Generic biomass functions for Norway spruce in Central Europe—a meta-analysis approach toward prediction and uncertainty estimation

CHRISTIAN WIRTH,[1–3] JENS SCHUMACHER[1] and ERNST-DETLEF SCHULZE[1]

[1] *Max-Planck-Institute for Biogeochemistry, Jena, Hans-Knöll-Str. 10, 07745 Jena, Germany*

[2] *Present address: Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA*

[3] *Author to whom correspondence should be addressed (cwirth@bgc-jena.mpg.de or cwirth@princeton.edu)*

Summary   To facilitate future carbon and nutrient inventories, we used mixed-effect linear models to develop new generic biomass functions for Norway spruce (*Picea abies* (L.) Karst.) in Central Europe. We present both the functions and their respective variance–covariance matrices and illustrate their application for biomass prediction and uncertainty estimation for Norway spruce trees ranging widely in size, age, competitive status and site. We collected biomass data for 688 trees sampled in 102 stands by 19 authors. The total number of trees in the "base" model data sets containing the predictor variables diameter at breast height (D), height (H), age (A), site index (SI) and site elevation (HSL) varied according to compartment (roots: $n = 114$, stem: $n = 235$, dry branches: $n = 207$, live branches: $n = 429$ and needles: $n = 551$). "Core" data sets with about 40% fewer trees could be extracted containing the additional predictor variables crown length and social class. A set of 43 candidate models representing combinations of $\ln D$, $\ln H$, $\ln A$, SI and HSL, including second-order polynomials and interactions, was established. The categorical variable "author" subsuming mainly methodological differences was included as a random effect in a mixed linear model. The Akaike Information Criterion was used for model selection. The best models for stem, root and branch biomass contained only combinations of D, H and A as predictors. More complex models that included site-related variables resulted for needle biomass. Adding crown length as a predictor for needles, branches and roots reduced both the bias and the confidence interval of predictions substantially. Applying the best models to a test data set of 17 stands ranging in age from 16 to 172 years produced realistic allocation patterns at the tree and stand levels. The 95% confidence intervals (% of mean prediction) were highest for crown compartments (~ ± 12%) and lowest for stem biomass (~ ± 5%), and within each compartment, they were highest for the youngest and oldest stands, respectively.

Keywords: allocation, allometry, competition, expansion factor, mixed-effect models, model selection, Picea abies, uncertainty.

Introduction

Historically, studies on whole-tree biomass in Europe were motivated by interest in (1) assessing forest resources other than stem wood (Grundner and Schwappach 1952, Fiedler 1987, Marklund 1987, Hakkila 1989) and (2) understanding how stem increment is determined by the mass and surface area of the assimilatory apparatus (e.g., Schmidt 1949, Burger 1953, Schöpfer 1961). Since the mid-1960s, the reasons for tree biomass studies have become more diverse. With the rise of ecosystem research, which was catalyzed by the International Biological Program, detailed allometric studies of tree crowns and root systems were needed to scale up branch-level gas exchange (Droste zu Hülshoff 1969, Schulze et al. 1977), quantify nutrient and energy flow associated with growth and mortality of tree organs (Duvigneaud and Denaeyer-De Smet 1970, Heller and Göttsche 1986), and parameterize allocation rules in forest growth models (Running and Gower 1991).

Recently, the need for information on whole-tree biomass allocation patterns has increased with the advent of the Kyoto Protocol under the United Nations Framework Convention on Climate Change (UNFCCC). Articles 3.3 and 3.4 of the protocol require measurement of carbon exchange as part of the greenhouse gas accounting by nations claiming biospheric sinks to meet their commitments to reduce CO_2 emissions (Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen (WBGU) 1998, Intergovernmental Panel on Climate Change (IPCC) 2000). One way of quantifying carbon exchange in forests is to follow the change in carbon stocks over time. Tree biomass, an important component of a forest's carbon pool, may be estimated using stem volume provided by national forest inventories as input (Burschel et al. 1993, Brown 2002) or by linking biomass functions directly to tree-level data of forest inventories (Hall et al. 2001). In the first case, stem volume is converted to whole-tree biomass based on biomass expansion factors that are calculated as the stand-level ratio of whole-tree biomass to stem volume. Although the denominator of this ratio may be provided by standard inventories, the numerator must be estimated by biomass functions. Thus, in both calculations, generic biomass functions are a prerequisite for carbon inventories in forests.

Biomass allocation patterns can be measured only at the tree level. In a subsequent step, biomass functions are established to scale up tree-level biomass to the whole stand (Satoo 1970). These functions relate the biomass of sample trees to dimensions that are easy to measure for a large number of trees such as diameter at breast height (*D*), tree height (*H*) and crown length (*C*). There are comprehensive collections of stand-specific biomass functions (e.g., Ter-Mikaelian and Korzukhin 1997). However, such stand-specific functions cannot be used for scaling-up biomass to the regional level where several age classes and structural types of stands coexist. For example, if the regression method is used, as in most studies, sample trees are selected to represent the whole range of the diameter distribution of a stand to avoid extrapolation errors. In Norway spruce, almost all biomass studies were conducted in even-aged stands, and hence any difference in size was related to the competitive status rather than the age of a tree. However, a suppressed tree in an older stand having the same approximate size as a dominant tree in a younger stand exhibits significantly different allocation patterns. In this situation, progress is possible only by pooling data from many stand-level studies and by using a larger set of predictors that, if combined, may discriminate between a range of growth situations. This versatility and comprehensiveness may make the resulting models suitable also for quantification of biomass in highly structured multi-species stands and for providing constraints for process-based stand growth models (Mäkelä et al. 2000).

In Central Europe, Norway spruce (*Picea abies* (L.) Karst.) covers about 35% of the forested area (Körner et al. 1993, Polley 1994). Based on a meta-data set comprising biomass data of 688 trees harvested in 102 stands in five Central European countries by 19 research groups, we developed new generic allometric models. A comparable database of tree-level biomass for Norway spruce exists only for Scandinavia and has been analyzed by Marklund (1987). For the biomass compartments needles, branches, dry branches, stem and roots, we developed allometric functions using mixed-effect models as a meta-analysis tool. We outline how these functions can be applied to calculate confidence intervals for individual tree predictions as well as for stand-level estimates.

Methods

Constructing the database

The database comprises both a stand-level table containing information on location, climate, substrate and biometric stand characteristics and a tree-level table containing data on biomass compartments and various tree-level predictors. Currently, the database holds entries for 688 trees, originating from 102 stands located in five Central European countries (Belgium, Denmark, Germany, Czech Republic and Switzerland), which were analyzed for total or compartment biomass by 19 authors (Figure 1, Table 1). We considered only studies that provided the following minimum set of five predictors: three basic variables, diameter at breast height (*D*; cm), height (*H*; m) and age (*A*; years); and two site variables, site index

(SI; mean height of trees at 100 years in m) and height above sea level (HSL; m). For each of the five compartments, smaller subsets could be extracted that, in addition to the basic and site variables, contained two variables that represented competition, namely crown length (*C*; m) and the position of the crown in relation to the main canopy or social status (SOC; categorical, see Burschel and Huss 1997). We call the large data sets "base" data sets and the smaller but more comprehensive data sets "core" data sets.

Climate data were not available for all sites, but elevation (HSL) was used as a proxy for climate, because it represents a dominant factor differentiating the temperature and precipitation regime in Central Europe. For the subset of sites with available climate data (*n* = 48), HSL was correlated positively with annual precipitation (*r* = 0.85) and negatively with mean annual temperature (*r* = –0.81). Further, only trees from sites less than 1400 m a.s.l. were included. In some cases, the social status of a sample tree was inferred by comparing its dimensions with the height and diameter distribution of the sample stands. If not specified, SI was reconstructed from stand age and mean stand height based on the yield table for Norway spruce (M-System) published by Wenk et al. (1985).

The studies differed with respect to the set of biomass compartments considered, the methods of destructive sampling and the strategy of tree selection. However, they all met the following criteria: (a) the mass of needles and woody compartments (excluding stem wood) or subsamples of these were
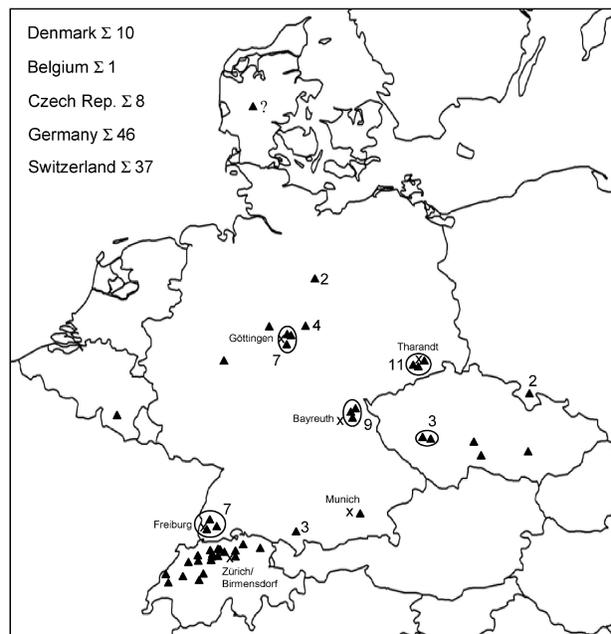

Figure 1. Map of Central Europe showing locations of study sites (▲). Figures next to sites or site clusters indicate how many stands per study site were investigated. Locations of some forest faculties are shown as crosses (×) to illustrate the clumped distribution of study sites around research facilities. The exact locations of the Danish stands were not reported.

Table 1. Methods used and compartments sampled by 19 authors to determine biomass at the tree and stand level. Abbreviations: n = needles; br = branches (often further divided into green and non-green parts); dbr = dead branches; st = stem including bark; r = roots including all compartments down to the size class in (mm) given in column no. 7; r* = only a few trees were analyzed for root biomass; excav. = excavated; extract. = extracted; nn = not necessary; nd = not determined; and ni = not indicated.

Study	No. of sample trees considered	Compartments measured	No. of crown layers	Method of branch sampling	Upscaling from branch level to crown level	Method of root sampling	Upscaling from tree level to stand level
Burger 1953	160	n, br	Undivided	Whole crown	nn	nd	–
Burger 1937	6	n	Undivided	Whole crown	nn	nd	–
Cerny 1990	26	n, br, dbr, r*	5	Several per layer	Mean per layer	Excav. = 2	Regression
Dietrich 1968	20	n, br, dbr	2	Branches of 5 whorls	Mean per layer	nd	Average tree
Drexhage 1994, Drexhage and Gruber 1999	15	n, r	5	Whole crown	nn	Excav. = 5	Regression
Droste zu Hülshoff 1969	5	n, br, dbr, st	3	Every 12th branch	Regression	nd	Stratified average
Duvigneaud et al. 1977	6	n, br, dbr, st	Undivided	Whole crown	nn	nd	Regression
Fiedler 1986, 1987	19	n, br, dbr, st	3	Every 2nd branch	Regression	nd	Regression
Heller and Göttsche 1986	18	n, br, st, r*	Undivided	Whole crown	nn	Excav. ni	Regression
Hesse 1990	3	n, br, dbr, st	4	Every 5th branch	Mean per layer	nd	Average tree
Kuhr 1999	25	n, br	Undivided	1 per tree (small trees)	Mean per tree	nd	–
Lee 1998	18	n, r	5	Whole crown	nn	Extract. = 5	Regression
Möller 1945	100	n	Undivided	Whole crown	nn	nd	–
Mund et al. 2002	38	n, br, dbr, st, r	3	Every 10th branch	Mean per layer	Extract. = 2	Regression
Poeppel 1989	17	n, br, dbr, st	3	One per layer	Mean per layer	nd	Regression
Raisch 1983	9	n, br, dbr, st	2	Several per layer	Mean per layer	nd	Average tree
Sharma 1992	45	n, br, dbr, st	3	Two per layer	Regression	nd	Regression
Vins and Sika 1981	3	n, br, dbr, st, r	5	Two per layer	Mean per layer	Excav. = 1	–
Vyskot 1981	45	n, br, dbr, st, r	3	One per layer	Mean per layer	Excav. < 1	Stratified average
Vyskot 1990	5	n, br, dbr, st	3	One per layer	Mean per layer	nd	Stratified average
Zimmermann 1985	4	n	3	Whole crown	nn	nd	–

directly measured (not the volume, cf. Grundner and Schwappach 1952); (b) needles and shoots were analyzed separately; and (c) stem mass was either directly measured or its volume was converted to mass based on wood density measurements on the same tree. Because the definition of size classes of roots and branches varied among studies, we had to subsume the various biomass compartments reported in the literature in one of five categories: (1) roots including the stump, (2) stem, (3) dry branches, (4) branches and (5) needles. All woody components include bark, because in many publications, bark biomass was not reported separately. Although the definitions are clear-cut for the aboveground categories, this is not the case for belowground biomass. In natural stands, fine roots are difficult to assign to individual trees. This is especially true for large trees where the root systems cannot be excavated as a whole, but have to be extracted in pieces with a cable winch. Also, among studies, definitions varied greatly and the lower diameter threshold was not always reported (see Table 1). Before model selection, data within the five groups were carefully compared according to authors and checked for outliers by graphical analysis. A file listing all sites with location names, geographical coordinates, climatic and edaphic data can be downloaded from ftp://panorama.bgc-jena.mpg.de/pub/science/cwirth/Wirth_ et_al_spruce_sites.doc.

Based on 17 fully inventoried stands ranging in age from 16 to 173 years, we compiled a test data set comprising 1985 trees with known diameter, height and age. In one of these, a 60-year-old stand called Wetzstein, crown length was also measured. This data set was used (1) to explore the ability of the functions to produce realistic allocation patterns, (2) to quantify the uncertainty of the predictions at the tree and stand levels, and (3) for comparison with Marklund's biomass functions for Sweden (Marklund 1987, 1988).

Rationale of model building

Candidate models   First, we established a candidate set of 43 models based on the minimum set of the five continuous predictors, i.e., the three basic variables D, H and A and the two site variables SI and HSL. The models differed in complexity, ranging from simple allometric models to more general model formulations with site-specific predictors, second-order polynomial terms or interaction terms added. We did not apply data-mining techniques such as automated procedures of best subset selection, but restricted the choice to models that had been successfully applied in the past or represented biologically meaningful extensions of such models. The tree-level predictors (D, H, A) were transformed using the natural logarithm, whereas the stand-level predictors SI and HSL were left untransformed. The transformations were performed to achieve linearization of the underlying allometric relationships and to ensure homogeneity of residual error variances. Second-order polynomial terms were considered for $\ln D$ and $\ln H$, but only in connection with the first-order terms. Because D is the variable with the highest predictive power in allometric models, $\ln D$ was contained in all tested models.

Incorporation of data sets produced by different authors into one comprehensive analysis introduces some heterogeneity that particularly affects the assessment of the accuracy of the resulting predictions (Crawley 2002). Estimates of the uncertainty of predicted values rely heavily on the assumption of independence of residuals from the fitted model, an assumption hardly met if data from different authors are combined. Rather, different measurement methodologies or peculiarities of the investigated stands presumably cause consistent deviation of observed data values of the same author from the fitted model.

One way to address this type of data heterogeneity adequately is the use of mixed models (Pinheiro and Bates 2000). Instead of assuming one global linear model:

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi} + \varepsilon_i \qquad (1)$$

with $p + 1$ unknown, but fixed, regression coefficients β_l ($l = 0, ..., p$) and p predictors x_{li} (with $l = 1, ..., p$ and $i = 1, ..., n$ trees), we additionally allow for slight random fluctuations in the regression coefficients from author to author as:

$$y_{ij} = (\beta_0 + b_{0j}) + (\beta_1 + b_{1j})x_{1ij} + ... + (\beta_p + b_{pj})x_{pij} + \varepsilon_{ij}$$
$$j = 1, ..., J,$$
$$i = 1, ..., n_j, \sum_{j=1}^{J} n_j = n$$

where the additional index j indicates the different authors and b_{lj} is the random deviation from the mean regression coefficient β_l due to author j. Separating fixed and random parts of the model gives:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + ... + \beta_p x_{pij} + b_{0j} + \beta_{1j} x_{1ij} + ... + b_{pj} x_{pj} + \varepsilon_{ij}$$

or in vector notation:

$$y_{ij} = \mathbf{x_{ij}}\boldsymbol{\beta} + \mathbf{x_{ij}}\mathbf{b_j} + \varepsilon_{ij}$$

with $\mathbf{x_{ij}} = (1, x_{1ij}, ..., x_{pij})$. If random fluctuations are incorporated only for a subset of predictor variables (because, e.g., for some predictors the differences are considered negligible), this subset is subsumed in a vector $\mathbf{z_{ij}}$, resulting in the linear mixed model (LMM, Pinheiro and Bates 2000)

$$y_{ij} = \mathbf{x_{ij}}\boldsymbol{\beta} + \mathbf{z_{ij}}\mathbf{b_j} + \varepsilon_{ij} \qquad j = 1, ..., J, \qquad (2)$$
$$i = 1, ..., n_j$$

where $\boldsymbol{\beta}$ is the $(p + 1)$-dimensional vector of fixed effects and $\mathbf{b_j}$ the q-dimensional ($q < p + 1$) vector of random effects associated with author j. Besides the usual assumptions of independence, normality and variance-homogeneity of the residual errors ($\varepsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$), the random effects are considered to be independent of the residual error, independent for the different authors, and normally distributed with covariance matrix $\boldsymbol{\Psi}$ ($\mathbf{b_j} \sim \mathrm{N}(0, \boldsymbol{\Psi})$). Thus, random effects for the same author are not assumed to be independent. This is a reasonable approach, because even in the case of a simple linear regression,

a slight change in the slope will affect the corresponding intercept. It is important to note that the LMM is not equivalent to separate linear fits for each author. When fitting an LMM, the individual estimates for each author ($\beta_{lj} + b_{lj}$) tend to be pulled toward the fixed-effect estimates β_{lj} compared with estimated coefficients from separate regressions (Pinheiro and Bates 2000). The diagonal elements of Ψ are the variances of the random effects, thus describing the degree of variability of the respective regression coefficient from author to author. The off-diagonal elements are covariances between different random effects. A large number of random effects implies a large number of variance-covariance parameters, thus increasing model complexity. Preliminary inspection of our candidate models revealed that random effects in higher-order terms and interaction terms were negligible. We therefore decided to include random effects only for the intercept and first-order terms ($\ln D$, $\ln H$, $\ln A$). Each of the 43 models was fitted separately for each biomass compartment using the maximum-likelihood method as implemented in the lme-procedure in S-PLUS 6.0 (Insightful, Hampshire, U.K.).

Model selection   Selecting a model is always a compromise between (1) including as many predictors as possible to reduce bias in the predictions and (2) keeping the set of predictors as small as possible to reduce the variability of predictions. The ultimate validation of a fitted model is the comparison of predicted values with new data. This ideal is seldom achieved, because the difficulty in obtaining data is often the reason for searching for a prediction model. Cross-validation techniques, where the available data set is split into a training data set for model fitting and an assessment data set for evaluating the predictive power of the fitted model, offer an alternative approach. The random data split is repeated thousands of times, resulting in estimates of mean prediction error for all the candidate models. In view of the large number of candidate models, we decided to use a less computer-intensive criterion, the Akaike Information Criterion (AIC). We opted for AIC because it explicitly addresses a highly relevant question in our prediction problem: it estimates the expected information loss when the unknown truth is approximated by a prediction model (Burnham and Anderson 1998). The value of AIC can easily be calculated from the results of model fitting, because AIC = -2 log-likelihood $+ 2P$, where P is the number of parameters in the fitted model (including σ^2 and the elements of the variance-covariance matrix Ψ). The best model is the one with the smallest value of AIC. This definition clearly supports the interpretation of AIC as a trade-off between model fit and complexity. We note that model comparison using AIC is possible only if the model fits are based on exactly the same data set and the parameters are estimated with the maximum-likelihood method.

Best models   We report best models ("best") at three levels of complexity. (1) The best candidate of models containing only combinations of the basic variables $\ln D$, $\ln H$ and $\ln A$ (termed DHA models). (2) The best candidate of models containing combinations of the basic variables and site variables SI and

HSL (termed DHAS models). (3) The best candidate DHA and DHAS models with the competition variable crown length C (as $\ln C$) added (termed DHA+C or DHAS+C model). (The effect of SOC was also tested by including it as a dummy-coded categorical variable, distinguishing four groups (dominants, co-dominants, intermediates and suppressed). However, the coefficients are not reported. Note that the dummy representation of SOC in four groups adds three new parameters to the models.)

For the sake of applicability, we also present simplified models ("simp") containing only $\ln D$ as a predictor (simp D) or only first-order terms of the basic variables $\ln D$, $\ln H$ and $\ln A$ (simp DHA). The latter form of the models should be used if extrapolation beyond the range of the predictors in the regression design matrices is necessary.

To illustrate the differences in predictive performance of different models, we estimated the aggregate prediction error (APE, Davison and Hinkley 1997) of the best models by cross-validation. The APE quantifies both precision and bias. For each cross-validation run, the core data set was randomly split into a training set containing 90% of the data and a validation set containing the remaining 10%. The best models were then fit based on the training data set and used to predict the responses in the validation set. An estimate of APE was obtained by averaging the mean quadratic prediction error over 100 cross-validation runs:

$$\text{APE} = \frac{1}{100}\sum_{i=1}^{100}\frac{1}{n_{\text{valid}}}\sum_{\text{validation set } i}(y_{\text{observed}} - y_{\text{predicted}})^2$$

based on ln-transformed y values.

Application of models for prediction and uncertainty assessment

Our starting point is the prediction of biomass and quantification of the prediction error for single compartments of an individual tree. To this end, the best model has to be applied to the set of new predictors, and the result must be back-transformed to the original scale. Among the different approaches to cope with the bias induced by nonlinear back-transformation, the nonparametric smearing estimate (Duan 1983) involves the fewest assumptions. Unfortunately, this estimate is not directly applicable to the case of mixed linear models. We therefore developed a modified smearing estimate that could be applied in this slightly more complex situation. The derivation is presented in Appendix 1.

To assess the precision of the modified smearing estimate, we start by estimating the variance of newly predicted values on the ln-transformed scale. According to Model 2 (Equation 2), this variance is given by:

$$\begin{aligned}\text{Var}(\hat{y}_{\text{new}} - y_{\text{new}}) &= \text{Var}(\mathbf{x}_{\text{new}}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) - \mathbf{z}_{\text{new}}\mathbf{b}_{\text{new}} - \varepsilon_{\text{new}}) \\ &= \mathbf{x}_{\text{new}}\text{Var}(\hat{\boldsymbol{\beta}})\mathbf{x}_{\text{new}}^T + \mathbf{z}_{\text{new}}\Psi\mathbf{z}_{\text{new}}^T + \sigma^2\end{aligned}$$

which shows partitioning resulting from the different sources of uncertainty. The first summand on the right-hand side de-

scribes the variance component due to the estimation of the fixed effects, the second is caused by uncertainties associated with the differences between authors, and the third accounts for the residual error. Estimates of the variance components resulting from maximum likelihood fits are known to underestimate true variances (Crawley 2002). We therefore refitted the best models using the restricted maximum likelihood method (REML) to obtain estimates of $\text{V\^ar}(\hat{\boldsymbol{\beta}})$, $\hat{\boldsymbol{\Psi}}$ and $\hat{\sigma}^2$. These estimates are available by anonymous ftp from http://panorama.bgc-jena.mpg.de/pub/science/cwirth. The estimated variance of a new response is thus given by:

$$\text{V\^ar}(\hat{y}_{\text{new}}) = \mathbf{x}_{\text{new}}\text{V\^ar}(\hat{\boldsymbol{\beta}})\mathbf{x}_{\text{new}}^T + \mathbf{z}_{\text{new}}\hat{\boldsymbol{\Psi}}\mathbf{z}_{\text{new}}^T + \hat{\sigma}^2 \quad (3)$$

and an approximate 95% confidence interval for the new response is:

$$\hat{y}_{\text{new}} - 1.96\sqrt{\text{V\^ar}(\hat{y}_{\text{new}})},\ \hat{y}_{\text{new}} + 1.96\sqrt{\text{V\^ar}(\hat{y}_{\text{new}})} \quad (4)$$

exploiting asymptotic normality of the prediction $\hat{y}_{\text{new}}$. Back-transformation to the original scale results in an asymmetric 95% confidence interval for the predicted biomass compartment for one individual tree. Bias correction is unnecessary because the nonlinear transformation has no effect on the coverage probability of the interval. An example of how to calculate confidence limits is shown in Appendix 2. A similar approach is necessary to arrive at the 95% confidence interval for stand-level predictions of biomass. To this end, variance at the logarithmic scale $\text{V\^ar}(\hat{y}_{\text{new}})$ must be converted to the original scale $\text{V\^ar}(\exp(\hat{y}_{\text{new}}))$, which can be additively propagated to the stand level. Consequences of possible overestimation of the variances (e.g., too conservative confidence intervals) are less severe than prediction bias, so we used the formula:

$$\text{V\^ar}(\exp(\hat{y}_{\text{new}})) = \exp(2\hat{y}_{\text{new}} + \text{V\^ar}(\hat{y}_{\text{new}}))(\exp(\text{V\^ar}(\hat{y}_{\text{new}})) - 1)$$

(Parresol 2001), which relies on normality of residuals on the logarithmic scale. Assuming independence of predictions for individual trees, the variance of the stand-level sum of the biomass predictions $\text{V\^ar}(\text{total})$ is calculated by adding the estimated variances for all K trees of a stand according to:

$$\text{V\^ar}(\text{total}) = \text{V\^ar}\left(\sum_{k=1}^{K}\exp(\hat{y}_{\text{new}})_k\right) = \sum_{k=1}^{K}\text{V\^ar}(\exp(\hat{y}_{\text{new}}))_k$$

An approximate 95% confidence interval is calculated analogously to Equation 5. Because the variances of the five biomass compartments could not be considered independent, the calculation of the 95% prediction interval for whole-tree biomass required knowledge of the correlation structure of the errors. This was obtained based on the 78 trees for which measurements were taken for all biomass compartments. The error propagation was done by Monte Carlo simulation.

Results

The range of the basic predictors D, H and A and the dependent variable W varied among compartments (Table 2). The largest ranges of predictor values were present in the base data sets of the crown compartments. Here, diameters ranged from 1.8 to 67.6 cm, height from 2.1 to 42.8 m and age from 13 to 157 years. Smaller ranges of predictor values were present in the data sets of dry branches and stems ($D = 3.5$–52.8 cm, $H = 4.2$–33.4 m and $A = 13$–148 years), although the medians exceeded those of the crown compartments by 10 to 20%. Lowest ranges for D and H were present in the root data set ($D = 5.1$–52.8 cm, $H = 5.7$–32.4 m), which therefore defines the limits for whole-tree biomass prediction.

There was considerable spread in the relationships between basic predictor variables (Figure 2). For example, 100-year-old trees exhibited diameters ranging from 15 to 60 cm, and

Table 2. Summary statistics of the basic variables (D, H and A) and the dependent variable (biomass) of the data sets used for the model fitting. Abbreviations: n = number of individual trees in the data set; Min = minimum; Max = maximum; Med = median; and Skew = skewness. All woody compartments include bark.

Compartment	Data set[1]	n	Diameter at 1.3 m, D (cm)				Height, H (m)				Age, A (years)				Biomass, W (kg dry mass)			
			Min	Max	Med	Skew	Min	Max	Med	Skew	Min	Max	Med	Skew	Min	Max	Med	Skew
Needles	Base	551	1.8	67.6	17.2	1.16	2.1	42.8	17.2	0.31	13	157	52	1.04	0.38	131.9	9.0	2.3
	Core	388	1.8	67.6	17.0	1.18	2.1	42.8	18.1	0.27	14	157	51	1.17	0.50	131.9	8.9	2.3
Branches	Base	429	1.8	67.6	17.2	1.12	2.1	42.8	17.9	0.25	13	157	53	1.02	0.30	372.4	11.2	3.7
	Core	384	1.8	67.6	17.4	1.14	2.1	42.8	18.4	0.24	15	157	53	1.18	0.35	372.4	11.6	3.6
Dry branches	Base	207	3.5	52.8	18.6	0.71	4.2	33.4	19.3	−0.08	13	148	58	0.72	0.05	50.8	5.0	2.5
	Core	167	5.1	52.8	19.0	0.82	5.9	33.4	19.5	0.14	15	148	60	0.90	0.72	50.8	5.3	2.4
Stem	Base	235	3.5	52.8	21.0	0.63	4.2	33.4	20.7	−0.28	13	148	62	0.35	1.05	1353.8	136.6	2.0
	Core	173	5.1	52.8	21.2	0.74	5.7	33.4	20.8	−0.13	15	166	58	1.06	2.69	1353.8	140.2	2.0

[1] The base data set comprises all entries that contain at least the minimum set of predictors (i.e., basic variables and site variables), and the smaller core data set comprises all entries that in addition contain the variables related to competition, crown length and the social status of the trees.
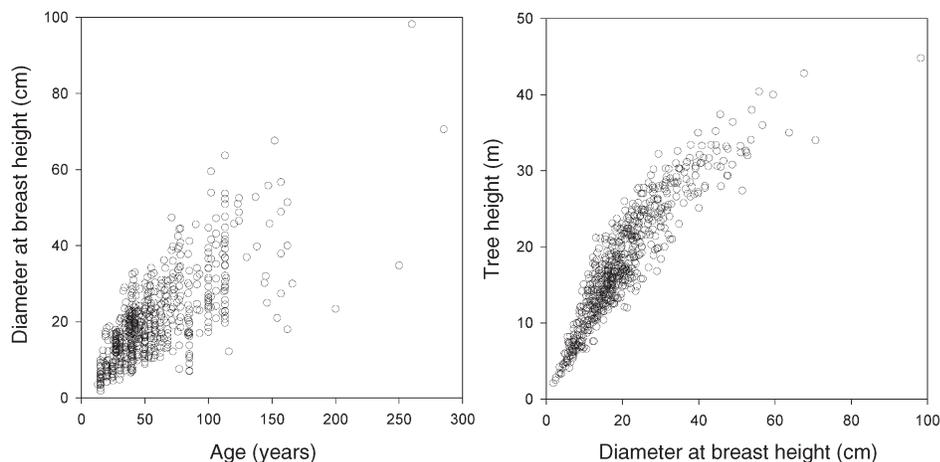
Figure 2. Diameter–age and height–diameter relationships for 688 trees in the database.

trees of 20 cm diameter could have reached a height of 12 to 25 m. It is evident that multicollinearity is present; however, as discussed later, this does not affect the precision and accuracy of the prediction.

Although all functions are multidimensional, it is instructive to look at plots of biomass versus diameter (Figure 3). Variability in the data of the crown compartments was much higher than in the data of stem and roots. The largest scatter was observed for dry branch biomass, which varied by almost an order of magnitude at a given diameter. In general, as tree size increases, the number of data points becomes less and the variance increases. This is especially true for needle and branch biomass. Taking the logarithms of D and W, one obtains a near-linear relationship and reduces the heterogeneity of variance along the predictor axis (insets in Figure 3). We note that, in the case of needles and branches, complete homogeneity of the variances cannot be achieved if only D is used as a predictor.

For the branch and root compartments, the highest ranking models (i.e., those with the lowest value of AIC) contained only four tree-specific parameters, whereas the best model for predicting needle biomass was more complex and contained eight parameters including both tree- and site-specific parameters (Table 3). For all compartments except dry branches, the simplest allometric model with only diameter as predictor ("simp D" in Table 3) exhibited much higher values of AIC and RMSE than the six best candidate models. The range of AIC values covered by the best-subset models was generally small except for the needle compartment, where the inclusion of SI and HSL improved the model substantially. For the needle, stem and root compartments, D also entered the model as a quadratic term, indicating that the allometric exponent of D is not constant over the whole range of ages and sizes in our regression design matrix. Visual inspection of biomass of these compartments versus diameter on a log–log scale suggested lower allometric exponents with decreasing D for needles and higher allometric exponents with decreasing D for bole and roots (see insets in Figure 3).

The finding that the D coefficient was always positive, whereas the coefficients for A and H were negative in all except two models (Table 4), underscores the role of D as the most important predictor. When combined with D, the information added by H and A is mainly a characterization of the competitive environment of a tree. For two trees with the same diameter, the taller and therefore more slender tree favored height growth in relation to diameter growth and crown development. Trees competing for light usually exhibit a poorly developed crown (Vanninen and Mäkelä 2000, Wirth et al. 2002). The negative sign of the A coefficient if D and H are also present as predictors can be explained in the following way. For two trees having the same diameter and height, the older tree took longer to attain the given stature, which probably reflects development under competitive pressure or in a less favorable environment. Predictor H did not enter the model for roots.

The smaller core data set also contained C as a quantitative predictor and SOC as a qualitative predictor. Unlike the basic variables, C directly accounts for crown geometry, whereas SOC directly accounts for the competitive effect. For the crown compartments (needles and branches), the best models having the lowest AIC (numbers with asterisks in Table 5) were those that included both variables describing competition. The same was also true for the root compartment, suggesting a functional balance during root and crown development. Thus, for these three compartments, the information contained in the combination of diameter, height and age was insufficient to capture the effect of competition on allocation patterns. However, including one or both variables to account for competition led to a deterioration of the model for the stem, but had very little effect on the dry branch compartment. The effect of adding crown length as a predictor was generally much greater than that of adding social status.

Model evaluation

Cross-validation showed that, irrespective of the compartment, the model containing only D as a predictor (the allometric model) exhibited the largest aggregate prediction errors
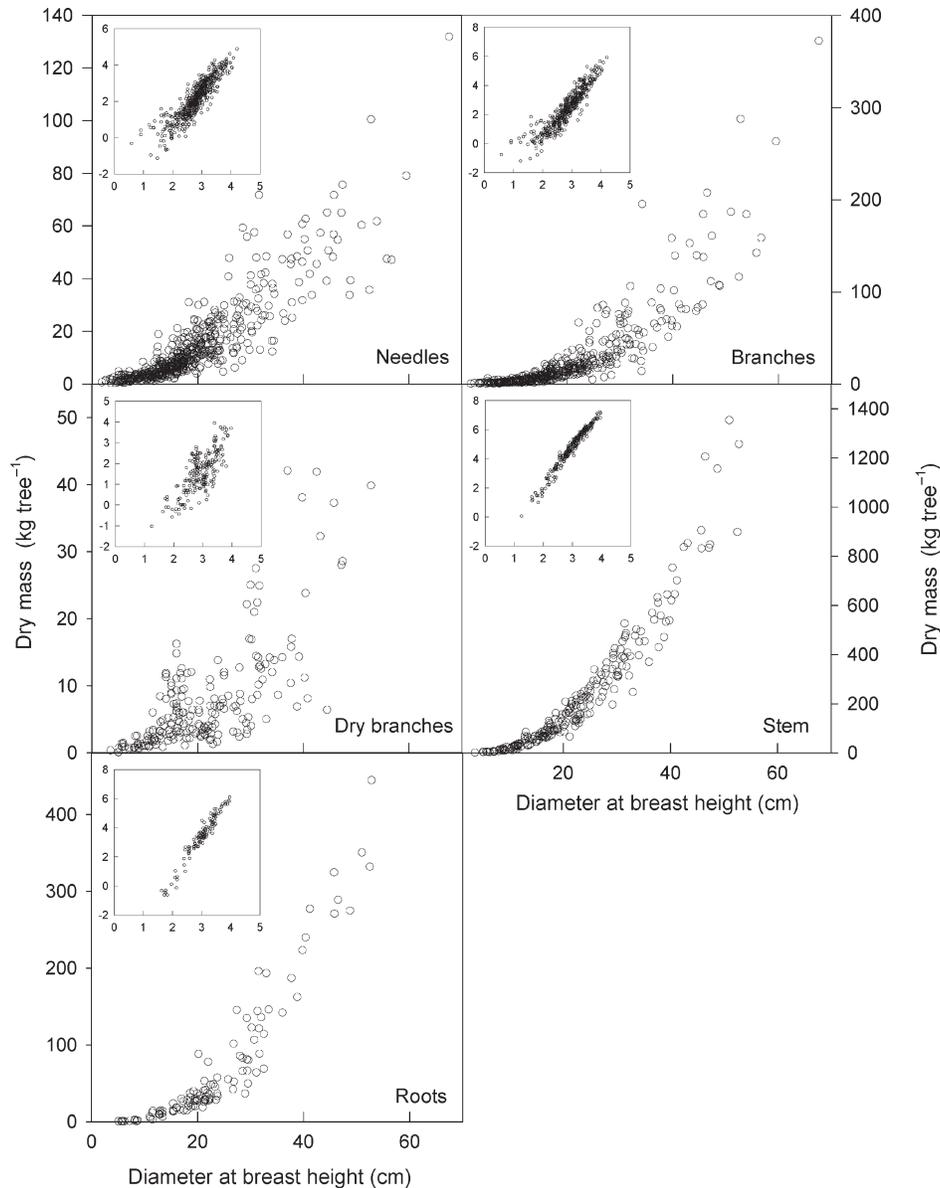
Figure 3. Biomass per tree of needles, branches, dry branches, stem and roots against diameter at breast height on the original scale. Note that the size of the data sets varies among compartments. Insets show log–log plots.

(APE) and thus the least predictive power (Table 6). With the best DHA model, the APE for needles and branches was about 30% lower than for the simple allometric model. Compared with crown compartments, the overall APE for the stem was much smaller and the relative improvement (66%) by using the best DHA model instead of the *D* model was higher. A further decrease in APE was achieved for needles if the best model containing crown length as a predictor was used. In contrast to the results presented in Table 5, crown length did not reduce APE for branches.

In a different approach to model evaluation, we calculated the 95% prediction intervals for stand-level biomass of needles, branches and stems in a 60-year-old Norway spruce stand of 108 trees with a known distribution of diameter, height and crown length. The prediction intervals were calculated using different models and different model data sets. Fig-

ure 4 illustrates the superiority of the models including crown length for predicting needle and branch biomass. Compared with the full DHA+C model, the simplest model with only the predictor *D* (*D* model) yielded a 50% higher prediction interval for needle and branch biomass, whereas the DHA models took an intermediate position. In contrast to the results presented in Tables 3 and 6, the DHAS and DHAS+C models, which included site-related predictors, had higher prediction intervals for needle biomass than their tree-level counterparts DHA and DHA+C. As expected, lower prediction intervals resulted for both needle and branch biomass if models were fit with the larger base data set (Figure 4 hatched bars).

Analysis of allocation patterns

Of 18 stands with data available for *D*, *H* and *A* for every tree, we chose the smallest 10% and largest 10% of the population

Table 3. The six best candidate models for each biomass compartment (DHA and DHAS models not separated). To illustrate the importance of adding variables other than diameter, statistical indicators of the simplest allometric model ($\ln W = \beta_0 + \beta_1 \ln D$), termed "simp D," are listed as well.

Score	Compartment	Type[1]	AIC[2]	RMSE[3]	Model structure[4]
1	Needles	DHAS	440.6	0.330	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2 + \beta_5 \ln A + \beta_6 \mathrm{SI} + \beta_7 \mathrm{HSL}$
2		DHAS	443.7	0.331	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2 + \beta_5 \ln A + \beta_6 \mathrm{SI}$
3		DHAS	456.2	0.336	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2 + \beta_5 \ln A + \beta_6 \mathrm{HSL}$
4		DHAS	459.2	0.336	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 \ln A + \beta_5 \mathrm{SI} + \beta_6 \mathrm{HSL}$
5*		DHA	462.2	0.338	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2 + \beta_6 \ln A$
6		DHAS	462.7	0.337	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 \ln A + \beta_5 \mathrm{SI}$
		simp D	663.7	0.420	$\ln W = \beta_0 + \beta_1 \ln D$
1*	Branches	DHA	474.9	0.395	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln H + \beta_3 (\ln H)^2$
2		DHAS	474.9	0.395	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln H + \beta_3 (\ln H)^2 + \beta_4 \mathrm{HSL}$
3		DHAS	475.9	0.395	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln H + \beta_3 (\ln H)^2 + \beta_4 \mathrm{SI} + \beta_5 \mathrm{HSL}$
4		DHAS	476.9	0.395	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln H + \beta_3 (\ln H)^2 + \beta_4 \mathrm{SI}$
5		DHA	476.9	0.395	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2$
6		DHAS	476.9	0.394	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2 + \beta_5 \mathrm{HSL}$
		simp D	648.1	0.495	$\ln W = \beta_0 + \beta_1 \ln D$
1	Dry branches	DHAS	339.7	0.487	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln A + \beta_3 \mathrm{SI} + \beta_4 \mathrm{HSL}$
2		DHAS	340.5	0.484	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln A + \beta_3 \mathrm{HSL}$
3		DHAS	340.8	0.481	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln A + \beta_4 \mathrm{SI} + \beta_5 \mathrm{HSL}$
4		DHAS	342.3	0.482	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln A + \beta_4 \mathrm{HSL}$
5		DHAS	342.3	0.473	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln H + \beta_3 \ln A + \beta_4 \mathrm{HSL}$
6*		DHA	342.4	0.481	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln H + \beta_3 (\ln A \times \ln D)$
		simp D	350.4	0.500	$\ln W = \beta_0 + \beta_1 \ln D$
1*	Stem	DHA	−300.6	0.108	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2 + \beta_5 \ln A$
2		DHAS	−300.1	0.109	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2 + \beta_5 \ln A + \beta_6 \mathrm{SI}$
3		DHAS	−298.8	0.109	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 (\ln H)^2 + \beta_5 \ln A + \beta_6 \mathrm{HSL}$
4		DHAS	−298.4	0.110	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln H + \beta_3 \ln A + \beta_4 \mathrm{SI}$
5		DHAS	−297.1	0.110	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 \ln H + \beta_3 \ln A + \beta_4 \mathrm{SI} + \beta_5 \mathrm{HSL}$
6		DHAS	−296.8	0.110	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln A + \beta_4 \mathrm{SI} + \beta_5 \mathrm{HSL}$
		simp D	−102.7	0.173	$\ln W = \beta_0 + \beta_1 \ln D$
1*	Roots	DHA	70.4	0.288	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln A$
2		DHA	70.6	0.273	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 \ln A$
3		DHAS	71.2	0.287	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln A + \beta_4 \mathrm{HSL}$
4		DHAS	71.5	0.310	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln A + \beta_4 \mathrm{SI}$
5		DHAS	71.9	0.285	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln A + \beta_4 \mathrm{SI} + \beta_5 \mathrm{HSL}$
6		DHAS	72.0	0.273	$\ln W = \beta_0 + \beta_1 \ln D + \beta_2 (\ln D)^2 + \beta_3 \ln H + \beta_4 \ln A + \beta_5 \mathrm{SI}$
		simp D	92.3	0.331	$\ln W = \beta_0 + \beta_1 \ln D$

[1] Models containing only combinations of the basic variables D, H and A are termed 'DHA' models. Models containing site variables SI and HSL in addition are called "DHAS" models. The * indicates the best DHA model. The simplest allometric model containing only D as predictor is termed "simp D."

[2] AIC = Akaike Information Criterion.

[3] RMSE = Root mean square error in logarithmic units (for comparison only).

[4] Abbreviations: W = dry mass of biomass component; D = diameter at breast height; H = height of tree; A = age of tree; SI = site index; and HSL = height above sea level.

to represent suppressed and dominant trees and predicted their biomass with our best DHA models. As expected, with increasing tree size, the relative share of crown biomass compartments decreased, whereas the proportion allocated to stem and root increased (Figure 5). At a given stem biomass, suppressed trees allocated relatively less biomass to crown com-

partments but more to the root system than dominant trees. Suppressed trees exhibited a higher proportion of dry branches, whereas the proportion of biomass allocated to the stem was unaffected by social status.

To illustrate the performance of our models, we applied the best DHA models to 17 stands with known distributions of

Table 4. Parameters of the best models of the DHA and DHAS type as shown in Table 3 (best). Simplified functions that include only diameter at breast height (D) as a predictor (simp D) or that do not contain polynomial terms, interactions and site-related variables (DHA simp) are also presented. The latter may be used for extrapolation, i.e., if biomass must be predicted for trees whose dimensions are outside the range of the predictors in the regression design matrices. The column "data set" refers to regression design matrices summarized in Table 2. The parameters were fitted with the restricted maximum likelihood method. Abbreviations: β_0 = intercept; D = diameter at breast height (cm); H = height of tree (m); A = age of tree (years); SI = site index (mean stand height after 100 years) (m); HSL = height above sea level (m); and C = crown length (m). All woody compartments include bark. Abbreviations and symbols: σ = estimate of the residual standard deviation; R^2 = coefficient of determination; $cf1$ = correction factor 1 for smearing estimate; and $cf2$ = mean of correction factor 2 calculated for the underlying data set. Note that exact computation of $cf2$ requires knowledge of the estimated random effects.

Compartment	Type	Data set	β_0	$\ln D$	$(\ln D)^2$	$\ln H$	$(\ln H)^2$	$\ln A$	SI	HSL	$\ln A \times \ln D$	$\ln C$	σ	R^2	$cf1$	$cf2$
Needles	DHAS	Best	Base	1.20480	3.57870	−0.21852	−2.93621	0.55666	−0.72043	−0.02765	0.00021			0.3315	0.9069	1.0509	1.0182
	DHA	Best	Base	−0.58133	3.63845	−0.21336	−2.77755	0.46540	−0.42940					0.3382	0.9029	1.0533	1.0183
	DHA	Simp	Base	−2.30440	2.55072		−0.43127		−0.38556					0.3447	0.8986	1.0559	1.0202
	D	Simp	Base	−3.19632	1.91620									0.4198	0.8473	1.0849	1.0226
	DHAS+C	Best	Core	0.09416	3.36327	−0.25202	−3.44418	0.55054	−0.34168				0.83931	0.3049	0.9365	1.0397	1.0277
	DHA+C	Best	Core	−1.18863	3.33792	−0.24482	−3.31885	0.49368	−0.13463				0.85797	0.2972	0.9343	1.0413	1.0315
Branches	DHA	Best	Base	−0.64565	2.85424		−2.98493	0.41798						0.3963	0.9179	1.0831	1.0138
	DHA	Simp	Base	−3.61106	2.99487		−0.87590							0.4207	0.9065	1.0967	1.0188
	D	Simp	Base	−3.96201	2.25520									0.4947	0.8709	1.1332	1.0103
	DHA+C	Best	Core	0.61063	2.40589		−3.65994	0.43980					0.91027	0.3390	0.9468	1.0613	1.1343
Dry branches	DHAS	Best	Base	−1.65561	1.36040				−0.15198	−0.02299	0.00125			0.4868	0.8022	1.1105	1.3520
	DHA	Best	Base	−1.21969	1.49138		−1.25928					0.18222		0.4831	0.8086	1.1067	1.1616
	DHA	Simp	Base	−3.090620	2.04823		−1.286761		0.62836					0.4760	0.8136	1.1037	1.1135
	D	Simp	Base	−3.22406	1.67320									0.4995	0.7942	1.1146	1.1107
Stem	DHA	Best	Base	−2.83958	2.55203	−0.14991	−0.19172	0.25739	−0.08278					0.1084	0.9947	1.0052	1.0045
	DHA	Simp	Base	−3.63141	1.73771		1.19218		−0.07560					0.1107	0.9859	1.0055	1.0053
	D	Simp	Base	−2.50602	2.44277									0.1735	0.9859	1.0142	1.0238
Roots	DHA	Best	Base	−8.35049	4.56828	−0.33006			0.28074					0.2921	0.9668	1.0406	1.0061
	DHA	Simp	Base	−5.98132	2.32428		0.834968							0.3078	0.9630	1.0470	1.0092
	D	Simp	Base	−5.37891	2.92111									0.3324	0.9563	1.0554	1.0098
	DHA+C	Best	Core	−8.15491	4.08262	−0.28378			0.34963				0.24520	0.3050	0.9722	1.0435	1.0088

Table 5. Exploring the effect of adding variables related to competition, crown length (C) and social status (SOC) to selected best models. The DHA and DHAS models were recalculated using the smaller core data set containing the variables related to competition. Abbreviations: +C = ln$C$ (crown length) was added to the model; no SOC classes = the social status was not added as categorical predictor; 4 SOC classes = a categorical variable was added indicating the four crown classes (1 = dominant, 2 = codominant, 3 = intermediate and 4 = suppressed); AIC = Akaike Information Criterion (asterisks indicate the smallest value); and RMSE = root mean square error in logarithmic units (for the purpose of comparison only). All woody compartments include bark.

Compartment	Model type	n	Score	No SOC classes		4 SOC classes	
				AIC	RMSE	AIC	RMSE
Needles	DHA	388	5	349.1	0.3506	336.2	0.3408
	DHAS	388	1	334.9	0.3409	331.1	0.3358
	DHA+C	388	5	249.4	0.2969	244.7	0.2923
	DHAS+C	388	1	243.1	0.2910	240.9*	0.2878
Branches	DHA	384	1	420.8	0.3900	422.0	0.3896
	DHA+C	384	1	332.3	0.3390	330.0*	0.3369
Dry branches	DHA	176	6	289.1	0.4892	287.2	0.4777
	DHAS	176	1	280.8	0.5105	278.5*	0.4895
	DHA+C	176	6	297.9	0.4885	295.4	0.4763
	DHAS+C	176	1	286.1	0.5065	283.8	0.4948
Stem	DHA	173	1	−223.6*	0.1037	−222.1	0.1033
	DHA+C	173	1	−212.1	0.1038	−210.5	0.1036
Roots	DHA	85	1	69.1	0.3095	66.3	0.2925
	DHA+C	85	1	65.7	0.2973	58.5*	0.2748

age, diameter and height ranging from stand age 16 to 172 years (Table 7). There was a clear age trend of the relative confidence interval (RCI = half the confidence interval expressed in percent of the mean prediction). Within a biomass compartment, young stands with mean D < 10 cm had the

Table 6. Aggregate prediction errors (APE) from 100 cross-validation runs for models of varying complexity for three exemplary biomass compartments (needles, branches and stem). For each predictor set (D, DHA, DHAS, DHA+C), the structure of the best models as presented in Table 4 was adopted. For each cross-validation run, the data set was randomly split into a model data set containing 90% of the data and a validation data set containing the remaining 10%. The APE were calculated based on the ln-transformed data.

Compartment	Model type	APE
Needles	D	0.3088
	DHA	0.1917
	DHAS	0.1871
	DHA+C	0.1523
Branches	D	0.3185
	DHA	0.2162
	DHA+C	0.2420
Stem	D	0.0705
	DHA	0.0239

highest RCI. The RCI was lowest in stands of intermediate age and diameter and increased again in older stands. This trend is explained by the presence of very small trees (in young stands) or large trees (in old stands) whose diameters or heights fall outside the ranges of D and H covered by the trees in the respective regression design matrix (cf. min and max values in Table 2 and column P in Table 7). In the case of dry branches in young stands, this phenomenon created extreme variances. Comparing biomass compartments, the highest median RCIs were found for dry branch biomass (28%) and the lowest median RCIs were observed for stem biomass (5%). Branch, needle and root biomass exhibited intermediate median RCIs of 12, 13 and 10%, respectively. Combining the uncertainties of predictions for individual biomass compartments at the stand level and considering the correlation structure of the errors, whole-tree biomass at the stand level may be predicted with an RCI of about 30% in stands younger than 40 years and only 4% in stands older than 40 years.

The percentage of stand-level biomass allocated to stems was 45% at a stand age of 16 years and reached a plateau of about 60% after 60 years. Parallel to this trend, the percentage of stand-level biomass allocated to roots increased from 10 to 20%. At the same time, the percentage of stand-level biomass allocated to needles decreased from 18 to 6%, whereas the percentage allocated to branches remained remarkably constant at about 12%. The share of dry branch biomass was only 3% except for young (up to 30-year-old) stands (Figure 6).
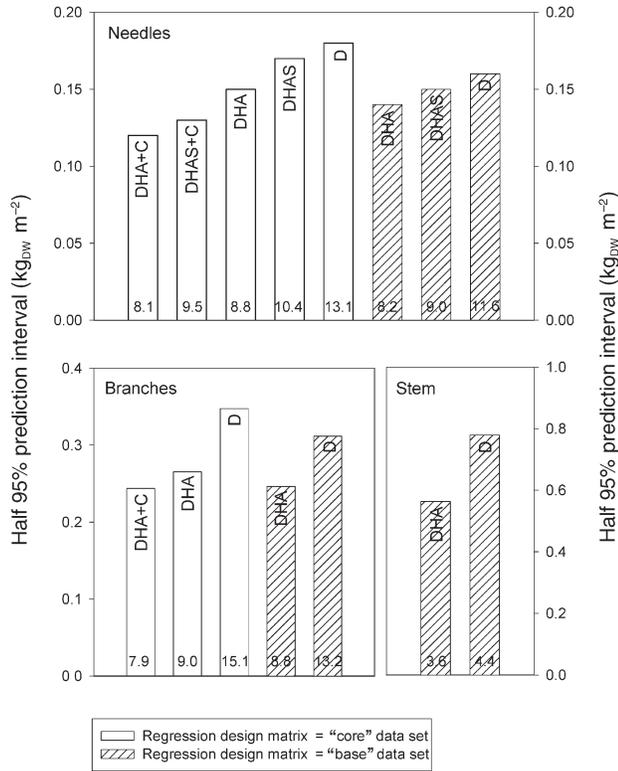
Figure 4. Comparison of half 95% prediction intervals for needle, branch and stem biomass in a 60-year-old Norway spruce stand based on different model types. In the case of needles and branches, DHA, DHAS and 'D simp' models were fit for both the larger base and the smaller core data sets (number of trees are given in Table 2) to allow a comparison based on the same core data set between all model types including those with crown length added. Figures on the bottom of each bar indicate the half 95% prediction interval in percent of the mean prediction.

Discussion

Our meta-analysis provides the first comprehensive system of functions for predicting whole-tree and compartment biomass and respective uncertainties of Norway spruce in Central Europe. Our study is based on data generated over a period of almost 100 years by 19 authors of which a substantial part was either published in doctoral theses, or directly communicated to the authors (M. Mund, R. Zimmermann, Max-Planck-Institute for Biogeochemistry, Jena, Germany, personal communication). Although the incorporation of data reflecting a wide range of regional variability is essential for the derivation of generic biomass functions, the heterogeneity of the database inevitably introduces variance that has to be taken into account. Differences among data from the various sources may be partially attributed to different methods used, characteristics of the investigated stands and regional effects. We decided to subsume all of these different aspects in one categorical variable "author" because their effects could not be disentangled. The contribution of this variable to the total variability in the data set is demonstrated by comparing one global linear fit using the predictors from our best DHA model (ignoring differences between authors) and a model with author-specific coefficients for the most important predictor $\ln D$ (F-test: $F_{17,528} = 4.822$, $P < 0.001$ for the needle compartment). The better fit of the latter model does not lead to better generic biomass functions, because "author" cannot be used as a predictor. Incorporating "author" as a source of random effects in a mixed linear model thus provides a way to account for heterogeneity in the data set and to obtain realistic estimates of the uncertainty of predictions derived from our generic biomass functions.

We note that it is generally desirable to develop additive biomass functions, i.e., that the sum of the separate predictions for
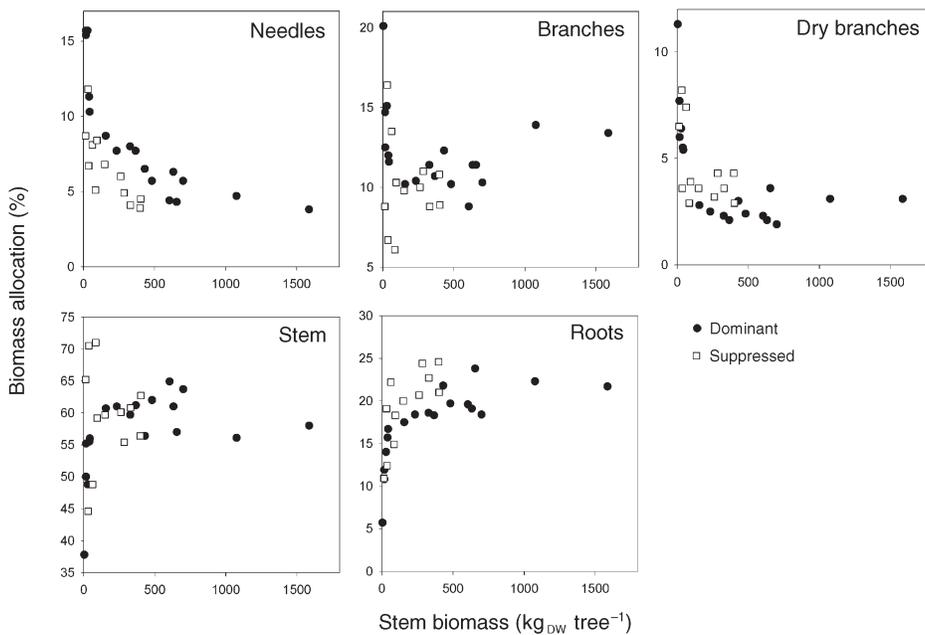

Figure 5. Predicted allocation patterns in dominant and suppressed trees at 17 test sites as a function of their individual stem biomass. Biomass allocation percentages and stem biomass are based on mean dimensions of the 10% largest trees (dominant trees, ●) and 10% smallest trees (suppressed trees, □) of the test site populations.

Table 7. Stand-level biomass predictions ± 95% confidence intervals for the 17 stands of the test data set (kg dry mass m⁻²). Stand names are composed of a regional indicator and the mean stand age in years. All woody compartments include bark. Abbreviations: RCI = relative confidence interval calculated as half the confidence interval as percentage of the mean prediction; P = percentage of trees in a stand that are out of the diameter and height range of the model data sets presented in Table 2; l = falling below the range; h = exceeding the range; SI = site index (m); and Σ = total biomass comprising all compartments. The higher the value of P, the more extrapolation errors occur.

Stand	Standing volume (m³ ha⁻¹)	SI (m)	Mean D (cm)	Mean H (m)	Needles (kg m⁻²)	RCI	P	Branches (kg m⁻²)	RCI	P	Dry branches (kg m⁻²)	RCI	P	Stem (kg m⁻²)	RCI	P	Roots (kg m⁻²)	RCI	P	Σ (kg m⁻²)	RCI
F-16	83	40	6.3	6.0	1.7 ± 0.3	19	l 12	1.5 ± 0.4	24	l 12	1.0 ± 2.1	203	l 20	4.4 ± 0.6	13	l 20	0.8 ± 0.1	13	l 37	9.4 ± 2.7	29
W-20	23	27	7.9	6.6	1.4 ± 0.3	23	l 10	1.3 ± 0.4	32	l 10	0.8 ± 1.1	147	l 20	3.7 ± 0.6	16	l 20	1.0 ± 0.2	17	l 37	8.1 ± 1.8	23
F-20	79	29	6.0	5.2	2.1 ± 0.3	13	l 10	2.1 ± 0.4	19	l 10	1.6 ± 4.9	311	l 37	5.1 ± 0.5	9	l 37	1.2 ± 0.1	10	l 52	12.0 ± 5.3	44
F-30	132	29	7.9	7.7	2.6 ± 0.4	14		2.9 ± 0.6	21		2.1 ± 2.0	94	l 15	10.7 ± 0.8	7	l 15	2.7 ± 0.3	12	l 29	21.0 ± 5.4	26
F-35	115	40	8.8	7.5	1.6 ± 0.3	19	l 15	1.9 ± 0.5	26	l 15	1.4 ± 3.1	223	l 18	6.3 ± 0.6	9	l 18	1.9 ± 0.3	16	l 18	13.0 ± 3.6	28
K-35[1]	367	34	15.4	15.4	2.2 ± 0.1	5		2.4 ± 0.1	5		0.9 ± 0.2	8		15.6 ± 0.4	2		3.9 ± 0.1	5		25.0 ± 0.6	2
W-40	333	33	21.4	17.0	1.9 ± 0.2	12		2.4 ± 0.4	15		0.7 ± 0.2	28		13.3 ± 1.0	7		4.0 ± 0.6	16		22.3 ± 1.5	7
K-42[1]	398	34	20.7	19.9	1.9 ± 0.1	7		2.3 ± 0.2	8		0.6 ± 0.1	13		16.1 ± 0.6	4		4.2 ± 0.3	8		25.1 ± 0.8	3
F-43	339	36	23.3	17.8	2.4 ± 0.3	13		3.1 ± 0.5	15		0.9 ± 0.3	28		16.7 ± 1.2	7		5.4 ± 0.8	15		28.5 ± 1.9	7
W-60	414	32	32.0	23.8	1.7 ± 0.1	8		2.8 ± 0.3	9		0.7 ± 0.1	19		16.0 ± 0.6	4		5.1 ± 0.5	10		26.2 ± 1.1	4
W-67	339	27	28.2	19.2	1.6 ± 0.1	7		2.7 ± 0.2	9		0.9 ± 0.2	18		12.3 ± 0.4	3		5.0 ± 0.4	8		22.4 ± 0.8	4
K-68[1]	611	32	26.3	25.3	2.1 ± 0.2	9		3.1 ± 0.3	10		0.8 ± 0.1	16		23.6 ± 0.7	3		6.3 ± 0.5	8		35.9 ± 1.1	3
F-72	465	34	32.8	24.7	1.9 ± 0.2	9		3.3 ± 0.3	9		0.9 ± 0.2	20		19.2 ± 0.6	3		6.3 ± 0.6	9		31.5 ± 1.2	4
F-112	605	28	32.8	28.0	1.6 ± 0.2	11		3.1 ± 0.3	10		0.9 ± 0.2	20		21.0 ± 0.7	3		6.8 ± 0.5	7		33.3 ± 1.2	4
W-120	474	24	41.7	27.0	1.5 ± 0.2	13		3.6 ± 0.4	11		1.1 ± 0.3	30	h 50	16.7 ± 0.6	4	h 50	6.8 ± 0.7	10	h 12	29.7 ± 1.4	5
F-142	526	26	36.5	26.7	1.4 ± 0.2	14		3.3 ± 0.4	11		1.2 ± 0.3	25	h 1	17.7 ± 0.8	4	h 1	7.1 ± 0.6	8	h 1	30.7 ± 1.4	5
F-173	543	26	48.7	31.4	1.4 ± 0.2	17		4.1 ± 0.5	12		1.3 ± 0.5	37	h 37	18.8 ± 1.0	5	h 37	7.6 ± 0.8	11	h 43	33.1 ± 2.0	6

[1] In these stands, only stem numbers per 4 cm diameter class were recorded, obscuring the real variability. Confidence intervals, therefore, are markedly smaller than for tree-wise inventoried stands.
[2] Errors were propagated to the whole-stand biomass using Monte Carlo simulation. The correlation structure of errors was obtained from the 78 trees for which all biomass compartments were measured. The individual correlation coefficients r were $r_{N\times B} = 0.69$, $r_{N\times D} = 0.30$, $r_{N\times S} = 0.31$, $r_{N\times R} = 0.18$, $r_{B\times D} = 0.20$, $r_{B\times S} = 0.44$, $r_{B\times R} = 0.22$, $r_{D\times S} = 0.32$, $r_{D\times R} = 0.30$, and $r_{S\times R} = -0.13$, where N = needles, B = branches, D = dry branches, S = stem and R = roots.
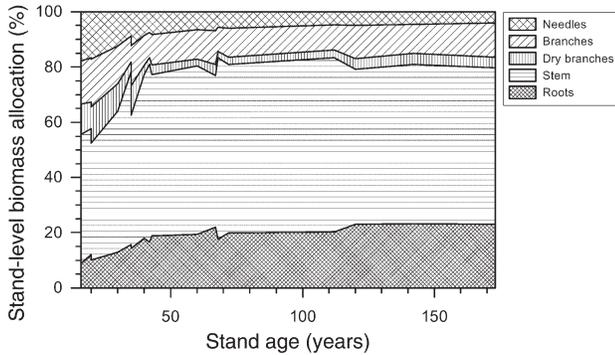
Figure 6. Stand-level allocation patterns as a function of stand age as predicted by the DHA model family for a set of 17 test sites. The test sites range in stand age from 16 to 172 years (cf. Table 7). The wiggles are due to slight differences in allocation patterns between the test stands.

biomass compartments is equal to the prediction of a model taking the sum of all compartments as an independent variable (Kozak 1970, Parresol 2001). However, this is possible only if information on all biomass compartments is available for every tree in the regression design matrix. Of the 688 trees in our database, only 78 (from just four studies) were completely sampled for biomass. We considered this subset too small to serve as a basis for developing generic functions for a region as large as Central Europe.

For biomass functions that use age and several dimensional predictors, multicollinearity always exists among these highly intercorrelated predictors. Multicollinearity has several adverse effects. It tends to increase the standard errors of the regression coefficients and it conceals the causal relationships; however, it does not affect the precision of the mean response (Neter et al. 1996). Because we are primarily interested in predicting the mean response, multicollinearity does not pose a problem to our work.

In recent publications, weighted nonlinear regression was used to develop biomass functions (Parresol 1999, Fonseca and Parresol 2001, Parresol 2001). Based on this approach, Parresol (2001) obtained a slightly smaller residual mean square error (RMSE) compared with use of linear regression of ln-transformed data. Although principally possible in combination with mixed-effect models (Pinheiro and Bates 2000), the complexity of most of our candidate models resulted in major numerical problems if nonlinear mixed-effect modelling was tested. To correct for bias introduced by the ln-transformation (Pinheiro and Bates 2000), we adapted the non-parametric smearing estimate for linear mixed models. The dependence of the additional correction factor $cf2$ on the values of the predictors for a new response at first sight seems to be a shortcoming (knowledge of the estimated random effects is required to compute the value of $cf2$). On the other hand, it clarifies that variability caused by methodological differences unevenly affects the resulting uncertainty of predictions for different values of the predictors. However, variability of $cf2$

was generally small over the regression design matrix used for the model fit, so that the use of the mean values given in Table 4 should be a pragmatic compromise.

An important shortcoming in our study is the lack of biomass data for young trees. There are no trees younger than 13 years in our model data set, and the smallest tree measured for root biomass had a diameter of 5.1 cm (Table 2), which limits the applicability of our functions to stands beyond the regeneration phase. Even in 20- to 30-year-old stands, up to 40% of the trees may have a diameter or height that falls below the range of validity of the model, i.e., they are smaller than the smallest trees in the regression design matrices. This may inflate the 95% prediction intervals, as shown in Table 7, which may lead to extreme values for compartment models with a high residual error, as is the case for dry branches. The same is true to a lesser extent for large trees. However, even if data were available for trees less than 10 years old, this would require the development of a separate set of functions using base diameter D_0 as a predictor instead of D. This is because D is either unavailable or is a poor predictor of biomass for small trees.

Our set of predictors misses some important sources of variation in biomass allocation patterns. We assume that the genetic variability of trees represents a large component of the unexplained variance. Numerous studies have revealed differences in allocation patterns, crown geometry and wood density among different tree provenances (Schmidt-Vogt 1987). A second unknown source of variance is probably the health of the trees sampled during the period of forest decline (Schulze et al. 1989). Another source of variation specific for needle biomass is starch content, which varies markedly over the course of a year (Bauer et al. 1997). Dry mass of needles is therefore also expected to vary with sampling date, being higher during winter months. Further, our data set does not allow for discrimination among the effects of different silvicultural practices.

The most comprehensive work on biomass functions for Norway spruce was carried out by Marklund (1987, 1988) for Swedish material. His landmark study was based on regression design matrices containing 281 to 544 trees depending on the biomass compartment. His sampling design for stands and trees was such that the trees were well distributed with respect to the main predictor variables, e.g., diameter, age, site index and geographic location. Figure 7 explores the performance of Marklund's functions for predicting biomass of Central European spruce trees. The solid lines represent the ratio of the measured component dry mass of the Central European sample trees and the respective Marklund predictions for the same trees (ratio 1 = measured/Marklund prediction). Although the Marklund functions accurately predicted stem biomass over the whole range of diameters (ratio close to 1), they overestimated biomass of needles, dry branches and roots by about 25% and slightly underestimated branch biomass within the central diameter range of 10 to 40 cm in which about 80% of the sample trees fall. For small trees ($D < 10$ cm), the biomass of the crown compartments tended to be largely underesti-
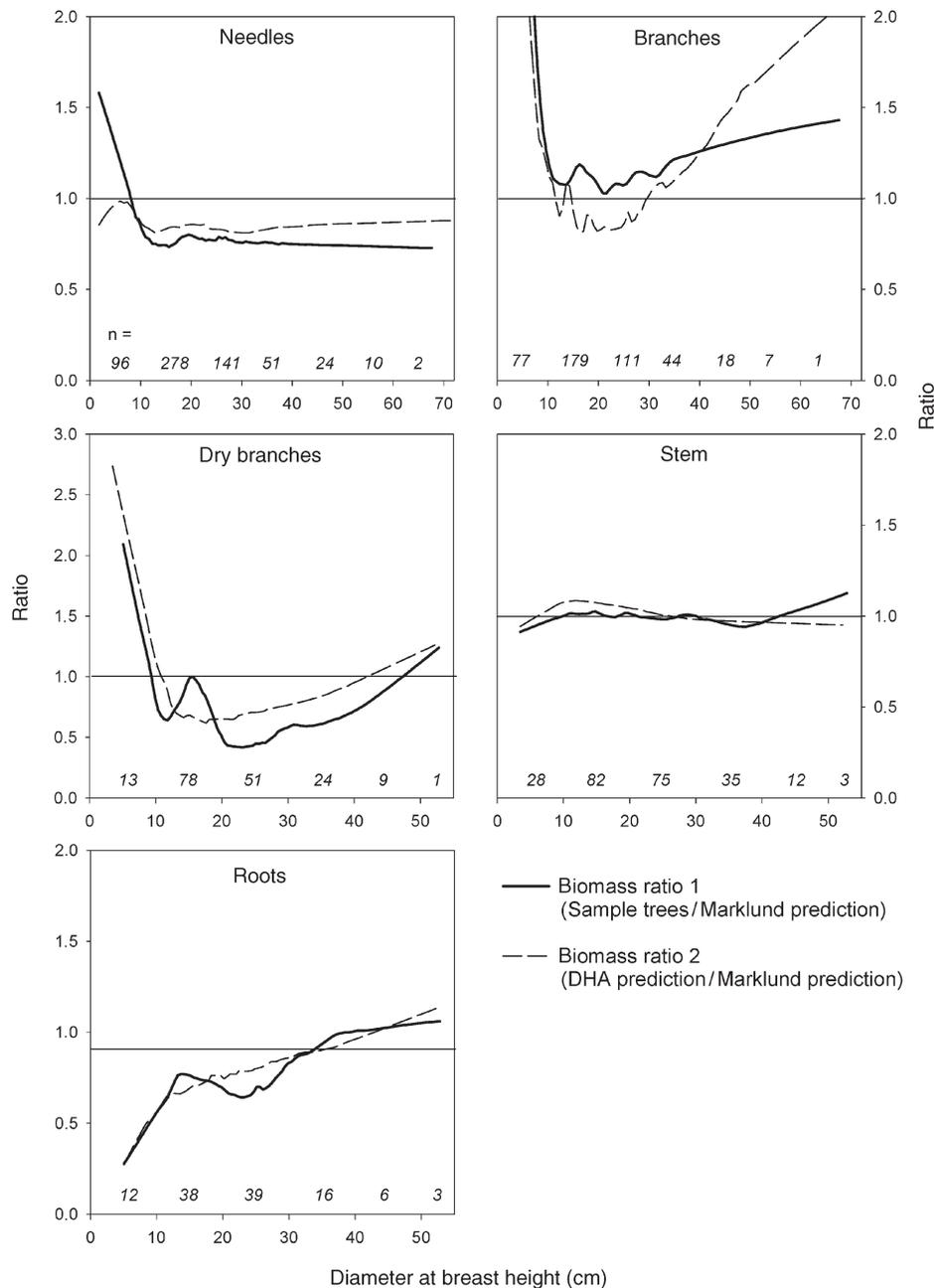
Figure 7. Comparison of our data with Marklund's functions (Marklund 1988). The solid line shows the ratio of measured component dry mass of the Central European sample trees contained in our design matrix and the dry mass predicted for these trees using the simplified Marklund functions (ratio 1 = measured/Marklund prediction). Numbers of sample trees falling in 10 cm diameter classes are given in italics at the bottom of each frame. For the test data set (n = 1985 trees), the dashed line shows the ratio of the dry mass predicted by our DHA models to the dry mass predicted by the same simplified Marklund functions (ratio 2 = our prediction/Marklund prediction). The lines result from smoothing of the data with a tricube weighted loess-regression based on a 20 to 30% local sampling proportion. We used the following Marklund functions: G-16 for needles, G-13 minus G-16 for branches, G-21 for dry branches, G-2 for stem and G-23 for roots.

mated by the Marklund functions, whereas the root biomass was overestimated. The dashed lines compare in the same manner the biomass predictions for the test data set using either our DHA models or the Marklund functions (ratio 2 = our DHA prediction/Marklund prediction). As expected, the deviation of the model predictions follows closely the deviation of the individual sample trees. The finding that Marklund's functions overestimate needle and root biomass, but predict stem biomass correctly for Central European trees, may be explained by the influence of more favorable growing conditions in Central Europe on allocation patterns. Palumets (1993) analyzed biomass partitioning in 99 spruce stands along a north–south gradient from boreal Russia to Central Europe

and found a decrease in percentage biomass in needles and roots from north to south. Because wood density of Norway spruce is known to decrease with site quality (Trendelenburg 1955), the agreement with respect to predicted stem biomass may indicate that differences in wood density are compensated for by differences in stem form. Underestimation of branch biomass by the Marklund functions may result from trees in the boreal zone exhibiting less branchiness because of lower evaporative demand (Berninger et al. 1995). Generally, the application of our DHA models to the test data set of 17 stands differing in age produced realistic age dynamics of allocation patterns that are consistent within a few percent with the results of the meta-analysis study of Palumets (1993) and recent

work by Mund et al. (2002).

We illustrated that, using the proposed functions and knowing only the diameter, height and age distribution of a stand, the biomass of needles, branches and roots can be predicted with relative 95% confidence intervals (RCI, i.e., half the interval in percent of the mean prediction) ranging from 5 to 32% with a median of 11%. Stem biomass can be predicted more precisely with a median RCI of 4%. Combining the uncertainties of predictions for individual biomass compartments at the stand level and considering the correlation structure of the errors, whole-tree biomass at the stand level may be predicted with an RCI of about 30% in stands younger than 40 years and only 4% in stands older than 40 years. The functions are capable of quantifying carbon stock changes over a relatively short time span and are therefore well suited for carbon accounting in the framework of the Kyoto Protocol. In future carbon accounting schemes, generic biomass functions like the ones we present here may be directly applied to dimensional tree-level data recorded in national inventory systems. This would improve the verification of carbon stock changes in forests and render the development of complicated systems of biomass expansion factors obsolete.

Because crown compartments exhibit higher nutrient concentrations and turnover rates than the stem compartment, the availability of separate functions for each biomass compartment is especially important for estimating nutrient budgets and productivity. In other words, nutrient and productivity allocation is different from biomass allocation. For example, needles and branches contain about 50% of the nitrogen (Scarascia-Mugnozza et al. 2000) and are responsible for 40% of the productivity (Mund et al. 2002), but comprise only 15% of the biomass. Because the models for needles and branches are associated with a higher prediction error, the prediction error of nutrient stocks and productivity rates will accordingly be higher.

References

Bauer, G., E.-D. Schulze and M. Mund. 1997. Nutrient contents and concentration in relation to growth of *Picea abies* and *Fagus sylvatica* along a European transect. Tree Physiol. 17:777–786.

Berninger, R., M. Mencuccini, E. Nikinmaa, J. Grace and P. Hari. 1995. Evaporative demand determines branchiness of Scots pine. Oecologia 102:164–168.

Brown, S. 2002. Measuring carbon in forests: current status and future challenges. Environ. Pollut. 116:363–372.

Burger, H. 1937. Holz, Blattmenge und Zuwachs: Nadelmenge und Zuwachs bei Föhren und Fichten verschiedener Herkunft. Mitt. Schweiz. Anst. Forstl. Versuchsw. 20:101–114.

Burger, H. 1953. Holz, Blattmenge und Zuwachs: Fichten im gleichaltrigen Hochwald. Mitt. Schweiz. Anst. Forstl. Versuchsw. 29: 38–130.

Burnham, K.P. and D.A. Anderson. 1998. Model selection and inference—a practical information-theoretic approach. Springer-Verlag, New York, 353 p.

Burschel, P. and J. Huss. 1997. Grundriß des Waldbaus. Parey Buchverlag, Berlin, pp 345–347.

Burschel, P., E. Kürsten and B.C. Larson. 1993. Die Rolle von Wald und Forstwirtschaft im Kohlenstoffhaushalt—eine Betrachtung für die Bundesrepublik Deutschland. Forst. Forschungsber. München 126:1–135.

Casella, G. and R.L. Berger. 2002. Statistical inference. 2nd Edn. Thomson Learning, Pacific Grove, CA, 660 p.

Cerny, M. 1990. Biomass of *Picea abies* (L.) Karst. in Midwestern Bohemia. Scan. J. For. Res. 5:83–95.

Crawley, M.J. 2002. Statistical computing—an introduction to data analysis using S-Plus. John Wiley and Sons, Chichester, U.K., 759 p.

Davison, A.C. and D.V. Hinkley. 1997. Bootstrap methods and their applications. *In* Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 582 p.

Dietrich, H. 1968. Untersuchungen zur Nährstoffdynamik eines Fichtenbestandes. I. Mitteilungen: Massewerte des Fichtenbestandes und Einfluß einer Bestandeskalkung. Arch. Forstwes. 17: 391–412.

Drexhage, M. 1994. Die Wurzelentwicklung 40-jähriger Fichten (*Picea abies* (L.) Karst.) in der Langen Bramke (Harz). Forschungszentrum Waldökosysteme Göttingen. Ber. Forschungsz. Waldökosysteme A, 111:1–165.

Drexhage, M. and F. Gruber. 1999. Above- and below-stump relationships for *Picea abies*: estimating root system biomass from breast-height diameters. Scan. J. For. Res. 14:328–333.

Droste zu Hülshoff, B. 1969. Struktur und Biomasse eines Fichtenbestandes aufgrund einer Dimensionsanalyse an oberirdischen Baumorganen. Ph.D. Diss., Ludwig-Maximilian-Univ., Munich, Germany, 209 p.

Duan, N. 1983. Smearing estimate: a nonparametric retransformation method. J. Am. Stat. Assoc. 78:605–610.

Duvigneaud, P. and S. Denaeyer-De Smet. 1970. Biological cycling of minerals in temperate deciduous forests. *In* Analysis of Temperate Forest Ecosystems. Ed. D.E. Reichle. Springer-Verlag, Berlin, pp 199–225.

Duvigneaud, P., P. Kestemont, J. Timperman and J.-C. Moniquet. 1977. La hêtraie ardennaise a Festuca altissima à Mirwart biomasse et productivité primaire. *In* Productivité Biologique en Belgique. Eds. P. Duvigneaud and P. Kestemont. Duculot, Paris-Gembloux, France, pp 107–154.

Fiedler, F. 1986. Die Dendromasse eines hiebsreifen Fichtenbestandes. Beitr. Forstwirtschaft 20:171–180.

Fiedler, F. 1987. Das Ökologische Meßfeld der Sektion Forstwirtschaft der TU Dresden. V. Die Verteilung der Dendromassekomponenten. Wiss. Z. Dresden 36:229–234.

Fonseca, T.J.F. and B.R. Parresol. 2001. A new model for cork weight estimation in Northern Portugal with methodology for construction of confidence intervals. For. Ecol. Manage. 152:131–139.

Grundner, F. and A. Schwappach. 1952. Massentafeln zur Bestimmung des Holzgehaltes stehender Waldbäume und Waldbestände. Ed. R. Schober. P. Parey Verlag, Berlin, 216 p.

Hakkila, P. 1989. Utilization of residual forest biomass. Springer-Verlag, Berlin, 568 p.

Hall, G.M.J., S.K. Wiser, R.B. Allen, P.N. Beets and C.J. Goulding. 2001. Strategies to estimate national forest carbon stocks from inventory data: the 1990 New Zealand baseline. Global Change Biol. 7:389–403.

Heller, H. and D. Göttsche. 1986. Biomasse-Messungen an Buche und Fichte. *In* Ökosystemforschung—Ergebnisse des Solling-Projekts. Ed. H. Ellenberg. E. Ulmer Verlag, Stuttgart, Germany, 507 p.

Hesse, C. 1990. Inventur der Bestandesbiomasse und ausgewählter chemischer Elemente in einem 63-jährigen Fichtenbestand im Sauerland. Ph.D. Diss., Lehrstuhl für Bodenkunde und Waldernährung, Georg-August-Univ., Göttingen, Germany, 230 p.

IPCC. 2000. Land use, land-use change, and forestry. Eds. R.T. Watson, I.R. Noble, B. Bolin, N.H. Ravindranath, D.J. Verardo, D.J. Dokken and R.T. Cambridge. Cambridge University Press, Cambridge, 377 p.

Körner, C., B. Schilcher and S. Pelaez-Riedl. 1993. Vegetation und Treibhausproblematik: eine Beurteilung der Situation in Österreich unter besonderer Berücksichtigung der Kohlenstoff-Bilanz. Verlag der Österreichischen Akademie der Wissenschaften, Vienna, 46 p.

Kozak, A. 1970. Methods for ensuring additivity of biomass components by regression analysis. For. Chron. 46:402–404.

Kuhr, M. 1999. Grobwurzelarchitektur in Abhängigkeit von Baumart, Alter, Standort und sozialer Stellung. Ph.D. Diss., Fakultät für Forstwissenschaften und Waldökologie, Georg-August-Univ., Göttingen, Germany, 121 p.

Lee, D.-H. 1998. Architektur des Wurzelsystems von Fichten (*Picea abies* (L.) Karst.) auf unterschiedlich versauerten Standorten. Ber. Forschungsz. Waldökosysteme A, Univ. Göttingen, Germany, 153:1–142.

Mäkelä, A., J. Landsberg, A. Ek, R. Burk, T.E. Ter-Mikaelian, M. Agren, G.I. Oliver and C.D. Puttonen. 2000. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol. 20: 289–298.

Marklund, L.G. 1987. Biomass functions for Norway spruce (*Picea abies* (L.) Karst.) in Sweden. Department of Forest Survey, Swedish Univ. Agric. Sci., Umeå, Sweden, 132 p.

Marklund, L.G. 1988. Biomassafunktioner för tall, gran och björk i sverige. Swedish Univ. Agric. Sci., Umeå, Sweden, 73 p.

Möller, C.M. 1945. Untersuchungen über Laubmenge, Stoffverlust und Stoffproduktion des Waldes. Kandrup and Wunsch's Bogtrykkeri, Copenhagen, Denmark, 287 p.

Mund, M., E. Kummetz, M. Hein, G.A. Bauer and E.-D. Schulze. 2002. Growth and carbon stocks of a spruce forest chronosequence in central Europe. For. Ecol. Manage. 171:275–296.

Neter, J., M.H. Kutner, C.J. Nachtsheim and W. Wassermann. 1996. Applied linear statistical models. 4th Edn. WCB McGraw-Hill, Boston, MA, 1408 p.

Palumets, J. 1993. Analysis of phytomass partitioning in Norway Spruce. Laboratory of Ecosystems, Univ. Tartu, Estonia, 48 p.

Parresol, B.R. 1999. Assessing tree and stand biomass: a review with examples and critical comparions. For. Sci. 45:573–593.

Parresol, B.R. 2001. Additivity of nonlinear biomass equations. Can. J. For. Res. 31:865–878.

Pinheiro, J.C. and D.M. Bates. 2000. Mixed-effects models in S and S-Plus. Springer-Verlag, New York, 521 p.

Poeppel, B. 1989. Untersuchungen der Dendromasse in mittelalten Fichtenbeständen. Forsteinrichtung und Forstliche Ertragskunde. Ph.D. Diss., Technische Univ. Dresden, Germany, 66 p.

Polley, H. 1994. Der Wald in den Bundesländern. Allg. Forst. 6: 318–322.

Raisch, W. 1983. Bioelementverteilung in Fichtenökosystemen der Bärhalde (Südschwarzwald). Freiburger Bodenkundl. Abh. 1: 1–239.

Running, S.W. and S.T. Gower. 1991. FOREST-BGC, A general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets. Tree Physiol. 9: 147–160.

Satoo, T. 1970. A synthesis of studies by the harvest method: primary production relations in the temperate deciduous forests of Japan. *In* Analysis of Temperate Forest Ecosystems. Ed. D.E. Reichle. Springer-Verlag, Berlin, pp 55–72.

Scarascia-Mugnozza, G., G.A. Bauer, H. Persson, C. Matteucci and A. Masci. 2000. Tree biomass, growth and nutrient pools. *In* Carbon and Nitrogen Cycling in European Forest Ecosystems. Ecological Studies 142. Ed. E.-D. Schulze. Springer-Verlag, Berlin, pp 49–62.

Schmidt, H. 1949. Die Verzweigungstypen der Fichte und ihre Bedeutung für die forstliche Pflanzenzüchtung. Ph.D. Diss., Ludwig-Maximilian-Univ., Munich, Germany, 156 p.

Schmidt-Vogt, H. 1987. Die Fichte. Vol. 1. 2nd Edn. Parey, Hamburg, Germany, 647 p.

Schulze, E.-D., M.I. Fuchs and M. Fuchs. 1977. Spatial distribution of photosynthetic capacity and performance in a mountain spruce forest of northern Germany. I. Biomass distribution and daily CO_2 uptake in different crown layers. Oecologia 29:43–61.

Schulze, E.-D., O.L. Lange and R. Oren. 1989. Air pollution and forest decline. A case study with spruce (*Picea abies*) on acid soils. Ecological Studies 77, Springer-Verlag, Heidelberg, 475 p.

Schöpfer, W. 1961. Beiträge zur Erfassung des Assimilationsapparates der Fichte. Freiburg, Baden-Württembergische Forstliche Versuchsanstalt, Abteilung Biometrie, Freiburg, Germany, pp 1–127.

Sharma, S.C. 1992. Untersuchungen über die Dendromasse der Baumart Fichte (*Picea abies* (L.) Karsten) im Tharandter Wald. Diss., Fakultät für Bau-, Wasser- und Forstwesen, Technische Univ. Dresden, Germany, 150 p.

Ter-Mikaelian, M.E. and M.D. Korzukhin. 1997. Biomass equations for sixty-five North-American tree species. For. Ecol. Manage. 97: 1–24.

Trendelenburg, R. 1955. Das Holz als Rohstoff. 2nd Edn. Ed. H. Mayer-Wegelin. Carl-Hanser Verlag, Munich, 527 p.

Vanninen, P. and A. Mäkelä. 2000. Needle and stem wood production in Scots pine (*Pinus sylvestris*) trees of different age, size and competitive status. Tree Physiol. 20:527–533.

Vins, B. and A. Sika. 1981. Biomasa smrkoveho porostu v chlumni oblasti. Prace VULHM 59:83–99.

Vyskot, M. 1981. Biomass of the tree layer of spruce forest in the Bohemian Uplands. Academia Publishing House of Czechoslovak Acad. Sci., Prague, 396 p.

Vyskot, M. 1990. Aboveground biomass of adult Norway spruce. Acta scientiarum naturalium academiae scientiarum Bohemoslovacae Brno 24:1–36.

WBGU. 1998. Die Anrechnung biologischer Quellen und Senken im Kyoto-Protokoll: Fortschritt oder Rückschlag für den globalen Umweltschutz. Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen, Bremerhaven, Germany, 76 p.

Wenk, G., K. Römisch and D. Gerold. 1985. DDR-Fichtenertragstafel 1984. Technische Universität Dresden—Sektion Forstwissenschaft, Agrarwissenschaftliche Gesellschaft der DDR, 64 p.

Wirth, C., E.-D. Schulze, V. Kusznetova, I. Milyukova, G. Hardes, M. Siry, B. Schulze and N.N. Vygodskaya. 2002. Comparing the influence of site quality, stand age, fire and climate on aboveground tree production in Siberian Scots pine forests. Tree Physiol. 22: 537–552.

Zimmermann, R. 1985. Untersuchungen zur Photosyntheseleistung, Transpiration und Nadelbiomasseverteilung von zwei geschädigten Fichtenforsten in den höheren Lagen des Fichtelgebirges. Univ. Bayreuth, Germany, 133 p.

Appendix 1

Modified smearing estimate

When predicting new values of the response variable these are considered as realizations of Model 2 (Equation 2):

$$y_{new} = \mathbf{x}_{new}\boldsymbol{\beta} + \mathbf{z}_{new}\mathbf{b}_{new} + \varepsilon_{new}$$

Although the best prediction on the ln-transformed scale is obtained by simply replacing the unknown model coefficients $\boldsymbol{\beta}$ by their estimated values $\hat{\boldsymbol{\beta}}$, the back-transformation to the original scale introduces a bias. The expected value of the back-transformed response is given by:

$$E(\exp(y_{new})) = E(\exp(\mathbf{x}_{new}\boldsymbol{\beta} + \mathbf{z}_{new}\mathbf{b}_{new} + \varepsilon_{new}))$$
$$= \int \exp(\mathbf{x}_{new}\boldsymbol{\beta} + \mathbf{z}_{new}\mathbf{b}_{new} + \varepsilon_{new})dF_{\mathbf{b},\varepsilon}$$

where $F_{\mathbf{b},\varepsilon}$ denotes the common distribution function of the random effects $\mathbf{b}$ and the residuals ε. Because the random effects are assumed to be independent of the residuals, the common distribution can be factorized:

$$E(\exp(y_{new})) = \int \exp(\mathbf{x}_{new}\boldsymbol{\beta} + \mathbf{z}_{new}\mathbf{b}_{new} + \varepsilon_{new})dF_{\mathbf{b}}dF_{\varepsilon}$$

where $F_{\mathbf{b}}$ and F_{ε} now denote the distribution function of the random effects and the residuals, respectively (Casella and Berger 2002). These distribution functions can be estimated by the empirical cumulative distribution functions of the estimated random effects and the estimated residuals:

$$\hat{F}_{\mathbf{b}}(b) = \frac{1}{J}\sum_{j=1}^{J}I(\hat{\mathbf{b}}_j \leq b) \text{ and}$$

$$\hat{F}_{\varepsilon}(e) = \frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{n_j}I(\hat{\varepsilon}_{ij} \leq e)$$

where $I(\ldots)$ denotes the indicator function and $\hat{\mathbf{b}}_j$ and $\hat{\varepsilon}_{ij}$ denote the estimated random effects and the residuals from the fitted model, respectively (Duan 1983). Substituting the empirical distribution functions for their unknown theoretical counterparts converts the integral into a sum:

$$\tilde{E}(\exp(y_{new})) = \exp(\mathbf{x}_{new}\boldsymbol{\beta})\frac{1}{J}\sum_{j=1}^{J}\exp(\mathbf{z}_{new}\hat{\mathbf{b}}_j)\frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\exp(\hat{\varepsilon}_{ij})$$

Finally, replacing the model parameters $\boldsymbol{\beta}$ by their estimates $\hat{\boldsymbol{\beta}}$,

we arrive at a modified smearing estimate for the new response:

$$\hat{E}(\exp(y_{new})) = \exp(\mathbf{x}_{new}\hat{\boldsymbol{\beta}})\underbrace{\frac{1}{J}\sum_{j=1}^{J}\exp(\mathbf{z}_{new}\hat{\mathbf{b}}_j)}_{\text{correction factor 2 }(cf2)}\underbrace{\frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\exp(\hat{\varepsilon}_{ij})}_{\text{correction factor 1 }(cf1)}$$

(A1)

which differs from the usual smearing estimate by the additional correction factor

$$cf2 = \frac{1}{J}\sum_{j=1}^{J}\exp(\mathbf{z}_{new}\hat{\mathbf{b}}_j)$$

reflecting the new error component. Note that there is no unique correction factor as in the standard case of multiple linear regression. The correction factor $cf2$ now depends on the vector $\mathbf{z}_{new}$. Replacing $\boldsymbol{\beta}$ by $\hat{\boldsymbol{\beta}}$ does not take into account the uncertainty caused by the estimation of parameters, but this error is small compared with the other sources of uncertainty.

Appendix 2

Calculation of confidence intervals for new predictions

In the following example, we show how to calculate a new prediction for branch biomass and its confidence interval for a single tree for which diameter (D = 18.8 cm), height (H = 16.9 m) and age (A = 40 years) are known. Consequently, we choose the simple DHA model for branches presented in Table 4, which requires only $\ln D$ (= 2.93), $\ln H$ (= 2.83) and $(\ln H)^2$ (= 7.99) as predictors. Recalling Equation A1:

$$\hat{E}(\exp(y_{new})) = \underbrace{\exp(\mathbf{x}_{new}\hat{\boldsymbol{\beta}})}_{\substack{\text{back-transformed}\\\text{prediction}}}\underbrace{\frac{1}{J}\sum_{j=1}^{J}\exp(\mathbf{z}_{new}\hat{\mathbf{b}}_j)}_{\text{correction factor 2 }(cf2)}\underbrace{\frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\exp(\hat{\varepsilon}_{ij})}_{\text{correction factor }(cf1)}$$

the mean prediction (in kg dry mass) on the original scale is calculated using the regression coefficients and the correction factors presented in Table 4 as:

$$= \exp\left[\underbrace{\begin{bmatrix} 1 & \overset{\ln D}{2.93} & \overset{\ln H}{2.83} & \overset{(\ln H)^2}{7.99} \end{bmatrix}}_{\mathbf{x}_{new}}\underbrace{\begin{bmatrix} -0.6456 \\ 2.8542 \\ -2.9849 \\ 0.4180 \end{bmatrix}}_{\substack{\boldsymbol{\beta}\text{-vector}\\\text{regression}\\\text{coefficients}}}\right]\underbrace{(1.0128)}_{cf2}\underbrace{(1.0831)}_{cf1} = 14.93$$

The variance of the new prediction is estimated according to Equation 3 as:

$$\hat{\text{Var}}(\hat{y}_{\text{new}}) = \underbrace{\begin{bmatrix} 1 & \overset{\ln D}{2.93} & \overset{\ln H}{2.83} & \overset{(\ln H)^2}{7.99} \end{bmatrix}}_{\substack{\mathbf{x}_{\text{new}} \\ \text{fixed component}}} \times \underbrace{\begin{bmatrix} 0.260 & 0.023 & -0.180 & 0.023 \\ 0.023 & 0.037 & -0.043 & -0.001 \\ -0.180 & -0.043 & 0.156 & -0.016 \\ 0.023 & -0.001 & -0.016 & 0.003 \end{bmatrix}}_{\substack{\hat{\text{Var}}(\hat{\boldsymbol{\beta}}) \\ \text{fixed component}}} \underbrace{\begin{bmatrix} 1 \\ 2.93 \\ 2.83 \\ 7.99 \end{bmatrix}}_{\mathbf{x}_{\text{new}}^T} + \underbrace{\underbrace{\begin{bmatrix} 1 & \overset{\ln D}{2.93} & \overset{\ln H}{2.83} \end{bmatrix}}_{\mathbf{z}_{\text{new}}} \underbrace{\begin{bmatrix} 0.668 & 0.256 & -0.492 \\ 0.256 & 0.244 & -0.357 \\ -0.492 & -0.357 & 0.558 \end{bmatrix}}_{\hat{\Psi}} \underbrace{\begin{bmatrix} 1 \\ 2.93 \\ 2.83 \end{bmatrix}}_{\mathbf{z}_{\text{new}}^T}}_{\text{random component}}$$

$$+ \underbrace{0.157}_{\text{MSE} = \sigma^2}$$

$$= \underbrace{0.002}_{\substack{\text{fixed} \\ \text{component}}} + \underbrace{0.016}_{\substack{\text{random} \\ \text{component}}} + \underbrace{0.157}_{\substack{\text{residual} \\ \text{component}}} = 0.176$$

The estimated variance–covariance matrices for the fixed ($\hat{\text{Var}}(\hat{\boldsymbol{\beta}})$) and the random effects ($\hat{\Psi}$) of the various models in Table 4 can be downloaded as an Excel file or S-Plus objects from ftp://panorama.bgc-jena.mpg.de/pub/science/cwirth/ Wirth_et_al._spruce_ v-cv-matrices.xls. A symmetric 95% confidence interval around the prediction in logarithmic units is calculated using the quantile of the standard normal distribution as $\pm 1.96\sqrt{\hat{\text{Var}}(\hat{y}_{\text{new}})}$. This expression becomes 0.822 and the confidence interval in logarithmic units is $\mathbf{x}_{\text{new}}\hat{\boldsymbol{\beta}} \pm$ 0.824 = 2.610 $\pm$ 0.824 $\Rightarrow$ lower boundary = 1.786, upper boundary = 3.434. Back-transforming the boundary values to the original scale, we obtain an asymmetric confidence interval around the unbiased prediction of 14.93 of [6.0, 31.0]. The alternative approach to calculate symmetric confidence intervals is explained in the method section. Fonseca and Parresol (2001) outline the construction of confidence intervals for simultaneous predictions.