from the error of the pine diameter distributions which will be dealt with in the following section.

#### 4.3.2.2 Evaluation of the Optimal Sample Size

A close examination of the error curves related to diameter distribution of pines shows that the error curve shape is quite similar in all stands and the level of error curve in terms of  $KS$  measure depends strictly on the density of trees. Consequently, the  $KS$  values of each sampling effort were

converted to commensurable relative  $KS$  values by dividing the  $KS$  value by the density of trees. The error curve was then derived by regression analysis. Owing to the nature of sampling, the error curve shape is best depicted with the form  $E_{KS} = a + bx^{-1}$ , where  $E_{KS}$  is the relative Kolmogorov-Smirnov test value,  $x$  is the number of sample trees and  $a$  and  $b$  are parameters. Leaving stand 3 out of calculations, the output of regression analysis for circular plot sampling and three nearest trees/BA sampling was as follows:

Circular plot sampling:

$$E_{KS} = 0.0982 + 2.25x^{-1} \quad (13)$$

$$n = 17, R^2 = 79.8$$

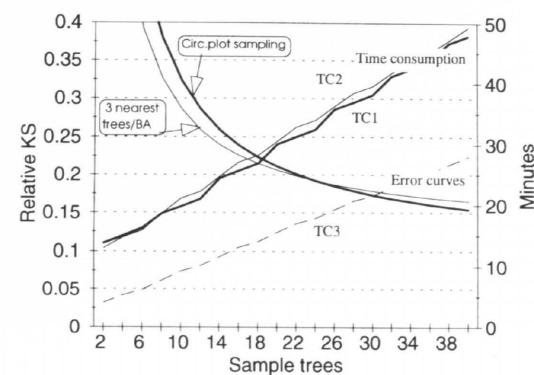
Tree nearest trees/BA sampling:

$$E_{KS} = 0.125 + 1.61x^{-1} \quad (14)$$

$$n = 17, R^2 = 62.8$$

The combination of error and time consumption curves in sampling related to pine diameter distribution is depicted in Figure 15. Error curves describe the formulas (13) and (14). Time consumption curves include the following work elements: moving, establishment of sample plots, measurement of dbh with calliper, measurement of quality parameters (chiefly dead branch height), measurement of basal area with relascope and recording of results. Two curves have been calculated for three nearest trees/BA sampling. In the first curve (TC2) it is supposed that sampling points are defined as in circular plot sampling by moving along a square grid and quality parameters are measured with a telescope pole. In the other curve (TC3) it is supposed that sampling points are defined by walking along a subjectively chosen route and quality parameters are estimated ocularly.

Judged from the curves of Figure 15 the differences between the sampling designs in terms of precise of diameter distribution are almost insignificant. As noted earlier, three nearest trees/BA sampling is slightly more precise with smaller sample sizes ( $n < 25$ ) and correspondingly circular plot sampling is slightly better with greater sample sizes ( $n > 25$ ). If the definition of sampling points has to be done by moving along square grid (TC2) as in circular plot sampling (TC1), the time consumption of these methods are about the



**Fig. 15.** Time consumption of a stand (one hectare) and average error ( $KS$  measure) in diameter distribution of pines by the number of sample trees in circular plot sampling and three nearest trees/BA sampling in pure pine stands. In time consumption curves TC1 and TC2, the cruiser is moving along a square grid while TC3 describes three nearest trees/BA sampling where sampling points are paced out along a subjectively chosen route.

same and are thus equally efficient sampling methods. However, if we can rely on the precision of the other task where sampling points are chosen by walking along a subjectively chosen route and sampling points are defined by pacing out, the three nearest trees/BA sampling tends to be more efficient than circular plot sampling.

As may be seen in Figure 15, the choice of a subjectively chosen route as moving technique (TC3) halves the time taken compared to the technique of defining sampling points by compass and thread measure (TC2). Determination of sampling points by subjectively chosen route has traditionally been avoided in forest mensuration since it is thought to provide biased predictions. When the nearest trees are selected from subjectively chosen points as sample trees there is always a risk that the measurer willingly or unwillingly chooses the most convenient routes for walking, thus obtaining biased estimates.

However, relascope sampling should not, at least in theory, be as prone to bias as the prediction of relative diameter distribution by using the nearest trees, since one measure covers quite a large area. In fact, basal area values are generally measured by relascope from several representative, subjectively chosen spots in preparing the regional for-

est management plans in Finland. Since the determination of basal area is based on the relascope measures in the nearest trees/BA sampling and the combination of basal area and relative diameter distribution seems to provide sufficiently accurate real diameter distributions, the measurer should be permitted to walk along a subjectively chosen route.

A free walking route enables the surveyor also to assess the stand from the logging operation point of view. Taking into account the purpose of the sampling – not to provide data for the wood trade but for production planning – I would recommend the nearest trees/BA sampling method, where the measurer walks along a subjectively chosen route and paces out the sample points as the most appropriate pre-harvest measurement method. However, the final justification of the method can not be done until the amount of bias occurring with this method has been investigated. A good prediction of diameter distribution as well as dead branch height distribution in even-aged pure pine stands of one hectare may be obtained with 25 trees, which means a time consumption of 19 minutes. Correspondingly, measuring 25 trees in similar stands of two hectares takes about 24 minutes.

Since the diameter distribution of pines is regarded as the most important characteristic of the stand, the need to raise the sample size in mixed stands may also be assessed by the error curve related to pine diameter distribution. A prerequisite for determining the “right” sample size is, however, a good estimate of the proportion of the other species in the stand which should not be a big problem, at least in Finnish conditions. Once the proportion of the other species has been estimated, the determination of sample size may be done simply by multiplying the changing part of the error curve by the inverse of proportion of pines in the stand. Hence, the error curve of pine diameter distribution may be written as:

Circular plot sampling:

$$E_{KS} = 0.0982 + 2.25(xp)^{-1} \quad (15)$$

Three nearest trees/BA sampling:

$$E_{KS} = 0.125 + 1.61(xp)^{-1} \quad (16)$$

where  $p$  is the proportion of pines in a stand and

the others are as explained earlier. Correspondingly, the sample size at the desired error level may be calculated from the equations:

Circular plot sampling:

$$x = 2.25(p(E_{KS} - 0.0982))^{-1} \quad (17)$$

Three nearest trees/BA sampling:

$$x = 1.61(p(E_{KS} - 0.125))^{-1} \quad (18)$$

In order to clarify the idea of this approach the error curves of three nearest trees/BA sampling incorporating the pine proportions of 100 %, 75 %, 50 % and 25 % are depicted in Figure 16. The time-consumption curve of three nearest trees/BA sampling is also included in the figure. This figure may be considered as the main result of this study and as a clear instruction in assessing the appropriate sample size in pre-harvest measurement for sawing production planning. The 25 trees as a "good" estimate of pine diameter distribution in a stand with proportion of pines of 75 % requires about 35 sample trees with an elapsed time of about 24 minutes in a stand of 1 ha and about 29 minutes in a stand of 2 ha. Further, the equivalent numbers for a stand with 50 % of pines are 50 trees with an elapsed time of 35 minutes (1 ha) or 40 minutes (2 ha).

#### 4.4 Tree Height and Crown Height Model Studies

##### 4.4.1 Material and Methods

As a result of some earlier studies (e.g. Lappi 1991a, Lemmetty 1993) the traditional approach in which the parameters of the height model are derived merely from sample trees of the stand in question by linear regression analysis was left out of the analysis. Since the traditional approach tends to achieve the error level of calibrated height models after 5...7 trees, it was thought to be too laborious in practice. The appropriateness of the mixed model and calibrated fixed general model should be studied instead. Accordingly, the mixed height model proposed by Lappi (1991a) and the calibration technique for Näslund's (1936) height model proposed by Lemmetty (1993) were tested in four (1, 2, 3 and 5) of the study stands used in previous sub-studies (Table 1). The appropriate-

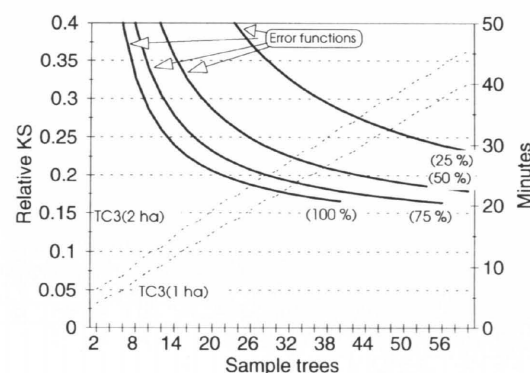


Fig. 16. Time consumption (TC3) and error curves in diameter distribution of pines by the number of sample trees and proportion of pines in mixed stands.

ness of the mixed crown height model proposed by Ärölä (1995) was also studied.

Parameters for Näslund equation were calculated by linear regression analysis from the data of the four study stands (stands 4, 8, 9 and 10, Table 1) that were left out of the height and crown height studies. The reestimates of parameters obtained from the data were  $a = 1.95$ ;  $b = 0.163$  for pine and  $a = 2.12$ ;  $b = 0.154$  for spruce. The residual errors ( $\sigma^2$ ) of the models were 0.178 for pine and 0.139 for spruce. The error resulting from linearization of Näslund's equation was reduced by multiplying the height equation by the transformation coefficient  $1 + 3\sigma^2/(a + bd)^2$ . This coefficient is an approximation of an infinite series resulting when Näslund's equation is substituted in Taylor series expansion. The whole equation can then be written as:

$$h - 1.3 = (d^2 / (a + bd)^2)(1 + 3\sigma^2 / (a + bd)^2) \quad (19)$$

Stand-specific height curves were obtained by substituting the regional estimates of parameters in Equation 19 and then calibrating the equation with the sample trees. The calibration coefficient  $k_i$  was set by using the equation proposed by Lemmetty (1993):

$$k_i = \frac{\sum_{j=1}^n \left( \frac{h_{ij} - 1.3}{\hat{h}_{ij} - 1.3} \right)}{n_i} \quad (20)$$

where  $n_i$  is the number of trees tallied,  $h_{ij}$  is the

height of tree  $j$  in stand  $i$  and  $\hat{h}_{ij}$  is the estimate for tree  $j$  in stand  $i$  calculated with Equation 19. The height estimate for each tree was then obtained from the equation

$$h - 1.3 = k_i(d^2 / (a + bd)^2)(1 + 3\sigma^2 / (a + bd)^2) \quad (21)$$

The mixed linear model proposed by Lappi (1991a) is in general form

$$\ln(H_{ki}) = A_0 + A_1 D_{ki}^{-1} + a_{0k} + a_{1k} D_{ki}^{-1} + e_{ki} \quad (22)$$

where  $D = dbh + 7$  cm,  $A_0$  and  $A_1$  are fixed population parameters,  $a_{0k}$  and  $a_{1k}$  are random stand parameters with zero expectations, and  $e_{ki}$  is the random residual error. The Lappi's study (1991a) was conducted in 26 pine-dominated stands in southern and central Finland. Since the population of that study is parallel to this, the estimates of the fixed parameters as well as the variances and covariances of the random stand parameters were obtained from it.

The height equation can be calibrated for a given new stand  $k$  by predicting the random stand parameters of the height equation (22) from the diameter and height measurements of the sample trees. Suppose that  $\mathbf{y}_k$  is an observed random vector and is in general form generated according to the random parameter model:

$$\mathbf{y}_i = \boldsymbol{\mu}_1 + \mathbf{Z}\mathbf{b} + \mathbf{e} \quad (23)$$

where  $\boldsymbol{\mu}_1$  is a fixed mean vector,  $\mathbf{b}$  is a random parameter vector to be predicted with  $E(\mathbf{b}) = \mathbf{0}$  and  $\text{var}(\mathbf{b}) = \mathbf{D}$ , and  $\mathbf{e}$  is a vector of random errors with  $E(\mathbf{e}) = \mathbf{0}$  and  $\text{var}(\mathbf{e}) = \mathbf{R}$ , and  $\text{cov}(\mathbf{b}, \mathbf{e}') = \mathbf{0}$ . In this case the random stand parameters  $a_{0k}$  and  $a_{1k}$  (random parameter vector  $\mathbf{b}$ ) were predicted using equation

$$\hat{\mathbf{b}} = [\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1}]^{-1}\mathbf{Z}'\mathbf{R}^{-1}(\mathbf{y}_1 - \boldsymbol{\mu}_1) \quad (24)$$

and the variance-covariance matrix of the prediction errors were computed using equation

$$\text{var}(\hat{\mathbf{b}} - \mathbf{b}) = [\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1}]^{-1} \quad (25)$$

The symbols of the previous equations are interpreted as follows (the values of the fixed parameters and variances and covariances being from Lappi 1991a):

$$\mathbf{y}_1 = \begin{pmatrix} \ln(H_{k1}) \\ \vdots \\ \ln(H_{kn_k}) \end{pmatrix}$$

$$\boldsymbol{\mu}_1 = \begin{pmatrix} A_0 + A_1 D_{k1}^{-1} \\ \vdots \\ A_0 + A_1 D_{kn_k}^{-1} \end{pmatrix} = \begin{pmatrix} 3.410 + 18.58 D_{k1}^{-1} \\ \vdots \\ 3.410 + 18.58 D_{kn_k}^{-1} \end{pmatrix}$$

$$\mathbf{Z} = \begin{pmatrix} 1 D_{k1}^{-1} \\ \vdots \\ 1 D_{kn_k}^{-1} \end{pmatrix}$$

$$\mathbf{b} = \begin{pmatrix} a_{0k} \\ a_{1k} \end{pmatrix}$$

$$\mathbf{R} = \text{var}(e_{ki})\mathbf{I} = 0.01113 \mathbf{I}$$

$$\mathbf{D} = \text{var} \begin{pmatrix} a_{0k} \\ a_{1k} \end{pmatrix} = \begin{pmatrix} 0.04739 & -0.3887 \\ -0.3887 & 20.64 \end{pmatrix}$$

The height/dbh function in stand  $k$ , denoted as  $\ln(H_{ki}) | D_{ki}$  is according to equation 22:

$$\ln(H_{ki}) | D_{ki} = 3.410 - 18.58 D_{ki}^{-1} + a_{0k} + a_{1k} D_{ki}^{-1} \quad (26)$$

The stand curve  $\ln(H_{ki}) | D_{ki}$  was predicted by replacing the random stand parameters  $a_{0k}$  and  $a_{1k}$  by their predicted values  $\hat{a}_{0k}$  and  $\hat{a}_{1k}$ .

The variances of the prediction errors were calculated as:

$$\text{var} \left\{ \left[ \ln(\hat{H}_{ki}) | D_{ki} \right] - \left[ \ln(H_{ki}) | D_{ki} \right] \right\} = \text{var}(\hat{a}_{0k} - a_{0k}) + 2 \text{cov}(\hat{a}_{0k} - a_{0k}, \hat{a}_{1k} - a_{1k}) D_{ki}^{-1} + \text{var}(\hat{a}_{1k} - a_{1k}) D_{ki}^{-2} \quad (27)$$

For a tree  $j$  in stand  $k$  of known diameter and unknown height,  $\ln(H_{kj})$  was predicted using equation 26 and  $\text{var}(\ln(H_{kj}) - \ln(\hat{H}_{kj}))$  was obtained by adding  $\text{var}(e_{kj}) = 0.01113$  to equation 26. The logarithmic predictions were transformed to unbiased predictions in the arithmetic scale by adding the factor  $0.5[\text{var}(\ln(H_{kj}) - \ln(\hat{H}_{kj}))]$  to the logarithmic equation before transformation back to arithmetic scale.

The mixed linear model proposed by Ärolä (1995) for determining the crown height of each tree in the stand is in general form:

$$\ln(C_{ij}) = a_0 + a_1 H_{ij} + a_2 \ln(K_{ij}) + a_3 \ln(HM_i) + b_i + e_{ij} \quad (28)$$

where  $C_{ij}$  is the ratio between the height and the crown height of a tree  $j$  in a stand  $i$ ,  $H_{ij}$  is the height of a tree  $j$  in a stand  $i$  (dm),  $K_{ij}$  is the ratio between the height and dbh of a tree  $j$  in a stand  $i$  (dm/mm),  $HM_i$  is the height of the tallest tree in a stand  $i$  (cm),  $b_i$  is the random stand effect of a stand  $i$ ,  $e_{ij}$  is the random error of a tree  $j$  in a stand  $i$  and  $a_0 \dots a_3$  are the fixed parameters.

Ärolä's (1995) data were collected from different regions of southern Finland located about 100–500 km from Parkano. Due to a general assumption that pines in western Finland (excluding the coast) do not differ markedly from those in southern and central Finland, the estimates of the fixed parameters as well as the variances of the random stand parameters were obtained from Ärolä's study (1995). Hence, the model may be written as:

$$\ln(C_{ij}) = -3.193 - 0.001 H_{ij} + 0.299 \ln(K_{ij}) + 0.370 \ln(HM_i) + b_i \quad (29)$$

and the estimate of the between-stand variance  $\hat{\sigma}_b^2$  is 0.006 and the estimate of the within-stand variance (= random error)  $\hat{\sigma}_e^2$  is 0.008. The estimate of the random stand parameter was computed from the equation

$$b_i = \left( \hat{\sigma}_b^2 / \left( \hat{\sigma}_b^2 + \hat{\sigma}_e^2 / n_i \right) \right) (\bar{y}_i - \hat{y}_i) \quad (30)$$

where  $\hat{\sigma}_b^2$  and  $\hat{\sigma}_e^2$  are as mentioned before,  $n_i$  is the number of sample trees,  $\bar{y}_i$  is the mean height of the sample trees and  $\hat{y}_i$  is the mean height of the sample trees estimated by the fixed part of the model. The logarithmic predictions were transformed to unbiased predictions in the arithmetic scale by adding the factor  $0.5\hat{\sigma}_e^2$  to the logarithmic equation before transformation back to arithmetic scale. Since Equation 29 gives only the ratio between the height and the crown height, the final estimate of the crown height can not be calculated without a height estimate. In the tests, crown height was estimated both from the real height of the tree as well as from the height estimate calculated by Lappi's height model (1991a).

Sample trees for analyzing the height and crown height models were chosen by using the sampling simulator described in section 4.3.1.2. Fifty replications of each sampling design were done in each of four stands. In each replication, the desired number of sample trees was selected by systematic sampling, choosing one tree of each species randomly from each circular plot. The models were then calibrated from sample trees and height and crown height estimates were derived for each tree in the stand. The appropriateness of the models was assessed by the average root-mean-square-error (RMSE) of the fifty replications.

#### 4.4.2 Results

As expected, the mixed linear model (method 2) proved to be clearly better than the calibrated fixed height model (method 1) (Fig. 17). The difference is clear but not very great. Both methods seem to be practicable. A close analysis proved that no significant amount of bias occurred. The shape of the curves shows that the accuracy of the model improves little after three calibration trees, so that three is probably a sufficient number of sample trees when the mixed height model or calibrated fixed height model is applied in practice.

As may be noticed, the mixed model gives in some stands better accuracy without calibration than after calibration with one or two trees. This points out that the calibration curves presented without separating the between-stand and within-stand variance, as done in Lappi's (1991a) and Ärolä's (1995) studies give too good picture of the effect of calibration in the stand level.

The accuracy of the calibrated fixed height model with spruce data in stands 1 and 2, which seems to be about the same as with pines, are shown in Figure 18. Apparently the heights of spruce may vary greatly in pine-dominated stands. The increase of the number of spruce sample trees may, however, be questioned since the accuracy of height estimates of spruce may not be essential if they are clearly in a minority. Altogether, three sample trees might be a sufficient number with spruce as well. Until an appropriate mixed height model for spruces is developed, the calibrated fixed height model may be used as substitute method.

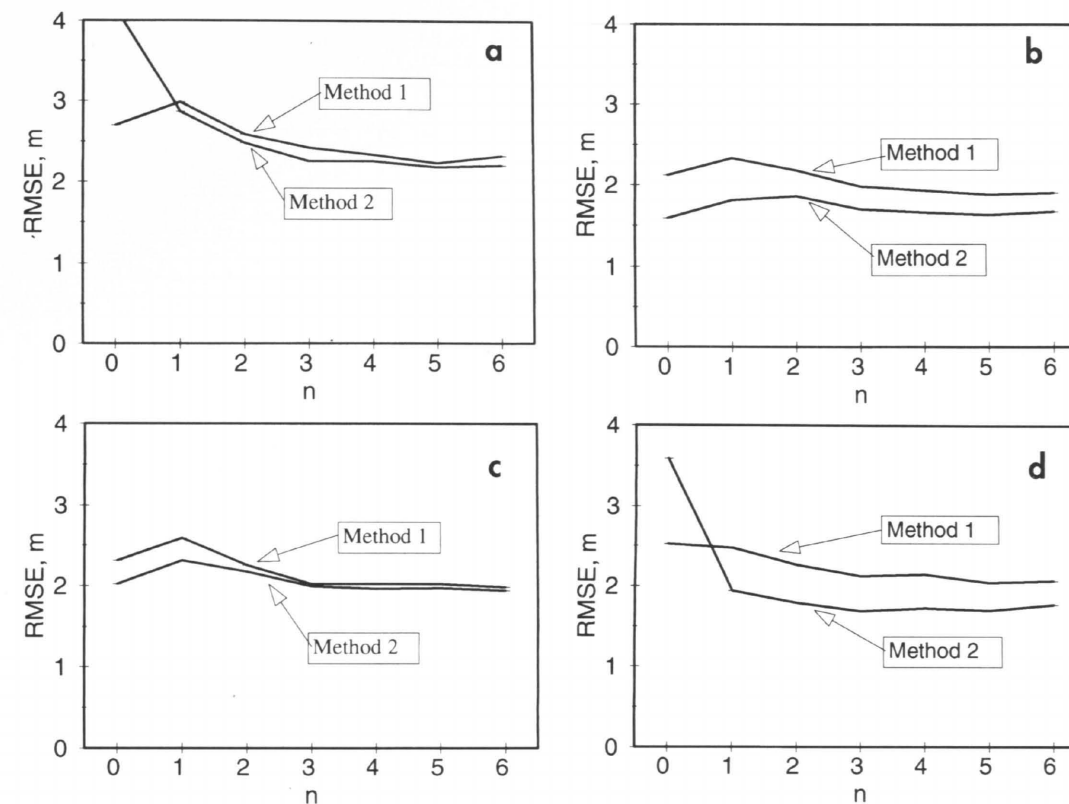


Fig. 17. Average RMSE on 50 replications for the tree heights of pines by the number of calibration trees in stands 1 (a), 2 (b), 3 (c) and 5 (d) with calibrated fixed model (method 1, Equations 20–21,  $a = 1.95$ ,  $b = 0.163$ ,  $\sigma^2 = 0.178$ ) and with Lappi's (1991a) mixed linear model (method 2).

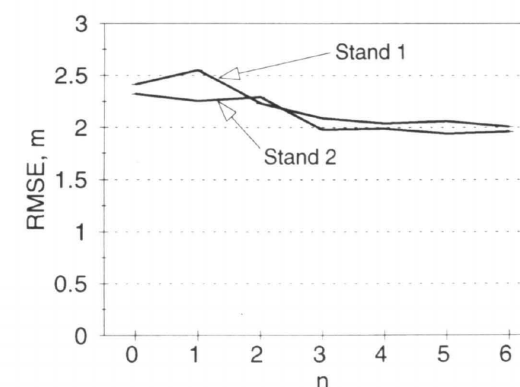


Fig. 18. Average RMSE on 50 repetitions for the tree heights of spruces by the number of calibration trees in stands 1 and 2, applying calibrated fixed model (method 1; Equations 20–21,  $a = 2.12$ ,  $b = 0.154$ ,  $\sigma^2 = 0.139$ ).

The accuracy of Ärolä's (1995) crown height model for pine is shown in Figure 19. As mentioned, this gives only the ratio between the height and the crown height which means that the final estimate for the crown height has to be derived from the height of the tree.

The accuracy of the mixed crown height model seems to improve little after two calibration trees. In this sense, the results are quite similar to Ärolä's (1995). However, the level of accuracy in terms of RMSE is clearly poorer than that obtained in Ärolä (1995), since the RMSE of the crown height estimates in that study was only about 90 cm. In assessing the great difference between the accuracy of these studies, the following reasons emerge. Firstly, the heights as well as the crown heights in this study were measured to an accuracy of 1 meter, while in Ärolä's study they were

measured to an accuracy of 10 cm. Secondly, the number of sample trees in Ärolä's study was only about one tenth of the numbers in this study. Thirdly, the data in Ärolä's study were collected from sample plots while in this study crown height was estimated for every tree in the stand. Moreover, the average crown length ratio (the ratio of crown length to tree height) of the stands used in this study was about 40 % while it was about 30 % in Ärolä (1995) and Lönnroth's (1925) studies. The study stands used by Ärolä (1995) were located in unmanaged forests owned by the Finnish Forest Research Institute. The density of unmanaged forest tends to be considerably higher than in privately-owned managed stands which leads to lack of light in the lower parts of the crowns and thus to higher crown heights. It seems that calibration of the random part in Ärolä's (1995) mixed model is obviously insufficient to achieve a good level of accuracy in managed stands. Consequently, the appropriateness of the mixed height model proposed by Ärolä for our purposes is questionable. The relative error of the crown height in terms of RMSE seems to be about 17 % while it stays as low as about 10 % in the case of heights. It was also noticed that in applying Ärolä's (1995) crown height model, some small bias occurred in some stands. All in all, the prediction of crown height seems to be quite difficult. It proved that crown height varies quite a lot within the stand and the

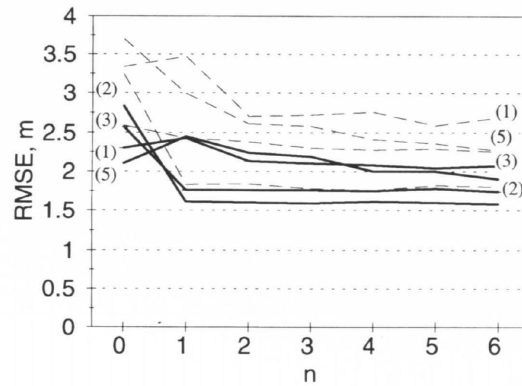


Fig. 19. Average RMSE on 50 repetitions for the crown heights of pines by the number of calibration trees estimated with Ärolä's (1995) crown height model in stands 1, 2, 3 and 5. Solid lines describe the alternatives where crown height estimates are derived from real height values while dashed lines describe the alternatives where crown height estimates are derived from the height estimates calculated by Lappi's (1991a) height model.

derivation of the estimate from three heights and diameter/height relations do not achieve the same level of accuracy as the height estimates derived from diameter/height relationships. Therefore, it seems that the prediction of crown height should be studied more intensively in the future.

## 5 Pre-harvest Measurement Application

The pre-harvest measurement method developed will be applied primarily in providing data for a comprehensive stem-optimization sawing model. The stands will be described by a total enumeration of trees or alternatively by 30...100 "element trees" characterized by the chosen variables. The method must be quick and efficient since it is meant to be used in connection with forest managers' daily timber buying activities. Ideally the measurement work should be done in 20–30 minutes and the results stored directly to a hand-held field computer.

Analysis resulted in choosing dbh, dead branch height, crown height and tree height as the most appropriate stem characteristics of Scots pine. Spruces (Norway spruce) and birches (pubescent and silver birches) will be described only by dbh and tree height. The remaining trees (aspen, alder, etc.) are not described for the stem optimization model but the total volumes of these species are naturally estimated by sample data.

Sample plots for the measurement may be defined by pacing out a subjectively chosen route. Roughly 25 sample trees seems to be appropriate in pure pine stands. In mixed stands equation 18 may be used to determine the sample size that equals the same level of accuracy ( $E_{KS} \sim 0.18...0.20$ ) in predicting the pine diameter distribution. Each sample tree in the plot is measured for dbh. Every pine sawlog stem is also measured for dead branch height. There seem to be no significant differences in accuracy choosing either three or five nearest trees as samples. From the efficiency point of view three nearest trees is better in pure stands and five nearest trees is better in mixed stands since there seems to be no reason to establish more than ten sample plots. The measurer also needs to take relascope measures if the total volumes of the tree species are not known exactly beforehand. The relation of 1/5 between basal-area measures and number of sample trees seems to be appropriate when BAF = 1 m<sup>2</sup>/ha is applied.

Pine and spruce diameter distribution as well as dead branch height distribution are most effectively predicted by the kernel function. A bin width (smoothing parameter) of 3 cm is adequate for pine and spruce diameter distribution when the sample size is under 30 trees, otherwise narrower bin widths should be used. Real hectare diameter distribution may be calculated once the relative distribution is related to basal area of the stand. In constructing the dead branch height distribution 1 m may be used as a suitable bin width unless the scale of the observation is extremely narrow as in very dry soils (CT). In these cases, narrower bin widths should be used.

As a result of their measurement difficulty, height and crown height are derived from dbh measures once the dbhs have been first related to the height and crown height measures of height sample trees. According to my findings, tree height and crown height sample trees should be enough to calibrate the height and the crown height models used in this study. This means that in mixed stands we usually need to measure six or nine tree height sample trees and in pure pine stands three height sample trees since tree height and crown heights are naturally measured from the same trees. The principle of the mixed models requires that height sample trees should be selected randomly within the sample plot. Depending on the number of sample plots and proportions of the tree species, height sample trees should be chosen from all, every second, every third or suchlike so that sample trees are measured from different parts of the stand.

A realistic description of the stem population characterized by the chosen four variables is attained once dbh and dead branch heights are linked together as a two-dimensional diameter-dead branch height distribution and some variation is incorporated in the height and crown height estimates derived from dbh measures. There are basically two approaches to linking diameter distri-



bution and dead branch distribution. One is to calculate the correlation between these variables in the stands and link the distribution together by following the principles used in simulating diameter/height distribution in artificial stands (e.g. Kilkki and Siitonen 1975). In this approach the upper and lower boundaries of dead branch height distribution are defined for each diameter and the dead branch height value for each tree is derived by random number. The other, perhaps more attractive approach is to construct multivariate distribution from these variables. Even though a credible extension of the kernel function to multivariate data has been presented (Silverman 1986), the application of this approach was left out this study.

Correlation coefficient is a good auxiliary tool for assessing the shape of the two-dimensional distribution. The lower the value of correlation coefficient is the more the shape of that distribution remains a circle, while the greater the value is

the more the shape remains a line. The present findings suggest that the correlation between diameter and dead branch height varies a lot from stand to stand. Once the correlation coefficient has been calculated, we have enough building blocks to construct a realistic estimate of the two-dimensional diameter/height distribution.

To attain a more realistic description we need to include variation in the height values estimated with the height model. We may assume that height is normally distributed with the parameters of mean and standard deviation. The mean is obtained from the height model and 10 % as calculated from the data of this study may be used as the value of standard deviation. Actually, the distribution of the heights does not necessarily follow the normal distribution, but it is probable that no great error is caused by this presumption. Naturally, the same principle described here with height may also be applied in including variation in crown height estimates.

## 6 General Discussion

### 6.1 Materials and Methods

Study materials were collected from ten Scots pine stands (*Pinus sylvestris*) located in North Häme and South Pohjanmaa, in southern Finland. The data comprise test sawing data on 314 pine stems, dbh and height measures of all trees and measures of the quality parameters of pine saw log stems in all ten study stands as well as the locations of all trees in six stands. Despite the smaller number of test sawing stems compared to earlier similar Finnish studies (Kärkkäinen 1980, Uusitalo 1994a) the results obtained may be considered reliable. This study material was collected with unparalleled accuracy since as distinct from the earlier studies the location of every sawn good was registered properly. Although the test sawing was done by circular saw the results may be considered to be comparable with the results that would have been obtained from the bandsaw process of Aureskoski. Logistic regression proved to be a satisfactory mean of assessing and modeling the lumber qualities of the stems. The appearances of the quality grades were formed as polytomous dependent variables instead of several dichotomous variables as in earlier studies (Uusitalo 1994a, 1994b). This approach proved to be as easily used as the earlier approach and penetrates the character of the quality distribution obtained from the sawing process better. The finding that some features of the models are almost identical to the earlier models in Uusitalo (1994a) implies that grading rules as well as the quality of the pines do not necessarily differ as significantly as has been presumed earlier.

The area of the mapped stands varied from 0.8 to 2.4 hectares, matching the average stand sizes in southern Finland quite well. The structures of the stands in terms of forest type, stand density, proportion of tree species and cutting method also varied considerably. Overall the results may be considered to be rather valid in most of southern Finnish final cut stands. The results may not how-

ever be valid in thinnings because they were wholly restricted to final cut areas. Applying the results in stands that are considerably larger than those studied is questionable. In larger stands the variation in characteristics tends to increase, forcing us to increase the sample size. If there are clear silvicultural sections within the stands they ought to be measured separately.

The use of a simulator in assessing the accuracy of sampling has its own risks. The advantage of simulating is that one sampling design may be repeated often which gives us at least in theory a tool to analyze the differences between the methods. The major disadvantage in simulation of sampling is that the human effects on each sampling design cannot be assessed.

### 6.2 Results

The choice of dbh, dead branch height and crown height as the main quality variables of Scots pine stems are in harmony with earlier Finnish studies (e.g. Heiskanen 1954a, Kärkkäinen 1980, Uusitalo 1994a). Dead branch height evidently has a clear connection with the thickness of the branches during the early growth of the pine. The thicker the branches grow early in life, the longer the natural pruning takes resulting in low dead branch height in later stages and poor quality. Dbh affects the probability of u/s grade in the butt logs in two ways. Firstly, the grading rules are to a some extent dependent on the dimension of the sawn good in that the bigger the pieces the thicker the knots permitted are. Secondly, large amounts of less knotty or even knotless u/s boards may be obtained only from large butt logs which means that the probability of u/s grade increases with increasing dbh.

The significance of the crown height in predicting the quality of the upper logs has been disregarded in earlier studies since almost all the atten-

tion has been directed at butt logs. The results obtained demonstrate that crown height is crucial especially in predicting the appearance of knotty pine in upper logs. Altogether, no big surprises concerning the influence of the quality parameters is found in this study. However, the mathematical connections between the quality grades and the main stem characteristics are presented more precisely than in earlier studies which gives a clearer understanding of this phenomenon.

In addition to the chosen stem characteristics, early growth rate measured from the stump height and second cut height was found to be a good quality parameter for Scots pine. The growth rate proved to be a good auxiliary variable in predicting the probability of high quality u/s grade in the butt logs. This is not surprising since the correlation between the early growth rate and good butt log quality has been noticed before (e.g. Heiskanen 1954b, 1965, Orver 1970, Weslien 1983). It seems apparent that growth rate is more influenced by the quality of the site while dead branch height is more a matter of the thickness of the branches which is also affected by the stand density. Accordingly both dbh, dead branch height and early growth rate should be incorporated into a good quality prediction model for butt logs. Unfortunately, early growth rate had to be omitted from the method since the drilling of standing trees and calculation of annual rings is too laborious with present available tools. Since the information about the early growth rate is crucial in some stands in obtaining a reliable prediction of the butt log quality, there is an urgent need to develop a more advanced tool for determining the growth rate.

Second cut height together with crown height and dead branch height is an important auxiliary characteristic in determining the qualities of the second and third logs. Therefore, there is no sense in applications predicting the quality of the upper logs if the second cut height is not determined. The height of the first cut however has not been seen to affect the quality of the butt logs in this study nor in my earlier study (Uusitalo 1994a).

Smoothing a sample-based distribution provides a more precise prediction of diameter and dead branch height distribution. The kernel method proved to be a very flexible and easily used function in the construction of diameter distribution, proving to be the best of the techniques consid-

ered. The use of prior knowledge of the general shapes of the diameter distribution applied by predicting the parameters of Weibull function with general regression models seemed not to be beneficial. However, since no comprehensive comparison between the parametric and non-parametric methods in constructing the diameter distribution was done, there are no opportunities to draw any conclusions about the superiority of these techniques, even though the non-parametric methods like the kernel method appear to have more potential in applications where great accuracy is needed.

From the comparative smoothing values in applying the kernel function, 3 cm seems best when the sample size is less than 20, 2 cm and 3 cm are equally good when sample size varies between 20...30 and 2 cm is most satisfactory when sample size varies between 30...50. There were no marked differences between pine and spruce, even though diameter distribution of spruce tends to form into bimodal or multimodal shapes when it is in a clear minority. These results are in accordance with the earlier findings of Drossler and Burk (1989). In their study, the optimal smoothing parameter of the kernel function was found to be 2.54...3.05 cm with samples of size 10 and 2.16...2.41 with samples of size 30 in hypothetical, bimodally shaped red pine (*Pinus resinosa* Ait.) stand. The simulations proved that the error in diameter distribution when the most efficient smoothing is applied, decreases quite noticeably up to 20...30 sample trees, but starts to decrease with clearly smaller steps after that. The usefulness of the adaptive kernel method (i.e. smoothing parameters may vary from point to another) was not analyzed in this study. It is however obvious that this approach may enhance the construction of diameter distribution in uneven-aged stands.

The kernel function also proved to be appropriate in constructing the dead branch height distribution. A smoothing value of 1 m was found to be most precise if the dead branch height was measured to an accuracy of 1 m. The normal distribution seemed poorly suited to constructing the dead branch height distribution of pine. The shape of the dead branch height error curve implies that there is no need to increase the sample size for the sake of the dead branch height distribution, but

20...30 also seems to be an adequate number of sample trees.

The sampling simulations showed that there are no big differences in precision between circular plot sampling and nearest trees/BA sampling. Circular plot sampling was found to be slightly more precise with the samples of more than 30 trees, while nearest trees/BA sampling was slightly more precise with smaller sample sizes. It was proved that even though nearest trees/BA sampling has some weakness in theory since the measurement of basal area and diameter distribution is done partly from different trees, the method may be applied with relatively small sample sizes in constructing the real density estimate of the diameter. Circular plot sampling is theoretically more precise but is inaccurate in practise in clustered stands. The more the sampling intensity is increased the more obvious is the superiority of the circular plot sampling in precision.

This is probably the first study where comparison of the sampling designs is based on KS measures of diameter distributions. However, the real diameter distribution per hectare is virtually calibrated by basal area measures in nearest trees/BA sampling. In this sense, I have, in fact, compared ACS and circular plot sampling. Accordingly, the results obtained from sampling simulations may be considered to be in accordance with the earlier studies that have dealt with the comparison of ACS and systematic sampling with fixed plots (e.g. Avery and Newton 1965, O'Regan and Arvanitis 1966, Wensel and John 1969).

The results obtained from the simulations of the calibration of height models supports the theoretical presumption that the mixed height model when applied in the right population provides height estimates most efficiently. The mixed height model for pines proposed by Lappi (1991a) seemed to work well in southern Finnish conditions. However, until an appropriate mixed height model for spruces is developed, the calibrated fixed height model may be used as substitutive method. The appropriateness of the crown height model proposed by Ärölä (1995) proved to be fairly poor. This is not perhaps wholly due to the weakness of the proposed model but rather to the difficulty of predicting the crown height.

### 6.3 Evaluation of the Proposed Pre-Harvest Measurement Method

The proposed pre-harvest measurement method meets the time requirements imposed. It seems that in most of the privately-owned pine stands in southern Finland the pre-harvest measurement can be carried out in less than one hour since the size of these stands seldom exceed three hectares. In pure pine stands the measurement usually should not take more than half an hour. The prerequisite of simplicity may also be regarded as fulfilled. Only ordinary measuring instruments, calliper and hypsometer are needed, which enables the extension of this method to every manager involved with wood procurement activities.

The results provided by the method are supposed to be linked with a comprehensive production planning system for sawing. A stem-optimization model based on linear programming constitutes a keystone of the production planning system. This model is intended to provide optimal instructions for wood procurement managers to choose the right stands to use and direct cross-cutting within each stand. The linear optimization pertains to optimal solutions only when information provided can be presumed to be accurate. The impact of inaccurate information on the appropriateness of the proposals provided by optimization is not obvious. In fact, the effect of inaccurate information on the financial benefit of sawmilling differs from case to case. Therefore it is almost impossible to give any calculated limits at the moment on the accuracy of data provided by the proposed pre-harvest measurement method. Once the whole integrated system is installed and has been proved in operational use we will be able to analyze the effect of accuracy on prior information about a stand.

The distribution predictions are naturally prone to bias if the sampling points are defined by walking a subjectively chosen route as proposed instead of using systematic sampling. The bias is the greater the more heterogeneous the stand is. In dense stands especially the measurer tends willingly or unwillingly to choose the most convenient route which might lead to biased estimates. The use of relascope in measuring the basal area as well as smoothing of the diameter distribution should however diminish this risk. Since there

might also be considerable differences between the measurers, the use of measurement method should be subjected to control activities in order to guide the managers in their measurement work.

#### 6.4 Future Perspectives and Final Remarks

A prerequisite to any major innovation in the area of wood procurement requires that we have accurate prior knowledge about the structure of the forests which are going to be exploited. It is obvious that in the future considerably more research activities will be directed at this field. As distinct from earlier research there will be a great need to emphasize the character of each process. In sawmilling, for example, there is already an urgent need to investigate the quality of spruce in a similar way to this study. The quality of spruce has until now been disregarded since the value relations of the grades has not traditionally varied as much as with pine.

Although the error factors in measurement are known to play a significant role in forest inventory there are quite few studies dealing with that problem (Hyppönen and Roiko-Jokela 1978, Gertner 1984, Lappi 1990b, Päivinen et al. 1992). Since the dead branch height has proved to be crucial in predicting the quality of Scots pine there is an urgent need to study the different source of errors related to measure. If we cannot rely on the accuracy of measuring the dead branch height the usefulness of the proposed pre-harvest measurement method is poor. Obviously, the accuracy of meas-

urement may be improved by practise and consultation.

At present, there is an apparent tendency to stress the end-user's requirements in the quality of structural timber (Johansson et al. 1994). This means that instead of general grading rules the sawmills will have to generate customer-specific grading rules in future. If timber is to hold its own or even increase its market share as a construction material, as is generally wanted, the sawmill industry will have to produce products which match the end-user's requirements. Turning to more customer-oriented grading does not, however, make the general quality models as created in this study unnecessary. Rather the contrary – applying customer grading rules in sawing production planning increases the need for prior knowledge of the raw material. The extension of the general models to customers' grades naturally requires modeling of some kind of converter.

For many decades, there has been a great desire in Finland to extend the quality requirements of sawn timber to silvicultural guides (e.g. Vuoristo 1936, Heiskanen 1965, Uusvaara 1974, Vuokila 1982, Kellomäki and Väisänen 1986). The principles elucidated by this research have also been applied widely in practice. Recently, Johansson et al. (1994) have required that the end-users' point of view should more actively be taken into account in silvicultural activities. The exaggeration of some view points in forestry has, however, in many cases proved to be a mistake in the long run. Desire for high value timber in wood production may generally decrease forest biodiversity (Uusitalo 1995).

## References

- Ärölä, E. 1995. Männyn elävän latvuksen alarajan ja oksan neulasmassan ennustaminen. Manuscript. University of Helsinki. 46 p.
- Avery, G. & Newton, R. 1965. Plot sizes for timber cruising in Georgia. *Journal of Forestry* 63(10): 930–932.
- Bailey, R.L. & Dell, T.R. 1973. Quantifying diameter distributions with the Weibull function. *Forest Science* 19(2): 97–104.
- Bickford, C.A. 1952. The sampling design used in the forest survey of the north east. *Journal of Forestry* 50(4): 290–293.
- 1961. Stratification for timber cruising. *Journal of Forestry* 59(10): 761–763.
- Bitterlich, W. 1949. Die Winkelzählprobe. *Allgemeine Forst- und Holzwirtschaftliche Zeitung* 59(1/2): 4–5.
- Bliss, C.I. & Reinker, K.A. 1964. A log-normal approach to diameter distributions in even-aged stands. *Forest Science* 10: 350–360.
- Borders, B.E., Souter, R.A., Bailey, R.A. & Ware, K.D. 1987. Percentile-based distributions characterize forest stand tables. *Forest Science* 33(2): 570–576.
- Brundage, M.R. 1936. Comparative lumber selling values of different grades and sizes of redwood logs. *Californian F & R Experiment Station. Research Note* 8.
- Burk, T.H. & Newberry, J.D. 1984. A simple algorithm for moment-based recovery of Weibull distribution parameters. *Forest Science* 30(2): 329–332.
- Cajander, A.K. 1926. The theory of forest types. *Acta Forestalia Fennica* 29(3). 108 p.
- Cajanus, W. 1914. Über die Entwicklung gleichaltriger Waldbestände. Eine statische Studie. I. *Acta Forestalia Fennica* 3. 142 p.
- Chang, S.J. 1990. Application of medical imaging technology to improve timber resource utilization. *Proc. Resource Technology 90. Second International Symposium on Advanced Technology in Natural Resources Management. November 12–15, 1990, Washington, D.C., USA.*
- Curtis, R.O. 1967. Height-diameter and height-diameter-age equations for second-growth Douglas-fir. *Forest Science* 13(4): 365–375.
- de Licourt, F. 1898. De l'aménagement des Sapinières. *Bull de la Société Forestière de Franche Comte et Belfort. Besancon* 4(6): 396–409.
- Dilworth, J.R. & Bell, J.F. 1973. Variable probability sampling. Variable plot and three-p. Corvallis, Oregon, O.S.U. Book Stores Inc. 130 p.
- Droessler, T.D. & Burk, T.E. 1989. A test of nonparametric smoothing of diameter distributions. *Scandinavian Journal of Forest Research* 4: 407–414.
- Epanechnikov, V.A. 1969. Nonparametric estimators of a multivariate probability density. *Theory of Probability and its Applications* 14: 153–158.
- Fox, J. 1984. Linear statistical models and related methods with applications to social research. John Wiley & Sons. 449 p.
- Freese, F. 1961. Relation of plot size to variability: An approximation. *Journal of Forestry* 59(9): 67.
- Füldner, K. 1995. Strukturbeschreibung von Buchen-Edellaubholz-Mischwäldern. PhD thesis. University of Göttingen. Germany. 87 p.
- Gertner, G.Z. 1984. Control of sampling error and measurement error in a horizontal point cruise. *Canadian Journal of Forest Research* 14: 40–43.
- Gregory, G.R. & Pearson, H. L. 1949. Lumber values for East Texas pine logs. *Southern Experiment Station. Occ. Paper* 113.
- Grosenbaugh, L.R. 1952. Plotless timber estimates – new, fast, easy. *Journal of Forestry* 50(1): 32–37.
- 1965. Tree-pee sampling theory and program 'THRP' for computer generation of selection criteria. *USDA Forest Service, Research Paper PSW-21*: 1–53.
- 1967. The gains from sample tree selection with unequal probabilities. *Journal of Forestry* 65(3): 203–206.
- Hafley, W.L. & Schreuder, H.T. 1977. Statistical distributions for fitting diameter and height data in even aged stands. *Canadian Journal of Forest Research* 7: 481–487.
- Hakala, H. 1992. Mäntytukkien sahausken järeyden mukainen taloudellinen tulos ja siihen vaikuttavia tekijöitä. Summary: Financial results of sawing pine logs as influenced by top diameter and other asso-

- ciated factors. *Acta Forestalia Fennica* 226. 74 p.
- Hanks, L.F. 1976. How to predict lumber-grade yields for graded trees. USDA Forest Service. General technical report NE-20: 1–9.
- Heiskanen, V. 1951. Mäntysahatukkien laadun mukaisista arvosuhteista. *Silva Fennica* 69: 92–103.
- 1954a. Tutkimuksia mäntytukkipuun laatuluokitustavoista ja niiden tarkkuudesta. Summary: Investigations into pine tree grading methods and their accuracy. *Communicationes Instituti Forestalis Fenniae* 44(1). 132 p.
- 1954b. Vuosiluston paksuuden ja sahatukin laadun välisestä riippuvuudesta. Summary: On the interdependence of annual ring width and sawlog quality. *Communicationes Instituti Forestalis Fenniae* 44(5). 31 p.
- 1965. Puiden paksuuden ja nuoruuden kehityksen sekä oksaisuuden välisistä suhteista männiköissä. Summary: On the relations between the development of the early age and thickness of trees and their branchiness in pine stands. *Acta Forestalia Fennica* 80 (2). 62 p.
- & Siimes, F.E. 1959. Tutkimus mäntysahatukkien laatuluokituksista. Paperi ja puu 41(8): 359–368.
- Hirata, T. 1955. Height estimation through Bitterlich's method-vertical angle-count sampling. (In Japanese). *Journal of the Japanese Forestry Society* 37(11): 479–480.
- Hosmer, D.W. & Lemeshow, S. 1989. Applied logistic regression. John Wiley & Sons. 309 p.
- Howard, A.F. 1991. Timber quality classes from cruise data for lodgepole pine sawn at a board mill in British Columbia. *Canadian Journal of Forest Research* 21: 498–503.
- & Gasson, R. 1989. A recursive multiple regression model for predicting yields of grade lumber from lodgepole pine sawlogs. *Forest Products Journal* 39(4): 51–56.
- & Yaussy, D.A. 1986. Multivariate regression model for predicting yields of grade lumber from yellow birch sawlogs. *Forest Products Journal* 36(10/11): 56–60.
- Hyink, D.M. & Moser, J.W. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. *Forest Science* 29(1): 85–95.
- Hyppönen, M. & Roiko-Jokela, P. 1978. Koemittauksen tarkkuus ja tehokkuus. Summary: On the accuracy and effectivity of measuring sample trees. *Folia Forestalia* 356. 25 p.
- Iles, K. & Fall, M. 1988. Can an angle gauge really evaluate "borderline trees" accurately in variable plot sampling? *Canadian Journal of Forest Research* 18: 774–781.
- Ilvessalo, Y. 1935. Käytännöllisistä metsänarvioimistavoista. Referat: Über die praktischen Forsttaxationsmethoden. *Silva Fennica* 39: 1–12.
- Johansson, G., Kliger, R. & Persorper, M. 1994. Quality of structural timber-product specification system required by end-users. *Holz als Roh- und Werkstoff* 52: 42–48.
- Johnson, F.A. & Hixon, H.J. 1952. The most efficient size and shape of plot to use for cruising in old-growth douglas-fir timber. *Journal of Forestry* 50(1): 17–20.
- Jonsson, B., Holm, S. & Kallur, H. 1992. A forest inventory method based on density-adapted circular plot size. *Scandinavian Journal of Forest Research* 7: 405–421.
- Kangas, Y. 1959–60. Koealan muoto ja koko koealuarvioinnissa. Summary: Influence of plot shape and size on sample plot survey. *Metsätaloudellinen Aikakauslehti*. p. 419–420, 17–20, 22 and 54–55.
- Kellog, L. 1941. Log grades of timber in eastern Ohio. I. Eastern pine on Mohican River Gorge, Holmes County, Ohio. Central Forest Experiment Station. Technical Note 38.
- Kellomäki, S. & Tuimala, A. 1981. Puuston tiheyden vaikutus puiden oksikkuuteen taimikko- ja riukuvaiheen männiköissä. Summary: Effect of stand density on branchiness of young Scots pines. *Folia Forestalia* 478. 27 p.
- & Väisänen, H. 1986. Kasvatustiheyden ja kasvu- paikan viljavuuden vaikutus puiden oksikkuuteen taimikko- ja riukuvaiheen männiköissä. Summary: Effect of stand density and site fertility on the branchiness of Scots pines at pole stage. *Communicationes Instituti Forestalis Fenniae* 139. 38 p.
- Kilkki, P., Maltamo, M., Mykkänen, R. & Päivinen R. 1989. Use of the Weibull function in estimating the basal area dbh-distribution. *Silva Fennica* 23(4): 311–318.
- & Siitonen, M. 1975. Metsikön puuston simulointimenetelmä ja simuloituun aineistoon perustuvien puustotunnusmallien laskenta. Summary: Simulation of artificial stands and derivation of growing stock models from this material. *Acta Forestalia Fennica* 145. 33 p.
- Kulow, D.L. 1966. Comparison of forest sampling designs. *Journal of Forestry* 64(7): 469–474.

- Kuusela, K. 1966. A basal area-mean tree method in forest inventory. *Seloste: Pohjapinta-alakeskipuun menetelmä metsäinventoinnissa*. *Communicationes Instituti Forestalis Fenniae* 61(2). 32 p.
- Kärkkäinen, M. 1980. Mäntytukkirunkojen laatuluokitus. Summary: Grading of pine sawlog stems. *Communicationes Instituti Forestalis Fenniae* 96(5). 152 p.
- 1986. Malli männyn, kuusen ja koivun puuaineen oksaisuudesta. Abstract: Model of knottiness of wood material in pine, spruce and birch. *Silva Fennica* 20(2): 107–115.
- & Uusvaara O. 1982. Nuorten mäntyjen laatuun vaikuttavia tekijöitä. *Folia Forestalia* 515. 28 p.
- Lappi, J. 1991a. Calibration of height and volume equations with random parameters. *Forest Science* 37(3): 781–801.
- 1991b. Estimating the distribution of a variable measured with error: stand densities in a forest inventory. *Canadian Journal of Forest Research* 21: 469–473.
- Leinonen, K. & Rita, H. 1995. Interaction of prechilling, temperature, osmotic stress and light in *Picea abies* seed germination. *Silva Fennica* 29(2): 95–106.
- Lemmetty, J. 1993. Leimikon puutavaralajiosuukien arviointi systemaattisella koealotannalla. Thesis for M.Sc. graduation. University of Helsinki. Department of Forest Resource Management. 47 p.
- & Mäkelä, M. 1992. Suunnittelumittauksen perusteet ja toteutus. Summary: Methods for measurement of a stand for harvest planning. *Metsätehon katsaus* 11. 4 p.
- Lindgren, O. 1984. A study on circular plot sampling of Swedish forest compartments. Swedish University of Agricultural Sciences. Section of Forest Mensuration and Management. Umeå. Sweden. Report 11. 153 p.
- Lönnroth, G. 1985. Integrationsmöjligheter i kedjan skogsäg-marknad. Summary: Integration possibilities in the forest-sawmill-market chain. The Swedish University of Agricultural Sciences. Department of Forest products. Uppsala. Sweden. Report 164. 130 p.
- Lönnroth, E. 1925. Untersuchungen über die innere Struktur und Entwicklung gleichaltriger naturnormaler Kiefernbestände basiert auf Material aus Südhälfte Finnlands. *Acta Forestalia Fennica* 30(1). 269 p.
- McCullagh, P. & Nelder, J.A. 1989. Generalized linear models. 2nd edition. Chapman and Hall, London. 511 p.
- Mesavage, C. & Grosenbaugh, L.R. 1956. Efficiency of several cruising designs on small tracts in North Arkansas. *Journal of Forestry* 54(9): 569–576.
- Meyer, H.A. 1949. Cruising intensity and accuracy of cruise. *Journal of Forestry* 47: 646–649.
- Mykkänen, R. 1986. Weibull-funktion käyttö puuston läpimittajakauman estimoinnissa. Thesis for M.Sc. graduation. University of Joensuu. Faculty of Forestry. 80 p.
- Mäkinen, H. & Uusvaara, O. 1993. Lannoituksen vaikutus männyn oksikkuuteen ja puuaineen laatuun. Summary: Effect of fertilization on the branchiness and wood quality of Scots pine. *Folia Forestalia* 801. 23 p.
- Nelson, T.C. 1964. Diameter distribution and growth of loblolly pine. *Forest Science* 10: 105–115.
- Niemelä, J. 1993. Marketing oriented strategy concept and its empirical testing with large sawmills. *Seloste: Markkinointilähtöinen strategiäkäsite ja sen empiirinen testaus suurilla sahoilla*. *Acta Forestalia Fennica* 240. 102 p.
- Nyysönen, A. 1954. Metsikön kuutiomäärän arviointi relaskoopin avulla. Summary: estimation of stand volume by means of the relascope. *Communicationes Instituti Forestalis Fenniae* 44(6). 31 p.
- 1963. The relascope in the determination of thinning needs. *Journal of Forestry* 61(10): 759–760.
- & Kilkki, P. 1965. Sampling a stand in forest survey. *Acta Forestalia Fennica* 79(4). 20 p.
- & Kilkki, P. 1966. Estimation of strata areas in forest survey. *Acta Forestalia Fennica* 81(3). 29 p.
- , Kilkki, P. & Mikkola, E. 1967. On the precision of some methods of forest inventory. *Seloste: Eräiden metsänarvioimismenetelmien tarkkuudesta*. *Acta Forestalia Fennica* 81(4). 60 p.
- , Roiko-Jokela, P. & Kilkki, P. 1971. Studies on improvement of the efficiency of systematic sampling in forest inventory. *Acta Forestalia Fennica* 116. 26 p.
- & Vuokila, Y. 1963. The effect of stratification on the number of sample plots of different sizes. *Selostus: Metsiköiden luokituksen vaikutus erisuuruisten koealojen lukumäärään*. *Acta Forestalia Fennica* 75(2). 17 p.
- Näslund, M. 1936. Skogsförsöksanstaltens gallingsförsök i tallskog. *Meddelanden från Statens Skogsförsöksanstalt* 28(1).
- O'Regan, W.G. & Arvanitis, L.G. 1966. Cost-effect-

- tiveness in forest sampling. *Forest Science* 12(4): 406–414.
- Orver, M. 1970. Klassifier av tallsågtimmer med objektiva mätbara faktorer. Summary: Grading of Scots pine saw timber using objective measurable factors. The Swedish University of Agricultural Sciences. Department of Forest Products. Report 66: 1–54.
- Payendah, B. & Paine, D.P. 1971. Relative precision of two dimensional systematic sampling as function of nonrandomness index. *Canadian Journal of Forest Research* 1: 167–173.
- Penttinen, A., Stoyan, D. & Henttonen, H. 1992. Marked point processes in forest statistic. *Forest Science* 38(4): 806–824.
- Petterson, H. 1955. Barrskogens volymproduktion. Die Massenproduktion des Nadelwaldes. Meddelanden från Statens Skogsforskningsinstitut. Band 45(1): 1–189.
- Pohjoismainen sahatavara. Mänty- ja kuusisahatavaran lajitteluohjeet. 1994. Suomen Sahateollisuusmiesten Yhdistys. Helsinki. 64 p.
- Prodan, M. 1958. Untersuchungen über die Durchführung von repräsentativaufnahmen. *Allgemeine Forst- und Jagdzeitung* 129: 15–33.
- Pukkala, T. 1990. Within-stand variation in forest management planning. *Scandinavian Journal of Forest Research* 5: 263–275.
- Päivinen, R. 1980. Puiden läpimittajakauman estimointi ja siihen perustuva puustotunnusten laskenta. Summary: On the estimation of the stem-diameter distribution and stand characteristics. *Folia Forestalia* 442. 25 p.
- , Nousiainen, M. & Korhonen, K.T. 1992. Puutunnusten mittaamisen luotettavuus. Summary: Accuracy of certain tree measurements. *Folia Forestalia* 787. 18 p.
- Reynolds, M.R., Burk, T. E. & Huang, W.-C. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. *Forest Science* 34(2): 373–399.
- Rinn, F. 1994. One minute pole inspection with RESISTOGRAPH-micro drillings. Proc. International Conference on wood poles and piles. Fort Collins, Colorado, USA.
- Samson, M. 1993a. Modelling of knots in logs. *Wood Science and Technology* 27: 429–437.
- 1993b. Method for assessing the effect of knots in the conversion of logs into structural lumber. *Wood and Fibre Science* 25(3): 298–304.
- SAS 1989. SAS/STAT user's guide. Version 6. Volume 2. SAS Institute, Cary, North Carolina; USA. 846 p.
- Schabenberger, O. 1995. The use of ordinal response methodology in forestry. *Forest Science* 41(2): 321–336.
- Screuder, H.T., Li, H.G. & Hazard, J.W. 1987. PPS and random sampling estimation using some regression and ratio estimators for underlying linear and curvilinear models. *Forest Science* 33(4): 997–1009.
- & Ouang, Z. 1992. Optimal sampling for weighted linear estimation. *Canadian Journal of Forest Research* 22: 239–247.
- Searle, S.R. 1971. *Linear models*. Wiley, New York. 532 p.
- Shiue, C.-J. 1960. Systematic sampling with multiple random starts. *Forest Science* 6(1): 42–50.
- & John, H.H. 1962. A proposed sampling design for extensive forest inventory: Double systematic sampling for regression with multiple random starts. *Journal of Forestry* 60(9): 607–610.
- Shiver, B.D. 1988. Sample sizes and estimation methods for the Weibull distribution for unthinned slash pine plantation diameter distributions. *Forest Science* 34(3): 809–814.
- Silverman, B. W. 1986. *Density estimation for statistics and data analysis*. Chapman and Hall. 184 p.
- Stayton, C.L., Marden, R.M. & Gammon, G.L. 1971. Predicting lumber grade yields for standing hardwood trees. USDA Forest Service. Research Paper NC-50: 1–8.
- Strand, L. 1957. Virkningen av flatestørrelsen på nøjaktigheten ved prøveflatetakster. Summary: The effect of the plot size on the accuracy of forest surveys. *Meddelelser fra det norske skokforsøksvesen* 14: 621–633.
- 1964. Determination of volume by means of the relascope. *Forest Science* 10(4): 410–414.
- Swindel, B.F., Smith, H.D. & Grosenbaugh, L.R. 1987. Fitting diameter distributions with a Hand-held, programmable calculator. *Scandinavian Journal of Forest Research* 2: 325–334.
- Tomppo, E. 1986. Models and methods for analysing spatial patterns of trees. *Communicationes Instituti Forestalis Fenniae* 138. 65 p.
- Turkia, K. & Kellomäki, S. 1987. Kasvupaikan viljavuuden ja puuston tiheyden vaikutus nuorten mäntyjen oksien läpimittaan. *Folia Forestalia* 705. 16 p.
- Unterdörfer, C. 1955. Die bestandesstammzahl nach der Winkelzählprobe. *Allgemeine Forst- und Jagdzeitung* 126(8/9): 176–179.

- Uusitalo, J. 1989. Sahatavaran arvonnustaminen mäntytukkileimikosta. Thesis for M.Sc. graduation. University of Helsinki. Department of Business Economics of Forestry. 45 p.
- 1994a. Sahatavaran laadun ennustaminen mäntytukkirungoista. Summary: Predicting lumber quality of pine stems. University of Helsinki. Department of Forest Resource Management. Publications 3: 1–53.
- 1994b. Using logistic regression for predicting lumber quality of pine stems. Proc. Decision Support – 2001, combined events of the 17th annual Geographic Information Seminar and the Resource Technology '94 Symposium, Toronto, Ontario, September 12–16, 1994. p. 628–633.
- 1995. Coexistence between the needs of forest owners, wood procurement and biodiversity of forest. Proc. Multiple use and environmental values in forest planning. Tohmajärvi, Finland, June 5–10, 1995. *EFI Proceedings* 4. p. 177–183.
- & Kivinen V.-P. 1994. Leimikon ennakkomittauksen ajanmenekki. *Folia Forestalia* 1994(2): 123–139.
- Uusvaara, O. 1974. Wood quality in plantation-grown Scots pine. Lyhennelmä: Puun laadusta viljelymänniköissä. *Communicationes Instituti Forestalis Fenniae* 80(2). 105 p.
- Varmola, M. 1982. Taimikko ja riukuvaiheen männikön kehitys harvennuksen jälkeen. Summary: Development of Scots pine stands at the sapling and pole stages after thinning. *Folia Forestalia* 524. 27 p.
- Vientisahatavaran lajitteluohjeet 1979. Suomen Sahateollisuusmiesten Yhdistys. Helsinki. 52 p.
- Vuokila, Y. 1982. Metsien teknisen laadun kehittäminen. Summary: The improvement of technical quality of forests. *Folia Forestalia* 523. 55 p.
- Vuoristo, I. 1936. Tutkimuksia sahatukien arvostusteista. I. Tukkien latvaläpimitan vaikutus sahaus-tuloksen myyntihintaan. Referat: Untersuchungen über Werthhältnisse der Sägebläcke. Einwirkung des Zopfdurchmessers des Sägebläcke auf die Verkaufspreise des Sägeergebnisses. *Communicationes Instituti Forestalis Fenniae* 23(3). 152 p.
- Wensel, L.C. & John, H. 1969. A statistical procedure for combining different types of sampling units in a forest inventory. *Forest Science* 15(3): 307–317.
- Weslien, H. 1983. Värdeklassificering av sågtimmer med objektiva mätbara faktorer. Del 1. Klassificering av talltimmer. Summary: Value grading of saw timber using objectively measurable factors. Part 1. Grading of pine timber. The Swedish University of Agricultural Sciences. Department of Forest Products. Report 140: 1–16.
- Wiant, H.V.Jr. & Yandle, D.O. 1980. Optimum plot size for cruising sawtimber in eastern forests. *Journal of Forestry* 78(10): 642–643.
- Yaussy, D. A. 1986. Green lumber grade yields from factory grade logs of three oak species. *Forest Products Journal* 36(5): 53–56.
- Zarnoch, S.J. & Dell, T.R. 1985. An evaluation of percentile and maximum likelihood estimations of Weibull parameters. *Forest Science* 31(1): 260–268.
- Zeide, B. 1980. Plot size optimization. *Forest Science* 26(2): 251–257.
- Zöhler, F. 1969. Ausgleich von Häufigkeitsverteilungen mit Hilfe des beta funktion. *Forstarchiv*. p. 37–42.
- 1978. Fundamentale stichprobenkonzepte der forst-iventur I. *Allgemeine forstzeitung* 89(5): 180–185.
- 1979. Fundamentale stichprobenkonzepte der forst-iventur II. *Allgemeine forstzeitung* 90(5): 112–115.

Total of 129 references

**Appendix 1.** Diameter class-specific cant and re-saw sawing patterns used in test sawing in the circular sawmill Hietaniemen Saha Ky located in Viljakkala, southern Finland.

Diameter class mm	Cant saw sawing pattern mm	Re-saw sawing pattern mm
120–129	19-75-19	19-38-38-19
130–149	19-75-19	19-38-38-38-19
150–169	19-19-100-19-19	19-19-50-50-19-19
170–189	19-19-125-19-19	19-19-50-50-19-19
190–209	19-19-150-19-19	19-19-50-50-19-19
210–229	19-19-150-19-19	19-19-19-50-50-19-19
230–249	19-19-150-19-19	19-19-75-75-19-19
250–269	19-19-175-19-19	19-19-75-75-19-19
270–289	19-19-200-19-19	19-19-19-75-75-19-19-19
290–309	19-19-19-200-19-19-19	19-19-50-50-50-50-19-19
310–	19-19-19-225-19-19-19	19-19-19-50-50-50-50-19-19-19

**Appendix 2.** Time consumption of each work element used in this study. For a more comprehensive examination see Uusitalo and Kivinen (1994).

Work element	Time consumption of work element, (min)	
	<i>Defining the sampling point by moving along square grid</i>	<i>Walking along a subjectively chosen route</i>
Moving, $T_m$	$T_m = 6.18 + 8.64 A$	$T_m = 0.540 + 4.170 A$
Walking speed (S)	$S = 35 \text{ m/min}$	$S = 73 \text{ m/min}$
Walking distance (L)	$L = 215.7 + 301.7 A$	$L = -100.9 + 275.3 A + 59.04 F$
	<i>Choosing sample trees with telescope pole</i>	<i>Choosing the nearest three/five trees as sample trees</i>
Establishment of sample plots, $T_c$	$T_c = (0.402 + 0.00160 D) N_s$	$T_c = 0.05 N_s$
	<i>Measurement of DBHs in circular plot</i>	<i>Measurement of DBHs of nearest three/five trees</i>
Measurement of DBH with diameter caliper, $T_d$	$T_d = 0.167 N$	$T_d = 0.117 N$
	<i>Measurement with telescope pole</i>	<i>Ocular estimation</i>
Measurement of the quality parameters, $T_q$	$T_q = 0.267 N_q$	$T_q = 0.150 N_q$
	<i>Measurement with Suunto-hypsometer</i>	<i>Ocular estimation</i>
Measurement of tree height, $T_h$	$T_h = 1.284 N_h$	$T_h = 0.117 N_h$
Recording of results on a form, $T_r$	$T_r = 0.0833 (N + N_q + N_h + N_b)$	
Measurement of basal area with relascope, $T_b$	$T_b = 0.667 N_b$	
Delays, $T_{del}$	2 %	1 %

**Submission of Manuscripts** Manuscripts should be submitted in triplicate to Acta Forestalia Fennica, Unioninkatu 40 A, FIN-00170 Helsinki, Finland. Detailed instructions to authors have last been printed in Acta Forestalia Fennica 244, are available from the editorial office, and can be found on Acta's WWW pages at <http://www.metla.fi/publish/acta/>

**Publication Schedule** Acta Forestalia Fennica is published intermittently, 3 to 6 numbers per year.

**Subscriptions and Exchange** Subscriptions and orders for back issues should be addressed to Academic Bookstore, Subscription Services, P.O. Box 108, FIN-00101 Helsinki, Finland, Phone +358 9 121 4322, Fax +358 9 121 4435. Subscription price for 1998 is 70 FIM per issue. Exchange inquiries should be addressed to the Finnish Society of Forest Science, Unioninkatu 40 A, FIN-00170 Helsinki, Finland, Phone +358 9 8570 5235, Fax +358 9 8570 5 677, E-mail [sms@helsinki.fi](mailto:sms@helsinki.fi)

**Statement of Publishers** Acta Forestalia Fennica has been published since 1913 by the Finnish Society of Forest Science. In 1989 Acta Forestalia Fennica was merged with Communicationes Instituti Forestalis Fenniae, started in 1919 by the Finnish Forest Research Institute. In the merger, the Society and Forest Research Institute became co-publishers of Acta Forestalia Fennica. The Finnish Society of Forest Science is a nonprofit organization founded in 1909 to promote forest research. The Finnish Forest Research Institute, founded in 1917, is a research organization financed by the Ministry of Agriculture and Forestry.

**Abstracting** Articles in Acta Forestalia Fennica are abstracted and indexed in Agrindex, Biological Abstracts, Current Advances in Ecological Sciences, Current Advances in Plant Sciences, Ecological Abstracts, Forest Products Journal, Forestry Abstracts, International Bibliography of Periodical Literature, Life Sciences Collection.

# ACTA FORESTALIA FENNICA

The Finnish Society of Forest Science  
The Finnish Forest Research Institute

---

- 251 **Jyrki Kangas, Teppo Loikkanen, Timo Pukkala and Jouni Pykäläinen:** A participatory approach to tactical forest planning.
- 252 **Pekka Ripatti:** Factors affecting partitioning of private forest holdings in Finland. A logit analysis.
- 253 **Vesa Kaarakka:** Management of bushland vegetation using rainwater harvesting in eastern Kenya.
- 254 **Pertti Hari, Juhan Ross and Marja Mecke (eds.):** Production process of Scots pine; geographical variation and models.
- 255 **Euan G. Mason and A. Graham D. Whyte:** Modelling initial survival and growth of radiata pine in New Zealand.
- 256 **Juha Nurmi:** Heating values of mature trees.
- 257 **Timo Hartikainen:** Late-industrial sawmill customers. The concept of industrial operating mode and its testing in secondary wood processing industry.
- 258 **Jari Varjo:** Change detection and controlling forest information using multi-temporal Landsat TM imagery.
- 259 **Jori Uusitalo:** Pre-harvest measurement of pine stands for sawing production planning.
- 260 **Wending Huang:** Productive coexistence and gain in agroforestry systems.
- 

ISBN 951-40-1599-1  
ISSN 0001-5636

