Dissertationes Forestales 158

Modelling tree biomasses in Finland

Jaakko Repola

Department of Forest Sciences
Faculty of Agriculture and Forestry
University of Helsinki

Academic dissertation

To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public criticism in Auditorium XIV at the University Main Building, Unioninkatu 34, Helsinki, on May 17th, 2013 at 12 o'clock noon.

Title of dissertation: Modelling tree biomasses in Finland
Author: Jaakko Repola

Dissertationes Forestales 158

Supervisor:
Dr. Risto Ojansuu
Finnish Forest Research Institute (Metla), Vantaa, Finland

Pre-Examiners:
Dr. Hans Petersson
Department of Forest Resource Management, Swedish University of Agricultural Sciences, Sweden

Prof. Matti Maltamo
School of Forest Sciences, University of Eastern Finland, Finland

Opponent:
Docent Lauri Mehtätalo
School of Computing, University of Eastern Finland, Finland

ISSN 2323-9220 (Print)
ISBN 978-951-651-404-1 (paperback)

ISSN 1795-7389 (Online)
ISBN 978-951-651-405-8 (PDF)

2013

Publishers:
Finnish Society of Forest Science
Finnish Forest Research Institute
Faculty of Agriculture and Forestry of the University of Helsinki
School of Forest Sciences of the University of Eastern Finland

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

Repola, J. 2013. Modelling tree biomasses in Finland. Dissertationes Forestales 158. 40 p. Available at http://www.metla.fi/dissertationes/df158.htm

ABSTRACT

Biomass equations for above- and below-ground tree components of Scots pine (*Pinus sylvestris* L), Norway spruce (*Picea abies* [L.] Karst) and birch (*Betula pendula* Roth and *Betula pubescens* Ehrh.) were compiled using empirical material from a total of 102 stands. These stands (44 Scots pine, 34 Norway spruce and 24 birch stands) were located mainly on mineral soil sites representing a large part of Finland. The biomass models were based on data measured from 1648 sample trees, comprising 908 pine, 613 spruce and 127 birch trees. Biomass equations were derived for the total above-ground biomass and for the individual tree components: stem wood, stem bark, living and dead branches, needles, the stump, and roots, as dependent variables. Three multivariate models with different numbers of independent variables for above-ground biomass and one for below-ground biomass were constructed. Variables that are normally measured in forest inventories were used as independent variables. The simplest model formulations, multivariate models (1) were mainly based on tree diameter and height as independent variables. In more elaborated multivariate models, (2) and (3), additional commonly measured tree variables such age, crown length, bark thickness and radial growth rate were added.

Tree biomass modelling includes consecutive phases, which cause unreliability in the prediction of biomass. First, biomasses of sample trees should be determined reliably to decrease the statistical errors caused by sub-sampling. In this study, methods to improve the accuracy of stem biomass estimates of the sample trees were developed. In addition, the reliability of the method applied to estimate sample-tree crown biomass was tested, and no systematic error was detected. Second, the whole information content of data should be utilized in order to achieve reliable parameter estimates and applicable and flexible model structure. In the modelling approach, the basic assumption was that the biomasses of the tree components on the same site and in the same tree are dependent. This statistical dependency was taken into account when simultaneously estimating parameter estimates for all biomass components, by applying a multivariate procedure. Based on the verified statistical dependence between the biomass components, the multivariate procedure had a number of advantages compared to the traditionally independently estimated equations, by enabling more flexible application of the equations, ensuring better biomass additivity, and giving the more reliable parameter estimates.

The generalization and applicability of the models may be restricted by the fact that the study material was not an objective, representative sample, and some tree components were poorly represented. Despite these shortcomings, the models provided logical biomass predictions for individual tree components and were comparable with other functions used in Finland and Sweden.

Keywords: Tree biomass, biomass equations, pine, spruce, birch

ACKNOWLEDGEMENTS

This work has been carried out in the Finnish Forest Research Institute (Metla), started in Vantaa Unit and finished off in Rovaniemi Unit. I thank Metla for financially supporting and providing me with good working facilities. The work was also funded by Academy of Finland of which I acknowledged.

I am indebted to my advisor Dr. Risto Ojansuu for support and invaluable guidance especially for statistical expertise throughout the study. I am grateful to Mikko Kukkola for providing the study material and helping me patiently in data management. I am very grateful for constructive comments on my work and support to Hannu Hökkä and Hannu Salminen. I am grateful to Dr. Pentti Hakkila and Dr. Pekka Saranpää and Dr. Risto Sievänen for supporting providing the study material of paper I. Special thanks to Dr. Jari Hynynen who has eased my tasks and given me the opportunity to concen¬trate on this thesis for a while. I want to also thank my co-author Anu Kantola. The authors would like to thank Pekka Välikangas, Pasi Aatsinki, Jouni Väisänen, Jukka Lahti and other assistants at the Salla Experimental Office for their laborious field and laboratory work. I was also give thanks to the personnel of Rovaniemi and Vantaa Units for co-operation.

I would like to sincerely thank Professor Annika Kangas for serving as the custos, and for valuable comments on my work and the encouraging support in my post-graduate studies. I am grateful to Dr. Hans Petersson and Prof. Matti Maltamo who reviewed the thesis. I want also thank to Dr. Timothy G. Gregoire, Dr. Lauri Mehtätalo and anonymous referees made who made valuable comments on the manuscripts of this thesis. I am grateful to Niemi Foundation for a grant I received to finish of my thesis.

Finally, but not the least, I want to thank my wife Jaana and children Eero, Aino-Kaisa and Elina for giving so much love and meaning to my life.

Rovaniemi, April 2013

Jaakko Repola

LIST OF ORIGINAL ARTICLES

This thesis is based on the original papers listed below, which are referred to in the text by their Roman numerals. These papers **I, III** and **IV** are reprinted with the kind permission of the publishers, while the study **II** is the author version of the submitted manuscript.

I Repola, J. 2006. Models for Vertical wood Density of Scots pine, Norway spruce and Birch stems, and Their Application to Determine Average Wood Density. *Silva Fennica* 40(4): 673–685.
http://www.metla.fi/silvafennica/full/sf40/sf404673.pdf

II Repola, J., Ojansuu, R. & Kantola, A. 2013. Design-based and Model-based estimators for Norway spruce (*Picea abies* L. Karst.) crown biomass. Submitted manuscript.

III Repola, J. 2008. Biomass equations for birch in Finland. *Silva Fennica* 42(4): 605–624.
http://www.metla.fi/silvafennica/full/sf42/sf424605.pdf

IV Repola, J. 2009. Biomass equations for Scots pine and Norway spruce in Finland. *Silva Fennica* 43(4): 625–647.
http://www.metla.fi/silvafennica/full/sf43/sf434625.pdf

Jaakko Repola was the person responsible for data analysis and writing of all four articles. Risto Ojansuu and Anu Kantola participated in the planning and working of paper II ("Design-based and Model-based estimators for Norway spruce (*Picea abies* L. Karst.) crown biomass"). Risto Ojansuu participated in the specification of the regression model, and commented on paper II. Anu Kantola was responsible for the field works and participated in writing of Material of paper II.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Background

Interest in estimating tree and stand biomass has increased during the last decades. Currently, reliable estimates for the biomass of a tree and its components are needed for practical forestry as well as for research purposes. Tree and stand biomass estimates are a prerequisite for, e.g., the assessment of nutrient cycling and fluxes, energy wood potentials, and carbon storage of forests.

The importance of the estimation of forest biomass has increased since the Kyoto Protocol was adopted in 1997. Based on this protocol, the objective of the climate politics is to reduce the emissions of greenhouse gases such as carbon dioxide (CO_2). Countries that have ratified the Kyoto Protocol are obligated to report emissions and removals of greenhouse gases to the United Nations Framework Convention on Climate Change (UNFCCC). Forest carbon sinks were included in the Kyoto Protocol as one of the mechanisms for mitigating climate change. Forest carbon sinks are known to play an important role in national and global greenhouse gas balances. The assessment of the forest carbon stock, in turn, is based on estimates of above- and below-ground tree biomass.

The aim of reducing global emissions of greenhouse gases has led to an increased desire to use renewable energy resources in energy production. Forest biomass (energy wood) has played an important role in increasing the use of renewable and carbon neutral energy resources. In Finland, an increasing amount of energy wood has been used in energy production during the last decades (Finnish Statistical Yearbook of Forestry 2011). Estimates of the biomass of tree components are required for assessing the energy-wood potential from different sources at national or regional scales. Biomass estimates of the crown (branches and foliage), the unmerchantable part of a stem, the stump and roots are needed when assessing the amount of logging residues in final cuttings. In addition, stem biomass and the whole above-ground biomass are needed when estimating the energy-wood availability from young stands.

1.2 Methods to estimate and predict tree biomass

1.2.1 Overview of the methods

The biomass of a whole tree or its components can be expressed as dry mass, fresh mass or volume. Dry mass, determined as the mass of dried organic matter, has mostly been used as an expression of tree biomass. Dry mass is also the most appropriate measure for the determination of forest carbon sinks and the energy content of tree biomass, because around fifty percent of the dry mass constitutes carbon, and due to the rather strong correlation between dry mass and energy content. In this study, tree biomass refers to the dry mass.

Biomass estimates of a tree or its components can be obtained using different methods. These methods can be divided into two approaches; biomass estimation and biomass prediction. Biomass estimation includes the process of determining tree biomass. Biomass estimation is commonly defined as the process of determining tree biomass by sub-sampling (Parresol 1999). But also biomass determination by direct measurement can be included in this approach, which is also called the measurement of biomass (Parresol 1999). In turn, biomass prediction can be defined as the process of obtaining tree biomass by utilizing measured tree dimensions and compiled biomass models, biomass tables, or expansion factors. In this study biomass estimation refers to the process of determining tree biomass by sub-sampling, followed by biomass prediction as defined above.

1.2.2 Biomass estimation with measurements and sub-sampling

Biomass estimates at regional and stand scales are commonly based on single-tree biomass estimates. Several approaches to determine the biomass of a tree or its components have been applied. The most comprehensive approach is a direct measurement based on the dry or fresh weight of a whole tree or an entire tree component. However, such a direct measurement of tree fresh or dry weight is in practice too expensive and time consuming, especially for individual tree components and for large trees (See Briggs et al. 1987). In biomass studies, determination of the biomass of tree components or total tree biomass is commonly based on sub-sampling (Parresol 1999, 2001). In sub-sampling, small samples are selected from tree components by a specific procedure (usually a random procedure) and measurements (fresh and/or dry weight) of the samples are then used for estimating the biomass of the entire tree component. Different statistical estimators, such as a design-based or a model-based estimator, have been applied in estimating the biomass of entire components (Briggs et al. 1987, Monserud and Marshall 1999, Parresol 1999).

The theoretical approaches of a design-based and a model-based method differ substantially. In the design-based method, the population is regarded as fixed, whereas the sample is regarded as a realisation of a random process. The reference distribution is a consequence of all possible samples under the sampling design. The inference is based on the distribution of estimates generated by the sampling design. Therefore, the sampling design is crucial for inference, and the inference is independent of any assumption about population structure and distribution (Gregoire 1998). In the model-based methods, the population is regarded as a realisation of a stochastic process, and values generated by the sampling are realisations of random variables. In model-based estimation, the characteristics of a population are described with a model, and inferences about the population depend on the assumed models, not on sampling design, as in design-based methods (Kangas 1994, Gregoire 1998).

Design-based estimators, such as ratio estimators, are commonly used for biomass estimation of all tree components (e.g. Marklund 1988, Hakkila 1991, Monserud and Marshall 1999, Claesson et al. 2001). The strategy is to measure the total fresh weight of a tree component in the field. Some samples are selected and the fresh and dry weight of the samples is measured. The ratio of the dry and fresh weight of samples is then used to estimate the biomass (dry weight) of the entire tree component. The other strategy is to first measure (or determine) the volume and average density of a tree component, and then biomass can be calculated by multiplying the volume by the average density. This approach has mainly been applied only for determining stem biomass (wood and bark). Stem volume can be reliably determined by applying tree dimensions and volume functions, but the determination of average stem density is much more troublesome for many reasons. Firstly, measurements of stem density are laborious and, secondly, many factors (tree species, environmental factors, tree age and size, growth rate and genetic factors) cause variation in wood density. Stem density varies in the radial and vertical directions of the stem according to a species-specific pattern (Tamminen 1962, Hakkila 1966, Knigge and Shultz 1966, Uusvaara 1974, Hakkila 1979, Björklund 1984). For tree species with a high vertical dependence of wood density in particular, an inappropriate sampling design may lead to biased estimates of average wood density and subsequent biomass estimates.

Regression is commonly applied in the model-based framework, especially for estimating tree crown biomass. The regression method is based on easily measured basic variables, such branch diameter, and their vertical position along the stem of all the branches of a tree. Some samples of the tree component are taken for dry-weight measurements. The samples are used to model dry weight as a function of the basic variables. The regression model is then used to determine the dry weights of the whole-tree components, e.g. the sum of the predicted biomass of all branches is the total crown biomass of a tree. The reliability and applicability of such an equation depends on how well the basic assumptions of the model are met, i.e. how efficiently the information and structure of the study material is utilised in the model estimation.

1.2.3 Biomass expansion factors and biomass tables

Sampling or direct measurement have commonly been applied in biomass studies, e.g., when collecting biomass data for constructing biomass equations. In practice, biomass estimates of a tree or its components are obtained by using biomass expansion factors, weigh tables, or regression models. Biomass expansion factors (BEFs) are used at tree and stand level to convert the stem volume into whole tree biomass or the biomasses of different tree components. In general, constant BEFs have been applied, although it is known that BEFs may vary depending on growth conditions and the phase of stand development (Satoo and Madgwick 1982, Hakkila 1991, Lehtonen et al. 2004). The BEFs are easy to apply because they need only stem volume as an input variable. The problem with using BEFs is that they produce only coarse biomass estimates, and may in the worst case lead to biased estimates (Kärkkäinen 2005). Biomass tables are based on one or more tree dimensions, such as diameter, height and stem taper (Baskerville 1965, Hakkila 1979). Normally, neither estimates of biomass based on BEFs nor on biomass tables take between-tree variation into account.

1.2.4 Regression models

Nowadays biomass estimates of a tree or its components are commonly obtained with regression models. The biomass models predict biomass as a function of easily measurable tree dimensions such as diameter and height. Biomass regression models are normally constructed for individual tree components such as stem, stem bark, crown (branches and foliage), stump and roots.

Several biomass models have been published in different countries since Kittrede (1944) applied tree allometry in the study of biomass. In the Nordic countries many studies on tree biomass, especially on above-ground tree components, have been published, but only a few of the functions are widely applicable and include all of the main tree components. Marklund's (1988) biomass functions, most widely applied in Scandinavia, are valid for predictions of different above-ground components of pine, spruce and birch. These functions are based on a large and representative material from the Swedish national forest inventory. In Finland there has been a lack of widely applicable (general) individual-tree biomass models, but according to Kärkkäinen (2005), Marklund's (1988) functions can be applied also in Finland. Kärkkäinen (2005) concluded also that Marklund's (1988) functions are primarily applicable for the calculation of biomass estimates at a large scale, and more applicable for the estimation of carbon sequestration than for the estimation of energy-wood resources. These conclusions (Kärkkäinen 2005) were based on the evaluation and comparison of tree-level biomass models, but not on empirical biomass data. In Finland, Hakkila's (1972, 1979 and 1991) functions have been also applied for predicting biomass of above- and below-ground tree components. However, the disadvantage of these functions is that the equations for the main tree components: stem, crown, and stump including roots, were not based on the same sample trees. Repola et al. (2007) published general biomass equations for pine, spruce and birch, in which the biomass of the above-ground and the below-ground tree components are modelled mainly on the basis of the same sample trees. All previously mentioned biomass functions are primarily applicable for trees growing on mineral soil, but according to Kärkkäinen (2005) Marklund's (1988) and Hakkila's (1979, 1991) functions can also be applied for trees growing on peatlands. Compared to equations for above-ground biomass, functions for the below-ground biomass components published in Nordic countries are generally based on a more limited material (Hakkila 1972, Marklund 1988, Finer 1991, Petersson and Ståhl 2006, Repola et al. 2007), which restricts their applicability in practice.

Biomass models should meet specific requirements before they can be incorporated into forest management planning systems or applied on large forest areas, e.g., when assessing forest carbon pool and energy-wood potentials at the national scale (Kärkkäinen 2005). First of all, at application the models should provide reliable biomass estimates of tree components for all growing stock, with a prerequisite that the derived models span a wide diameter range and a wide range of stand and site conditions in the whole country. In order to obtain reliable biomass estimates at the national scale (large scale), the biomass models should be based on a representative sample of the stands in which the results are to be applied (Parresol 1999). A representative sample based an objective sampling, such as national forest inventory data, has generally been considered to be a prerequisite for valid and unbiased models for national-scale application. In addition, a logical model specification throughout the range of the material, and unbiased biomass estimation (determination) of the sample trees are both also prerequisites for reliable biomass estimation. Secondly, the biomass models should be based on variables that are normally measured in forest inventories, or which can be estimated easily and reliably from inventory data. Thirdly, the models of individual tree components should be based on the same sample trees, in order to give a reliable description of the relationships between the tree components. In addition, one desirable feature of the tree-component equations is biomass additivity, which means that the sum of the predictions for the tree components equals the prediction for the whole tree (Kozak 1970 Cunia and Briggs 1984, Parresol 1999, 2001). Information on biomass accumulation in different parts of the stem and crown are needed, e.g., when assessing the amount of energy wood in stands or trees. For energy-wood estimation purposes, the biomass model should also be able to predict the vertical biomass distribution of the tree components along a tree.

1.3 The process of modelling tree biomass with regression analyses

The common procedure for estimating tree biomass is through the use of regression analysis. The process of developing a regression model for tree biomass contains three main phases, all of which have an effect on the reliability of the models. In the first phase, a number of stands are selected, from which sample trees for biomass determination are chosen through an appropriate selection procedure. (e.g. by using random selection). Secondly, biomass of tree components is determined on the basis of within-tree sampling (Parresol 1999, 2001). Finally, for biomass regression models, quantitative relations between tree dimensions and tree biomass are constructed.

The reliability of biomass prediction depends on the study material and also on how efficiently the study material is utilized in the model estimation. Ideal modelling data would be a representative sample of the forests in the region where the models will be applied. But at the same time, the data should include a wide range of site conditions, in order to have a high variation of independent variables. This is a prerequisite for a reliable regression model, i.e., the higher the variation in independent variables, the higher the reliability of parameter estimates (lower standard errors) usually obtained in the model estimation (Lappi 1993).

The dependent variable, biomass of a tree component of sample trees, is usually estimated on the basis of within-tree sampling (Parresol 1999). Small samples are selected from the tree components by a specific procedure and information about the samples is then used to estimate the biomass of the entire tree component. This process, determination of the biomass by sub-sampling, produces an error in the biomass estimate (Parresol 2001), and it is therefore important to apply sub-sampling methods that minimize such errors. Biomass estimation of sample trees should be as reliable as possible, because errors in biomass estimation can cause problems also in the subsequent analysis (e.g., when constructing biomass models) or at application (when applying biomass models to predict biomass values). As long as the error in a biomass estimate can be interpreted as random, uncorrelated with independent variables, it does not cause a problem with respect to the assumption of a linear model (Parresol 2001). However, this source of error increases random errors of the model (Parresol 2001). If the errors are correlated with independent variables, it may produce a trend of increasing variance, and the assumption of homoscedastic variance is not valid. The ignored heteroscedasticity, in turn, can generate bias in parameter estimates and in the reliability of the parameter estimates (Parresol 1999, 2001).

In the model specification different facts, such as unbiased fixed effects, the correlation structure of data, and biomass additivity, should be taken into consideration in order to have an applicable and reliable biomass model. To have reliable parameter estimates and thereby biomass predictions, the model specification should be correct and the data structure should be considered when choosing appropriate correlation structure in model estimation. Biomass equations have generally been constructed by applying linear or nonlinear model forms (Parresol 1999). In many biomass studies, a nonlinear model form with multiplicative error has been used as the basis of the model formulation. In this form, a logarithmic transformation has commonly been used to obtain homoscedastic variance, and to transform the equation into a linear form.

In order to obtain efficient parameter estimates for constructed biomass models, it is important to correctly address the correlation structure of the data. Both spatial and temporal correlation is known to be common in all types of forestry data, including biomass data. Modelling these correlations provides considerable gains in the efficiency of estimation (Parresol 1999, Gregoire et al. 1995). Spatial correlation (e.g. hierarchical data) exists when biomasses of trees are more strongly correlated within stands than between stands. Contemporaneous correlation means that biomasses of tree components of the same tree or the same stand are correlated (Parresol 1999, 2001). Contemporaneous correlation also means that the errors in the different equations of tree components can not be assumed to be independent. Hierarchically structured data and contemporaneous correlation have commonly been ignored in model estimation, which has meant a loss of precision in the parameter estimates. If ordinary least squares (OLS) is applied to hierarchically structured data, parameter estimates are unbiased, but standard errors of these estimates are commonly underestimated (Lappi 1993). Hierarchically structured data can be analysed more precisely, e.g., by the generalized least squares (GLS) estimation method, which permits analysis of the between-stand and within-stand variation (Claesson et al. 2001, Repola et al. 2007). Contemporaneous correlation can be taken into account by using a multivariate procedure, i.e. by constructing a set of linear or non-linear models, the parameters of which are estimated simultaneously (linear or non-linear seemingly unrelated regression) (Zellner 1962, Srivastava and Giles 1987, Parresol 1999 and 2001, Carvalho et al. 2003, Bi et al. 2004, Návar et al. 2004). Advantages of utilizing the multivariate procedure instead of independently estimated equations have been found. Parresol (1999 and 2001) concluded that, by utilizing contemporaneous correlation in the model estimation, parameter estimates were also more reliable. Lappi (1991) showed that the across-equation correlation in hierarchical data could be utilized to flexibly calibrate the model to a new stand. The across-equation covariance is also needed when calculating the prediction reliability for any combination of tree components. This information is not available for independently estimated equations.

At application, a desirable feature in the equations of tree components is that the sum of the predicted individual tree components equals the prediction of whole tree biomass (Kozak 1970 Cunia and Briggs 1984, Parresol 1999, 2001). The multivariate procedure (seemingly unrelated regression), has also been used to ensure biomass additivity by estimating across-equation correlation at the tree level, and by setting linear restrictions, i.e., across-equation constraints to the regression coefficients (Parresol 1991 and 2001, Carvalho et al. 2003, Bi et al. 2004, Návar et al. 2004). However, in many studies the equations for the total above-ground biomass have not been formulated. In those cases, the sum of individual tree-component models might produce biased estimates for whole tree biomass.

1.4 Research aims

The first and principal objective of this study is to develop tools that can be utilized for estimating and predicting tree biomass. This is achieved by compiling individual-tree biomass equations for above- and below-ground tree components of Scots pine, Norway spruce and birch, which produce reliable biomass predictions over a wide range of stand and site conditions in Finland, and which can be utilized flexibly in practical forestry and research. The second aim is to develop methods to determine the biomass of sample trees, and to evaluate the reliability of tree biomass estimates. This is important because a prerequisite for a reliable biomass model is that the error of a response variable (biomass of a tree component) caused by biomass estimation is minor and that no systematic error exists, i.e. biomass estimation of sample trees should be as reliable as possible.

In paper I, models were constructed for the vertical dependence of the basic density of Scots pine, Norway spruce, and birch stems. These models can be utilized in determining the biomass of the whole stem, a stem section, or biomass distribution within the stem. The purpose is to construct a model that can be flexibly calibrated to new trees or stands by utilizing measurements of wood density and information on between-tree variation. The linear mixed model technique with both fixed and random parameters was used in the model estimation.

The purpose of paper II was to compare methods for estimating the living crown biomass of individual Norway spruce trees by applying both design-based and model-based estimators. The aim was also to assess the reliability of these estimates at tree level, i.e., assess the amount of statistical error generated by the process of determination of the crown biomass. In the model-based method, two different approaches of constructing a regression model were compared. In the design-based method, ratio estimation with two different sampling designs, objective and subjective sample branch selection, were compared. Of particular interest was the comparison of the design-based methods with each other, because subjective sampling was applied to biomass data in papers III and IV.

The aim of papers III and IV was to develop biomass equations for above- and below-ground tree components of Scots pine, Norway spruce and birch by applying a statistical method (multivariate procedure) that effectively utilizes the information of the biomass data as well as the information produced by national forest inventories. This enables more flexible application of the biomass models. Another aim of papers III and IV was to study whether the multivariate procedure gives more reliable parameter estimates than the separately (independently) estimated equations published by Repola et al. (2007). In paper IV, the purpose was to study the advantages of applying a multivariate procedure instead of estimating the models independently using data with unbalanced response variables, i.e. when not all the biomass components have been measured on all the sample trees.

2 MATERIAL AND METHODS

2.1 Study materials

The study material consisted of two main data sets: data on stem-wood density (paper I) and biomass data (papers II-IV). Data on wood density were used for modelling the vertical dependence of the basic density of stems. Biomass data were used for modelling the biomass of tree components. The study stands were mainly located on mineral soil sites representing a large part of Finland (Fig. 1). The distribution of forest site types was typical for Scots pine, Norway spruce and birch on mineral soil. The spruce and birch stands were growing on fertile (the mesic *Myrtillus* type) or highly fertile (the fertile *Oxalis-Myrtillus* type) sites, and the pine stands on poor to fertile sites (the dry *Calluna* type, dryish *Vaccinium* type or mesic *Myrtillus* type) (Cajander 1949).

The material for paper I was obtained from sample trees collected during 1993-2000 in connection with the National Bioenergy Research Program, carried out by the Finnish Forest Research Institute. The material consisted of a total of 90 stands, comprising 38 Scots pine (*Pinus sylvestris* L.) stands, 39 Norway spruce (*Picea abies* [L.] Karst.) stands and 13 birch stands (*Betula pendula* Roth and *Betula pubescens* Ehrh.) located in southern Finland (Fig. 1). The study material was composed of trees removed in cuttings from stands at different stages of development; first and second commercial thinning and final cutting. The total number of sample trees was 1365: 585 pine, 585 spruce and 195 birch (A in Table 1).

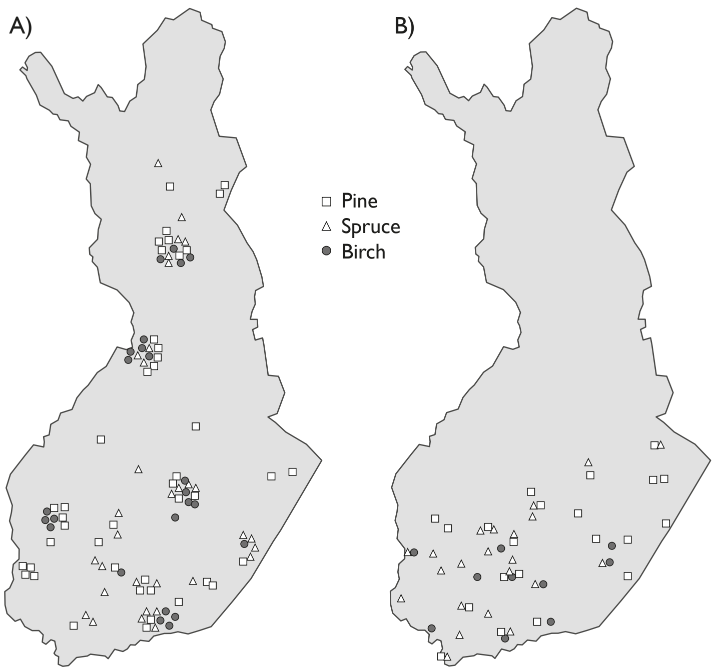

Figure 1. The location of the study sites of biomass studies (A, papers III-IV) and wood density study (B, paper I).

Table 1. Sample tree characteristics in the two data sets A and B.

Data	Number of stands	Number of sample trees	Age, years	d, cm	h, m
Density data (A):					
Pine	38	585	60 (23-139)	17.1 (12.0-37.8)	16.4 (7.2-30.0)
Spruce	39	585	60 (24-153)	18.0 (7.4-47.2)	17.8 (7.9-31.2)
Birch	13	195	30 (16-48)	11.9 (7.0-21.2)	15.6 (10.0-21.8)
Biomass data (B):					
Pine	44	908	56 (11-146)	13.1 (1.5-35.8)	11.2 (2.9-28.6)
Spruce	34	613	52 (15-164)	17.9 (1.7-41.7)	15.9 (2.1-35.0)
Birch	24	127	44 (11-134)	16.5 (2.5-38.0)	17.1 (3.9-29.0)

The empirical material of papers III and IV consisted of a total of 102 stands: 44 Scots pine, 34 Norway spruce and 24 birch stands, representing a large part of Finland (Fig. 1). The average annual effective temperature sum (dd, >5ºC) varied between 705 and 1385 dd. The stands were even-aged, and ranged from young to mature growing stands (B in Table 1). The birch stands were mainly dominated by *Betula pendula* or *Betula pubecens*, with a varying admixture of conifer trees. The study material was gathered between 1983 and 2003.

The whole data set consisted of three sub data sets: 53 temporary sample plots, nine thinning experiments, and the control plots of 39 fertilization experiments. The majority of the study material for pine and spruce was from the fertilization experiments. The temporary plots and thinning experiments were selected to obtain higher variation in independent variables of the modelling data. The temporary plots were established subjectively in five of the Finnish Research Institute's research areas, and the plots were located subjectively in representative parts of the stands. At the establishment the locations of fertilization and thinning experiments were selected subjectively. Hence the study material was not based on objective sampling.

The total number of sample pine, spruce and birch trees was 908, 613 and 127 respectively (B in Table 1). The majority of the sample trees were from the control plots of fertilization experiments. In the thinning experiments the sample trees were taken from unthinned, moderately thinned and heavily thinned plots. The diameter and age distribution of the sample trees was broad (B in Table 1).

For paper II the additional field data were gathered from three (subjectively selected) Norway spruce stands from the material used in paper IV.

2.2 Field and laboratory measurements

The studies (papers I–IV) consisted of field and laboratory measurements. Field work consisted of the measurement of tree attributes (height, diameter and bark thickness at several points along the stem, living crown length, age and diameter increment), fresh weighing, and sampling (stem, crown, stump and roots). In the laboratory, dimensions as well as fresh and dry weigh of the samples were measured for the determination of biomass or wood density.

In order to assess wood density (Paper I), knot-free sample disks were taken at one-meter intervals from the pulpwood and unmerchantable top sections, and from the saw logs at the base and top of the logs. The number of disks per tree varied from 7 to 20. In the laboratory, the disks were separated into wood and bark sections and, after 2–3 days drying at a temperature of 106 ^{o}C, the basic density of the wood and bark was determined.

For the biomass study (papers III–IV) the samples were taken from the stem, crown, stump and roots. Because the proportion and properties of tree components is not constant along the length of a tree, the samples were taken systematically from different parts of a tree. Sample disks of a stem were taken at breast height and at a relative height of 70% for stem biomass determination. For crown biomass determination, the sampling was done by crown sections, since the proportion of needles and branches varies throughout a tree crown. First, the living crown was divided into four sections of equal length, and one living sample branch was selected subjectively from each section. One dead sample branch per tree was taken from the lowest crown section. The fresh weight of all the branches was measured by crown sections in the field. Fresh and dry weights of the sample branches were measured in the laboratory. The stump and root biomasses were measured for a sub-sample of the trees on the temporary plots. The minimum coarse-root diameter varied from 2–5 cm, depending on tree diameter. To facilitate field works, the minimum coarse-root diameter was taken to be 5 cm for trees with a diameter > 20 cm. In addition, the root biomass was determined for roots with a diameter larger than 1 cm for some of the trees. The fresh weight of the stump and roots was measured in the field. One sample (stump sector) was taken from the stump, plus two root discs for moisture content determination.

In addition, for the material of paper II, branch diameter over bark at the stem junction, and distance of the branch whorl from the tree top were measured in three Norway spruce stands. The living crown of the tree was divided into ten equally long sections and one living branch was randomly sampled from each section.

2.3 Methodology

2.3.1 Modelling the vertical dependency of stem-wood density (Paper I)

Data of stem-wood density were used to fit the models for the vertical dependence of the basic density of the stem. The material was hierarchically structured at stand, tree and within-stem levels. Therefore the linear mixed model technique, with fixed and random effects, was used in the model estimation (McCulloch and Searle 2001). In the final model, the stand- and tree-level effects were combined at the tree level because the stand level was not significant. The final model structure was:

$$\hat{\mathbf{y}}_{ik} = \mathbf{x}_{ik}^T \mathbf{b} + \mathbf{z}_{ik}^T \mathbf{u}_k + \mathbf{e}_{ik} \quad (1)$$

where $\hat{\mathbf{y}}_{ik}$ = basic density at stem position *i* in tree *k*

$\mathbf{x}_{ik}$ = vector of the fixed regressors for position *i* in tree *k*

$\mathbf{b}$ = vector of the fixed effects

$\mathbf{z}_{ik}$ = vector of the independent random regressors for tree *k*

$\mathbf{u}_k$ = vector of the random effects for tree *k*

$\mathbf{e}_{ik}$ = random error term for position *i* in tree *k*

The random effects ($\mathbf{u}_k$) of the different trees are assumed to be uncorrelated. Random errors $\mathbf{e}_{ik}$ are assumed to be uncorrelated and also assumed to be normally distributed with a mean of zero and variance σ^2.

2.3.2 Methods to estimate Norway spruce crown biomass (Paper II)

The needle and branch (including wood and bark) biomasses of an individual tree were estimated by applying both model-based and design-based approaches. A regression estimator was applied in model-based approaches, and a ratio estimator in design-based approaches. Two different variants of both methods were studied (Table 2). In the ratio estimation method, the estimates for tree crown biomasses were based on the fresh weight of four crown sections and on two different sampling designs for branches: objective (N=10) and subjective (N=4) sample branch selection. In the regression method, biomass estimates of a tree crown were based on the measurements of the tally branches and the compiled branch-level regression models.

The regression models were estimated in two ways: separately for each tree (TREE-SPECIFIC MODELS) and for all sample trees (OVERALL MODEL) by using objectively selected sample branches (N=10 per tree). TREE-SPECIFIC MODELS were based on the ordinary least squares method (OLS), and the OVERALL MODEL, on the generalized least squares (GLS) method. The reliability of the biomass prediction of a tree was examined on the basis of the prediction errors. The basic assumption in both regression models was that the branch and needle biomasses of the same branch and tree are dependent, i.e., the errors of the branch and needle biomass equations are correlated (contemporaneous correlation). This statistical dependence was taken into account by applying linear seemingly unrelated regression (SUR) in the estimation of both models (Zellner 1962, Srivastava and Giles 1987, Parresol 1999, 2001,). This procedure enables to calculate the prediction errors for branch and needle biomass and also for whole crown biomass.

Table 2. The description of the methods applied for estimating tree crown biomass.

	MODEL-BASED ESTIMATOR Regression methods		**DESIGN-BASED ESTIMATOR Ratio estimation methods**	
Measurements:				
Sample branch	Dry weight		Dry and fresh weight	
Whole crown	Tally branches		Fresh weight	
Methods name	TREE-SPECIFIC MODELS	OVERALL MODEL	RATIO OBJECTIVE	RATIO SUBJECTIVE
Estimation	OLS (sur)	GLS (sur)	Ratio estimator	Ratio estimator
Sample branches:				
Number	10	290	10	4
Selection	Objectively	Objectively	Objectively	Subjectively

* OLS = ordinary least square, GLS = generalized least square, SUR = seemingly unrelated regression

2.3.3 Modelling biomass of Scots pine, Norway spruce and birch (Papers III and IV)

2.3.3.1 Biomass estimation for the sample trees

Biomass data were used for modelling tree biomass. The biomass was estimated for individual tree components; stem wood, stem bark, living and dead branches, foliage, stump and roots. The branch biomass included both branch wood and bark, and the living branch biomass included the cones. Not all the biomass components were measured on all sample trees (Table 3).

The branch biomass of a tree was predicted by applying ratio estimation methods based on subjective sample branches (RATIO SUBJECTIVE method). The ratio of the dry and fresh weight of the sample branches was used to estimate the branch and needle dry weight from the fresh mass. Ratio estimates for living branch biomass were calculated first by crown sections. The total living branch biomass was the sum of the crown sections. Constant moisture content, based on the mean moisture content of dead sample branches on the plots, was used for dead branches.

The biomass of stem wood was calculated by multiplying the stem volume by the average stem-wood density. Stem volume, both under-bark and over-bark, was calculated by applying Laasasenaho's (1982) taper curve equations, calibrated with diameter measurements at six points along the stem. Owing to the risk of bias in the estimates of average wood density, which was determined on the basis of only two sample disks per tree (breast height and a height of 70%), the average wood density was determined by applying equations for the vertical dependence of wood density presented in paper I and the two sample disk measurements and the stem taper curve. These equations (paper I) were calibrated with the measurements performed on the two disks, in order to obtain the tree level density curve, which predicted the wood density at different points along the stem. The corresponding stem diameters, which were used as a weight in estimating the average wood density, were obtained from the taper curve. The average wood density was then calculated from the density curve and taper curve. The obtained estimates for stem-wood density were, on average, 411 kgm^{-3} (SD 29.6), 379 kgm^{-3} (SD 34) and 478 kgm^{-3} (SD 33.2) for pine, spruce and birch, respectively.

The biomass of stem bark was obtained from the average bark density and bark volume of the tree. The bark volume of the stem was calculated as the difference between the under-bark and over-bark stem volume. Bark volume was based on measured bark dimensions of the sample discs. The average bark density of the tree was the mean of the bark density

Table 3. Number of measured biomass components by tree species.

Tree component	Scots pine	Norway spruce	Birch
Stem wood	626	366	127
Stem bark	311	170	127
Living branch	892	611	127
Dead branch	892	609	127
Foliage	892	611	21
Stump	36	31	39
Roots: > 2-5 cm	35	31	39
> 1 cm	6	5	6

measurements made on the two sample disks (breast height and a height of 70%). Disk level bark density was obtained by dividing the bark dry mass by the bark volume.

The stump and root biomasses were measured on a sub-sample of the trees on the temporary plots. The minimum determined coarse-root diameter varied from 2 to 5 cm, depending on tree diameter. In addition, the root biomass was also determined on roots with a diameter larger than 1 cm on some of the trees (Table 3). The fresh weight of the stump and roots was determined in the field. For moisture content determination, one sample was taken from the stump (sector) and two from the roots (discs). The stump and root biomasses of the tree were estimated by applying ratio estimation methods. First, simple regression equations (2–4) were constructed for the dependence of > 1cm root biomass on the biomass of coarse roots (2–5 cm). The >1 cm root biomass was then predicted for the whole root material by applying equations (2–4).

Scots pine $y = 0.103+1.525x$ $R^2 = 0.99,\ \hat{\sigma} = 1.471$ kg (2)

Norway spruce $y = 0.842+1.306x$ $R^2 = 0.99,\ \hat{\sigma} = 2.332$ kg (3)

Birch $y = 1.068+1.364x$ $R^2 = 0.99,\ \hat{\sigma} = 1.698$ kg (4)

where y is the >1 cm root biomass, x, the coarse-root biomass (minimum root diameter 2–5 cm), R^2, the coefficient of determination and $\hat{\sigma}$, the random error.

2.3.3.2 Modelling approach

The basic assumption in the modelling approach was that biomasses of individual tree components in the same site and in the same tree are dependent. This statistical dependency between the equations means that the errors of the individual biomass equations are correlated. Thus, multivariate procedures with random parameters were applied to take into account the across-equation correlation at both the stand and tree level. The multivariate procedure has a number of advantages compared to the independently estimated equations, if across-equation correlation is detected. The multivariate procedure enables more flexible model calibration and produces more reliable parameter estimates (Lappi 1991, Parresol 1999, 2001). It also enables biomass additivity to be ensured and the calculation of the prediction reliability for any combination of the tree components

Because we currently need biomass estimates not just for the total tree, but also for the tree components, the biomass equations for above- and below-ground tree components were compiled. The models for the above-ground tree components consisted of the equations for wood, stem bark, foliage, living and dead branches and total above-ground tree biomass. Equations for below-ground biomass components were estimated for stump and root (> 1cm) biomass. The equations for individual tree components and total above-ground biomass were first fitted independently, and a set of linear models was then constructed to form a multivariate linear model (Lappi 1991). The parameters of the multivariate models were estimated simultaneously, separately for the above-ground and below-ground biomasses. The compiled multivariate model was written as follows:

$$
\begin{aligned}
y_{1ki} &= \mathbf{b}_1\mathbf{x}_{1ki} + u_{1k} + e_{1ki} \\
y_{2ki} &= \mathbf{b}_2\mathbf{x}_{2ki} + u_{2k} + e_{2ki} \\
&\vdots \\
y_{nki} &= \mathbf{b}_n\mathbf{x}_{nki} + u_{nk} + e_{nki}
\end{aligned}
\qquad (5)
$$

where y_{1ki}, y_{2ki}... y_{nki} = dependent variables of biomass components 1, 2, ... n for tree i in stand k

n = number of biomass components

$\mathbf{x}_{1ki}$, $\mathbf{x}_{2ki}$... $\mathbf{x}_{nki}$ = vectors of the independent variables of biomass components 1, 2, ... n for tree i in stand k

$\mathbf{b}_1$, $\mathbf{b}_2$... $\mathbf{b}_n$ = vectors of the fixed effects parameters

u_{1k}, u_{2k}... u_{nk} = random effects of biomass components 1, 2, ... n for stand k

e_{1ki}, e_{2ki}...e_{nki} = random effects of biomass components 1, 2, ... n for tree i in stand k (residual error)

The covariance components, cov(u_{jk}, u_{j+1k}) and cov(e_{jki}, e_{j+1ki}), which addressed the dependency between the random effects of biomass components j, were estimated for both the stand and tree level. All the random parameters (u_{1k}, u_{2k}... u_{nk}) of the same stand are correlated with each other, and the residuals errors (e_{1ki}, e_{2ki}...e_{nki}) of the same tree are correlated. The random parameters and residual errors are assumed to be uncorrelated and are also assumed to be identically distributed Gaussian random variables with a mean of 0. In addition, the random parameters are assumed to have different variances.

The material had a hierarchical, 2-level (temporary plots) and 3-level (thinning and fertilization experiments), structure. To define the model, the study site was treated as a 2-level unit and the tree as a 1-level unit. In order to simplify the structure of the data, the plot level was ignored in the fertilization experiments. In the thinning experiments, the plots were assumed to be independent, i.e. treated as if they were from different stands.

3 RESULTS

3.1 Models for vertical dependence of wood density (Paper I)

According to the constructed model (Table 4), wood density was dependent on the vertical location along the stem in all species. This dependency was highest for pine. Wood density of pine decreased from the butt to the top, and the gradient of wood density decrease was steepest at the butt (the lowest part of the stem). The difference in wood density between the butt and the top was considerable, about 100 kg/m^3. The vertical dependence was similar in birch, but the density gradient was much smaller. For spruce the vertical dependence of the basic density was slight; the density first decreased slowly and then started to increase when approaching the top.

In additon, stem-wood density was correlated with tree height and tree growth rate for all tree species. Positive correlation with tree height indicated that taller trees tended to have higher wood density. Growth rate, expressed as the interaction between the tree diameter and age, improved the equation considerably for all tree species. The negative parameter estimates of growth rate suggested that fast-growing trees had low wood density. The effect of growth rate on wood density was highest for spruce and lowest for birch.

The random part of the model, related to the vertical tree-level variation in the density gradient, consisted of constant and random coefficients for relative height (Table 4). Random

Table 4. Parameter estimates of models for the vertical dependence of the basic density of the stem (*BD*, kgm^{-3}). For the fixed parameters, the standard error of estimates is given in parentheses.

Variable	Scots pine Estimate	Norway spruce Estimate	Birch Estimate
Intercept	467.073 (5.490)	431.339 (4.639)	470.098 (13.980)
h	0.228 (0.026)	0.075 (0.018)	0.359 (0.078)
d/t	-11.777 (0.829)	-16.973 (1.041)	-9.612 (1.750)
hr	-239.074 (11.999)	-41.801 (3.782)	-82.642 (8.915)
hr_2	-	63.583 (4.150)	180.597 (18.632)
hr_3	332.810 (6.502)	-	-136.931 (12.335)
hr_5	-231.238 (4.352)	-	-
$h*hr$	0.346 (0.135)	-	-
h_2*hr	-0.001 (0.000)	-	-
$var(u_{0k})$	1068.687	637.196	799.304
$var(u_{1k})*hr$	9496.149	6664.309	4760.085
$var(u_{2k})*hr_2$	8177.287	8398.160	5036.492
$cov(u_{0k},u_{1k})$	-2136.378	-324.087	-662.301
$cov(u_{0k},u_{2k})$	1430.737	-43.835	317.603
$cov(u_{1k},u_{2k})$	-8275.861	-6931.747	-4591.453
e_{ik}	92.740	114.407	115.461

Note: *h*, tree height (dm); *d*, tree diameter at breast height (mm); *t*, tree age (years); *hr*, relative height of position *i* (0…1); u_{0k}, u_{1k} and u_{2k}, random tree effects, e_{ik}, residual error.

tree effect proved to be applicable when the model was calibrated to a new tree. By utilizing the random effect and one or more wood density measurements at freely selected heights, the vertical dependence could be predicted more accurately. The advantages were greater for pine, which had a strong vertical dependence in basic density, than for spruce and birch. These models offered tools for obtaining reliable estimates of average stem-wood density, and decreased the risk of systematic errors when only a few wood density measurements have been carried out. A reliable estimate of average wood density is needed when, e.g., stem volume and wood density are used to determine stem biomass. An error in average wood density estimation also leads to an error in stem biomass estimates. This risk is most significant for pine, with a high vertical dependence of wood density.

3.2 Crown biomass obtained with different methods (Paper II)

The branch and needle biomass for the tree crown was calculated by applying the branch-level regression models (model-based) and ratio estimation (design-based) methods, with two different variations of both methods. In the ratio estimation methods, both the objective (RATIO OBJECTIVE) and subjective (RATIO SUBJECTIVE) sample branch selection were applied. The latter was used in crown biomass estimation of the sample trees in papers III and IV. Comparing the ratio estimation methods, estimates based on the subjective and objective sampling were, on average, quite similar. Estimates differed on average by < 3% (0.7 kg) and < 2% (0.4 kg) for branch and needle biomass (Table 5). Although no significant differences were found on a stand level, there were differences at a tree level between the methods, in predictions of both the needle and branch biomass (Figure 2). Despite tree to tree variation, there was no systematic trend in the results between the methods with regard to tree size or stand (Figure 2). Using the RATIO SUBJECTIVE design (applied in papers III and IV), it was not possible to estimate the accuracy of estimates by statistical theory. However, using the RATIO OBJECTIVE design, the average accuracies of estimates, SE_{pred}, were 5.3% (SD 2.8), 4.5% (SD 2.5) and 4.2% (SD 1.9) for needle, branch and total crown biomass, respectively (Table 5).

Both model-based approaches, the TREE-SPECIFIC MODELS (*OLS*-models) and the OVERALL MODEL (mixed models), on average produced similar biomass estimates (Table 5). The branch biomass estimates using the TREE-SPECIFIC and the OVERALL MODELS were almost identical (38.1 kg and 38.5 kg), but the OVERALL MODEL on average produced a 1.4 kg (7.5%) larger needle biomass. No significant differences or major trends in estimates were identified between methods. In terms of prediction errors (SE_{pred}), independently of the method used, estimates of branch biomass were found to be more accurate than estimates of needle biomass (Table 5). Estimates based on the OVERALL MODEL method were found to

Table 5. Average crown biomass (kg) estimated using the model-based methods (the OVERALL and TREE-SPECIFIC MODELS) and design-based methods (the RATIO SUBJECTIVE and RATIO OBJECTIVE). Relative SE_{pred} (%) in parenthesis.

Methods	Needle, kg	Branch, kg	Total crown, kg
RATIO OBJECTIVE	20.99 (5.3)	44.64 (4.5)	65.63 (4.2)
RATIO SUBJECTIVE	21.49 (-)	43.99 (-)	65.39 (-)
OVERALL	18.99 (12.0)	38.46 (7.2)	57.45 (7.8)
TREE-SPECIFIC	17.59 (18.1)	38.13 (7.8)	55.72 (10.7)

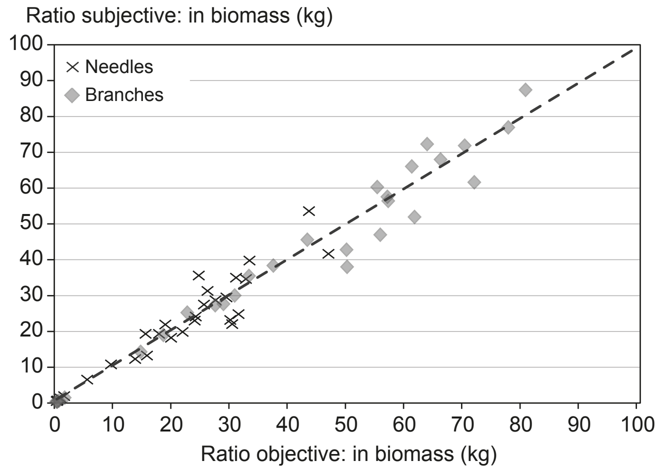

Figure 2. The tree needle and branch biomass obtained by RATIO OBJECTIVE and RATIO SUBJECTIVE.

be more accurate, especially for needle biomass, than estimates based on TREE-SPECIFIC MODELS (Table 5). The relative SE_{pred} obtained by the OVERALL MODEL was, on average, 12.0% and 7.2% for needle and branch biomass. The predictions were also stabile, i.e., the reliability of the biomass values (SE_{pred}) varied only a little from tree to tree. Estimates using the TREE-SPECIFIC MODELS were almost as accurate, with an estimated average accuracy of 7.8%. However, the needle biomass predictions obtained by TREE-SPECIFIC MODELS were found to be less reliable and stabile; the relatve SE_{pred} was on average 18.1%, with a range of 7.5–55.5 %.

3.3 Biomass equations for Scots pine, Norway spruce and birch in Finland (Papers III–IV)

Multivariate models were constructed separately for the above-ground and below-ground biomass. Owing to the different number of observations of the above- and below-ground components, the model parameters could not be estimated simultaneously. The multivariate models for above-ground biomass contained the individual equations for stem wood, stem bark, foliage, living and dead branches, and total tree biomass (Tables 6, 7 and 8). The equation for foliage biomass of birch was estimated independently due to the limited amount of material. The multivariate model for below-ground biomass included the equations for stump and roots with a diameter > 1cm, and for birch, also the equation for total below-ground biomass (Table 6). The biomass equations had a multiplicative model form. Logarithmic transformation was used to obtain homogenous variance, and to transform the equations to a linear form.

Three multivariate models for above-ground biomass and one for below-ground biomass were constructed separately for each tree species. All the multivariate models were based only on the variables that are commonly measured in the national forest inventory. In the simplest model formulation, multivariate models (MV models 1) were based only on tree diameter at breast height (d) and tree height (h), as independent variables (Table 6). Tree age at breast height (t_{13}) and crown variables such crown length *(cl)* or crown ratio (cr), as independent variables, were added to more complex multivariate models (MV models 2) (Table 7). The most elaborate multivariate models (MV models 3) were based, in addition to the previously mentioned variables, also on bark thickness (*bt*) and radial increment (without bark) during the last five years (i_5) (Table 8).

Tree diameter proved to be the most significant independent variable in all equations (Tables 6, 7 and 8). Tree height was used as an independent variable in most cases. Only the equations for dead branches and stump and roots were mostly based only on *d*. The inclusion of more independent variables (cl, cr, t_{13}, bt, i_5) improved the multivariate models by reducing especially the between-stand variance.

Stem-wood biomass was correlated with tree dimensions (*d* and *h*) and growth rate. Negative correlation with growth rate indicated that fast-growing trees seemed to have low biomass. Adding variables depicting tree growth rate (t_{13} or $\frac{d}{t_{13}}$) to equations for stem-wood biomass (MV models 2) reduced the total error variance (the sum of random stand- and tree-level variance) by 13–34%. The inclusion of radial growth (i_{g5}) decreased the total error variance by a further 6–10%, but only for conifers. The equation for stem bark biomass had a similar form in MV models (1) and (2), and contained only *d* and *h* as independent variables. In MV models (3) bark thickness (*bt*) was added to bark equations, which decreased the total error variance by 10–15%.

The tree crown biomass (living branches and foliage) was correlated with tree dimensions; positively with *d* and negatively with *h*. Negative correlation of *h* may indicate that, at a given *d,* taller trees tended to have a lower crown biomass (Fig. 3). Furthermore, long crown length and fast growth rate were related to high crown biomass. The inclusion of crown variables, crown length or crown ratio, improved significantly the performance of equations for crown components (living branches and foliage) (MV models 2) by decreasing the total error variance by 29–45%, and between-stand variance even more. The total error variance was reduced by about a further 7–24% when variables describing tree growth rate (i_5, t_{13}, $\frac{d}{t_{13}}$) were included in the equations for crown components (MV models 3). In contrast, the equation for dead branches could be improved only marginally compared to the simplest equation formulation (MV models 1), and the error variance was considerable in all cases.

The total above-ground biomass was positively correlated with *d* and *h*. In addition, crown variables (*cr*) improved slightly the fit of the total above-ground biomass equation only for spruce and pine, but not for birch. Only the equation for pine showed better fit after adding bark thickness (*bt*) and variables indicating tree growth rate.

The assumed statistical dependence between the biomass equations was verified in the analysis at both the stand and tree levels. In general, the across-equation correlation at the stand level was higher than that at the tree level. The tree-level errors were not systematically correlated between the tree components, and no correlations over 0.5 were detected. The magnitude of across-equation correlation at stand level depended on the tree species and the MV models. Generally dead branch biomass showed a high correlation with other tree components. The random parameter of dead branches and needles showed a negative correlation in all the multivariate models of pine and spruce. In addition, dead-branch biomass was systematically correlated with stem-wood biomass in the MV models of pine and birch. Also uniform correlations between the needles and living branches as well as with stem bark occurred for the spruce models.

Table 6. The parameter estimates of multivariate models (1) for pine, spruce and birch biomasses (ln(kg)).

Variable	Stem wood, kg			Stem bark, kg			Living branches, kg			Foliage, kg			Dead branches			Total above-ground, kg		
	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch
Intercept	-3.721 (0.032)	-3.555 (0.067)	-4.879 (0.065)	-4.548 (0.111)	-4.548 (0.103)	-5.401 (0.150)	-6.162 (0.090)	-4.214 (0.128)	-4.152 (0.220)	-6.303 (0.524)	-2.994 (0.634)	-29.566 (3.881)	-5.201 (0.172)	-4.850 (0.261)	-8.335 (1.141)	-3.198 (0.038)	-1.808 (0.050)	-3.654 (0.053)
$\frac{d_s}{(d_s+n)}$	8.103 (0.106)	8.042 (0.183)	9.651 (0.162)	7.997 (0.402)	9.448 (0.589)	10.061 (0.460)	15.075 (0.260)	14.508 (0.417)	15.874 (0.580)	14.472 (0.350)	12.251 (0.400)	33.372 (4.201)	10.574 (0.293)	7.702 (0.924)	12.402 (1.966)	9.547 (0.107)	9.482 (0.243)	10.582 (0.146)
$\frac{h}{(h+m)}$	5.066 (0.107)	-	-	-	-	2.657 (0.504)	-2.618 (0.284)	-3.277 (0.425)	-4.407 (0.642)	-3.976 (0.789)	-3.415 (0.929)	-	-	-	-	3.241 (0.116)	-	3.018 (0.150)
ln(h)	-	0.869 (0.056)	1.012 (0.042)	0.357 (0.086)	0.436 (0.123)	-	-	-		-	-	-	-	0.513 (0.220)	-	-	0.469 (0.052)	-
h	-	0.015 (0.003)	-	-	-	-	-	-		-	-	-	-	-	-	-	-	-
n	14	14	12	12	18	12	12	13	16	6	10	2	16	18	16	12	20	12
m	12	-	-	-	-	20	12	5	10	1	1	-	-	-	-	20	-	22
$var(u_k)$	0.002	0.009	0.003	0.015	0.023	0.010	0.041	0.039	0.027	0.109	0.107	0.000	0.253	0.367	1.115	0.003	0.006	0.001
$var(e_{ki})$	0.009	0.009	0.005	0.061	0.041	0.044	0.089	0.081	0.077	0.118	0.089	0.077	0.362	0.352	2.679	0.010	0.013	0.007
c													0.911	1.343	2.074			

	Stump, kg			Roots > 1 cm, kg		
	Pine	Spruce	Birch	Pine	Spruce	Birch
Intercept	-6.753 (0.190)	-3.964 (0.248)	-3.574 (0.233)	-5.550 (0.178)	-2.294 (0.336)	-3.223 (0.472)
$\frac{d_s}{(d_s+n)}$	12.681 (0.312)	11.730 (0.575)	11.304 (0.528)	13.408 (0.315)	10.646 (0.575)	6.497 (0.853)
ln(h)	-	-	-	-	-	1.033 (0.273)
n	12	26	26	15	24	22
$var(u_k)$	0.010	0.065	0.022	0.000	0.105	0.048
$var(e_{ki})$	0.044	0.058	0.045	0.079	0.114	0.027

Note: d_S, 2 + 1.25 d (d = tree diameter at breast height, cm); h, tree height (m); u_k, random stand effects; e_{ki}, residual error; c, the empirical correction factor.

Table 7. The parameter estimates of multivariate models (2) for pine, spruce and birch biomasses (ln(kg)).

Variable	Stem wood, kg			Stem bark, kg			Living branches, kg			Foliage, kg			Dead branches			Total above-ground, kg		
	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch
Intercept	-4.018 (0.030)	-4.000 (0.060)	-4.886 (0.058)	-4.695 (0.108)	-4.437 (0.101)	-5.433 (0.152)	-5.166 (0.084)	-3.023 (0.121)	-5.067 (0.183)	-1.748 (0.476)	-0.085 (0.570)	-20.856 (4.015)	-5.318 (0.173)	-5.317 (0.235)	-7.996 (1.141)	-3.416 (0.039)	-2.141 (0.073)	-3.659 (0.053)
$\frac{d_s}{(d_s + n)}$	8.358 (0.099)	8.881 (0.203)	9.965 (0.164)	8.727 (0.395)	10.071 (0.588)	10.121 (0.467)	13.085 (0.246)	12.017 (0.403)	14.614 (0.580)	14.824 (0.431)	15.222 (0.753)	22.320 (4.628)	10.771 (0.295)	6.384 (0.886)	11.824 (1.966)	9.555 (0.098)	9.074 (0.250)	10.588 (0.157)
$\frac{h}{(h + m)}$	4.646 (0.008)	-	-	-	-	2.647 (0.509)	-5.189 (0.259)	-5.722 (0.401)	-5.074 (0.563)	-12.684 (0.723)	-14.446 (1.020)	-	-	-	-	3.592 (0.116)	-	2.966 (0.159)
h	-	0.728 (0.056)	0.966 (0.040)	0.228 (0.084)	0.261 (0.123)	-	-	-	-	-	-	-	-	0.982 (0.207)	-	-	0.570 (0.056)	-
t_{13}	-	0.022 (0.003)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$\ln(t_{13})$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.001 (0.000)
$\frac{d}{t_{13}}$	0.041 (0.008)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cl	-	-0.273 (0.040)	-0.135 (0.025)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$\ln(cl)$	-	-	-	-	-	-	-	-	0.092 (0.009)	-	-	-	-	-	-	-	-	-
cr	-	-	-	-	-	-	1.110 (0.050)	1.033 (0.071)	-	1.209 (0.062)	1.273 (0.076)	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-		2.819 (0.795)	-	-	-	0.395 (0.030)	0.403 (0.059)	-
n	14	12	12	12	18	12	12	14	12	4	4	2	16	18	16	12	20	12
m	10	-	-	-	-	20	8	5	12	1	1	-	-	-	-	24	-	22
$var(u_k)$	0.001	0.003	0.002	0.014	0.019	0.011	0.020	0.017	0.015	0.032	0.028	0.011	0.265	0.263	1.065	0.002	0.006	0.000
$var(e_{ki})$	0.008	0.008	0.005	0.057	0.039	0.044	0.063	0.068	0.057	0.093	0.087	0.044	0.347	0.356	2.691	0.009	0.013	0.007
c	-	-	-	-	-	-	-	-	-	-	-	-	0.913	1.208	2.149	-	-	-

Note: d_s, 2 + 1.25 d (d = tree diameter at breast height, cm); h, tree height (m); cl, length of living crown (m); cr, crown ratio (0…1); t_{13}, tree age at breast height; u_k, random stand effects; e_{ki}, residual error; c, the empirical correction factor.

Table 8. The parameter estimates of multivariate models (3) for pine, spruce and birch biomasses (ln(kg)).

Variable	Stem wood, kg			Stem bark, kg			Living branches, kg			Foliage, kg			Dead branches			Total above-ground, kg		
	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch	Pine	Spruce	Birch
Intercept	-4.590 (0.046)	-3.950 (0.060)	-4.915 (0.058)	-5.565 (0.165)	-4.626 (0.100)	-5.304 (0.303)	-4.833 (0.122)	-3.950 (0.301)	-5.918 (0.193)	-2.209 (0.512)	-4.258 (0.154)	-	-5.798 (0.199)	-0.140 (1.429)	-16.113 (1.983)	-3.259 (0.040)	-2.037 (0.072)	-3.713 (0.050)
$\frac{d_s}{(d_s+n)}$	8.520 (0.119)	8.534 (0.203)	9.984 (0.174)	9.691 (0.252)	9.638 (0.591)	8.498 (0.591)	13.126 (0.297)	12.014 (0.315)	12.867 (0.612)	9.347 (0.425)	9.200 (0.163)	-	17.820 (0.739)	11.293 (0.374)	37.902 (5.801)	9.337 (0.110)	9.146 (0.227)	10.616 (0.155)
$\frac{h}{(h+m)}$	5.013 (0.110)	-	-	-	-	3.380 (0.511)	-4.808 (0.304)	-1.296 (0.464)	-3.573 (0.571)	-6.364 (0.784)	-	-	-	-	-17.342 (4.654)	3.265 (0.122)	-	3.235 (0.164)
$\ln(h)$	-	0.743 (0.055)	0.981 (0.042)		0.266 (0.117)	-	-	-	-	-	-	-	-	-	-	-	0.543 (0.050)	-
h	-	0.022 (0.003)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
t_{13}	0.002 (0.000)	0.001 (0.000)	-	-	-	-	-	-	0.007 (0.001)	-	-	-	-0.017 (0.002)	-	-0.063 (0.013)	0.001 (0.000)	-	-
$\ln(t_{13})$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$\frac{d}{t_{13}}$	-	-	-0.180 (0.043)	-0.444 (0.103)	-	-	-	-0.461 (0.121)	-	-	-	-	-	-	-	-	-	-0.214 (0.039)
cl	-	-	-	-	-	-	-	-	0.095 (0.010)	-	-	-	-	-	-	-	-	-
$\ln(cl)$	-	-	-	-	-	-	0.727 (0.048)	-	-	0.611 (0.065)	-	-	-0.738 (0.139)	-	-	-	-	-
cr	-	-	-	-	-	-	-	1.528 (0.116)	-	-	0.967 (0.118)	-	-	-7.014 (1.492)	-	-	0.296 (0.059)	-
$\ln(cr)$	-	-	-	-	-	-	-	-	-	-	-	-	-	3.058 (0.960)	-	-	-	-
i_{g5}	0.002 (0.000)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$\ln(i_{g5})$	-	-	-	-	-	-	0.098 (0.019)	-	-	0.309 (0.023)	-	-	-0.461 (0.056)	-	-	-	-	-
i_5	-	-0.071 (0.017)	-	-	-	-	-	0.112 (0.046)	-	-	-	-	-	-	-0.166* (0.041)	0.124 (0.011)	-	0.007* (0.001)
$\ln(i_5)$	-	-	-	-	-	-	-	-	0.238* (0.035)	-	0.287 (0.024)	-	-	-0.189 (0.058)	-	-	-	-
bt	-	-	-	0.068 (0.014)	0.084 (0.027)	-	-	-	-	-	-	-	-	-	-	-0.006 (0.003)	-	-
$\ln(bt)$	-	-	-	-	-	0.382 (0.057)	-	-	-	-	-	-	-	-	-	-	-	-
n	9	12	12	8	16	8	10	18	10	6	12		16	14	6	12	20	12
m	16	-	-	-	-	22	4	2	10	1	-	-	-	-	10	18	-	22
$var(u_k)$	0.001	0.003	0.001	0.008	0.013	0.011	0.018	0.011	0.012	0.027	0.022	-	0.140	0.196	0.578	0.003	0.007	0.000
$var(e_{ki})$	0.008	0.008	0.005	0.058	0.042	0.035	0.059	0.067	0.043	0.082	0.068	-	0.345	0.278	2.570	0.009	0.013	0.007
c	-	-	-	-	-	-	-	-	-	-	-	-	0.918	1.091	1.788	-	-	-

Note: d_s, 2 + 1.25 d (d = tree diameter at breast height, cm); h, tree height (m); cl length of living crown (m); cr, crown ratio (0…1); t_{13}, tree age at breast height; i_{g5}, cross-sectional area increment at breast height during the last five years (cm^2); i_5, breast height radial increment during the last five years (cm, *mm for birch); bt, double bark thickness at breast height (cm); u_k, random stand effects; e_{ki}, residual error; c, the empirical correction factor.

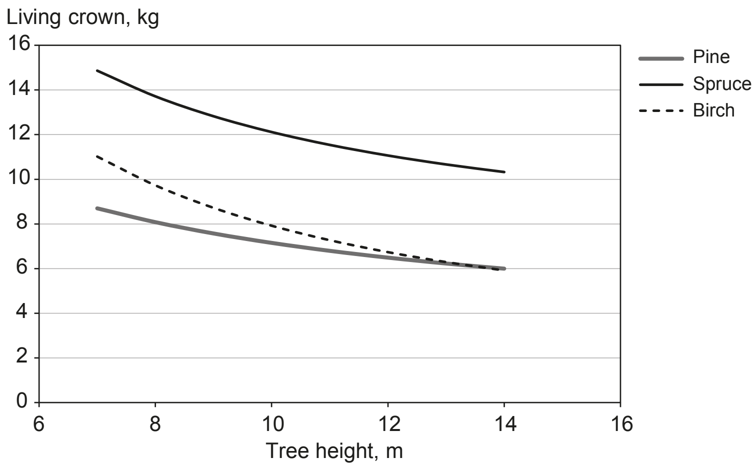

Figure 3. The effect of tree height on living crown biomass (living branches and foliage) at a given diameter (*d* = 12 cm) using MV models 1.

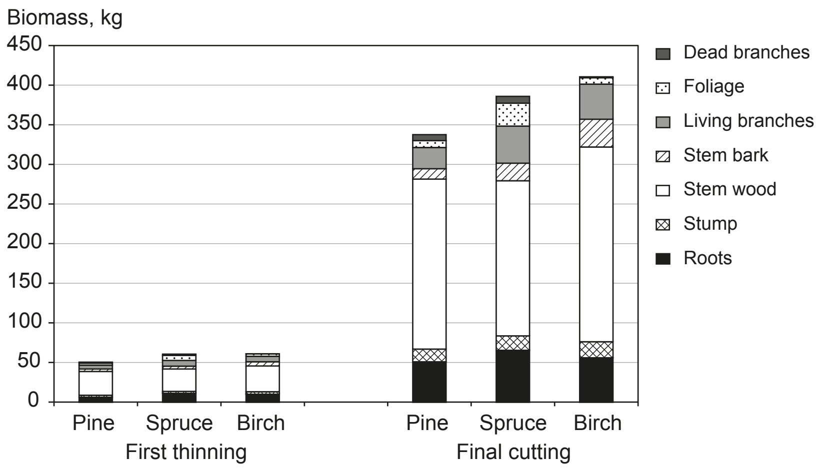

Figure 4. The expected biomass of tree component at the stage of first thinning (*d*=13 and *h*=12) and final cutting (*d*=27 and *h*=22) using MV models 1.

Biomass allocation to the tree components was illustrated by applying the MV model (2) to the biomass data. Biomass allocation varied by tree species, and the differences between tree species also depended on the stage of stand development, which is demonstrated in Fig. 4 with two example trees. At given tree dimensions the highest whole stem biomass (stem wood and bark), especially at the stage of final cutting, was detected for birch, and the lowest for spruce. The highest crown and below-ground biomass was obtained for spruce, and the lowest for pine.

Allometric relationships between tree-component biomasses changed with tree size (Figs. 5–7). The stem is the greatest biomass component and its proportion of whole tree biomass showed an increasing trend with tree size, especially for spruce and birch. The stem proportion in mature stands was mainly between 60–70%; it was highest for birch and lowest for spruce. The proportion of the whole crown biomass was highest for spruce and lowest for birch. The crown proportion showed a decreasing tendency with age for all tree species. For spruce the crown proportion dropped from 30% in young stands to 17% in mature stands. This trend was less clear for pine and birch; the crown proportion dropped from 19% to 11% for birch and from 24% to 12% for pine. The relative share of the below-ground biomass of the whole tree biomass is ca. 20%; it is higher in young spruce and birch stands and lower in young pine stands.

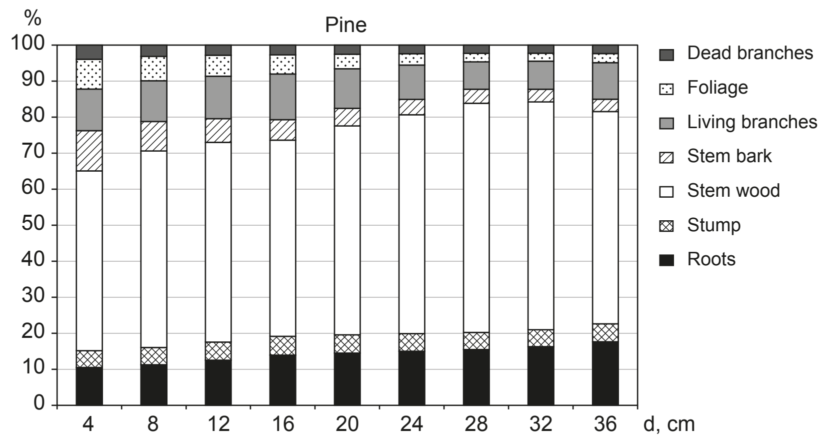

Figure 5. The biomass allocation of pine by applying MV models 2 to the biomass data of this study.

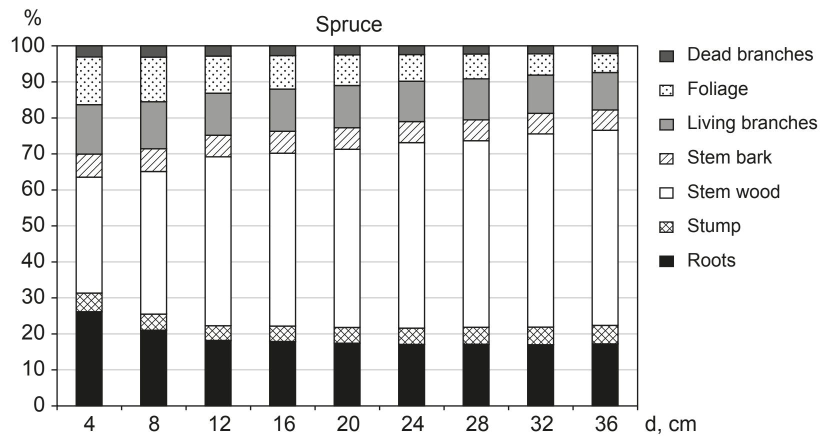

Figure 6. The biomass allocation of spruce by applying MV models 2 to the biomass data of this study.

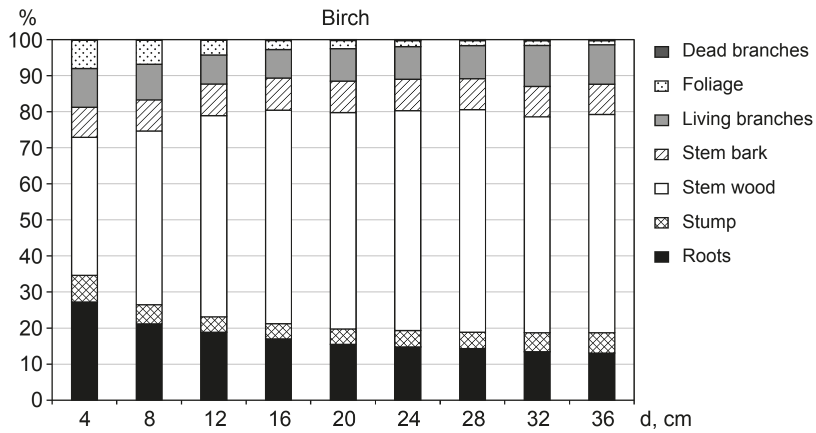

Figure 7. The biomass allocation of birch by applying MV models 2 to the biomass data of this study.

4 DISCUSSION

4.1 General

In this study, individual-tree biomass equations for above- and below-ground tree components of Scots pine, Norway spruce and birch were compiled. The constructed biomass equations were based on the same material, and mainly also on the same independent variables, that were used in a previous study (See Repola et al. 2007). However, the parameters of the new equations were estimated using a statistical method that utilizes the information of the biomass data more efficiently and produces a more flexible and exact model structure. The applied model structure enables the model to be calibrated more flexibly and the prediction reliability for any combination of the tree components to be assessed more exactly. The aim was to construct models that are applicable for wider range of purposes (e.g. for assessing forest carbon-pool and energy-wood recourses), and which fulfill the current requirements for the biomass models.

4.2 Material

In order to obtain reliable biomass estimates at the national scale, the biomass models should be based on a representative sample of the stands in which the results are to be applied, e.g., national forest inventory data (Parresol 1999). However, sampling and analysis of biomass of different tree components is laborious and costly, which effectively reduces the possibilities to obtain data. Therefore, existing biomass data collected during previous projects was utilized, and only a limited amount of new material could be collected for the current study. Hence, the majority of the pine and spruce study material originated from a series of long-term fertilization experiments (only unfertilized plots included), which were primarily not designed for constructing general biomass equations. When deriving regression models it might be better to subjectively select sites and stands rather than use objective sampling. It is essential to have a high variation in the independent variables. The selection is made such as to represent a variety of different site and stand conditions of the population. However, within stands, trees are usually sampled in an objective way. The main purpose of the collection of new biomass material (temporary sample plots and thinning experiments) was to reflect a wide range in site and stand conditions in the modelling data, i.e., to have higher variation in the independent variables. This design is assumed to have improved the reliability of the model, because high variation in the independent variables is a prerequisite for reliable parameter estimates (Lappi 1993). Small variation in independent variables along with a weak correlation between the response and independent variables produces unstable parameter estimates and reliable predictions only for observations with near to average values. It is also important that the models include the most essential variables and that the dependencies are delineated correctly. These dependencies should be described with common and easily measurable variables. To obtain a more applicable model and to decrease the bias caused by the unrepresentative data (subjective selected material), the equations were based only on the variables commonly measured in Finnish National Forest Inventories, and the equations based only on tree diameter are not presented here. Variables such as stem diameter and tree height are assumed to be easy to measure with a negligible measurement error. By allometry these variables have a natural correlation with volume and biomasses. Thus, despite some minor shortcomings, the material was found to be suitable and representative for its purpose.

The number of sample trees used in this study varied by tree species; the volume of pine material was the largest and the volume of birch material the smallest. All the tree components were relatively well represented in the material, apart from below-ground tree components and birch foliage. In addition, pine and spruce material was unbalanced in terms of the response variables for above-ground tree components, i.e., not all tree components were measured on all the sample trees. In pine and spruce data, only the crown components were measured for all sample trees, and stem bark data was the least complete. The underlying number of units used for modelling was similar to that used for Marklund's functions (1988). In fact we had data on more pines, about the same number of spruces and less birches Marklund's data (1988) were more comprehensive in terms of above-ground tree components (except birch foliage), i.e., there were less missing values for tree components. In addition, Marklund (1988) clearly had more data on the below-ground biomass of pine and spruce, but no data at all for the below-ground biomass of birch. The shortcoming of Marklund's data is that the minimum diameter for roots of < 5 cm diameter was poorly defined, which makes the biomass predictions vague.

In order to give a reliable description of the relationships between the tree components, the equations for individual tree components should be based on the same sample trees. In our data this requirement was partly fulfilled, but not all tree components were measured on all the sample trees, especially in pine and spruce data. Consequently, the equations for individual tree components were based on different numbers of observations, which can cause distortion in the relations of tree components. In addition, the equations for below-ground tree components were based on relatively limited material, which decreases the reliability of the predictions of these components, and can also produce an error in the relationship between below- and above-ground tree biomass. Despite this, the compiled equations produced predictions for the below-ground biomass and the share of below- and above-ground biomass which are in line with earlier studies (Marklund 1988, Vanninen et al. 1996, Helmisaari 2001 and Petersson and Ståhl 2006). Also the relationships between the above-ground tree components showed a similar trend with regard to tree size in this study and in Marklund's (1988) study, although the equations in this study predicted a lower stem proportion for conifers in mature stands.

In addition to the biomass of entire tree components, there is currently also a need to predict biomass accumulation along a tree (Kärkkäinen 2005). This information is requested especially when assessing the amount of energy wood. The compiled biomass equations give predictions for entire tree components, but not for the vertical biomass distribution along a tree. However, the equations for the vertical dependence of wood density (Paper I) can be utilized for assessing the biomass distribution along a stem. The total biomass accumulation for crown and below-ground components was not directly addressed in this study, which can restrict the use of the equations, e.g., in calculating energy wood removal of a stand.

4.3 Modelling tree biomass

The reliability and applicability of biomass equations depends on the study material and also on how efficiently the study material is utilized in the model estimation in order to obtain reliable parameter estimates. In addition, the reliability of the predicted biomass value is affected by the statistical errors of the dependent variable, caused by a sub-sampling (Parresol 1999). Therefore it is important that biomasses of sample trees are estimated reliably. In this study, the estimation errors in the sample tree biomasses could not be estimated reliably, and no exact estimate for the magnitude of this error was presented.

In biomass data (papers III and IV), the determination of stem biomass was based on tree volume and average stem-wood density. For wood density measurements, only two sample disks per tree (breast height and a height of 70%) were taken. The low number of sample disks can lead to a biased estimate of average wood density, and consequently of biomass, especially for tree species with a high vertical dependence of wood density. Therefore models for the vertical dependence of the wood density of pine, spruce and birch stems (paper I) were constructed. These models can be calibrated for any stem using one or more wood density measurements at a freely chosen height. Hence, they were applied for determining the average wood density of the sample trees in our biomass data. Compared to other studies (e.g. Hakkila 1979) the estimates for average values and SD were similar, 411 kgm^{-3} (SD 29.6), 379 kgm^{-3} (SD 34.0) and 478 kgm^{-3} (SD 33.2) for pine, spruce and birch, respectively. Hence, it can be concluded that the applied method improved the accuracy of wood density estimates and decreased the risk of systematic errors. The advantage was most significant for pine, which has a high vertical dependence of wood density.

The constructed models for the vertical dependence of wood density were based on hierarchically structured data. The correlation structure of the observations was not properly addressed in the model specification. The compiled models were specified as linear mixed models by addressing the random effects on two levels; between-tree and within-tree levels. The random errors (within-tree variation) were assumed to be uncorrelated, but in fact spatial autocorrelation of the successive measurements of the stems obviously exists. This did not affect the parameter estimates of the fixed effects, but it affects the reliability of the test by producing too low a standard error, i.e. the reliability of the parameter estimates were probably overestimated.

In biomass data, crown biomass of each sample tree was based on the ratio estimation method, with four subjectively selected sample branches. A number of factors caused uncertainty in the results obtained by the applied method. First, subjective sample branch selection, with the aim of selecting representative sample branches from each crown stratum, can lead to biased estimates of the crown biomasses, which depend on the observer. The results of paper II, the subjective sub-sampling applied to spruce data produced similar results on average and caused no systematic bias with regard to tree size compared to the objective sub-sampling. Therefore, an error in crown biomass caused by the applied sub-sampling can be interpreted as a random error, which is not a problem in the linear model (Parresol 2001). The results of the paper II showed also that the statistical error of the dependent variable caused by sub-sampling was clearly higher in the needle biomass estimates than in the branch biomass estimates. This error in the objective sub-sampling design was, on average, 5.3% and 4.5% for needle and branch biomass. However the magnitude of this error in the subjective sub-sampling design could not be estimated, which was a disadvantage of the method applied in papers III and IV. Despite this, the error can be assumed to be at least at the same level as that in the objective sampling design. In addition, it is a well-known fact that ratio estimators are biased, especially if the sample size per stratum in stratified sampling is small and the number of strata is large, like it was in our data (Cochran 1977, Valentine et al. 1984, Cunia 1979, Parresol 1999). An alternative ratio estimate with a small sample size is a single combined estimate, i.e. the mean ratio estimator of total crown (Hansen et al. 1946). The combined ratio estimate is applicable if the sample size in different strata is small and the ratio estimate can be assumed to be constant among the strata (Cochran 1977). Despite the small sample size, we used a separate ratio estimator for each crown section, because the ratio estimates of both the needle and branch varied systematically between crown sections, i.e., the assumption of constant ratio estimate was not valid.

The reliability and applicability of biomass equations depends partly on how the model has been formulated. The compiled equations were based only on the variables commonly measured in forest inventories, and were formulated so that the predictions would be logical throughout the range of the material, i.e., nonnegative values (small trees) or overestimates (big trees) are not obtained even in cases where the functions are extrapolated. Furthermore, whole information of data has been utilized in order to produce reliable parameter estimates and an applicable and flexible model structure. For an unbiased test of the parameters, the correlation structure of observations must be addressed in the model specification. To avoid a too complicated random part of the model and the problem in the model estimation, the correlation structure of the data was not totally addressed in the model specification. Biomass data was hierarchically, 2-level (temporary plots) and 3-level (thinning and fertilization experiments) structured. In the thinning experiments, based on the different thinning treatments, the plots were assumed to be independent (treated as if they were from different stands). In the fertilization experiments, the treatment in the control plots did not differ from each other and stand and plot levels were combined, i.e. plot level was ignored. Temporal autocorrelation existed in some fertilization experiments; the sample trees had been removed at two different times (with a 5-year interval). This temporal correlation was ignored and the sampling time of the same plot was assumed to be independent (treated as if they were from different stands). These simplifications may decrease the reliability of the parameter test when the standard error of parameter estimates could be underestimated.

Generally, equations for the biomass of individual tree components have been estimated separately and ignoring the correlation between the biomass components of the same tree or stand. In this study, this across-equations correlation (contemporaneous correlation) was taken into account in the model estimation by applying the multivariate procedure. Based on the verified statistical dependence between the biomass equations, especially at the stand level, the multivariate procedure had a number of advantages compared to the independently estimated equations. First, the across-equation correlations of the random parameters enable information to be transferred from one equation to another, which is especially useful in calibrating the model for a new stand (Lappi 1991). In the model calibration, the determination of one biomass component, e.g., stem biomass as a result of stem volume and average wood density, also enables the prediction of random stand effect for the other tree components, which results in more reliable predictions for all tree components in a stand. Second, the multivariate models also produced across-equation covariance of the fixed parameters, which enables the calculation of the prediction reliability for any combination of tree components. This information is not available for independently estimated equations. Third, the multivariate model usually produces more reliable parameter estimates when contemporaneous correlations occur (Parresol 1999, 2001). This advantage was, however, of only minor importance in this study (see paper III).

The applied statistical method enables biomass additivity to be ensured by setting across-equation constraints (Briggs 1984, Parresol 1999, Carvalho et al. 2003, Bi et al. 2004, Návar et al. 2004). Across-equation constraints were not applied because of the unbalanced data and to avoid unnecessary complexity in the total tree equation. The unbalanced data (pine and spruce), i.e., the equations for the total tree biomass were clearly based on a lower number of observations compared to the equations for the biomass of individual tree components, was partly responsible for some shortcomings in terms of biomass additivity. In our study, logarithmic transformation was applied to the dependent variables. This caused biases in the back-transformed value, and also problems with biomass additivity. Despite this, the compiled equations ensured better biomass additivity compared to the independently estimated equations.

4.4 Application area of the biomass models

Representativeness of the modelling data is a criterion of the relevant application area for the models. From that point of view, the compiled biomass equations are applicable for living trees on mineral soil over the whole country. However, due to the lack of material the validity of the equations is uncertain in fertilized stands, in the northernmost parts of Finland, and in peatland forest, especially if the dependency between the response and independent variables or the combination of tree dimensions deviates from that on mineral soil. The equations can be applied to the whole growing stock over a wide range of stand and site conditions, from young to mature stands. Despite this, the equations are primarily applicable for trees growing on normally managed stands. Hence the validity for the trees growing in sparse stands or in open space is uncertain. The study material includes sample trees from all tree classes over a wide diameter range, from 1 to 42 cm. Nevertheless, the applicability of the equation to trees with a height < 1.3 m is uncertain, especially for crown components.

This study resulted in three multivariate models for above-ground tree components. The simplest models (MV models 1) were based on tree diameter and height, and the more elaborate models (MV models 2 and 3), on additional commonly measured tree variables. In contrast, the material on below-ground tree components was relatively limited and therefore only one model formulation per tree species was constructed. The reliability of biomass prediction varied across models, but generally the most robust (stable) predictions were obtained by using the simplest models (MV models 1). The more elaborate models (MV models 1 and 2) give more reliable predictions and also decrease the risk of systematic errors caused by the study material. The inclusion of independent variables decreased between-stand variance in particular, which indicated the high correlation of these variables and the more reliable predictions at stand level. But in many cases, the applicability of these models (MV models 1 and 2) is restricted by the availability of the necessary tree variables.

Stem-wood biomass consisted of two components: stem volume and average wood density of the stem. The reliability of the models depended on how these components were depicted in the models. Growth rate is highly (negatively) correlated with wood density in conifers and birch (Mergen et al. 1964, Hakkila 1979, Saranpää 1983, MacPeak et al. 1990, Mäkinen and Uusvaara 1992), and stem volume, with the tree dimensions: diameter, height and stem taper (Laasasenaho 1982). The variation in wood density caused by growth rate was taken into account in multivariate models (2) and (3). In turn, the variation in stem form caused by stem tapering (diameter at a height of 6 meters) was not taken into account in the multivariate models, because it would have restricted the validity of the equations only to trees with a height > 6 m. When the upper diameter is available, the stem biomass can be calculated more reliably by applying an applicable volume function (e.g. Laasasenaho 1982) and the models for average wood density produced by Repola et al. (2007) or the equations presented in paper I.

The most significant independent variables in equations for the crown components (living branches and needles) proved to be tree diameter, height and crown variables (crown length or crown ratio). In addition, the variable describing tree growth rate had an impact on the predicted crown biomass. These variables – diameter, height, crown length and ratio, and growth rate – have commonly been used as independent variables in the crown biomass equations (Marklund 1988, Hakkila 1991, Parresol 1999). In this study, the simplest equations for crown components included no variables related directly to crown dimensions, but height-diameter ratio had a clear impact on crown biomass; increasing height-diameter ratio produced lower crown biomass. This is consistent with the previous studies, which have shown that height-diameter ratio captures effectively the competition status of a tree, which in turn has a strong effect on the dimension and biomass of the tree crown (Holdaway 1986, Marklund 1988, Hakkila 1991, Mäkelä and Vanninen 1998, Mäkinen and Colin 1998). Despite this, the predictions based only on tree diameter and height may be unreliable on sites with a deviated height to diameter ratio, e.g., in pristine mires (a low ratio), and in typical young stands for energy wood thinning (a high ratio), where the pre-commercial thinning has been neglected. The equations also including crown variables such as crown length gave significantly more reliable predictions by decreasing the between-stand error variance. This also implies more reliable predictions for the stand-level crown biomass and for the amount of logging residues as well, which were most unreliably predicted, e.g., by Marklund's equations (1988) equation based on diameter and height was applied (Kärkkäinen 2005). Different treatments such as thinning and fertilization increase tree growth and the proportion of needle biomass within a few years. The effect of these treatments is not directly described in the models, but the positive correlation between tree diameter growth (5-years growth) and needle biomass was utilized in MV-models (3). Hence these models are recommended to be applied to fertilized stands. In terms of random stand and tree error variance, the equations for branch biomass were more reliable compared to the equations for needle biomass. This was the result of the higher inherent variation in needle biomass between stand and tree levels, but also of within-tree variation, i.e., the dependent variable includes higher statistical error caused by sub-sampling (see Paper II).

The applicability of the equations for below-ground biomass is restricted by the limited amount of material. The validity of these equations is more restricted than that of the equations for above-ground biomass. These facts should be kept in mind when applying the equations for root biomass, especially for trees with a diameter > 30 cm, and for trees growing on peatlands, where the root biomass is usual higher than that on mineral soil (Hakkila 1972, Marklund 1988). In addition, the compiled equations predict only coarse-root biomass (diameter > 10 mm), i.e. fine roots (< 2 mm) and part of the coarse roots (2-10 mm) are excluded. This fact can restrict the applicability of the equations for scientific purposes, but not that much for practical forestry, e.g., in the assessment of the amount of stump biomass for energy. Despite these drawbacks, the compiled equations produced logical predictions for the below-ground biomass when compared with the functions of Marklund (1988) and Petersson and Ståhl (2006). Also the share of coarse roots in whole pine tree biomass is comparable with the results produced by Helmisaari (2001) and Vanninen et al. (2006). However, Helmisaari (2001) reported a lower share of the coarse roots in mature pine stands. In our study, trees of two mature stands had pole roots, which partly explains the higher root biomass.

5 CONCLUSIONS

In this study, individual-tree biomass models were derived for Finland. The models produce reliable biomass predictions of the different above- and below-ground tree components in a wide range of site and stand conditions in Finland. The biomass equations for the individual tree components were derived from the same sample trees and estimated simultaneously by applying the multivariate procedure. This approach took account of across-equation correlation (contemporaneous correlations), which had a number of advantages compared to the traditional independently estimated equations, by enabling more flexible application of the equations, ensuring better biomass additivity, and giving more reliable parameter estimates.

Even though the amount of study material was quite large and all the tree components were represented, the validity of the models may be restricted by a deficiency of material. The deficiency of data may cause unreliability in the predictions for birch foliage and for below-ground tree components.

The reliability of the compiled biomass models was improved by constructing the tools to decrease and assess the statistical error of the dependent variables, which were caused in the biomass determination of the sample trees by sub-sampling. The models of paper I offered tools to estimate reliably the average stem-wood density when only a few wood density measurements have been carried out. These models improved the accuracy of wood density estimates in our biomass data. The results of paper II showed that the design-based estimator applied to determine the tree crown biomasses in our biomass data did not produce any systematic trend in errors. Thus the error in crown biomasses could also be interpreted as a random error, which is not a problem in the linear model.

The challenge of further biomass modeling is to expand the applicability of models to more diverse growing conditions. A current need is to test the applicability of the models on peatlands, where the relationships between the tree components may be different; root biomass in particular has been shown to be higher than that on mineral soil. Similarly, the effect of fertilization on biomass allocation should be tested.

REFERENCES

Baskerville, G. L. 1965. Estimation of dry weight of tree components and total standing crop in conifer stand. Ecology, Volume 46, Issue 46:867–869.

Bi, H., Turner, J. and Lambert, M.J. 2004. Additive biomass equations for native eucalyptus forest trees of temperate Australia. Trees 18:467–479.

Björklund, L. 1984. Massavedens torr-rådensitet och dess beroende av olika faktorer. Summary: Pulpwood basic density and it's dependence on different factors. Rapp. Instn. Virkeslära Sveriges Lantbruksuniv. 155:1–29.

Briggs, R.D., Cunia, T., White, E.H. and Yawney H.W. 1987. Estimating sample tree biomass by subsampling: some empirical results. P.119–127 in Estimating tree biomass regressions and their error. Proc. Of the Workshop on Tree biomass regression functions and their contribution to the error of forest inventory estimates. Wharton, E.H. and T. Cunia (comps.) ASDA For Serv. Gen. Tech. Rep. NE-117. 303 p.

Cajander, A. K. 1949. Forest types and their significance. Acta Forestalia Fennica. 56:1–71.

Carvalho, J.P. and Parresol, B.R. 2003. Additivity in the tree biomass components of Pyrenean oak (*Quercus pyrenaica* Willd.). Forest Ecology and Management 179:269–276.

Claesson, S., Sahlén, K. and Lundmark, T. 2001. Functions for biomass estimation of young *Pinus sylvestris*, *Picea abies* and *Betula* spp. from stands in northern Sweden with high stand densities. Scandinavian Journal of Forest Research 16: 138–146.

Cochran, W.G. 1977. Sampling techniques. Third edition. Wiley & Sons. 427 p.

Cunia, T., and R.D. Briggs 1984. Forcing additivity of biomass tables: some empirical results. Canadian Journal of Forest Resources 14: 376–384.

Finér, L. 1991. Effect of fertilization on dry mass accumulation and nutrient cycling in scots pine and on an ombrotrophic bog. Acta Forestalia Fennica 208. 63 p.

Finnish Statistical Yearbook of Forestry. 2011. SVT Agriculture, forestry and fishery 2011. Finnish Forest Research Institute. 470 p.

Gregoire, T.G., 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Canadian Journal of Forest Research 28: 1429–1447.

–, Schabenberger, O. and Barrett, J.P. 1995. Linear modeling of irregularly space, unbalanced, longitudinal data from permanent-plot measurements. Canadian Journal of Forest Resources 25:137–156.

Hakkila, P. 1966. Investigation on the basic density of Finnish pine, spruce and birch wood. Communicationes Instituti Forestalis Fenniae 61(5):1–98.

– 1972. Mechanization harvesting of stumps and roots. Communicationes Instituti Forestalis Fenniae 77(1), 71 p.

– 1979. Wood density surveys and dry weight tables for pine, spruce andbirch stems in Finland. Communicationes Instituti Forestalis Fenniae 96(3), 59 p.

– 1991. Crown mass of trees at the harvesting phase. Folia Forestalia 773, 24 p.

Hansen, M.H., Hurwitz, W.N. and Gurney, M. 1946. Problems and methods of sample survey of business. Journal of American Statistical Assosiation 41: 173–189.

Helmisaari, H. S., Makkonen, K., Kellomäki, S., Valtonen, E. and Mälkönen, E. 2001. Below- and above-ground biomass, production and nitrogen use in Scots pine stands in eastern Finland. Forest Ecology and Management 5695: 1–11.

Holdaway, M.R. 1986. Modeling tree crown ratio. The Forestry Chronicle. 62:451–455.

Kangas, A. 1994. Classical and model-based estimators for forest inventory. Silva Fennica 28(1): 3–14.

Kärkkäinen, L. 2005. Evaluation of performance of tree-level biomass models for forestry modelling and analyses. Finnish Forest Research Institute, Research Papers 940. 123 p.
Kittredge, J. 1944. Estimation of the amount of foliage of trees and stands. Journal of Forestry. 42: 905–912.
Knigge, W. and Schultz, H. 1966. Grundriss der Forstbenutzung. Paul Parey Hamburg, 584 pp.
Kozak, A. 1970. Methods of ensuring biomass additivity of biomass components by regression analysis. Forest Chronology 46(59): 402–404
Laasasenaho, J. 1982. Taper curve and volume functions for pine, spruce and birch. Communicationes Institute Forestalis Fenniae 108. 74 p.
Lappi, J. 1991. Calibration of Height and Volume Equation with Random Parameter. Forest Science, 37(3):781–801.
– 1993. Biometrical Methods for Forest Science (in Finnish). Silva Carelica 24. 182 p.
Lehtonen, A., Mäkipää, R., Heikkinen, J., Sievänen, R. and Liski, J. 2004. Biomass expansion factors (BEFs) for Scots pine, Norway spruce and birch according to stand age for boreal forests. Forest Ecology and Management 188: 211–224.
MacPeak, M.D., L.F. Burkart and D. Weldon. 1990. Comparison of grade, yield, and mechanical properties of lumber produced from young fast-grown and old slow- grown planted slash pine. Forest Products Journal 40 (1), 11–14.
Mäkelä, A. and Vanninen, P. 1998. Impacts of size and competition on tree form and distribution of aboveground biomass in Scots pine. Canadian Journal of Forest Research. 28:216–227.
Mäkinen, H. and Uusvaara, O. 1992. Effect of fertilization on the branchiness and wood quality on Scots pine. Folia Forestalia 801:1–23
– and Colin, F. 1998. Predicting branch angle and branch diameter of Scots pine from usual tree measurements and stand structural information. Canadian Journal of Forest Research. 28:1686–1696
Marklund, G. 1988. Biomass functions for pine, spruce and birch in Sweden. Swedish University of Agricultural Sciences. Department of Forest Survey Report 45. 71 p.
McCulloch, C. E. and Searle, S. R. 2001. Generalized, Linear and Mixed Models. Wiley, New York.
Mergen, F., Burley, J. and Yeatman, C.W. 1964. Variation in growth characteristics and wood properties of Norway Spruce. Tappi 47(8): 499–504.
Monserud, R.A. and Marshall J.D. 1999. Allometric crown relations in three northern Idaho conifer species. Canadian Journal of Forest Research. 29:521–535.
Návar, C., González, B., Graciano, L., Dale, V. and Parresol, B.P. 2004. Additive biomass equations for pine species of forest plantations of Durango, Mexico. Madera y Bosques 10(2):17–28.
Parresol, B.R. 1999. Assessing tree and stand biomass: a review with examples an critical comparisons. Forest Science 45(4): 573–593.
– 2001. Additivity of nonlinear biomass equations. Canadian Journal of Forest Research 31: 865–878.
Petersson, H. and Ståhl, G. 2006. Functions for below-ground biomass of *Pinus sylvestris*, *Picea abies*, *Betula pendula* and *Betula pubescens* in Sweden. Scandinavian Journal of Forest Research. 21(7):84–93.
Repola, J., Ojansuu, R. and Kukkola, M. 2007. Biomass functions for Scots pine, Norway spruce and birch in Finland. Working Papers of the Finnish Forest Research Institute 53. 28 p.
Saranpää, P. 1983. The influence of basic density and growth ring width on the impact strength of spruce wood from south and north Finland. Silva Fennica 17(4):381–388.

Satoo, T. and Madgwick, H.A.I. 1982. Forest biomass. Forestry Sciences. Martinus Nijhoff/ Dr. W. Junk Publisher, The Hague. United Nations Framework Convention on Climatic Change, UNCCC 1993.

Srivastava, V.K. and Giles, D.E.A. 1987. Seemingly unrelated regression equations models: Estimation and inference. Marcell Dekker, New York. 374 p.

Tamminen, Z. 1962. Fuktighet, volymvikt, m.m. hos ved och bark. I. tall. Summary: Moisture content, density and other properties of wood and bark. I Scots pine. Rapp. Uppsats. Instn. Virkeslära Skogshögsk. 41:1–118.

Uusvaara, O. 1974. Wood quality in plantation-grown Scots pine. Communicationes Instituti Forestalis Fenniae 80(2):1–105.

Valentine, H.T., Tritton, L.M. and Furnival, G.M. 1984. Subsampling trees for biomass, volume or mineral content. Forest Science 30:673–681.

Vanninen, P., Ylitalo, H., Sievänen, R. and Mäkelä, A. 1996. Effects of age and site quality on the distribution of biomass in Scots pine (*Pinus sylvestris* L.). Trees 10, 231–238.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated relations and tests for aggregation bias. Journal of the American Statistical Association 57, 348–367.