Generic biomass functions for Common beech (*Fagus sylvatica*) in Central Europe: predictions and components of uncertainty

Thomas Wutzler, Christian Wirth, and Jens Schumacher

Abstract: This study provides a comprehensive set of functions for predicting biomass for Common beech (*Fagus sylvatica* L.) in Central Europe for all major tree compartments. The equations are based on data of stem, branch, timber, brushwood (wood with diameter below 5 or 7 cm), foliage, root, and total aboveground biomass of 443 trees from 13 studies. We used nonlinear mixed-effects models to assess the contribution of fixed effects (tree dimensions, site descriptors), random effects (grouping according to studies), and residual variance to the total variance and to obtain realistic estimates of uncertainty of biomass on an aggregated level. Candidate models differed in their basic form, the description of the variance, and inclusion of various combinations of additional fixed and random effects and were compared using the Akaike information criterion. Model performance increased most when accounting for between-study differences in the variability of biomass predictions. Further, performance increased with the inclusion of age, site index, and altitude as predictor variables. We show that neglecting variance partitioning and the fact that prediction errors of trees are not independent with respect to their predictor variables will lead to a significant underestimation of prediction variance.

Résumé : Cette étude fournit un ensemble complet de fonctions qui permettent de prédire la biomasse de toutes les composantes principales pour le hêtre commun (*Fagus sylvatica* L.) en Europe centrale. Les équations sont basées sur des données de biomasse de la tige, des branches, du bois d'œuvre, des broussailles (bois dont le diamètre est plus petit que 5 ou 7 cm), du feuillage, des racines et de la biomasse aérienne totale de 443 arbres répartis dans 13 études. Nous avons utilisé des modèles non linéaires à effets mixtes pour évaluer la contribution des effets fixes (dimensions de l'arbre, descripteurs de la station), des effets aléatoires (regroupement par étude) et de la variance résiduelle à la variance totale et pour obtenir des estimations réalistes de l'incertitude de la biomasse lorsque les études sont regroupées. Les modèles potentiels différaient par leur forme de base, la description de la variance et l'inclusion de différentes combinaisons d'effets fixes et aléatoires additionnels et ils ont été comparés en utilisant le critère d'information d'Akaike. La performance des modèles augmentait le plus lorsqu'ils tenaient compte des différences entre les études dans la variabilité des prédictions de la biomasse. De plus, la performance a augmenté avec l'inclusion de l'âge, de l'indice de qualité de station et de l'altitude comme variables explicatives. Nous montrons que la variance de prédiction est grandement sous-estimée lorsqu'on ne tient pas compte du fractionnement de la variance et du fait que, dans le cas des arbres, les erreurs de prédiction ne sont pas indépendantes des variables explicatives.

[Traduit par la Rédaction]

Introduction

The estimation of biomass at the tree level and the subsequent step of scaling up biomass to the stand and eventually the regional level using forest inventory data is an essential component of monitoring carbon storage in forests (Kauppi et al. 1992; Nabuurs et al. 2003; Liski et al. 2006). Advances in the quality and the efficiency of carbon monitoring will affect decisions on climate politics and energy politics (Raupach et al. 2005). Furthermore, accurate forest carbon stocks are important to validate models (Vanclay and Skovsgaard 1997; Thurig and Schelhaas 2006) and for validating spatial extrapolations based on remote sensing (Lu 2006). The basis for the assessment of forest carbon stocks are biomass equations. These equations relate variables that are commonly measured during forest inventories, such as tree diameter at breast height, to dry mass of biomass compartments. They may be applied directly at the tree level or as a component of biomass expansion factors, which operate on aggregated data at the stand level (Wirth et al. 2004*a*; Jalkanen et al. 2005).

There are numerous studies on biomass equations of different species for different regions (Marklund 1987; Jenkins et al. 2003; Zianis et al. 2005) and also several studies for Common beech (*Fagus sylvatica* L.) (Table 1) (Hochbichler et al. 1994; Lebaube et al. 2000; Hochbichler 2002; Zianis and Mencuccini 2003, 2005). With the exception of Burger (1949–1950) and Joosten et al. (2004), the biomass equations presented in these studies are limited with respect to the number of trees (median 20 trees, minimum 7 trees, maximum 38 trees), the size range of the sample trees, and the extent of the study area (mostly single stands or catchments). None of these studies is sufficiently representative

Received 14 March 2007. Accepted 4 October 2007. Published on the NRC Research Press Web site at cjfr.nrc.ca on 27 May 2008.

T. Wutzler[1] and C. Wirth. Max-Planck-Institute for Biogeochemistry, Hans-Knöll-Str. 10, 07745 Jena, Germany. **J. Schumacher.** Friedrich Schiller University, Institute for Stochastic, Ernst-Abbe-Platz 2 07743 Jena, Germany.

[1]Corresponding author (e-mail: twutz@bgc-jena.mpg.de).

 doi:10.1139/X07-194

Table 1. Methods used and compartments sampled by 13 studies to determine tree biomass of Common beech.

ID	Study	n_{tree}	n_{plot}	Country	Compartments	Altitude (m)	Site index (m)	Variable(s)	Comments
BAR	Bartelink 1997	38	6	NL	agr, st, br, l	23	34–36	sr, cl, cp	
BUR	Burger 1949–1950	91	18	SZ	bw, l	480–1360	24.5–36	sr	br included thin stem parts and was not considered here
CIE	Cienciala et al. 2006*a*	20	4	EZ	agr, st, br	450–750	24–28	cl, cp	Reconstructed tree heights
DU1	Duvigneaud et al. 1971	7	1	BE	agr, st, br, t, bw	330	26		
DU2	Duvigneaud et al. 1977	13	1	BE	agr, st, br	250–250	24	cl	
HEI	Krauß and Heinsdorf 1996	14	3	GM	agr, st, br, l	42	28.5–30		Also quantified bark
HEL	Heller and Göttsche 1986	29	3	GM	agr, t, bw, r	500	22.5–24	sr	r only for four trees
JO	Joosten et al. 2004	116	28	GM	agr, t, bw	30–500	18–45.5		
LGO	Le Goff et al. 2004	23	2	FR	agr, st, br, l, r	300	36	sr, cl	r only for a subset
MAS	A. Masci, personal communication within project FORCAST; Schulze et al. 2003	28	1	IT	agr, st, br, l, r	1560	18	sr, cl, cd	r only for a subset
MAT	G. Matteucci, personal communication within project FORCAST; Schulze et al. 2003	30	3	GM	agr, st, br, l	440–450	32–35		l only for a subset
PEL	Pellinen 1986	19	5	GM	agr, t, bw, l, r	420	29–30.5	sr	r only for a subset, age uncertain
VYS	Vyskot 1990	15	1	EZ	agr, st, br, l, r	505	36	sr, cl, cd, cp	Also quantified twigs and stump; provides further data on root
Total		443	76	7		23–560	18–45.5		

Note: n_{tree}, number of sample trees considered; n_{plot}, number of plots where trees have been samples. Country: NL, Netherlands; SZ, Switzerland; EZ, Czech Republic; BE, Belgium; GM, Germany; FR, France; IT, Italy. Compartment: agr, aboveground wood; st, stem including bark; br, branches; t, timber (agr with diameter >7 cm); bw, brushwood (agr with diameter <7 cm); l, leaves; r, roots. Variable, in addition to diameter at breast height, tree height, age, site index, and altitude: sr, social rank; cl, crown length; cd, crown diameter; cp, crown projection.

for a nationwide monitoring of forest carbon stocks and covers a large enough environmental gradients to be applicable to such a large and diverse area as Central Europe. Furthermore, none of the above studies provides the statistical background information to allow a straightforward variance estimation of stand- and regional-level biomass (for other species, see Phillips et al. 2000; Wirth and Schumacher 2004). Nor can this information be provided by meta-analysis studies that are based on published equations instead of tree measurements (Zianis and Mencuccini 2003; Muukkonen 2007). However, uncertainties of biomass predictions are as important as the predictions themselves (e.g., Raupach et al. 2005). And finally, most studies only report data and equations for a subset of the biomass compartments. Compartments considered usually include the economically relevant aboveground woody compartments (timber, stem), less often branches and leaves, and rarely belowground compartments. However, monitoring and modelling changes of carbon stocks require the estimation of all of the biomass compartments.

In our study, we compiled available biomass data for Common beech and developed generic biomass equations applicable to a broad range of sites and situations and to all major biomass compartments. In addition, we provide tools for a realistic estimation of uncertainties of biomass predictions that account for the heterogeneity of the underlying data. We used nonlinear mixed-effects models to make inference on variance components. We outline how functions of biomass and functions of prediction variance were developed and how they can be used to estimate variances and confidence intervals for individual tree predictions as well as for estimates of biomass stocks on an aggregated level, e.g., stands, from forest inventory data. While the aggregation of stand-level predictions and uncertainties to the regional level involves further steps, our generic equations for the widespread Common beech, covering 15% of the forest area in Germany, will serve as an important component in improved national carbon monitoring.

Methods

Data

We collected biomass of tree compartments from sample trees of Common beech originating from 13 studies (Table 1), which cover the extent of Central Europe quite well (Fig. 1). The tree-level entries included biomass m of a tree compartment (kg) and as predictor variables the diameter at breast height (1.3 m) d (cm), tree height h (m), and tree age (years). Stand-level predictors are the site index "si" (mean height of trees at age 100 years (m)) and altitude "alt" (height above sea level (m)).

The biomass compartments considered are foliage ("leaves"), coarse roots ("root"), and aboveground woody biomass ("agr"). Due to the weak apical dominance of Common beech, there is a gradual transition between stem and branch wood, which renders the separation into these two compartments ambiguous (Fig. 2). In this study, we report functions for both schemes and thus for all four compartments: brushwood, timber, stem, and branches. The number of sample trees and the range of predictor values differed between compartments (Table 2).

Fig. 1. Map showing the location of study sites. The plot labels correspond to the studies listed in Table 1. The size of the symbols increases with the number of trees sampled at the location.

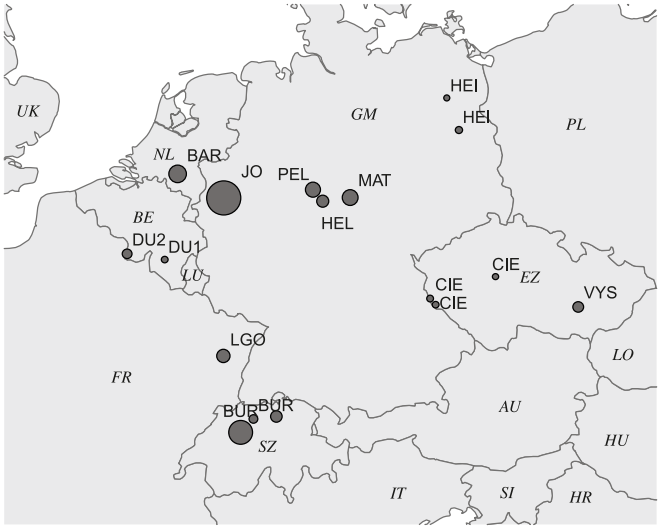
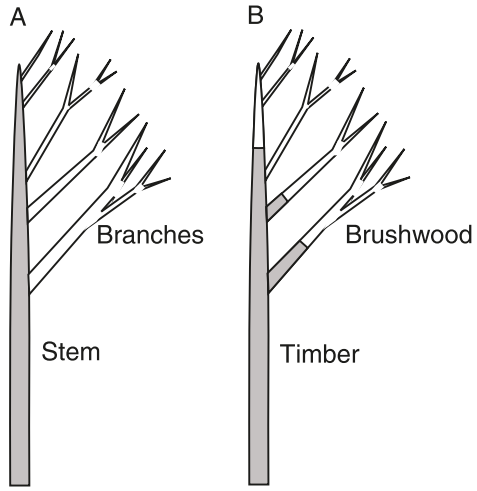


Fig. 2. Schemes of different definitions of aboveground woody biomass compartments. Scheme A distinguishes between stem and branch wood with the assumption that the main stem can be clearly identified all the way to the top. Scheme B, which is less subjective and most commonly applied in forest sciences, distinguishes between timber and brushwood based on a fixed-diameter threshold (usually 5 or 7 cm).



Basic model forms

We used three basic model forms as starting points of our model selection. First, the simplest allometric equation for predicting the biomass of a tree compartment m is a function of its diameter at breast height d (eq. 1). It can be shown that the functional form of a power function arises from the assumption that the ratio of the relative growth rates of mass and diameter (here m and d) is constant (Wenk et al. 1990). Second, to improve the predictive power, this basic allometric equation is often extended to include the tree height h as an additional predictor. Equation 2 still can be viewed as an allometric equation that relates biomass to the volume of a cylinder defined by d and h (Cienciala et al. 2006*a*; Wirth et al. 2004*b*). Third, a multiple allometric equation for predicting the biomass of a tree compartment as a multiplicative function of d and h is given by eq. 3 (Widlowski et al. 2003):

$$[1] \quad m = c_0 d^{c_1} \quad \text{(d2)}$$

$$[2] \quad m = c_0 (d^2 h)^{c_1} \quad \text{(dh2)}$$

$$[3] \quad m = c_0 d^{c_1} h^{c_2} \quad \text{(dh3)}$$

Here, m is the biomass of a tree compartment for a sample tree, d is the diameter at breast height (cm), h is the tree height (m), c_s, i.e., c_0, c_1, and c_2, are model coefficients to be estimated. The number at the end of the equation labels (d2, dh2, and dh3) indicates the number of parameters.

Rational of using nonlinear mixed-effects models

We used nonlinear mixed-effects models (Lindstrom and Bates 1990) to directly fit the candidate models to the data. The main advantage of using mixed-effects models lies in their capability to account for groupings in residual variance due to random effects (Pinheiro and Bates 2000). In our case, the grouping variable is the study from which the data originate. The mere fact that sample trees from one study usually share a common provenance and were collected in the same environment by the same team of scientists with a specific set of methods often causes their residuals to be consistently lower or higher than the mean predictions of a fixed-effects model. This violates a fundamental assumption of independent residuals in conventional regression analysis (Crawley 2002) and will lead to an underestimation of variance. Mixed-effect models are one way to adequately address this type of data heterogeneity. Instead of assuming the same fixed effects across all groups (eq. 4*a*), mixed models allow the coefficients c_s of the model (eqs. 1–3) to vary between groups by adding a group-dependent random effect (eq. 4*b*). Additionally, they allow ncluding covariates that in part explain the deviation from generic coefficient value β_s (eq. 4*c*):

$$[4a] \quad c_s = \beta_s$$

$$[4b] \quad c_s = \beta_s + b_{s,i}$$

$$[4c] \quad c_s = \beta_s + b_{s,\,i} + \beta_{s,\,\text{age}} \times \text{age} + \beta_{s,\,\text{si}} \times \text{si} + \beta_{s,\,\text{alt}} \times \text{alt} + \beta_{s,\,\text{age}\times\text{alt}} \times \text{age} \times \text{alt} + \beta_{s,\,\text{si}\times\text{alt}} \times \text{si} \times \text{alt}$$

where β_s and $\beta_{s,\,\langle\text{covar}\rangle}$ are fixed effects, $b_{s,\,i}$ is the study-dependent random effect, and age, site index (si), and altitude (alt) are covariates.

For a single level of grouping to studies, the tree compartment biomass m_{ij} of tree j from the ith study can be expressed as the target by a single-level mixed-effects model (Lindstrom and Bates 1990):

$$[5] \quad m_{ij} = y_{ij} = f(\boldsymbol{x}_{ij}, \boldsymbol{\nu}_{ij}; \boldsymbol{\beta}, \boldsymbol{b}_i) + \varepsilon_{ij}$$

Table 2. Number of sample trees and range of predictors by tree compartment.

Compartment	n_{tree}	n_{study}	Diameter at breast height (cm)	Height (m)	Age (years)	Site index (m)	Altitude (m)
Aboveground woody biomass	350	12	1–79	2–37	8–173	18–46	23–1560
Stem	187	9	2–79	3–37	8–165	18–36	23–1560
Branches	175	8	2–64	3–37	8–165	18–36	23–1560
Timber	170	4	1–79	2–35	18–173	18–46	30–500
Brushwood	276	6	1–79	2–40	14–173	18–46	30–1360
Root	48	5	3–38	7–29	21–160	18–36	300–1560
Leaves	247	8	1–62	3–40	8–165	18–36	23–1560

$$\boldsymbol{b}_i \sim N(0, \boldsymbol{\Psi})$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

where f is a general, real-valued, differentiable function (one of eqs. 1–3, with coefficients expressed as eq. 4*c*) of a vector of predictors $\boldsymbol{x}_{ij}(d,h)$, a covariate vector $\boldsymbol{\nu}_{ij}$ (age, si, alt), the vector of fixed effects $\boldsymbol{\beta}$, and study-dependent vector of random effects $\boldsymbol{b}_i$ (Pinheiro and Bates 2000). In this study, we used an unconstrained symmetric positive-definite covariance matrix $\boldsymbol{\Psi}$ of random effects. We used extended mixed-effects models, where the assumption of the within-group residuals $\varepsilon_i = (\varepsilon_{i1}, ..., \varepsilon_{in})$ is relaxed by $\varepsilon_i \sim N(0, \sigma^2, \boldsymbol{\Lambda}_i)$ and where $\boldsymbol{\Lambda}_i$ are positive-definite matrices parameterized by a set of parameters δ_i. Typically, the variance of the residuals ε_j increases with compartment biomass m_j for tree j at the original scale. We thus explicitly modelled the variance of the residuals by a power function:

$$[6] \qquad \mathrm{Var}(\varepsilon_j) = g(\varepsilon_j; \sigma^2, \delta) = \sigma^2 |m_j|^{2\delta}$$

Not only the coefficients but also the residual variance may potentially vary between studies. We accounted for this by modifying the model of residual variance (eq. 6) to eq. 7 by replacing the single exponent δ by the group-dependent exponent δ_i:

$$[7] \qquad \mathrm{Var}(\varepsilon_{ij}) = g(m_{ij}; \sigma^2, \delta_i) = \sigma^2 |m_{ij}|^{2\delta_i}$$

for the ith study and the jth observation. The parameters σ^2 and δ_i were estimated by iteratively reweighted sum of squares simultaneously with the other coefficients in the model-fitting algorithm.

We used nonlinear models (Lindstrom and Bates 1990) for the following reasons. For model fitting, the biometric data are often log transformed to linearize the allometric equation and to homogenize the variance, which otherwise increases with size on the original scale (Baskerville 1972). However, the back-transformation of the predicted value to the original scale introduces biases in the expected values and the uncertainties (Smith 1993). The proper dealing with these biases introduces new assumptions, and the simple nonparametric correction using the smearing estimate (Duan 1983; Taylor 1986), as it is frequently employed (e.g., Joosten et al. 2004), is not directly applicable if mixed-effects models are used (Wirth et al. 2004*b*). In addition, the logarithmic form does not allow inclusion of the covariances between predictions errors at the original scale when calculating the variance of biomass prediction errors for several trees (Appendix B, eq. *B*1). In Appendix B, we show how the variance of new predictions can be propagated to aggregated levels.

Fixed-effects models partition the variance around the mean prediction into the variance arising from uncertainty in the parameter estimates and into residual variance. In mixed-effects models, a further component is added: the variance that is induced by the random effects, which represents groupings in the data (Appendix A, eq. *A*2)

In our case, the random effect accounts for all implicit differences between sites, provenances, methods, etc., associated with the sample material and sites of different studies that are not represented by any specific predictors. Besides the grouping of variances according to studies, there is potentially also a grouping according to stands. In addition to the presented results, we tried to fit two-level mixed models to account for this additional grouping level. However, the highly unbalanced design of the data, i.e., several studies include only one stand, and the differences in variances of random effects between the studies caused problems in the numerical algorithm to fit the two-level mixed model and we concluded that the available data were not sufficient to account for groupings at both the stand and study level at the same time. Diagnostic graphs of stand-level random effects and study-level random effects showed that groupings according to studies were much more pronounced than groupings according to stands. In the presented approach, the stand-level differences are partly accounted for by the study-level random effect. This approach can be seen as a pooling of the comparatively similar stands of each study to obtain enough within-group cases.

It shall be noted that the equations for the different compartments are based on different subsets of the whole data set. Hence, we separately fitted the models for the different biomass compartments and the derived biomass equations are not additive (Parresol 2001; Lambert et al. 2005).

Definition of the candidate model set

To find an appropriate model, we compared 246 models for each biomass compartment by the Akaike information criterion (AIC) (Akaike 1987). Table 3 summarizes the different dimension of constructing the candidate model set.

First, we tested which coefficients (c_s) differed randomly by study. We fitted the models (eqs. 1–3) with all eight combinations of random effects (either eq. 4*a* or eq. 4*b*) for the model coefficients and compared model performance using the AIC. The nonlinear fixed- or mixed-effects model

Table 3. Terms used to describe the dimensions of the candidate model set.

Symbol	Description	Equation(s)
d2, dh2, dh3	Basic model forms	1–3
ran	Inclusion of study random effects $b_{s,\,i}$	4*b*
c	Inclusion of covariate effects $\beta_{s,\,\langle\text{covar}\rangle}$	4*c*
ranres	Inclusion of the study effect in the residual model	7 instead of 6

after this step is called the "best dh model". In some cases with random effects in three or two coefficients, the model-fitting algorithm did not converge. We excluded these cases from the candidate model set.

Second, we included covariates in the model coefficients (eq. 4*c*). To reduce the number of possible combinations and to avoid overly complex models, we constructed our candidate models according to the following three rules: (*i*) Higher-order polynomial terms for any of the covariates were avoided because such models are difficult to extrapolate. (*ii*) A covariate effect was added to a coefficient c_s only if the same covariate effect was not already associated with another coefficient. This was done because including covariates on correlated coefficients caused unrealistically high counteracting effects (e.g., c_0 strongly increases with age, while c_2 decreases with age). (*iii*) Covariates were added to either c_1 or c_2 in the dh3 model form (eq. 3) because diameter and height were highly correlated and the AIC hardly changed when the covariate was included in either one of the two coefficients. If the dh model with random effects in c_0 and c_2 resulted in a better AIC than the dh model with random effects in c_0 and c_1, we selected c_2 else c_1. In total, this led to 41 combinations of the terms in eq. 4 across the three coefficients. Finally, the resulting best model, which included the random effects from the first step, was compared with the models with covariates but with fewer random effects. We call the resulting best model after this step the "best dhc model".

Third, we assessed if including a study-dependent formulation of the residual variance increased model performance by replacing eq. 6 by eq. 7.

The model fitting was done using the nlme and gnls functions using a general positive-definite variance-covariance matrix using the maximum likelihood (ML) method for model selection. The resulting best models were refitted using the restricted maximum likelihood (REML) method to improve estimates for the variance components. We used R-nlme library version 3.1-66 (Pinheiro and Bates 2000).

Calculation of confidence intervals

We computed symmetrical 95% confidence intervals around single tree biomass predictions with width $w_{\text{cf95}} = 2 \times 1.96 \times \sqrt{\text{Var}(m_{\text{new}})}$. The variances of the predictions were computed with eq. *A*2 and the residual variance component $\text{Var}(\varepsilon_{\text{new}})$ was determined with eq. *A*5. A numerical example for the tree-level calculation is given in Appendix C.

Table 4. Comparison of the model forms and inclusion of random components by the AIC for stem biomass.

Form	AIC	df
dh3_c_ran_ranres	1507.3	17
dh2_c_ranres	1509.0	16
dh2_ran_ranres	1512.8	15
dh3_ranres	1538.3	13
dh3_c_ran	1567.2	9
dh3_ran	1571.3	8
dh2_c	1596.5	7
dh2	1623.6	4

Note: Results of the other compartments are given in electronic supplement S2.

At an aggregated level, the 95% confidence intervals around the biomass predictions were calculated with the same equation as for the tree level. However, the variance of the prediction error of the sum of the biomass of several trees has to include covariances between the single tree-level prediction errors (eqs. *B*1 and *B*3). The covariance between predictions errors of two trees explicitly depends on the predictors and covariates of the two trees (eq. *B*2). All calculations were programmed with the statistical software R.[2] The calculation of variance at the stand level was exemplified[3] using data of an inventory of a chronosequence of shelterwood beech forests in Thuringia, Germany (Mund 2004).[4]

Cross-validation and comparison with previously published biomass functions

The validity of the presented model and its performance in comparison with published studies is demonstrated using cross-validation (Davison and Hinkley 1997). This was done by comparing the predictive performance of a range of published functions with the results of the cross-validation of our generic functions. In the cross-validation, the observations that were used for validation were not used to fit the model, i.e., the biomass of a tree from a given study was predicted with our generic model, but the parameterization of the respective model was based only on the data of all other published studies. For each model, we calculated the

[2] www.r-project.org. The programmed R-model objects including variance-covariance matrices, derivative functions, and additional functions to apply prediction and uncertainty calculation at the tree and stand level are provided as supplementary material S5. Appendices S1–S7 are available as supplementary data for this article and are available on the journal Web site (cjfr.nrc.ca) or may be purchased from the Depository of Unpublished Data, Document Delivery, CISTI, National Research Council Canada, Building M-55, 1200 Montreal Road, Ottawa, ON K1A 0R6, Canada. DUD 3734. For more information on obtaining material, refer to cisti-icist.nrc-cnrc.gc.ca/cms/unpub_e.shtml.

[3] R-code is provided with electronic supplementary material S6.

[4] Details of the inventory are provided with electronic supplementary material S1.

Table 5. Regression coefficients (β_i) and their standard errors (in parentheses) of the best models of diameter and height (dh models, eqs. 1–3, with coefficients as in eq. 4*b*).

Compartment	Form	β_0	β_1	β_2
Aboveground woody biomass	dh3_ranres	0.0523 (0.00330)	2.12 (0.0225)	0.655 (0.0394)
Stem	dh2_ran_ranres	0.0293 (0.00406)	0.974 (0.0163)	
Branches	dh3_ran_ranres	0.123 (0.0296)	3.09 (0.107)	–1.17 (0.175)
Timber	dh3_ranres	0.00775 (0.00125)	2.11 (0.0406)	1.21 (0.0841)
Brushwood	dh3_ran_ranres	0.466 (0.0862)	1.85 (0.100)	–0.349 (0.150)
Root	d2_ran_ranres	0.0282 (0.00263)	2.39 (0.0467)	
Leaves	dh3_ran_ranres	0.0377 (0.00686)	2.43 (0.0817)	–0.913 (0.120)

Note: This table is also available in electronic supplementary S2. Empty cells denote that the term is not included in the model.

root mean squared error (RMSE) with the modification of applying a weighted mean. The weights were the inverse of the square root of the expected variance according to eq. *A*5.

Results

Comparison of the models without additional covariates

The three-parameter model dh3 (eq. 3) fitted the data best for most compartments (see Table 4 for the example of stem biomass).[5] Only for the stem and timber biomass the dh2 model (eq. 2) and for the root biomass the d2-model (eq. 1) showed similar or better performance. In all cases, the inclusion of parameters that allow the variance to differ between studies (eq. 7) resulted in the largest improvement of model performance. In comparison, the mere inclusion of random components into the model coefficients improved the model performance only slightly for the compartments stem, branches, roots, and leaves. However, it did not improve the model performance at all for the compartments aboveground wood and timber. The coefficients of the best dh models are listed in Table 5.

Effect of additional covariates

The inclusion of the additional covariates age, site index, or altitude improved the model performance in all cases (see Table 4 for the example of stem biomass).[6] The highest improvement was achieved for the compartments aboveground wood, branches, brushwood, and leaves (see Fig. 4; compare neighboring graphs). The coefficients of the best dhc models are listed in Table 6. We will come back to the magnitude and sign of individual coefficients and their interpretation in the Discussion section.

Variance of predictions

The estimated parameters that are needed to calculate the residual variance (eq. *A*5) are listed in Table 7. The standard deviations of the random effects with the dhc models in Table 7 were smaller compared with those of the corresponding dh models. This indicates that a large part of the variance that was previously accounted for in the random effects component was now accounted for by the covariates.

The 95% confidence intervals of several single tree biomass predictions by the best dh models (Table 5) and dhc models (Table 6) are shown in Fig. 3. In all cases, the width of the confidence intervals strongly increased with the size of predicted biomass. This represented the increasing variability in the observed biomass that was modelled with a power model (eq. 6 or 7, respectively). The amount of prediction variance differed considerably between the different compartments. The confidence intervals for aboveground wood biomass and timber are comparatively narrow (both have a coefficient of variation (CV) = 12% for tree age of about 70 years). However, confidence intervals of single tree biomass predictions of branches, root, and leaves are very wide (CV = 59%, 35%, and 49%, respectively). The inclusion of additional covariates in the models resulted in narrower confidence intervals in most cases (compare neighboring graphs in Fig. 3). However, the confidence intervals for branches and leaves were still comparatively wide (CV = 45% and 43%). Inclusion of additional covariates also accounted for a large part of the variance that, before, in the dh models was attributed to unknown random effects between the studies (see background bars in Fig. 3).

The biomass equations were developed at the tree level. However, many applications make inference at the aggregated level, e.g., stands, by taking the sum of predicted biomass of all trees. At the stand level, the confidence intervals were much narrower (Fig. 4) (CV = 2.5% and 2.9% for aboveground wood and timber) compared with the corresponding confidence intervals at the tree level (Fig. 3). The effect of wider confidence intervals for higher biomass predictions was still present but not as pronounced as with the single tree biomass predictions. The relative contribution of model residuals to total variance was much smaller (background bars in Fig. 4). The decrease in total variance upon inclusion of covariates was much stronger than at the tree level (CV = 34% to 20%, 18% to 8%, and 20% to 10% for branches, root, and leaves, respectively, at stand age 69 years).

The different model forms are compared exemplarily for stem biomass at the tree level in Fig. 5. The predictions were very similar for average covariate values. However, the variability between studies added uncertainty to the biomass prediction. The mixed-effects dh model accounted for these differences by random effects, which resulted in a wider confidence band. The mixed-effects dhc model explained these differences in part by additional covariates and the width of the confidence band was smaller than without covariates. The fixed-effects model neglected the intergroup variability and underestimated the variance. Hence,

[5] For other compartments, see electronic supplementary material S2.
[6] For other compartments, see electronic supplementary material S2.

Table 6. Regression coefficients and their standard errors (in parentheses) for the models that include additional covariates (dhc models, eqs. 1–3, with coefficients as in eq. 4*c*).

	Compartment and form						
	Aboveground woody biomass, dh3_ran_ranres	Stem, dh3_ran_ranres	Branches, dh3_ran_ranres	Timber, dh2_ran_ranres	Brushwood, dh3_ranres	Root, d2_ranres	Leaves, dh3_ranres
β_0	0.0551 (0.00463)	0.00351 (0.00704)	0.122 (0.0294)	0.0106 (9.43×10^{-4})	0.805 (0.159)	0.0292 (0.00225)	0.0561 (0.00882)
β_1	2.11 (0.0242)	1.84 (0.0333)	3.09 (0.106)	1.08 (0.00795)	1.83 (0.113)	1.70 (0.0792)	2.07 (0.0770)
β_2	0.589 (0.0427)	1.04 (0.0548)	–0.151 (0.304)		–0.560 (0.149)		–1.09 (0.112)
$\beta_{s,\,\mathrm{age}}$	$s = 2$: 4.06×10^{-4} (1.07×10^{-4})	$s = 0$: 3.47×10^{-5} (2.43×10^{-5})			$s = 1$: 0.00134 (3.03×10^{-4})	$s = 0$: 4.36×10^{-5} (1.65×10^{-5})	
$\beta_{s,\,\mathrm{si}}$	$s = 0$: 2.39×10^{-4} (8.15×10^{-5})	$s = 0$: 6.72×10^{-4} (2.38×10^{-4})	$s = 2$: –0.0309 (0.00783)			$s = 1$: 0.0209 (0.00387)	$s = 1$: 0.0137 (0.00192)
$\beta_{s,\,\mathrm{alt}}$	$s = 0$: -4.68×10^{-6} (1.39×10^{-6})	$s = 0$: 8.11×10^{-6} (2.77×10^{-6})	$s = 2$: -9.87×10^{-4} (2.58×10^{-4})	$s = 0$: -1.54×10^{-6} (5.84×10^{-7})	$s = 1$: -1.68×10^{-4} (5.26×10^{-5})	$s = 1$: 7.43×10^{-4} (1.72×10^{-4})	$s = 0$: -3.29×10^{-6} (3.62×10^{-6})
$\beta_{s,\,\mathrm{sixalt}}$			$s = 2$: 3.06×10^{-5} (1.08×10^{-5})			$s = 1$: -2.70×10^{-5} (8.55×10^{-6})	

Note: See Table 5; see Appendix C for an example of constructing the full equation.

the confidence band for the fixed-effects dh model was too narrow. However, this underestimation of variance due to ignoring differences between studies was small compared with ignoring covariances between single tree prediction errors at the stand scale (Fig. 6).

Cross-validation

The advantages of generic models of the dh and dhc type can be evaluated by comparison with other published beech biomass functions for Central Europe that are based on far smaller data sets. The stem biomass predictions of the cross-validation of the dh models (including only diameter and height as predictors) were very similar to the predictions of the previously published equations by Cienciala et al. (2006*a*) and Bartelink (1997) across a range of sizes (Fig. 7*a*). For our generic functions, the inclusion of additional covariates in the dhc model improved the model fit slightly (weighted root mean squared error decreased from 21 to 18 kg). For foliage biomass, however, the larger data set and the inclusion of covariates led to a more pronounced improvement of the model fit and a reduction of estimated biomass (Fig. 7*b*) in comparison with the predictions according to Bartelink (1997) and Le Goff and Ottorini (2001).

Discussion

Our study provides the first comprehensive set of functions for predicting biomass for Common beech in Central Europe for all major tree compartments. Combining original tree biomass data from many sites across Central Europe, which varied in climate and soil characteristics, it was possible to develop generic equations that are representative for the great majority of beech sites in the study region. However, the obvious advantages of combining data from various sources comes at the cost of data heterogeneity, which can only be appropriately dealt with by using nonstandard statistical methods (Bates and Watts 1988; Wirth et al. 2004*b*). We used nonlinear mixed-effects models (Pinheiro and Bates 2000) that have been successfully applied in forestry studies for trunk circumference (Lindstrom and Bates 1990), tree height (Fang and Bailey 2001; Hall and Bailey 2001; Calama and Montero 2004; Calegario et al. 2005), stand and bole volume (Gregoire and Schabenberger 1996; Fang et al. 2001; Zhao et al. 2005), yield (Hall and Clutter 2004), and biomass (Wirth et al. 2004*b*). Using nonlinear mixed models allowed us to propagate different sources of variance and to analyse the structure of variance when scaling up from the tree level to the aggregated level.

Our study illustrates the continuum between fixed- and random-effects models. When we compare models with and without additional covariates (dh and dhc models), the contribution of the random effects to the total variance was much smaller in the dhc models than in the dh models (Figs. 4 and 5). The effects of the covariates were formerly accounted for in part by the random effects in the dh models. This finding corroborates the finding of Fang et al. (2001) where the inclusion of the predictor dominant height accounted for the differences between stands that were formerly accounted for by random effects. The effect of covariates was less pronounced for the biomass compartments branch, brushwood, and stem biomass (Fig. 3). We hypothe-

Table 7. Uncertainty coefficients of the best models.

Compartment	Form	σ^2	Mean(δ_i)	Var(δ_i)	sd($b_{0,i}$)	sd($b_{1,i}$)	sd($b_{2,i}$)
dh models							
Aboveground woody biomass	dh3	0.166	0.770	0.00965			
Stem	dh2	0.0708	0.873	0.0177	3.47×10^{-7}	0.0176	
Branches	dh3	0.249	0.863	0.0274			0.132
Timber	dh3	1.29	0.614	0.0104			
Brushwood	dh3	0.149	0.979	0.00448		0.0811	
Root	d2	0.0432	0.902	0.118		0.0660	
Leaves	dh3	0.179	0.854	0.117	0.00882		
dhc models							
Aboveground woody biomass	dh3	0.142	0.782	0.00912			
Stem	dh3	0.0903	0.842	0.0196	0.00316		
Branches	dh3	0.250	0.860	0.0260			0.0690
Timber	dh2	1.29	0.615	0.00779			
Brushwood	dh3	0.191	0.935	0.00434		0.0661	
Root	d2	0.0410	0.874	0.185			
Leaves	dh3	0.174	0.858	0.121			

Note: See Table 5. The estimated covariance matrices of the fixed and the random effects are provided in electronic supplementary S4 in csv format. σ, base variance; δ_i, power of variance for study i; sd($b_{s,i}$), estimated standard deviation of the random coefficients.

Fig. 3. Tree-level predictions and 95% confidence intervals of the best dh models (left panels) and the best dhc models (right panels) for each biomass compartment, respectively. The background bars represent the proportions of the different variance components (residual, random effects, and fixed effects). The predictors site index and altitude were kept constant (si = 30 m, alt = 470 m). The predictors height and age were chosen to be consistent with the site index and diameter.

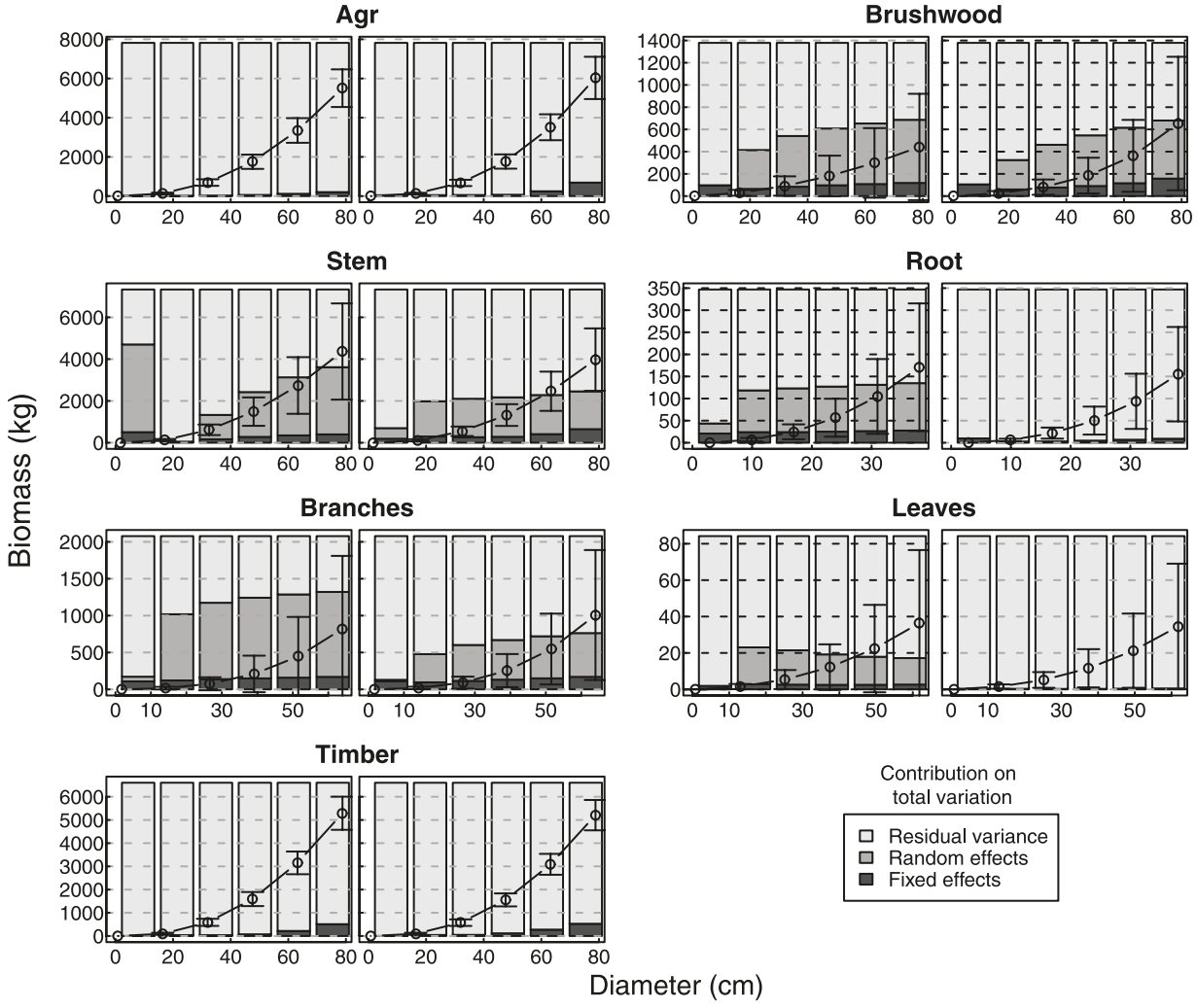
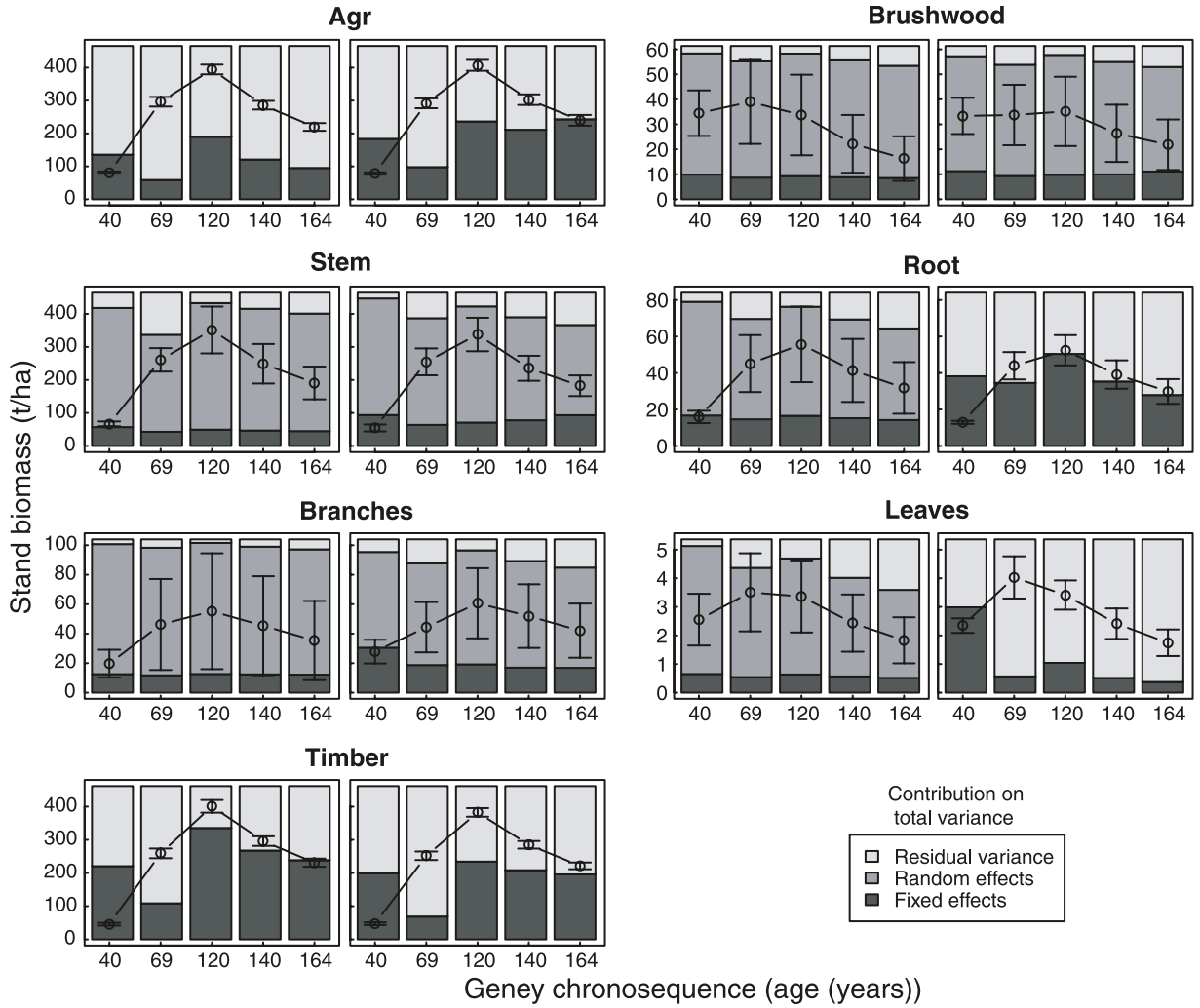


Fig. 4. Predictions and 95% confidence intervals of the best dh models (left panels) and the best dhc models (right panels) for biomass predictions at the stand level for five stands of a shelterwood chronosequence. The background bars represent the proportions of the different variance components as in Fig. 3.



size that this is in part caused by the subjectivity involved in the separation of the stem and branch compartments and that thus there are inherently large differences between studies that are not due to environmental conditions but unknown differences in sampling protocols of different teams. This is confirmed by the fact that the random effects almost disappear for the sum of the two compartments (aboveground woody biomass). In this context, it is important to realize that the inclusion of covariates is only possible if, as in our case, data from many stands covering a range of ages and site conditions are pooled.

We generally observed that the random effect of the variable "study" was small for biomass of stem, timber, and aboveground wood (Tables 4 and 5). This indicates that study-specific effects are relatively small compared with the dominating effect of the predictors diameter and height. Hence, the predictions of stem biomass did not vary much when we compared equations from different studies (Fig. 7*a*). However, we observed that models accounting for grouping effects in the residual variance performed better (eq. 7; Table 4). This implies that although the mean prediction was similar, the estimated variance of the biomass did vary between studies. The data, which were used in our study, do not allow us to distinguish whether this was an effect of differences in the sampling scheme between the studies or a real effect of differences in growth variability between the studies. For other tree compartments, the random effects associated with the study were larger. Hence, for a specific new inventory, the biomass predictions will be more strongly biased towards the mean across studies. If a few additional biomass measurements for the new inventory are available, it is possible to estimate the specific values of the random effects (e.g., Nothdurft et al. 2006). However, in most applications, additional measurements of tree biomass compartments are too expensive. Lappi (1991) provides methods to estimate the values of the random effects of linear mixed-effects models for volume equations by related equations that require only diameter and height measurements. To develop similar related equations for the nonlinear tree biomass equations presented here is beyond the scope of this paper and warrants further study.

Applying biomass functions from a single study outside the reference area will inevitable ignore the site influence on biomass allocation patterns (see discussion below) and

Fig. 5. 95% confidence intervals of stem biomass for different forms of modelling variation between groups. The predictors site index and altitude were given intermediate values (si = 30 m, alt = 470 m). The predictors age, diameter, and height were chosen to represent typical values for these conditions from the data set of sample trees.

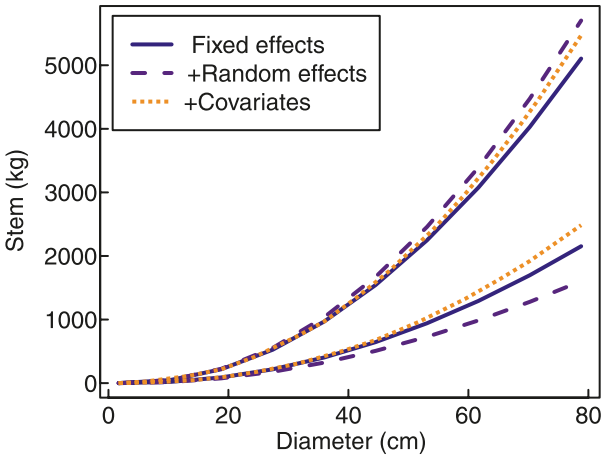
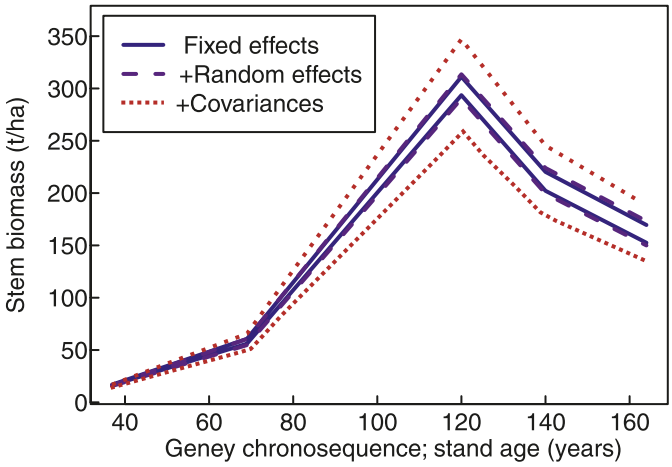


Fig. 6. 95% confidence bands of stem biomass at stands of the Geney beech forest chronosequence. None of the displayed model results used additional covariates.



will also underestimate variance (Fig. 5). This is because the biomass function does not account for the differences between the studies, neither implicitly by random effects nor explicitly by covariates.

In the statistical analysis, we assumed that the measurement error of the predictors does not have a profound effect on the estimation of the model coefficients (Table 6). Diameter, height, stand age, and altitude have been measured with high precision in the considered studies. For the covariate site index, which was in some cases estimated by age and height, we performed a Monte Carlo study for the aboveground biomass where we varied the site index randomly with a standard deviation of ±1 m and refitted the best dhc model. The additional uncertainty introduced in the estimates of the model coefficients ranged only from 3% to 24% of the standard errors in Table 6 for the for coefficients β_1 and $\beta_{0\,si}$, respectively. Hence, we conclude that the uncertainty in predictor site index has a sufficiently small effect on the results and does not change our interpretations.

Although certainly not all features of the models can be readily interpreted, some obvious biologically plausible patterns emerged that can be related to well-known allocation patterns, thus increasing our confidence in the model predictions.

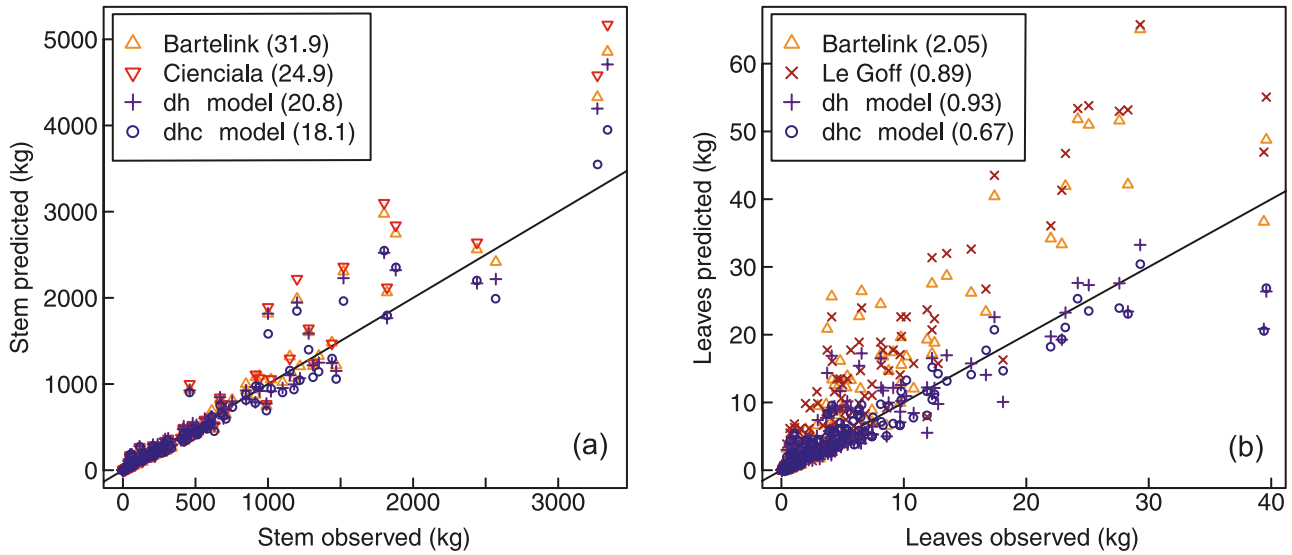
The three-parameter model fitted the data of most biomass compartments best (Table 4). The better performance of the dh models showed that the parameter tree height is an important additional predictor of biomass as observed in other studies for beech (Zianis and Mencuccini 2003; Joosten et al. 2004; Cienciala et al. 2006*a*) as well as in studies of other species (Cienciala et al. 2006*b*; Montagu et al. 2005). However, tree height did not significantly influence root biomass, as indicated by the best performance of the d2 model form (eq. 1). This corresponds to findings for Norway spruce (*Picea abies* (L.) Karst.) (Wirth et al. 2004*b*).

The parameter β_2 associated with the predictor tree height was negative for all crown compartments. This means, that (at a given diameter) higher trees tended to have a lower biomass of crown compartments. We think that this is most likely due to the fact that individuals with a high height to diameter ratio tend to be suppressed trees with an elevated allocation to stem growth at the expense of allocation to crown biomass (Nilsson and Albrektson 1993; Vanninen et al. 1996; Wirth et al. 2004*b*). Negative values of β_2 have also been observed in a similar study on Norway spruce (Wirth et al. 2004*b*)

At a given diameter and height, stem biomass increased with stand age in the best dhc model. This may be related to a negative correlation between wood density and ring width (Bouriaud et al. 2004). At a given diameter and height, older trees have more and thus smaller tree rings. This implies a higher wood density and hence higher biomass. The fact that the best model included also the site index and altitude as covariates with positive coefficients suggests additional environmental modulation of wood density that warrants further investigation.

Biomass equations are usually applied to make inferences at the aggregated level. For upscaling, the sum of the biomass predictions of many single trees, e.g., within one stand, is calculated. When calculating the variance of the sum, prediction errors of the single trees are usually regarded as independent of each other for simplicity's sake. Instead, our statistical approach accounts for covariances between prediction errors for several trees. The residuals of different trees are still considered independent. However, biomass predictions based on uncertain model coefficients deviate from the prediction that would result if the true (but unknown) model coefficients were used. The deviations of the predictions have the same direction for similar predictor values and therefore have a positive covariance (for a more formal description, see Appendix B). This issue is independent of using fixed-effects models, single-level random-effects models, or multilevel mixed-effects models. We showed how much the variance of biomass predictions at the aggregated level is underestimated when covariances between single tree prediction errors are neglected (Fig. 6). This was already shown by exploring different assumptions about the

Fig. 7. Predictions of the models fitted in the cross-validation of (a) stem and (b) leaves biomass and comparison with previously published biomass functions: $m_{stem} = 0.0109d^{1.951}h^{1.262}$ (Bartelink 1997), $m_{stem} = 0.014d^{2.053}h^{1.084}$(Cienciala et al. 2006a), $m_{leaves} = 0.0167d^{2.951}h^{-1.101}$ (Bartelink 1997), and $m_{leaves} = \exp(-4.8599 + 2.1935\ln(d))$ (Le Goff and Ottorini 2001). Values in parentheses indicate the variance-weighted root mean squared error.



covariances (Lehtonen et al. 2004, 2007); however, we provide the approach to actually quantify the covariances.

Further, we demonstrated that the reduction of variance upon scaling from the tree level to the aggregated level depends on the partitioning of variance. When differences between studies were represented explicitly by additional covariates instead of random effects, the relative contribution of fixed and random effects to total variance on tree level decreased (Fig. 3, background bars). Because of the linear scaling of the residual variance with the number of trees, the coefficient of variation of the predicted biomass at the stand level decreases with the square root of the number of trees if trees are regarded independent. However, variance attributed to the uncertainty of fixed and random effects scale in a quadratic manner with the number of trees (Appendix B); these partitions of variance became much more important at aggregated level. Hence, the decreased contribution of fixed and random variance at the tree level with the inclusion of additional covariates led to a large decrease of total variance at the aggregated level (Fig. 4, background bars). This finding highlights the importance of factoring out variance components at the original, i.e., not log-transformed, scale.

Conclusions

This study presents generic biomass equations of seven biomass compartments for Common beech trees in Central Europe. A meta-analysis of biomass measurements of 443 trees of 76 sites from 13 studies across Central Europe enabled the assessment of the effect of the covariates age, site index, and altitude on tree biomass. Further, our study illustrates for the first time the importance of separating variance components (residual, fixed, random) in the context of scaling up uncertainties from tree to the aggregated level.

(1) Leaves and branch biomass prediction varied considerably across Central European studies. Using our large data set for calibration improved model performance most for these compartments in comparison with previously published functions. Stem and above ground biomass did not vary as much, but model performance still improved slightly.

(2) In addition to mean predictions, the variability of tree biomassalso differed between studies in Central Europe. Biomass functions based on a data set of a single study did not account for the implicit differences between studies. Hence, using these functions outside the calibration area underestimates the variance of the prediction error for new biomass predictions.

(3) The covariates age, site index, and altitude modulated the effect of diameter and height. These additional variables accounted for a large part of the differences in biomass predictions between studies, which were otherwise accounted for by the random effects. Hence, the inclusion of these covariates increased model performance for several biomass compartments and reduced prediction variance.

(4) The prediction errors of trees are correlated because of uncertain model coefficients. Neglecting these correlations when scaling up biomass to the aggregated level underestimates prediction variance significantly. We developed equations and tools to quantify the covariances between single tree prediction errors as well as for upscaling.

Acknowledgements

We thank the contributors of the original tree biomass data for providing us with the results of their time-consuming and labour-intensive work. We especially thank Giorgio Matteuci and Emil Cienciala for additional feedback on the details of the data and the sampling. We thank three anonymous referees for their constructive comments.

References

Akaike, H. 1987. Factor-analysis and AIC. Psychometrika, **52**: 317–332. doi:10.1007/BF02294359.

Bartelink, H.H. 1997. Allometric relationships for biomass and leaf area of beech (*Fagus sylvatica* L). Ann. Sci. For. **54**: 39–50. doi:10.1051/forest:19970104.

Baskerville, G.L. 1972. Use of logarithmic regression in the estimation of plant biomass. Can. J. For. Res. **2**: 49–53. doi:10.1139/x72-009.

Bates, D.M., and Watts, D.G. 1988. Nonlinear regression analysis and its applications. Wiley, New York.

Bouriaud, O., Breda, N., Le Moguedec, G., and Nepveu, G. 2004. Modelling variability of wood density in beech as affected by ring age, radial growth and climate. Trees: Struct. Funct. **18**: 264–276. doi:10.1007/s00468-003-0303-x.

Burger, H. 1949/50. Holz, Blattmenge und Zuwachs (X): Die Buche. Mitt. Schweiz. Anst. Forstl. Versuchswes. **26**: 419–468.

Calama, R., and Montero, G. 2004. Interregional nonlinear height–diameter model with random coefficients for stone pine in Spain. Can. J. For. Res. **34**: 150–163. doi:10.1139/x03-199.

Calegario, N., Daniels, R.E., Maestri, R., and Neiva, R. 2005. Modeling dominant height growth based on nonlinear mixed-effects model: a clonal *Eucalyptus* plantation case study. For. Ecol. Manag. **204**: 11–20. doi:10.1016/j.foreco.2004.07.051.

Cienciala, E., Cerny, M., Apltauer, J., and Exnerova, Z. 2006*a*. Biomass functions applicable to European beech. J. For. Sci. **51**: 147–154.

Cienciala, E., Cerny, M., Tatarinov, F., Apltauer, J., and Exnerova, Z. 2006*b*. Biomass functions applicable to Scots pine. Trees: Struct. Funct. **20**: 483–495.

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

Davison, A.C., and Hinkley, D.V. 1997. Bootstrap methods and their application. Cambridge University Press, Cambridge, U.K.

Duan, N. 1983. Smearing estimate — a nonparametric retransformation method. J. Am. Stat. Assoc. **78**: 605–610. doi:10.2307/2288126.

Duvigneaud, P., Denaeyer, S., Ambroes, P., and Timperman, J. 1971. Recherches sur l'écosystème forêt. Inst. R. Sci. Nat. Belg. Mem. **164**: 101.

Duvigneaud, P., Kestemont, P., Timperman, J., and Moniquet, J.-C. 1977. La hetraie ardennaise à *Festuca Altissima* à mirwart biomasse et productivité primaire. *In* Productivité biologique en Belgique. *Edited by* P. Duvigneaud and P. Kestemont. Editions Duculot, Paris-Gembloux. pp. 107–154.

Fang, Z.X., and Bailey, R.L. 2001. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. For. Sci. **47**: 287–300.

Fang, Z.X., Bailey, R.L., and Shiver, B.D. 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. For. Sci. **47**: 550–562.

Gregoire, T.G., and Schabenberger, O. 1996. Nonlinear mixed-effects modeling of cumulative bole volume with spatially correlated within-tree data. J. Agric. Biol. Environ. Stat. **1**: 107–119. doi:10.2307/1400563.

Hall, D.B., and Bailey, R.L. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. For. Sci. **47**: 311–321.

Hall, D.B., and Clutter, M. 2004. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. Biometrics, **60**: 16–24. doi:10.1111/j.0006-341X.2004.00163.x. PMID:15032769.

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

Hilborn, R., and Mangel, M. 1997. The ecological detective: confronting models with data. Princeton Univesity Press, Princeton, N.J.

Hochbichler, E. 2002. VorläufigeErgebnissevonBiomasseninventuren in Buchen- und Mittelwaldbeständen. *In* Inventur von Biomasse- und Nährstoffvorräten in Waldbeständen. Forstliche Forschungsberichte, Heft 186. *Edited by* H.-P. Dietrich, S. Raspe, and T. Preushlert. LWF, München. pp. 37–46.

Hochbichler, E., Putzgruber, N., and Krapfenbauer, A. 1994. Biomassen- und Nährstoffuntersuchungen in einem 40 jährigen Buchenbestand (*Fagus silvatica* L.). Centralbl. Gesamte Forstwes. **111**: 1–22.

Jalkanen, A., Mäkipää, R., Stahl, G., Lehtonen, A., and Petersson, H. 2005. Estimation of the biomass stock of trees in Sweden: comparison of biomass equations and age-dependent biomass expansion factors. Ann. For. Sci. **62**: 845–851. doi:10.1051/forest:2005075.

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., and Birdsey, R.A. 2003. National-scale biomass estimators for United States tree species. For. Sci. **49**: 12–35.

Joosten, R., Schumacher, J., Wirth, C., and Schulte, A. 2004. Evaluating tree carbon predictions for beech (*Fagus sylvatica* L.) in western Germany. For. Ecol. Manag. **189**: 87–96. doi:10.1016/j.foreco.2003.07.037.

Kauppi, P.E., Mielikäinen, K., and Kuusela, K. 1992. Biomass and carbon budget of European forests, 1971 to 1990. Science (Washington, D.C.), **256**: 70–74. doi:10.1126/science.256.5053.70. PMID:17802594.

Krauß, H.-H., and Heinsdorf, D. 1996. Herleitung von Trockenmassen und Nährstoffspeicherung in Buchenbeständen. Forstliche Forschungsanstalt Eberswalde e. V. Abteilung Standort/Umwelt. Forschungsbericht im Auftrag der Landesforstverwaltung Brandenburg.

Lambert, M.C., Ung, C.H., and Raulier, F. 2005. Canadian national tree aboveground biomass equations. Can. J. For. Res. **35**: 1996–2018. doi:10.1139/x05-112.

Lappi, J. 1991. Calibration of height and volume equations with random parameters. For. Sci. **37**: 781–801.

Le Goff, N., and Ottorini, J.M. 2001. Root biomass and biomass increment in a beech (*Fagus sylvatica* L.) stand in North-East France. Ann. For. Sci. **58**: 1–13. doi:10.1051/forest:2001104.

Le Goff, N., Granier, A., Ottorini, J.M., and Peiffer, M. 2004. Biomass increment and carbon balance of ash (*Fraxinus excelsior*) trees in an experimental stand in northeastern France. Ann. For. Sci. **61**: 577–588. doi:10.1051/forest:2004053.

Lebaube, S., Le Goff, N., Ottorini, J.M., and Granier, A. 2000. Carbon balance and tree growth in a *Fagus sylvatica* stand. Ann. For. Sci. **57**: 49–61. doi:10.1051/forest:2000100.

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. For. Ecol. Manag. **188**: 211–224. doi:10.1016/j.foreco.2003.07.008.

Lehtonen, A., Cienciala, E., Tatarinov, F., and Mäkipää, R. 2007. Uncertainty estimation of biomass expansion factors for Norway spruce in the Czech Republic. Ann. For. Sci. **64**: 133–140. doi:10.1051/forest:2006097.

Lindstrom, M.J., and Bates, D.M. 1990. Nonlinear mixed effects models for repeated measures data. Biometrics, **46**: 673–687. doi:10.2307/2532087. PMID:2242409.

Liski, J., Lehtonen, A., Palosuo, T., Peltoniemi, M., Eggers, T., Muukkonen, P., and Mäkipää, R. 2006. Carbon accumulation in Finland's forests 1922–2004 — an estimate obtained by combination of forest inventory data with modelling of biomass, litter and soil. Ann. For. Sci. **63**: 687–697. doi:10.1051/forest:2006049.

Lu, D.S. 2006. The potential and challenge of remote sensing-based biomass estimation. Int. J. Remote Sens. **27**: 1297–1328. doi:10.1080/01431160500486732.

Marklund, L.G. 1987. Biomass functions for Norway spruce (*Picea abies* (L.) Karst.) in Sweden. Rapporter – Skog. 43. Department of Forest Survey, Swedish University of Agricultural Sciences, Uppsala.

Montagu, K.D., Duttmer, K., Barton, C.V.M., and Cowie, A.L. 2005. Developing general allometric relationships for regional estimates of carbon sequestration — an example using *Eucalyptus pilularis* from seven contrasting sites. For. Ecol. Manag. **204**: 113–127. doi:10.1016/j.foreco.2004.09.003.

Mund, M. 2004. Carbon pools of European beech forests (*Fagus sylvatica*) under different silvicultural management. University Göttingen, Göttingen.

Muukkonen, P. 2007. Generalized allometric volume and biomass equations for some tree species in Europe. Eur. J. For. Res. **126**: 157–166.

Nabuurs, G.J., Schelhaas, M.J., Mohren, G.M.J., and Field, C.B. 2003. Temporal evolution of the European forest sector carbon sink from 1950 to 1999. Glob. Change Biol. **9**: 152–160. doi:10.1046/j.1365-2486.2003.00570.x.

Nilsson, U., and Albrektson, A. 1993. Productivity of needles and allocation of growth in young Scots pine trees of different competitive status. For. Ecol. Manag. **62**: 173–187. doi:10.1016/0378-1127(93)90049-S.

Nothdurft, A., Kublin, E., and Lappi, J. 2006. A non-linear hierarchical mixed model to describe tree height growth. Eur. J. For. Res. **125**: 281–289.

Parresol, B.R. 2001. Additivity of nonlinear biomass equations. Can. J. For. Res. **31**: 865–878. doi:10.1139/cjfr-31-5-865.

Pellinen, P. 1986. Biomasseuntersuchungen im Kalkbuchenwald. *In* Forstwissenschaftlicher Fachbereich. Georg-August-Universität Göttingen, Göttingen. p. 145.

Phillips, D.L., Brown, S.L., Schroeder, P.E., and Birdsey, R.A. 2000. Toward error analysis of large-scale forest carbon budgets. Glob. Ecol. Biogeogr. **9**: 305–313. doi:10.1046/j.1365-2699.2000.00197.x.

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

Raupach, M.R., Rayner, P.J., Barrett, D.J., DeFries, R.S., Heimann, M., Ojima, D.S., Quegan, S., and Schmullius, C.C. 2005. Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Glob. Change Biol. **11**: 378–397. doi:10.1111/j.1365-2486.2005.00917.x.

Schulze, E.-D., Zimmermann, R., and Hofmann, Y. 2003. ForCast-Projektseite. Available from www.bgc-jena.mpg.de/public/carboeur/Forcast/home/index.html.

Smith, R.J. 1993. Bias in equations used to estimate fossil primate body mass. J. Hum. Evol. **25**: 31–41. doi:10.1006/jhev.1993.1036.

Taylor, J.M.G. 1986. The retransformed mean after a fitted power transformation. J. Am. Stat. Assoc. **81**: 114–118. doi:10.2307/2287976.

Thurig, E., and Schelhaas, M.J. 2006. Evaluation of a large-scale forest scenario model in heterogeneous forests: a case study for Switzerland. Can. J. For. Res. **36**: 671–683. doi:10.1139/x05-283.

Vanclay, J.K., and Skovsgaard, J.P. 1997. Evaluating forest growth models. Ecol. Model. **98**: 1–12. doi:10.1016/S0304-3800(96)01932-1.

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

Vyskot, M. 1990. Juvenile beech in biomass. Academia, Prague, Czech Republic. p. 167.

Wenk, G., Antanaitis, V., and Šmelko, Š. 1990. Waldertragslehre. Deutscher Landwirtschaftsverlag GmbH, Berlin.

Widlowski, J.-L., Verstraete, M.M., Pinty, B., and Gobron, N. 2003. Allometric relationships of selected European tree species. Tech. Rep. EUR 20855 EN. EC Joint Research Centre, Ispra, Italy.

Wirth, C., and Schumacher, J. 2004. Biomass functions for Norway spruce in central Europe. COST E21 Workshop, 4–5 July 2002, Besalu.

Wirth, C., Schulze, E.-D., Schwalbe, G., Tomczyk, S., Weber, G., and Weller, E. 2004*a*. Dynamik der Kohlenstoffvorräte in den Wäldern Thüringens- Abschlussbericht zur 1. Phase des BMBF-Projektes "Modelluntersuchung zu Umsetzung des Kyoto-Protokolls". Thüringer Landesanstalt für Jagd, Wald und Fischerei, Jena Gotha.

Wirth, C., Schumacher, J., and Schulze, E.D. 2004*b*. Generic biomass functions for Norway spruce in central Europe — a meta-analysis approach toward prediction and uncertainty estimation. Tree Physiol. **24**: 121–139. PMID:14676030.

Zhao, D.H., Wilson, M., and Borders, B.E. 2005. Modeling response curves and testing treatment effects in repeated measures experiments: a multilevel nonlinear mixed-effects model approach. Can. J. For. Res. **35**: 122–132. doi:10.1139/x04-163.

Zianis, D., and Mencuccini, M. 2003. Aboveground biomass relationships for beech (*Fagus moesiaca* Cz.) trees in Vermio Mountain, northern Greece, and generalised equations for *Fagus* sp. Ann. For. Sci. **60**: 439–448. doi:10.1051/forest:2003036.

Zianis, D., and Mencuccini, M. 2005. Aboveground net primary productivity of a beech (*Fagus moesiaca*) forest: a case study of Naousa Forest, northern Greece. Tree Physiol. **25**: 713–722. PMID:15805091.

Zianis, D., Muukkonen, D., Mäkipää, R., and Mencuccini, M. 2005. Biomass and stem volume equations for tree species in Europe. Silva Fenn. Monogr. **4**: 64.

Appendix A. Variance of single predictions

The prediction $\widehat{y}_{\text{new}}$ for a predictor vector $\boldsymbol{x}_{\text{new}}$, covariates $\boldsymbol{\nu}_{\text{new}}$, and an unknown group was done by applying the model formula to the vector of new predictors assuming zero random effects (expected value) and zero residual term.

In order to estimate the variance of the error of the non-linear prediction, we approximated the nonlinear f in eq. 1 (main text) by its first-order Taylor expansion in the parameter space around the estimated parameters $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{b} = \mathbf{0}$ (eq. *A*1) (Gregoire and Schabenberger 1996):

$$[A1] \qquad f(\boldsymbol{x}_{\text{new}}, \boldsymbol{\nu}_{\text{new}}; \boldsymbol{\beta}, \boldsymbol{b}_{\text{new}}) \approx f(\boldsymbol{x}_{\text{new}}, \boldsymbol{\nu}_{\text{new}}; \widehat{\boldsymbol{\beta}}, \mathbf{0}) + \boldsymbol{u}_{\text{new}}^{T}(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}) + \boldsymbol{w}_{\text{new}}^{T}(\boldsymbol{b}_{\text{new}} - \mathbf{0})$$

where

$$\boldsymbol{u}_{\text{new}} = \frac{\partial f(\boldsymbol{x},\ \boldsymbol{\nu};\ \boldsymbol{\beta},\ \boldsymbol{b})}{\partial \boldsymbol{\beta}}\Big|_{\boldsymbol{x}_{\text{new}},\ \boldsymbol{\nu}_{\text{new}},\ \widehat{\boldsymbol{\beta}},\ \mathbf{0}}$$

and

$$\boldsymbol{w}_{\text{new}} = \frac{\partial f(\boldsymbol{x},\ \boldsymbol{\nu};\ \boldsymbol{\beta},\ \boldsymbol{b})}{\partial \boldsymbol{b}}\Big|_{\boldsymbol{x}_{\text{new}},\ \boldsymbol{\nu}_{\text{new}},\ \widehat{\boldsymbol{\beta}},\ \mathbf{0}}$$

are vectors of partial derivates evaluated at the estimated parameters $\widehat{\boldsymbol{\beta}}$, $\boldsymbol{b} = \mathbf{0}$, and the values of predictors and covariate for the new observation.

Equation *A*1 describes the prediction that uses the true parameters $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_{\text{new}}$ by a prediction that uses the modified parameters $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{b} = \mathbf{0}$ plus some deviation depending on the model parameters and the predictors. With this approximation, the variance of the error of a nonlinear prediction evaluates to

$$\begin{aligned}
&\text{Var}(y_{\text{new}} - \widehat{y}_{\text{new}}) \\
&= \text{Var}(f(\boldsymbol{x}_{\text{new}}, \boldsymbol{\nu}_{\text{new}};\boldsymbol{\beta}, \boldsymbol{b}_{\text{new}}) + \varepsilon_{\text{new}} - f(\boldsymbol{x}_{\text{new}}, \boldsymbol{\nu}_{\text{new}};\widehat{\boldsymbol{\beta}}, \mathbf{0})) \\
[A2] \quad &= \text{Var}(\boldsymbol{u}_{\text{new}}^T(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}) + \boldsymbol{w}_{\text{new}}^T\boldsymbol{b}_{\text{new}} + \varepsilon_{\text{new}}) \\
&= \boldsymbol{u}_{\text{new}}^T\text{Var}(\widehat{\boldsymbol{\beta}})\boldsymbol{u}_{\text{new}} + \boldsymbol{w}_{\text{new}}^T\text{Var}(\boldsymbol{b}_{\text{new}})\boldsymbol{w}_{\text{new}} + \text{Var}(\varepsilon_{\text{new}}) \\
&= \boldsymbol{u}_{\text{new}}^T\text{Var}(\widehat{\boldsymbol{\beta}})\boldsymbol{u}_{\text{new}} + \boldsymbol{w}_{\text{new}}^T\boldsymbol{\Psi}\boldsymbol{w}_{\text{new}} + \sigma^2
\end{aligned}$$

Estimates for the unknown true covariance matrices $\text{Var}(\widehat{\boldsymbol{\beta}})$ and $\boldsymbol{\Psi}$ as well for the residual variance σ^2 are obtained as a by-product of the numerical optimization algorithm used for REML estimation of the unknown parameters. The three terms of eq. *A*2 correspond to three components of variance: the fixed effects, the random effects, and the residual variance.

To account for variance heterogeneity within groups, we modelled the residual variance as a power function of the predicted values:

$$[A3] \qquad \text{Var}(\varepsilon_{ij}) = \sigma^2|y_{ij}|^{2\delta}$$

If the coefficient δ of the power variance model additionally depended on the particular group ($\delta \rightarrow \delta_i$), the parameter δ_i for a new prediction was unknown because the group of the new prediction was unknown. The best estimate for the new prediction is the mean $\bar{\bar{\delta}}_i$ of the estimate for parameters δ_i. However, the mean function appears in a nonlinear term and a correction factor has to be applied. The general form of the expected value of a nonlinear equation $y = f(x)$ obtained by the delta method is (Hilborn and Mangel 1997, p. 58)

$$[A4] \qquad E(y) = f(E(x)) + \frac{1}{2}f''(E(x))\text{Var}(x)$$

This leads in the case of eq. *A*3 ($f(\delta) = \sigma^2|y|^{2\delta}$) to

$$\begin{aligned}
[A5] \quad \widehat{\text{Var}}(\varepsilon_{\text{new}}) &= \sigma^2|y_{\text{new}}|^{2\bar{\bar{\delta}}_i} \\
&\quad + 2\sigma^2(\ln|y_{\text{new}}|)^2|y_{\text{new}}|^{2\bar{\bar{\delta}}_i}\text{Var}(\widehat{\delta}_i) \\
&= \sigma^2|y_{\text{new}}|^{2\bar{\bar{\delta}}_i}(1 + 2(\ln|y_{\text{new}}|)^2\text{Var}(\widehat{\delta}_i))
\end{aligned}$$

Appendix B. Variance of the sum of predictions

Many applications of statistical models calculate the sum of several individual predictions that originate from the same group (e.g., sum of biomass of individual trees measured by the same team with the same measurement procedure). The expected value of the sum of n predictions is simply the sum of the single expected values, i.e., model predictions. However, the variance of the sum of prediction errors has to account for covariances between the individuals. The variance of the sum of prediction errors is given by

$$\begin{aligned}
&\text{Var}\left(\sum_j^n (y_j - \widehat{y}_j)\right) \\
[B1] \quad &= \sum_j^n \text{Var}(y_j - \widehat{y}_j) + 2\sum_j^n\sum_{k=j+1}^n \text{Cov}(y_j - \widehat{y}_j, y_k - \widehat{y}_k) \\
&= \sum_{j=1}^n\sum_{k=1}^n \text{Cov}(y_j - \widehat{y}_j, y_k - \widehat{y}_k)
\end{aligned}$$

where $\text{Cov}(y_j - \widehat{y}_j, y_k - \widehat{y}_k)$ denotes the covariance between the errors of the individual prediction for the new observations j and k.

In the following, we derive the covariance between two prediction errors for predictions based on model eq. 1 within the same group i. With approximating model eq. 5 by its first-order Taylor expansion (eq. *A*1) around $\widehat{\beta}$ and $\boldsymbol{b} = \mathbf{0}$, we derive eq. *B*2:

$$\begin{aligned}
&\text{Cov}(y_{\text{new},1} - \hat{y}_{\text{new},1}, y_{\text{new},2} - \hat{y}_{\text{new},2}) \\
&\approx \text{Cov}\begin{pmatrix} f(\boldsymbol{x}_{\text{new},1}, \boldsymbol{\nu}_{\text{new},1}, \hat{\boldsymbol{\beta}}, \mathbf{0}) + \boldsymbol{u}_{\text{new},1}^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + \boldsymbol{w}_{\text{new},1}^T\boldsymbol{b}_{\text{new}} + \varepsilon_{\text{new},1} - \boldsymbol{f}(\boldsymbol{x}_{\text{new},1}, \boldsymbol{\nu}_{\text{new},1}, \hat{\boldsymbol{\beta}}, \mathbf{0}) \\ f(\boldsymbol{x}_{\text{new},2}, \boldsymbol{\nu}_{\text{new},2}, \hat{\boldsymbol{\beta}}, \mathbf{0}) + \boldsymbol{u}_{\text{new},2}^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + \boldsymbol{w}_{\text{new},2}^T\boldsymbol{b}_{\text{new}} + \varepsilon_{\text{new},2} - \boldsymbol{f}(\boldsymbol{x}_{\text{new},2}, \boldsymbol{\nu}_{\text{new},2}, \hat{\boldsymbol{\beta}}, \mathbf{0}) \end{pmatrix} \\
[B2] \quad &= \text{Cov}(\boldsymbol{u}_{\text{new},1}^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + \boldsymbol{w}_{\text{new},1}^T\boldsymbol{b}_{\text{new}} + \varepsilon_{\text{new},1}, \boldsymbol{u}_{\text{new},2}^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + \boldsymbol{w}_{\text{new},2}^T b_{\text{new}} + \varepsilon_{\text{new},2}) \\
&= \text{Cov}(\boldsymbol{u}_{\text{new},1}^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}), \boldsymbol{u}_{\text{new},2}^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + \text{Cov}(\boldsymbol{w}_{\text{new},1}^T\boldsymbol{b}_{\text{new}}\boldsymbol{w}_{\text{new},2}^T\boldsymbol{b}_{\text{new}}) + \text{Cov}(\varepsilon_{\text{new},1}, \varepsilon_{\text{new},2}) \\
&= \boldsymbol{u}_{\text{new},1}^T\text{Var}(\hat{\boldsymbol{\beta}})\boldsymbol{u}_{\text{new},2} + \boldsymbol{w}_{\text{new},1}^T\text{Var}(\boldsymbol{b}_{\text{new}})\boldsymbol{w}_{\text{new},2} + \mathbf{0} \\
&= \boldsymbol{u}_{\text{new},1}^T\text{Var}(\hat{\boldsymbol{\beta}})\boldsymbol{u}_{\text{new},2} + \boldsymbol{w}_{\text{new},1}^T\boldsymbol{\Psi}\boldsymbol{w}_{\text{new},2}
\end{aligned}$$

where $\boldsymbol{u}_{\text{new},1}$, $\boldsymbol{u}_{\text{new},2}$, $\boldsymbol{w}_{\text{new},1}$, and $\boldsymbol{w}_{\text{new},2}$ are vectors of partial derivates as explained for eq. *A*1. Estimates for the unknown true covariance matrices $\text{Var}(\widehat{\boldsymbol{\beta}})$ and $\boldsymbol{\Psi}$ are obtained as a by-product of the numerical optimization algorithm.

If the two observations were of different groups, i.e., trees of different regions and measured by different teams, the covariance in prediction errors due to random effect would be zero, i.e., $\text{Cov}(\boldsymbol{w}_{\text{new},2}^T\boldsymbol{b}_{\text{new},1}, \boldsymbol{w}_{\text{new},2}^T\boldsymbol{b}_{\text{new},2}) = 0$.

Now we are ready to calculate and interpret the variance of a sum of prediction errors of model eq. 5. Inserting eq. *B*2 into eq. *B*1 leads to

$$[B3] \qquad \mathrm{Var}\left(\sum_j^n (y_j - \hat{y}_j)\right) = \sum_j^n \sum_k^n \left(\boldsymbol{u}_j^T \mathrm{Var}(\hat{\boldsymbol{\beta}})\boldsymbol{u}_k + \boldsymbol{w}_j^T \boldsymbol{\Psi} \boldsymbol{w}_k\right) + \sum_j^n \mathrm{Var}(\varepsilon_j)$$

Similar to the variance of single predictions (eq. *A*2), eq. *B*3 is composed of the three terms of the components of variance (fixed, random, and residuals). However, the residual variance occurs in a simple sum over all individuals, whereas the fixed and random terms occur within a sum of sums. Hence, the residual variance increases linearly with the number of individuals, whereas the random and fixed components of variance increase quadratically with the number of individuals.

The covariance terms in eq. *B*3 can be negative. Hence, they potentially cancel out each other. However, individuals of the same group often have similar predictor values and have positive covariances.

Appendix C. Application example

In the following example, we demonstrate how to use the presented models and equations to calculate a new prediction for stem biomass and its confidence interval for a single tree with diameter d = 18.8 cm, height h = 16.9 m, age = 40 years), site index si = 30 m, and altitude alt = 470 m. Consequently, we choose the dhc model because all of the additional covariates are known. The label dh3 in Table 6 indicates that the basic model form is $m_{\mathrm{new}} = c_0 d^{c_1} h^{c_2}$ (eq. 3). Site index, age, and altitude affect the coefficient c_0, and Table 7 indicates that the model also includes random effect in c_0: $c_0 = \beta_0 + b_{0,1} + \beta_{0,\,\mathrm{age}} \times \mathrm{age} + \beta_{0,\,\mathrm{si}} \times \mathrm{si} + \beta_{0,\,\mathrm{alt}} \times \mathrm{alt}$ (eq. 4*c*). The other coefficients include neither covariates nor a random effect. Hence, the full equation of aboveground biomass for tree j is given by

$$[C1] \qquad \begin{aligned} m_{\mathrm{new}} &= f(d, h, \mathrm{age}, \mathrm{si}, \mathrm{alt}; b) \\ &= (\beta_0 + b_{0,\,i} + \beta_{0,\,\mathrm{age}} \times \mathrm{age} + \beta_{0,\,\mathrm{si}} \times \mathrm{si} \\ &\qquad + \beta_{0,\,\mathrm{alt}} \times \mathrm{alt}) d^{(\beta_1)} h^{(\beta_2)} \end{aligned}$$

Hence, with setting the random effect $b_{0,i}$ to its expected value 0 for a general prediction, the stem biomass computes to $(0.00351 + 0 + 3.47 \times 10^{-5} \times 40 + 6.72 \times 10^{-4} \times 30 + 8.11 \times 10^{-6} \times 470) \times 18.8^{1.84} \times 16.9^{1.04} = 121$ kg. The variance of the prediction error is estimated according to eqs. *A*2 and *A*5 as

$$\begin{aligned} &\mathrm{Var}(\hat{m}_{\mathrm{new}} - m_{\mathrm{new}}) \\ &= \left[\underbrace{4250.244}_{\frac{\partial f(d,h,\mathrm{age},\mathrm{si},\mathrm{alt};b)}{\partial \beta_0}} \underbrace{127507.3}_{\frac{\partial f}{\partial b_{0,\mathrm{si}}}} \underbrace{170009.8}_{\frac{\partial f}{\partial \beta_{0,\mathrm{age}}}} \underbrace{1997615}_{\frac{\partial f}{\partial \beta_{0,\mathrm{alt}}}} \underbrace{360.0897}_{\frac{\partial f}{\partial \beta_1}} \underbrace{347.013}_{\frac{\partial f}{\partial \beta_2}}\right] \mathrm{Var}(\hat{\boldsymbol{\beta}}) \begin{bmatrix} 4250.244 \\ 127507.3 \\ 170009.8 \\ 1997615 \\ 360.0897 \\ 347.013 \end{bmatrix} \\ &+ \left[\underbrace{4250.244}_{\frac{\partial f(d,h,\mathrm{age},\mathrm{si},\mathrm{alt};b)}{\partial b_{0,\mathrm{i}}}}\right] \boldsymbol{\Psi} [4250.244] \\ &+ \underbrace{\overbrace{0.0903}^{\sigma^2} \left| \overbrace{122.7360}^{\hat{m}_{\mathrm{new}}} \right|^{\overbrace{2.0842}^{\mathrm{Mean}(\delta)}} \left(1 + 2\ln^2\left(\overbrace{122.7360}^{\hat{m}_{\mathrm{new}}}\right) \overbrace{0.0196}^{\mathrm{Var}(\delta)}\right)}_{\mathrm{Var}(\varepsilon_{\mathrm{new}})} \\ &= 30.3 + 179.9 + 568.6 = 778.9 \end{aligned}$$

where f is given by eq. *C*1, $\mathrm{Var}(\hat{\boldsymbol{\beta}})$ is the estimated 6 × 6 covariance matrix the estimated fixed effects, and $\boldsymbol{\Psi}$ is the 1 × 1 random effects covariance matrix.[7] The values of the coefficients of the third term are given in Table 7. A symmetric 95% confidence interval around the prediction is calculated using the quantile of the standard normal distribution as $\pm 1.96\sqrt{\mathrm{Var}(\hat{m}_{\mathrm{new}})}$. This becomes 54.7 kg and the 95% confidence interval ranges from 68 to 177 kg.

We provided R-objects[8] of all of the best models together with coefficients, gradient functions, the fixed and random effects covariance matrices, and functions to calculate the variance of single tree biomass predictions (nlVar) and the variance of a group of trees[9] (nlCovar). The results of this example are simply obtained with the R-command "nlVar(dhcme.beech$stem, data.frame(dbh=18.8, height=16.9, age=40, si=30, alt=470), pred=TRUE)".

[7] Both matrices are made available with electronic supplementary material S4.
[8] Electronic supplementary material S5.
[9] An example of calculating biomass and its variance at stand level is given in electronic supplementary material S6.