Jukuri, open repository of the Natural Resources Institute Finland (Luke)

This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.

Author(s): Annika Kangas, Timo P Pitkänen, Lauri Mehtätalo & Juha Heikkinen

Title: Mixed linear and non-linear tree volume models with regional parameters to main tree species in Finland

Year: 2022

Version: Published version

Copyright: The Author(s) 2022

Rights: CC BY 4.0

Rights url: http://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Annika Kangas, Timo P Pitkänen, Lauri Mehtätalo, Juha Heikkinen, Mixed linear and non-linear tree volume models with regional parameters to main tree species in Finland, Forestry: An International Journal of Forest Research, 2022;, cpac038, https://doi.org/10.1093/forestry/cpac038

All material supplied via *Jukuri* is protected by copyright and other intellectual property rights. Duplication or sale, in electronic or print form, of any part of the repository collections is prohibited. Making electronic or print copies of the material is permitted only for your own personal use or for educational purposes. For other purposes, this article may be used in accordance with the publisher's terms. There may be differences between this version and the publisher's version. You are advised to cite the publisher's version.

Forestry An International Journal of Forest Research

Institute of Chartered Foresters

Forestry 2022; 1–19, https://doi.org/10.1093/forestry/cpac038

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Mixed linear and non-linear tree volume models with regional parameters to main tree species in Finland

Annika Kangas[1], Timo P. Pitkänen[2], Lauri Mehtätalo[1] and Juha Heikkinen[3]

[1]*Bioeconomy and Environment, Natural Resources Institute Finland, Yliopistokatu 6, Joensuu 80100, Finland*
[2]*Bioeconomy and Environment, Natural Resources Institute Finland, Latokartanonkaari 9, Helsinki 00790, Finland*
[3]*Natural resources, Natural Resources Institute Finland, Latokartanonkaari 9, Helsinki 00790, Finland*

*Corresponding author Tel: +358295322461; E-mail: annika.kangas@luke.fi

Received 27 January 2022

The volume models that have been used in Finland for the last 40 years, while generally well thought-out, exhibit an illogical behaviour for small trees. In recent studies, tree stem form was observed to have changed in time and also involve spatial variation attributable to environmental factors. It is yet unclear how the stem taper has actually changed. To overcome these problems, we fitted a completely new set of volume and taper curve models and examined whether this change is attributable to the changes in management and environmental factors rather than to measurement errors in the previously used datasets. For the latter, we added a dataset into the analysis, which was smaller but of higher quality due to the destructive nature of the stem taper measurements. We aim at (1) developing a new non-linear variable form factor volume function that works with trees of all sizes, (2) improving the description of the variation of the stem form in time and space by including temperature sum and soil type as predictors, (3) understanding the changes in the stem form by fitting new taper curve models and (4) improving the statistical properties of the predictions by using mixed model techniques and by addressing the effect of parameter uncertainty. To assess the impact of renewing the models, we (5) predicted the mean volume and its confidence interval with each model for forest inventory data at country level. The results show that the tree stem form has a spatial trend that can be described with the temperature sum. Moreover, the changes in stem form also have a spatial trend, with largest changes in Lapland. The difference is mostly observable in the lowest part of the stem, and it is especially large in the largest pines. We conclude that environmental variables can help to improve national stem taper functions in countries with pronounced environmental gradients.

Introduction

Volume models are the most important allometric models in forestry, needed for any tree species and region, as measuring tree volume in the field is usually not possible. The models of Laasasenaho (1982), which are frequently used in Finland, include one-predictor models with diameter at breast height (dbh , diameter at a 1.3 m height above ground) as the sole predictor, two-predictor models with dbh and height (*h*) as predictors, and three-predictor models with dbh, *h* and diameter at height of 6 m (*d6*) as predictors. His set of models also include taper curve functions for average stem form with diameter at 20 per cent of the total height as a predictor and correction functions for predicting individual deviations from the average taper curve form with dbh and h as predictors. The two-predictor volume models are used in practically all forest planning systems in Finland, while the three-predictor model is mainly used in the National Forest Inventory (NFI) for calculating the sample tree volumes. The taper curve models are used to buck the stems into logs and pulpwood.

The models of Laasasenaho (1982), in general, have a geometrically well thought-out structure, and they have shown a very good performance with the original data used to fit the model. However, the models are old and have a set of problems that need to be tackled for future applications. For instance, the most-often used volume models based on dbh and h do not behave logically for the smallest trees (Laasasenaho, 1982, p. 43; Kangas *et al.*, 2020, see section 3.2). Different approaches have been used to overcome this problem in the NFI and in various forest simulators. Our hypothesis in this study is that by analysing the theoretical properties of the models, an improved solution can be found.

The classical starting point for modelling is the Schumacher function, which is based on the logarithmic transformation of the function $v = Ad^Bh^C$ (e.g. Schumacher & Hall 1933; Zianis *et al.*, 2005; Burkhart and Tomé, 2012; Vibrans *et al.*, 2015; Goussanou

Handling Editor: Dr. Rubén Manso
© The Author(s) 2022. Published by Oxford University Press on behalf of Institute of Chartered Foresters.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

et al., 2016). The model is justified by the shape of tree stems: using $B = 2$ and $C = 1$, the formula will provide the volumes of cylinder and cone as special cases with base diameter d and height h, depending on the value of A, which in turn determines the breast height form factor as $4A/\pi$. The form factors for cylinder and cone are 1 and 1/3, respectively. For most trees, the form factor should be between these values. For some species, the form factor is fairly constant, whereas for others, it varies with tree size (Vallet *et al.,* 2006; Kershaw *et al.,* 2016). When the powers B and C of diameter and height are estimated empirically, the differences from the theoretical values of 2 and 1 actually model the dependence of form factor on diameter and height. As we will show later, the two-predictor model of Laasasenaho (1982) can be seen as such an extension of the model of Schumacher, where the form factor depends on diameter and height and is restricted to be positive. However, the form factor has no upper bound, which causes the abovementioned unrealistic behaviour for small trees. Little work has been carried out in modelling the form factor explicitly (see, however, Vallet *et al.,* 2006). A reason for the lacking efforts to improve the model may be that they lead to models that cannot be linearized easily and require the use of non-linear model fitting techniques. However, methods are well developed nowadays and have been widely available already for decades (Mehtätalo and Lappi, 2020).

The data for the old models were collected during 1968–1972 from the plots of the fifth NFI by climbing the living trees and conducting diameter measurements at multiple heights (later referred to as climbed data). When these models were used in later NFIs to predict the sample tree volumes, the mean differences between Laasasenaho's two-predictor models (dbh, h) and three-predictor models (dbh, $h, d6$) exhibit an increasing trend (Kangas *et al.,* 2020), indicating that the shape of the stems has changed since the fifth NFI. This conclusion is based on the three-predictor model, including information of the changed shape ($d6$), while the two-predictor model does not include such information. To quantify this change, a new point cloud dataset, containing detailed information in individual trees' stem forms, was measured with terrestrial laser scanning (later referred to as scanned data; for details, see Kangas *et al.,* 2020) and was used in combination with the climbed data to re-calibrate the old models. The plots to be scanned were selected from the plots of the 12th NFI.

The results showed evidence for a change in the stem form for all tree species (Kangas *et al.,* 2020). It was noted that when conditioned on diameter, expected volumes in the scanned data were higher than those in the climbed data, and when conditioned on both diameter and height, the expected volumes in the scanned data were smaller than those in the climbed data. Thus, the trees were taller and slenderer in the scanned data than in the climbed data. In Kangas *et al.* (2020), the change was mainly attributed to changed forest management and denser forests.

Yet, there was also evidence of regional variation in those differences and therefore environmental or climatic differences may have further induced that change on top of forest management. Traditionally, volume models are fitted without any environmental factors (e.g. Zianis et al. 2005). However, recent studies have shown evidence of environmental factors affecting the allometry. Mean temperature and mean precipitation are most commonly used variables in the allometric models. For instance, Fortin *et al.* (2019), Qiu *et al.* (2021) and Xu *et al.* (2022) used those as predictors in height models, Schneider *et al.* (2018) used those as predictors in volume models and Chave *et al.* (2014) and Fu *et al.* (2017) used those as predictors in biomass models. Other climatic factors, such as temperature sums, mean solar radiation and length of frost-free period, have also been used (e.g. Xu *et al.,* 2021). Other environmental factors, such as soils, have also shown to have an effect. For instance, Cysneiros *et al.* (2021) included both climatic and soil variables in their allometric models.

In Finland, the temperature sums (or degree days, sum of degrees for days with mean temperature >5°C) at different parts of country have a variation from around 400 in the northernmost part of Finland to around 1300 in the southernmost part of Finland. That can be assumed to be the most important environmental factor affecting the stem form. Moreover, the trees located on peatlands and mineral lands were not separated when the old models were fitted, even though this may have an effect on the stem form. Thus, there is a need to address the potential effect of these factors in the new models. To take this on board, the original two-predictor and three-predictors models were also refitted in this study with environmental factors included.

It is also important to find out the relative height where the largest changes in the stem form have occurred since that will affect the volumes of different timber assortments. Therefore, there is also a need to re-fit the old taper curve models to address this question.

In the recent studies, Kangas *et al.* (2020) and Pitkänen *et al.* (2021) noted the possibility of the observed changes in stem form being due to the differences of the error structure rather than from actual change in time. To overcome this problem, we decided to add a third dataset, collected during 1988–2001, into the analysis. As in this dataset, the trees were felled and measured on the ground (later referred to as felled data), it can be assumed to be the most accurate of the three datasets. Therefore, we hypothesized that these data will help in confirming if the changes in stem form are due to changes in time rather than artefacts of the measurements and also give further insights into the effect of the environmental factors.

Finally, the volume models fitted by Kangas *et al.* (2020) were re-calibrations of old models, meaning the models were fitted as linearized OLS models. We deemed it important to introduce new techniques available, such as non-linear mixed models, into the modelling exercise. In addition, new techniques may also be used to address the parameter uncertainties of the models in addition to the residual errors. This is important since, even if the parameters are estimated using a method that gives unbiased predictions of the response variable, the expected values of prediction errors are zero only over repeated sampling of the model fitting dataset and not for any single model fitting dataset. Because the parameter estimates are based on one dataset only, the prediction errors induced by the errors in the parameter estimates do not cancel out. Instead, they behave like bias, i.e. remain constant irrespective of the number of predictions (see Mehtätalo and Lappi, 2020, p. 127). This differs from the effects of residual errors, which cancel out if predictions are done for a large number of trees. For example, if the parameter uncertainty results in a 1 per cent overestimate in the volume, all inventory

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Table 1 The number of observations in different datasets

Species	All observations			Observations with $d6 > 0$		
	Climbed	Felled	Scanned	Climbed	Felled	Scanned
Pine	2326	797	943	2013	722	925
Spruce	1864	479	623	1668	435	619
Birch	863	258	355	827	242	351

Table 2 The means, minima and maxima of tree characteristics by dataset (all observations included) and by species

Species	Dataset	Stem volume, v, dm^3			Breast height diameter, dbh, cm			Upper diameter, $d6$, cm			Height, h, m		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Pine	Climbed	312.4	0.4	1920.5	20.2	0.9	50.6	14.5	0.0	44.2	13.7	1.5	28.3
	Felled	212.0	4.7	1920.3	16.7	5.0	44.6	11.1	0.0	38.2	12.3	3.2	31.4
	Scanned	336.9	3.8	1947.7	20.8	3.9	49.7	16.0	0.0	40.4	16.5	4.1	29.6
Spruce	Climbed	264.5	0.6	3795.9	18.0	1.5	61.9	13.1	0.0	52.4	13.8	1.8	32.7
	Felled	316.4	5.8	2483.1	18.5	4.9	49.8	13.6	0.0	47.6	14.8	3.9	32.1
	Scanned	448.3	7.4	2856.4	22.2	5.6	57.0	17.9	0.0	47.7	18.2	4.5	32.7
Birch	Climbed	228.6	0.4	2026.0	16.7	1.2	49.7	12.5	0.0	39.0	15.3	2.4	29.5
	Felled	116.1	5.6	2112.4	12.4	5.0	45.3	7.9	0.0	37.8	12.2	4.5	32.0
	Scanned	211.2	4.3	1427.7	16.1	4.0	42.5	12.2	0.0	34.8	16.1	4.5	29.8

results and growth simulators give the respective overestimate. The exact value of this systematic error is unknown, but a confidence interval can be produced for it.

To overcome the problems discussed above, we decided to fit a completely new set of volume and taper curve models. We aim at (1) developing a new non-linear variable form factor volume function that works with trees of all sizes, (2) improving the description of the variation of the stem form in time and space by including temperature sum and soil type as predictors, (3) understanding the changes in the stem form by fitting new taper curve models and (4) improving the statistical properties of the predictions by using mixed model techniques and by addressing the effect of parameter uncertainty. To assess the impact of upgrading the models, we (5) predicted the mean volume and its confidence interval with each model for forest inventory data at country level. The new models are reported as an R-object and prediction function that produces confidence intervals to quantify the systematic error due to parameter uncertainty.

Material

Our study material consists of three datasets: (1) The climbed data that include 2326 Scots pine (*Pinus sylvestris* L.), 1864 Norway spruce (*Picea abies* L. Karst.) and 863 silver (*Betula pendula* Roth) and downy (*Betula pubescens* Ehrh.) birches collected between 1968 and 1972; (2) the felled data with 797 pines, 479 spruces and 258 birches collected between 1988 and 2001 and (3) the scanned data with 943 Scots pine, 623 Norway spruce and 355 downy or silver birch collected between 2017 and 2018 (Table 1). The trees were in all cases measured in a selection of NFI plots, with 4.1 plots from the same cluster on average. The main characteristics of the three datasets are shown in Table 2. For more details concerning the climbed data, see Laasasenaho (1982); for the felled data, see Korhonen and Maltamo (1990); and for the scanned data, see Pitkänen *et al.* (2019). For modelling the tree volume, the three datasets were merged into one dataset and the differences between datasets were parameterized into the models.

In the original publication of volume models based on the climbed data (Laasasenaho, 1982) and their re-calibration (Kangas *et al.*, 2020), only trees in productive forest land were included. In this study, trees in poorly productive stands were included, as the volumes for these trees are also needed in applications. Most of the observations are from mineral soils, ($N = 6711$) and the rest ($N = 1795$) are from peatlands (Table 3). For the volume models including upper diameter, only trees with $d6 > 0$ were considered.

In the climbed and felled data, the volumes of the trees were predicted using a natural cubic spline (splinefun, R Core Team, 2021) to interpolate the taper curve between measured diameters at given relative heights. The volumes were obtained by integrating the resulting taper curve using $v = \pi/4 \int_0^h s^2(x)dx$ (Lahtinen and Laasasenaho, 1979), where s denotes the diameter as the function of height, h is the height of the tree and 0 is the stump height level.

Taper curves for the scanned data were based on a large number of diameters, which were derived from circles fitted to

Table 3 Observations by soil types and land classes

	Soil type		Land class		
	Mineral	Peatland	Productive forest land	Poorly productive forest land	Unproductive land
Climbed	4115	938	4727	292	34
Felled	1133	401	1534	0	0
Scanned	1463	456	1906	15	0

the point cloud data using the random sample consesus algorithm (Fischler and Bolles, 1981). To prevent substantial over- or underestimates of diameters, earlier measurement data and taper curve models estimated with the method presented by Lappi (1986, 2006) were used as prior information to filter out expected gross errors resulting from poorly fitted circles (Pitkänen *et al.*, 2021).

The final taper curves were predicted using cubic splines like in the original study (Laasasenaho, 1982; Pitkänen *et al.*, 2021). In scanned data, however, a smoothing spline function (smooth.spline, R core team, 2021) was used instead of an interpolating spline, as there are much more diameter/height combinations available than in climbed or felled data. The flexibility of the curve was set with the 'spar' parameter in the algorithm. In scanned data stem, volumes above the stump height were predicted using the cubic spline and Huber formula (Husch *et al.*, 1972) with 1-cm slices.

In order to analyse the regional environmental variation, the coordinates of the original plots were used to get the temperature sum from the databases of Finnish Meteorological Institute. That data were also used to divide the area to four regions with an equal number of observations (Figure 1). Data collected in the 11th NFI (Korhonen *et al.*, 2017) were also used in our demonstration of the application potential of the models. That data included 31 237 pine, 18 885 spruce and 12 316 birch sample trees from the forest land and poorly productive forest land (national classification, see Tomppo *et al.*, 2011), with measured dbh and h and $d6$ for trees with $h > 8$ m and no forks between 1.3 and 6 m.

In the climbed data, a pre-defined number of diameter measurements along the stem were carried out. Since the measurements were taken in standing trees, it is possible that there are errors in the relative heights and diameters, especially in the upper part of the stem. In the scanned data, an arbitrary amount of diameters along the stem can be calculated from the laser point cloud data, but in each scan, only one side of the stem can be seen. While several scans can be combined to circumvent this problem, windy weather may make the merging challenging. In addition, branches and other trees may cause occlusion problems in the upper part of the stem. In the felled dataset, all measurements were carried out from felled trees, and therefore, this dataset can be assumed to be the most accurate of the tree datasets, specifically in the upper parts of the stem. The error structures (i.e. means, variances and correlations of errors at different parts of the stem) of these three datasets are very different, which needs to be taken into account when interpreting the results. To minimize the potential effects of the different error structures, all variables used as predictors (dbh, h, $d6$ and diameter at stump height (dsh)) of all datasets were measured in the field.

Methods

Volume functions

We fitted three forms of tree-level volume models to our datasets:

$$\frac{v}{\pi \cdot \mathrm{dbh}^2 \cdot \mathrm{h}/40} = a + b \cdot \mathrm{dbh} + c \cdot \mathrm{h} + d \cdot \frac{1}{\mathrm{h}} + e \cdot \frac{\left(\mathrm{dbh}^2 + \mathrm{dbh} \cdot d6 + d6^2\right)}{\mathrm{dbh}^2 \mathrm{h}} + f \cdot \frac{\left(\mathrm{d6}^2 \cdot (\mathrm{h}-6)\right)}{\mathrm{dbh}^2 \cdot \mathrm{h}} + \varepsilon_{\mathrm{cluster}} + \varepsilon_{\mathrm{plot}} + \varepsilon_{\mathrm{tree}} \quad (1)$$

$$\log(v) = a + b \cdot \log(\mathrm{dbh}) + c \cdot \log(\mathrm{h}) + d \cdot \log(\mathrm{h} - 1.3) + e \cdot \mathrm{dbh} + \varepsilon_{\mathrm{cluster}} + \varepsilon_{\mathrm{plot}} + \varepsilon_{\mathrm{tree}} \quad (2)$$

and

$$v = \mathrm{logit}^{-1}\left(a + b \cdot \mathrm{dbh} + c \cdot \mathrm{h} + d \cdot \frac{1}{\mathrm{h}} + e \cdot \mathrm{dbh} \cdot \mathrm{h} + f \cdot \frac{1}{\mathrm{dbh} \cdot \mathrm{h}} + \varepsilon_{\mathrm{cluster}} + \varepsilon_{\mathrm{plot}}\right) \cdot \frac{\pi \cdot \mathrm{dsh}^2 \cdot \mathrm{h}}{40} + \varepsilon_{\mathrm{tree}}, \quad (3)$$

where v is the stem volume in dm^3; and dbh, d6 and dsh, respectively, are the diameters at 1.3 m, 6 m, and stump height in cm and h is the height in m. Values of parameters $a - f$ were estimated from the data as explained below, and $\varepsilon_{\mathrm{cluster}}$, $\varepsilon_{\mathrm{plot}}$ and $\varepsilon_{\mathrm{tree}}$ are, respectively, the zero-mean cluster-, plot- and tree-level random effects that follow the standard assumptions of mixed-effects models (e.g. Mehtätalo and Lappi, 2020). Model (3) included the inverse logit transformation $\mathrm{logit}^{-1}(x) = 1/(1 + \exp(-x)) \in (0, 1)$ and the approximation

$$\mathrm{dsh} = \left[w(\mathrm{h}, \lambda) + (1 - w(\mathrm{h}, \lambda)) \frac{\mathrm{h}}{\mathrm{h} - 1.3}\right] \mathrm{dbh} \quad (4)$$

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

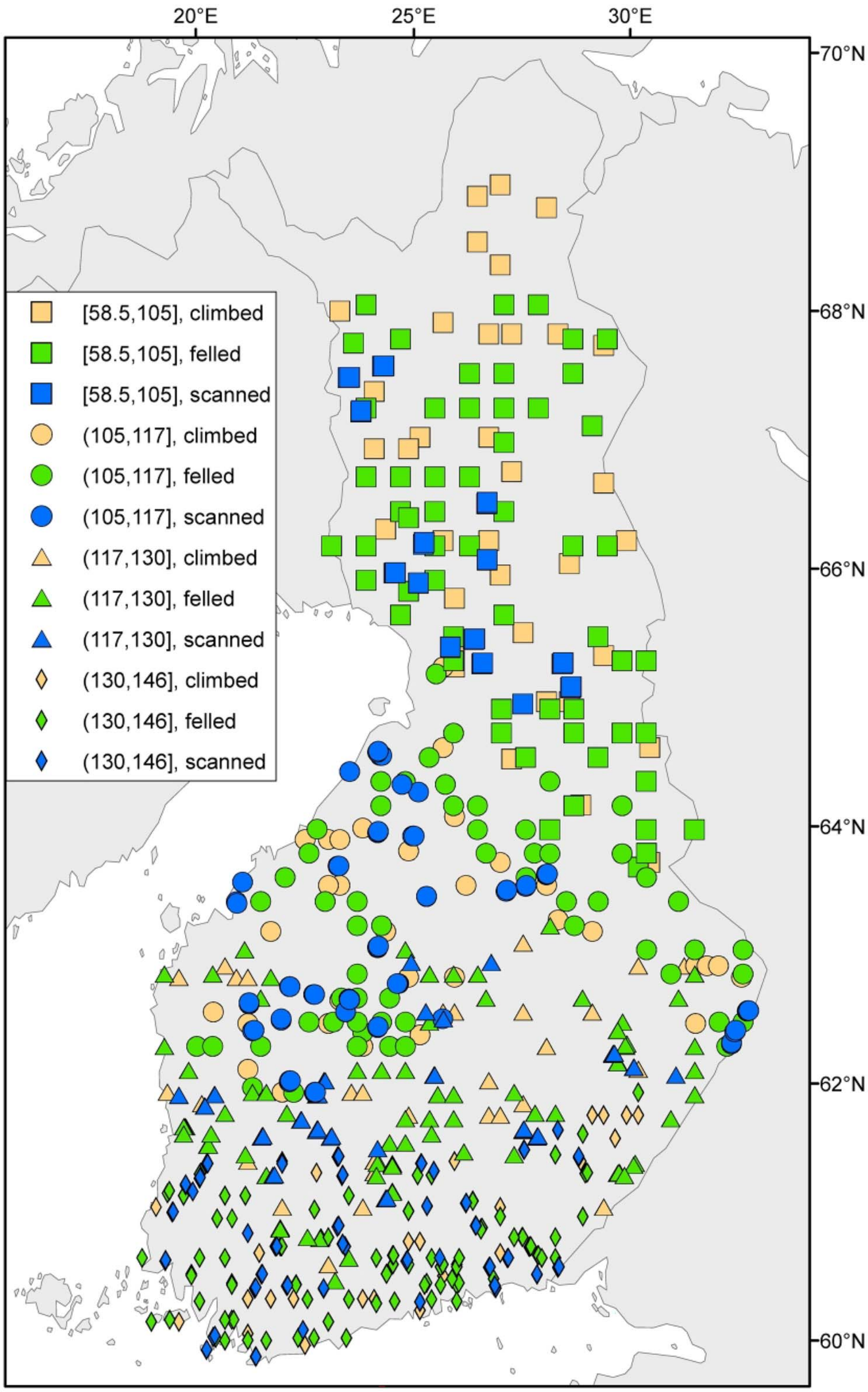

Figure 1 The four regions defined by the locations of the observations and quantiles of temperature sum. Background map: EuroGeographics/UN-FAO. The colours depict the datasets.

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

of the stump height diameter, which was assumed unobserved. Here,

$$w\left(\mathrm{h}, \lambda\right) = 2 - 2\frac{\exp\left(\left(\mathrm{h}-1.3\right)/\exp(\lambda)\right)}{1+\exp\left(\left(\mathrm{h}-1.3\right)/\exp(\lambda)\right)} \tag{5}$$

and λ is a parameter estimated from the data.

Models (1) and (2) are the two- and three-predictor models recommended by Laasasenaho (1982) (61.3 and 61.5). In this study, we introduce the temperature and soil type as well as the change due to the dataset to these models. Model (3) is our new non-linear 'variable form factor volume function' described in more detail in Variable form factor volume model section. Note that regression parameters $a - f$ have different interpretations in different models; the right-hand sides (RHS) of equations (1–3) with random effects set to zero give global uncorrected predictions of form factor, log-volume and volume, respectively (for bias corrections, see Volume predictions and model diagnostics section).

Variable form factor volume model

A problem with Laasasenaho's two-predictor volume function (2) is that it has a minimum as a function of h at $\mathrm{h} = 1.3c/(c+d)$. In all models presented by Laasasenaho (1982) (61.3), $c > 0, d < 0$ and $|c| > |d|$ so that the minimum occurs at $\mathrm{h} > 1.3$ m. Thus, model (2) can produce illogical predictions for the smallest trees. Specifically, the volume tends to infinity as tree height approaches 1.3 m. The aim in developing our new model (3) was to achieve equal flexibility to model (2) while avoiding its illogical behaviour for the smallest trees.

To justify the form of model (3), let us first note that the fixed part of model (2) can be expressed as $v = \exp(a + e \cdot \mathrm{dbh} + d \cdot \log(h - 1.3)) \cdot \mathrm{dbh}^b \cdot \mathrm{h}^c$,which shows that it is a variable form factor version of the Schumacher's model $v = A \cdot \mathrm{dbh}^B \cdot \mathrm{h}^C$, where the form factor depends on dbh and h but has no upper bound. Compared with model (3), the powers B and C are also estimated, and their departure from $B = 2$ and $C = 1$ also models the dependence of form factor on diameter and height. We regard model (3) as better justified because it models the form factor completely within the logit link, thus ensuring that the form factor remains within the range [0, 1].

Model (3) was developed from Schumacher's general form by replacing dbh with dsh, fixing $B = 2$ and $C = 1$ and searching for a sufficiently flexible form factor model. A linear model was specified to the logit-transformed form factor $\eta = 40A/\pi$, resulting in the restriction $0 < \eta < 1$. This is well justified, since $\eta = 1$ gives the volume of a cylinder with diameter dsh and height h. If $\eta = 1/3$, then the function gives the volume of a cone with base diameter dsh and height h. The shape of tree stem is usually between these two extremes but may be neiloid, especially for very short trees, which suggests $\eta < 1/3$. η can be characterized as the only first-order parameter of the model in the context discussed in Mehtätalo and Lappi (2020, Chapter 7). Assuming a common η for all trees would mean that the stem shape is the same regardless of tree size, which is unrealistic. Therefore, we included such parametric transformations of dbh and h in the linear predictor of η which empirically resulted in a good fit. The coefficients of these terms – parameters a, b, c, d, e and f in equation (3) – are the second-order parameters of the model. Notice that defining these second-order predictors at tree level is included in the model definition of Pinheiro and Bates (2000, Chapter 7.1).

The above discussion used stump height diameter, which is usually not measured and needs to be approximated using dbh. The shape of the bottom part of the stem is usually a fustum of a neiloid or cylinder (Kershaw *et al.,* 2016), but for robustness and simplicity, we use the approximation between cone and cylinder. A conical approximation for the lowest part of the stem gives $\mathrm{dsh} = \frac{h}{(h-1.3)}\mathrm{dbh}$, where 1.3 is the applied breast height. The conical approximation is well justified for large trees but is unrealistic for trees that are only slightly taller than the breast height. For those trees, a conical approximation would imply that dsh and v approach infinity as h approaches 1.3 m same way as in model (2). For short trees, a cylinder approximation $\mathrm{dsh} = \mathrm{dbh}$ is better and sufficiently realistic. Equations (4) and (5) were motivated by these considerations: equation (4) is a weighted mean of the two approximations and the weights of equation (5) were based on the right tail of a scaled logit-function. Parameter λspecifies how quickly the weight of a cylinder approximation approaches zero as height increases (the smaller the λ, the faster is the decrease). For example, if $\exp(\lambda) = 0.2$, the weight of cylinder approximation falls <0.01 already when height is 2.3 m. The form of equation (5) guarantees that $(\mathrm{h} - 1.3)/\exp(\lambda)$ is positive for any estimated value of λ.

To find appropriate transformations of dbh and h to the linear predictor of η, the empirical approximations of form factor were computed for each tree as $\eta = 40\mathrm{v}/\pi \cdot \mathrm{dsh}^2 \cdot h$, where the applied dsh was based on initial approximation $\exp(\lambda) = 0.2$. The behaviour of form factor as a function of dbh and h was explored by smoothing the surface of η's using a lowess smoother to find a sufficiently good function to model the surface (see supplementary material, Figure a).

Model fitting and simplification

Each of the three models was fitted separately to the subsets of pine, spruce and birch trees; and in each of the models, the parameters a, b and c of the models 1–3 were allowed to vary regionally and between datasets and soil types. The parameters are described using second-level parameters α (intercept), β (coefficient for the temperature sum, ts), γ (coefficients for the dataset factors, ds) and δ (coefficients for the soil type factor, s). So, for tree i,

$$\begin{aligned} a_i &= \alpha_a + \beta_a \cdot \mathrm{ts}(i) + \gamma_{a,\mathrm{ds}(i)} + \delta_{a,\mathrm{s}(i)}, \\ b_i &= \alpha_b + \beta_b \cdot \mathrm{ts}(i) + \gamma_{b,\mathrm{ds}(i)} + \delta_{b,\mathrm{s}(i)}, \\ c_i &= \alpha_c + \beta_c \cdot \mathrm{ts}(i) + \gamma_{c,\mathrm{ds}(i)} + \delta_{c,\mathrm{s}(i)}. \end{aligned} \tag{6}$$

where $\mathrm{ts}(i)$ is the effective temperature sum at the growing site of tree i, $\mathrm{ds}(i) \in \{\mathrm{climbed, felled, scanned}\}$ is the dataset that includes measurements of tree i and $\mathrm{s}(i) \in \{\mathrm{mineral, peatland}\}$ is the soil type on which tree i was growing. For identifiability, parameter values $\gamma_{\cdot,\mathrm{climbed}}$ and $\delta_{\cdot,\mathrm{mineral}}$ were set to 0 so that, for example, $\alpha_a + \beta_a \cdot \mathrm{ts}(i)$ is the a parameter for mineral soil trees i in climbed data and $\gamma_{a,\mathrm{felled}}$ is the difference in a between felled and climbed trees and so on.

Models (1) and (2) were fitted as linear mixed models with separate variance components $\sigma^2_{\text{cluster}}$, σ^2_{plot} and $\sigma^2_{\text{tree,ds}}$ for cluster-, plot- and tree-level intercepts $\varepsilon_{\text{cluster}}$, $\varepsilon_{\text{plot}}$ and $\varepsilon_{\text{tree}}$, respectively. The tree-level residual variance $\sigma^2_{\text{tree,ds}}$ was allowed to vary between datasets. The volume-transformations of these models led to residuals that were homoscedastic enough so that the assumption of constant variances was justified. Heteroskedasticity in the untransformed volumes of model (3) was modelled with prediction-dependent residual variance according to the power model: $\text{var}(\varepsilon_{\text{tree},i}) = \sigma^2_{\text{tree,ds}(i)}\hat{v}_i^{2\varphi}$. R packages lme and nlme (Pinheiro *et al.*, 2013; Bates *et al.*, 2015; R Core Team, 2021) were used in model-fitting.

The full models specified by equations (1–6) were simplified as far as possible without significant losses in model fit using backward elimination. Starting from the full model, each step of the elimination process consisted of dropping one of the remaining α or β parameters, one group of γ or δ parameters (i.e. categorical predictor dataset or soil from the model of a, b or c), or one of the tree-level parameters d, e or f. The effect that was the least significant according to the F-test ('type III ANOVA') was chosen for elimination. The process was continued until all remaining effects were significant at level 0.01. This strict limit was imposed to avoid over-fitting due to a large number of trees. The principle of marginality (Nelder, 1977) was followed in the elimination: $\alpha_.$ was not dropped, if any of plot-level effects for the same tree-level parameter a, b or c were still present, α_a was not dropped if either of α_b or α_c were present, and the same principle was followed for β, γ and δ.

Volume predictions and model diagnostics

All predictions reported in this paper were computed using only the fixed part of the model, and bias corrections were used when needed as follows (Mehtätalo and Lappi, 2020, Section 10.2). Linear back-transformation of predictions obtained from the right-hand side of equation (1) into volumes, $\hat{v} = \pi \cdot \text{dbh}^2 \cdot h \cdot (\text{RHS of } (1) \text{ with } \varepsilon\text{'s set to } 0)/40$, does not result in any need of bias correction. The back-transformed predictions from equation (2) were multiplied by bias correction $\exp\{(\sigma^2_{\text{cluster}} + \sigma^2_{\text{plot}} + \sigma^2_{\text{tree,ds}(i)})/2\}$. As the residual error in model (3) is additive (i.e. not affected by the non-linear transformation), it does not cause bias. However, the random plot and cluster effects within the model function (i.e. affected by the non-linear transformation) cause bias. That bias can be corrected by averaging the prediction over the distribution of random effects (Mehtätalo and Lappi, 2020). Because the model includes only random intercept in the logit-transformed form factor, such prediction can be easily computed using numerical integration algorithms, such as R-function integrate (R Core Team, 2021).

Model diagnostics and comparisons were based on relative prediction errors $r_{\text{pred}} = (v_{\text{pred}} - v_{\text{obs}})/v_{\text{obs}}$ and relative differences $d_{\text{pred}} = (v_{\text{pred}} - v_{\text{pred0}})/v_{\text{pred0}}$, where pred0 refers to a reference or baseline model. For well-fitting models, they should be approximately independent of v_{pred}. For model (3), also Pearson residuals (i.e. residuals divided by square root of the expected value of residual variance) were used for evaluating the model fit (Mehtätalo and Lappi, 2020).

In order to clarify the effects of dataset, soil and temperature sum, volume predictions were computed for all trees of the study material using their measured dbh, h and $d6$ and all three-way combinations of the levels of predictors 'dataset' and 'soil' and two rather extreme 'temperature sum' values (0.125 and 0.875 quantiles). The $3 \times 2 \times 2 = 12$ different volume predictions were thereby obtained for each tree. The effects were high-lighted by computing their relative differences d_{pred} from predictions v_{pred0} obtained from null models fitted without any temperature sum, dataset or soil effects.

Reasonability of the assumed linear effect of temperature sum was assessed by fitting alternative models that were otherwise similar to the final reduced models, but included, instead of temperature sum, a categorical region effect using the regions of Figure 1. The estimated region effects were plotted against the average temperature sum values of the regions and the linearity (or, at least, general monotonicity) of the resulting graphs was assessed visually.

Behaviour of the two-predictor models for the shortest trees was demonstrated by volume predictions computed at a series of h values between 1.35 and 3 m. The applied dbh values were obtained as a function of h from a simple model

$$\log(\text{dbh}) = a + b \cdot \log(\text{h}) + \varepsilon$$

with parameters a and b estimated from the trees of the study material with $\text{h} \leq 3$ m.

The nature of the change in the stem was addressed by estimating the average stem curve model separately in each dataset using the formula (Laasasenaho, 1982)

$$\frac{d_l}{d_{.2h}} = b_1x + b_2x^2 + b_3x^3 + b_4x^5 + b_5x^8 + b_6x^{13} + b_7x^{21} + b_8x^{34} + \varepsilon, \tag{7}$$

where $x = 1 - l/\text{h}$, i.e. the relative distance to the top, $d_{.2h}$ is the diameter at 20 per cent height and d_l is the diameter at height l. This model was then used for predicting the average stem form in the datasets in order to illustrate the change along the taper curve.

Evaluating the model predictions at area level

In order to test the relevance of updating the volume models currently in use, various comparable mean volume estimates from the NFI11 sample tree measurements based on the NFI division of the country into South and North Finland were computed. The tree-level volume predictions used by Korhonen *et al.* (2017) – to be called NFI volumes in this paper – were mainly based on the three-predictor model of Laasasenaho (1982). We computed alternative volume predictions for the NFI11 sample trees with the models developed in this paper. We mainly used the parameters associated to the scanned dataset as their measurement time is closest to NFI11 among our study material. To verify that the differences are, indeed, due to changes in stem form measurements rather than due to the updated modelling strategy, we also computed predictions of model (1) with parameters associated to the climbed dataset.

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

From each set of alternative volume predictions for the NFI11 sample trees, we computed estimates of mean volume per ha of forest land and poorly productive forest land using the usual NFI routines (e.g. Tomppo *et al.,* 2011) adapted to the use of sample trees only. The results computed with the NFI volumes will be slightly different from those reported in Korhonen *et al.* (2017) because the tally trees (i.e. trees with dbh as the only measured characteristics) were also used in the latter.

The estimates of mean volume based on the new models were affected by the uncertainty in parameter estimates, and therefore, we also provided the standard errors and prediction intervals of the estimates of the mean. This was done through repeated simulation of new parameter values from multivariate normal distribution with mean vector equal to the parameter estimates and covariance matrix equal to their estimated covariance (the Monte Carlo method, e.g. Robert and Casella, 2004). The mean volume estimates were computed using tree volume predictions obtained with each of the $T = 500$ simulated set of parameters, standard error was the standard deviation of these estimates and end points of the prediction interval were their 0.025 and 0.975 quantiles (the percentile method, e.g. Davison and Hinkley, 1997).

It should be noted that these uncertainty metrics are different from the sampling errors that are conventionally reported with the NFI results. The conventional sampling errors are not relevant in our comparison, where the different models are applied to the same set of sample tree measurements, i.e. our comparison is conditional on the realized sampling error of the NFI measurements. Parameter uncertainty reported here results from the sampling of modelling data and not from the sampling of the data to which the models are applied. In applications, this error component is systematic: the same realized errors in parameters are replicated in all applications of the model. The value and the sign of this systematic error remain unknown, but our standard errors and prediction intervals due to parameter uncertainty provide an estimate about the magnitude of this error component when the model is applied in forestry routinely.

Our model fitting dataset and the functions to fit the models of form (3) and to compute the predictions of a fitted model are available in R-package lmfor (Mehtätalo and Kansanen, 2020). The prediction function allows a bias correction based on averaging the prediction over the distribution of random plot and cluster effects using numerical integration. Also a simple approach based on a two-point distribution approximation was implemented (Mehtätalo and Lappi, 2020, p. 318). Prediction errors associated with parameter uncertainty can be estimated by two methods: using a Monte-Carlo simulation (see above) with the fitted non-linear model parameters, or applying the analytical formulas of linear model with a linearized model, which is based on the widely used first-order Taylor approximation of the non-linear model (e.g. Mehtätalo and Lappi, p. 217). The R-script for fitting the full model of form (3) (including routines for finding initial guesses for parameters) is also included in the supplementary material, Appendix.

Results

The currently used three-predictor (model type 1) and two-predictor (model type 2) volume models of Laasasenaho (1982) over-estimated the volumes of the scanned trees by 1.19–3.44 per cent on average, respectively, depending on species and the type of model (Table 4). In all cases, except for the three-predictor model for spruce, this tendency is also apparent for the felled trees. In this study, the original models showed lack of fit also with the climbed data. This is due to the fact that the trees from the poorly productive sites, which were not used by Laasasenaho (1982), were also included in our analysis.

Soil was not a significant predictor in any of the new three-predictor models (1). All parameters of the pine model, the intercept and either dbh or h coefficient of the spruce model, and the intercept of the birch model significantly depended on the dataset and temperature sum (supplementary material, Appendix E, Appendix A). A few more significant effects remained in the two-predictor models for spruce and birch. For example, significant differences were found between soil types in model (2) for spruce (supplementary material, Appendix E, Appendices B and C). In three-predictor models, trees from the felled dataset had the largest residual variation for all species (Appendix A), and in two-predictor models, trees from the scanned dataset had generally the smallest variance (Appendices B and C).

The dataset effects imply that stem volumes of trees with the same dbh, h and ***d*6** were generally greater in the oldest dataset (climbed) and smaller in the newest one (scanned; Figure 2). For pine and spruce, the relative differences between the datasets increased with tree size. Temperature sum was clearly an important predictor in two-predictor models, especially for pine and birch. The dependence of form factor on tree diameter and height for each dataset and species is shown in (supplementary material, Figure b).

The nature of the difference in the stem form between the datasets was illustrated with the average taper curve for each of the datasets (Figure 3). This confirms that the volume is smaller in the newest, i.e. the scanned dataset, and that the change is more obvious in the lower part of the stem (parameters for model (7) in Appendix D).

Except for the smallest pines in the felled and scanned datasets, the relative prediction errors of the two sets of two-predictor models, models (2) and (3), were quite similar (Table 5, see also supplementary material, Figure c). The improvement achieved by including ***d*6** was at the same level as reported in Laasasenaho (1982, p. 43). The respective values for the volumes based on integrating the taper curves, with h and diameter at 20 per cent as predictors, were 7.20 per cent for pine, 6.44 per cent for spruce and 7.71 per cent for birch.

When the temperature sum was replaced by region in the final forms of models (2), it was found that the region effect to parameters ***a***, ***b*** and ***c*** were relatively monotonic functions of the regional mean temperature sums (Figure 4). Thus, our more easily applicable model with a linear effect of the continuous temperature sum predictor seems acceptable.

The problem of models (2) is that their volume predictions often have a minimum at $h > 1.3$ m when considered as a function of tree height. In addition to the values of the model parameters, the location of this minimum depends on dbh − h relationship in the target data. Model (3) does not have that problem (Figure 5).

NFI11 mean volume estimates would have been up to 2 per cent smaller if the new three-predictor models for the scanned dataset had been used (Figure 6, Table 6). Considering uncer-

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Table 4 Mean %-difference of the predictions of Laasasenaho (1982) models (61.3) and (61.5) from the measured volumes for all trees in the modelling data

	2-var models (61.3)			3-var models (61.5)		
Dataset	Pine	Spruce	Birch	Pine	Spruce	Birch
Climbed	0.43	0.90	0.79	0.24	0.09	0.69
Felled	1.41	1.39	1.83	1.07	0.09	1.33
Scanned	1.19	2.53	3.44	1.40	1.38	2.14

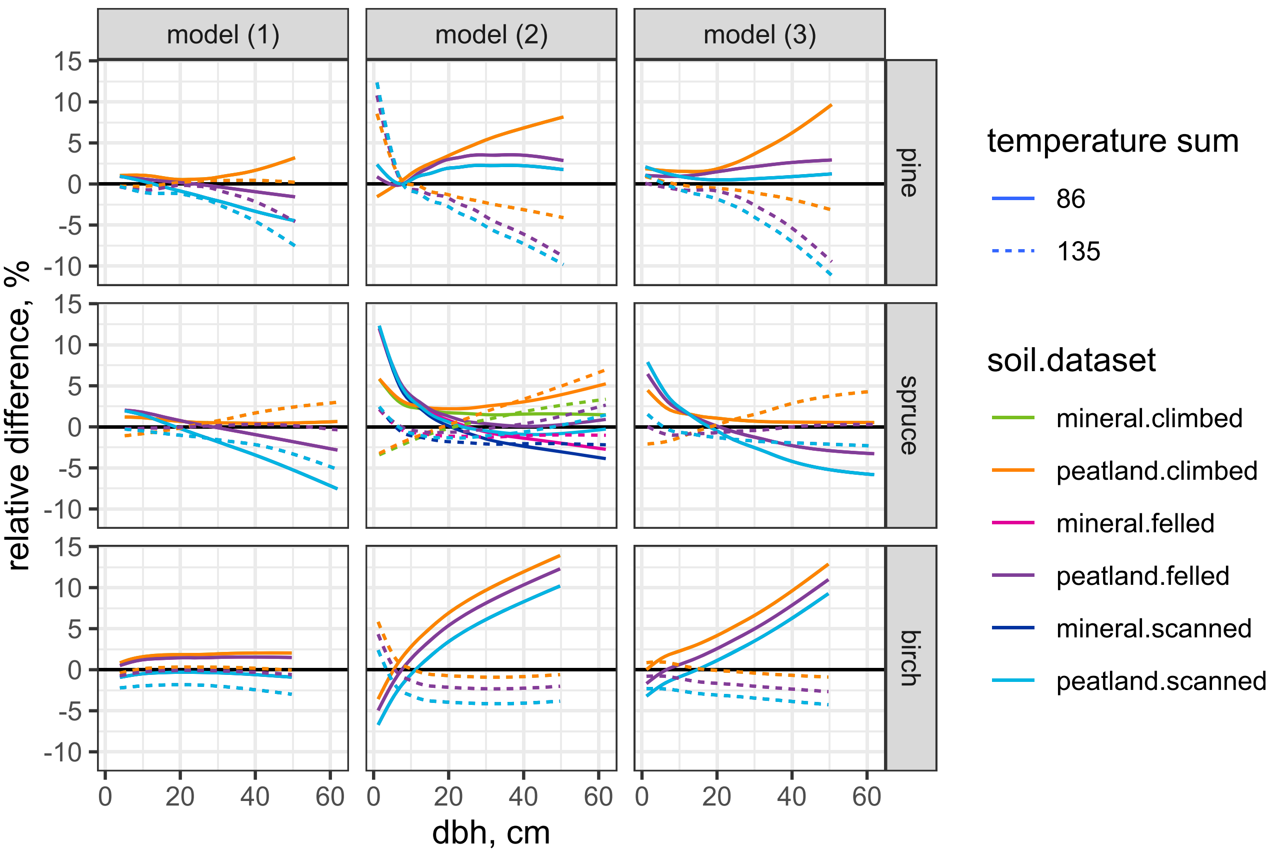

Figure 2 Difference in volume predictions for different values of predictors dataset, soil and temperature sum. Predictions were computed for all trees in the modelling data using the measured dbh, h and ***d6*** (when available) and applying 12 combinations of the other predictors to each tree. They are presented as % differences from the predictions of a null model fitted without any dataset, soil or temperature sum-effects.

tainty in parameter estimates of the new models, the difference was significant in all other strata except for spruces in North Finland. The results with models for climbed data were generally closer to those with sample tree volumes of Korhonen *et al.* (2017). This makes sense since the latter were based on models of Laasasenaho (1982) fitted to the climbed data. Both two-predictor models led to remarkably larger mean volume estimates in South Finland and for spruce also in the North. Note that the effect of parameter uncertainty does not decrease with increasing size of the target data (Table 6, whole study region *vs* South/North). However, the effect is generally smaller for all species combined than for individual species because errors in parameter estimates can be assumed to be independent between species.

The parameter uncertainty with model (3) was estimated using simulation with non-linear models, which is slow but reliable. Alternatively, a linearized model was considered for these purposes, which is fast but may be problematic due to the unaccounted errors of the linear approximation. In the datasets of this study, the results from these approaches were very similar. The effect of bias correction was negligible. The numerical integration and two-point distributions led to very similar results.

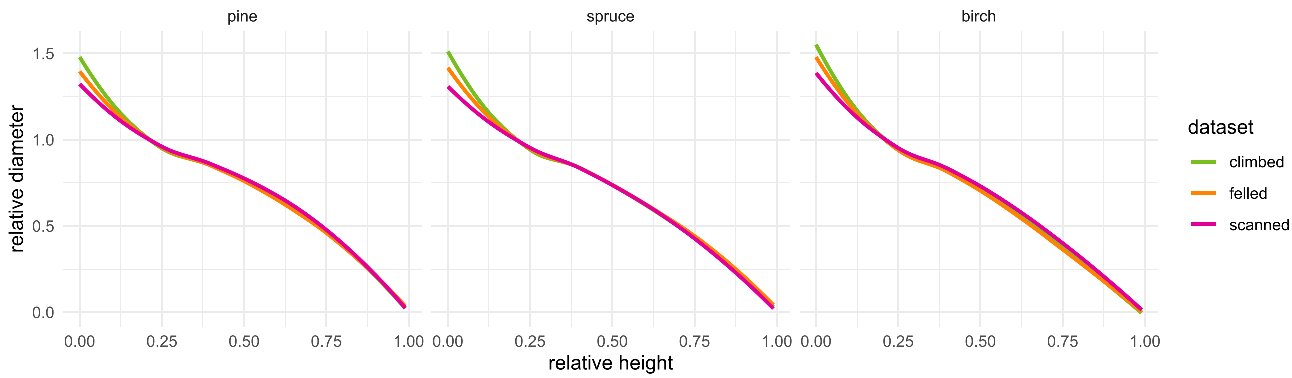

Figure 3 The average stem form in the three datasets by species.

Table 5 Root mean squares of relative prediction errors, r_{pred}, % of measured volume, within the study material. In the column 'models (1 & 2)', model (2) was used for trees with missing *d6*; values in column 'models (1)' were computed from those trees only that had non-missing ***d6***

Species	Model (3)	Model (2)	Models (1 & 2)	Model (1)
Pine	6.63	6.58	4.55	3.94
Spruce	6.98	6.95	4.34	3.61
Birch	8.13	8.21	5.59	5.29

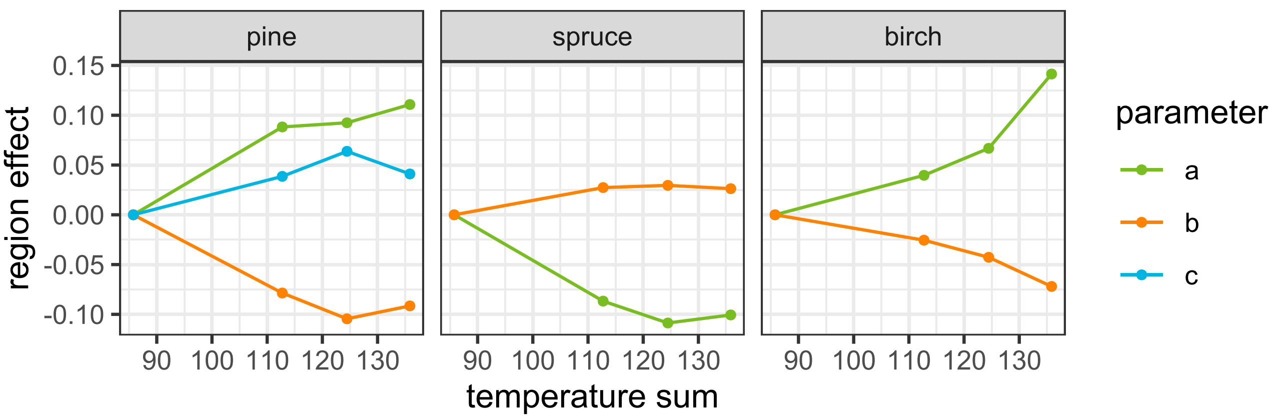

Figure 4 Parameter estimates for region effects (differences from the northernmost region) in alternative versions of model (2), where continuous predictor temperature sum, ts, in (6) was replaced by categorical region. The estimates were plotted against the regional mean of the temperature sum.

Discussion

Some of the model forms used in previous allometric studies are not theoretically sound and such model forms should be avoided (Sileshi, 2014). Laasasenaho (1982) motivated each model selection with a thorough theoretical analysis of the stem form. For instance, the three-predictor model was assumed to consist of a cylinder in the butt section, truncated cone in the middle section and a cone in the uppermost section. The thorough analysis has paid off: it proved to be difficult to improve the fit from the model formulations that Laasasenaho (1982) had originally used, and in this study, we ended up using two of the original model forms again.

Despite this, the original two-predictor model provides illogical results with the smallest trees, which has been problematic in many applications (e.g. Tomppo *et al.*, 2011). With the new formulation (model (3)), it was possible to improve the model behaviour with the smallest trees markedly, while maintaining the good performance with large trees (Figure 5). The non-linear approach allowed us to include a justified model formulation of the variable form factor function to approximate the stump diameter using dbh as weighted sum of conical and cylindrical approximations (equation (5)) and to apply restriction of the form factor to range [0, 1] using a logit link. This is a step forward from the original model forms.

Most volume models used do not include any variables describing the environment or the stand (cf. Zianis *et al.*, 2005). In this study, we included the temperature sum and soil as additional predictors into the basic equations. Thus, we treated the temperature sum and soil as secondary parameters of the

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

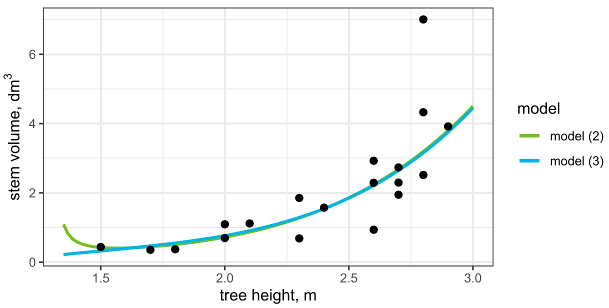

Figure 5 Volume predictions of the two-predictor models for the smallest pines obtained using dbh values predicted as a function of h and the parameters for mineral soils, scanned dataset and average temperature sum over the study plots. The dots show the measured heights and volumes of the smallest pines in the study material; the applied dbh prediction model was fitted to these 19 trees.

models (1–3). This implies that we assume that the primary parameters of the model vary as a function of some additional variable (Mehtätalo and Lappi, 2020). This approach is especially useful in the context of non-linear models, which generally have only a limited number of parameters (e.g. Mehtätalo et al., 2015). The temperature sum proved to be significant for all models (1–3) for at least one of the parameters *a*, *b* or *c* (Appendices A–C).

The temperature sums are given as 30-year averages. Thus, the temperature sums primarily reflect regional differences in the environment. To be able to separate the effect of environmental change from the regional effect, we would need temperature sums for more than one time period. New temperature sums for years 1991–2020 will be available soon, and these temperature sums are likely to reflect the environmental change from the years 1961–1990 temperature sums used in this study. However, even then it would be difficult to define what weight each set of temperature sums should get for trees of different age at measurement times. Obviously, larger living trees would have lived through changes occurring before the younger trees even started their growth, and many of the trees in the old datasets would have been harvested before any major effects of climate change. The fact that trees in northernmost part of Finland are older, when conditioned on size, would also complicate the analysis.

Separating the effect of forest management from other possible causes has been left to future studies. The regional differences in the stem form can be due to environment, the forest management applied, or both. Especially for pine and spruce, the northernmost part of Finland clearly differs from the other parts (Figure 4), which may be due to the fact that the southern part of Finland have been under heavy use much longer than the northern part of Finland (Henttonen *et al*., 2020). However, as the regional effect is dependent on the dataset, also the environmental change (Henttonen *et al.,* 2017) may have had an effect on the observed change.

The most commonly used environmental factors in previous studies were the mean annual temperature and mean precipitation (e.g. Chave *et al.,* 2014; Schneider *et al.,* 2018; Fortin *et al.,* 2019), also given as 30-year mean values. Thus, also these climatic factors share the problem that the 30-year period used has a different significance in the lives of the different trees in the data, while the effect of the climate is likely to be accumulative through time. In this study, the temperature sum was preferred over the mean annual temperature as it combines the latter and the length of the frost free season, which is also an important environmental factor in Finland. Length of winter varies from roughly 100 days in the southern part to 200 days in the northern part of Finland. Precipitation was not deemed as important, as it is rarely a limiting factor of growth.

Including the upper diameter into the model markedly reduced the need for the temperature sum as an explanatory variable, especially for birch but also for pine (Figure 2). This implies that this variable to some extent can reflect the changes in the forest management as well as those in environment, although tree breeding may also have had an effect. One possible explanation is that selective logging was banned in 1948. Since then, commercial forestry in Finland has been based on clear-cutting and planting. The mean volume per hectare in the clear-cut stands has more than doubled (from 98 m^3 ha^{-1} in 1970 to current 225 m^3 ha^{-1} based on NFI results). This may be reflected, for instance, in the mean upper diameters: the climbed data had larger mean upper diameters in the Northernmost region (Figure 1) for pine and birch than the other datasets but smaller for spruce. The change was especially important in the Northern part of Finland, where large areas of selectively logged stands were clear-cut (largest single clear-cut area being 18 000 hectares). The oldest age classes ($\geq$100 years) have largely been harvested in this region, and it means that forest management may have had a larger effect in this region than elsewhere. This also implies the need of updating the models used for predicting the upper diameter in the future.

In the previous study with the climbed and scanned data (Kangas *et al.,* 2020), the dataset effect proved to be highly significant. It needs to be noted, however, that the dataset effect includes the effect of measurement time as well as that of the measurement method. While Pitkänen *et al.* (2021) and Li *et al.* (2021) tested the scanned data to volume modelling with promising results, the error structures are different from the more traditional datasets. For instance, in the scanned data, the lowest and highest parts of the stem is often hard to detect due to the visibility problems caused by the understory vegetation or occlusion by other trees. In addition, with climbed data, the possible deviation from a round tree assumption can be accounted with two measured diameters; in laser scanned data, the algorithm assumes a circular cross-section of the stem and assuming an ellipsoid rather than a circle as the shape of the stem slices would have been difficult. Moreover, using spline smoothing rather than interpolation may have caused some error.

The potential problem of measurement errors was addressed through different approaches. First, an additional dataset was collected in the interval between the collection of the climbed data and the scanned data was included. This dataset, with all diameters measured from felled trees, has a yet different measurement error structure. The assumption was that if the felled data show a change in stem form that is consistent with the other datasets, the strength of the argument of change in time increases, even though the measurement error effect is still confounded with the time effect. When the average taper curve models for all species were calculated, it turned out that

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

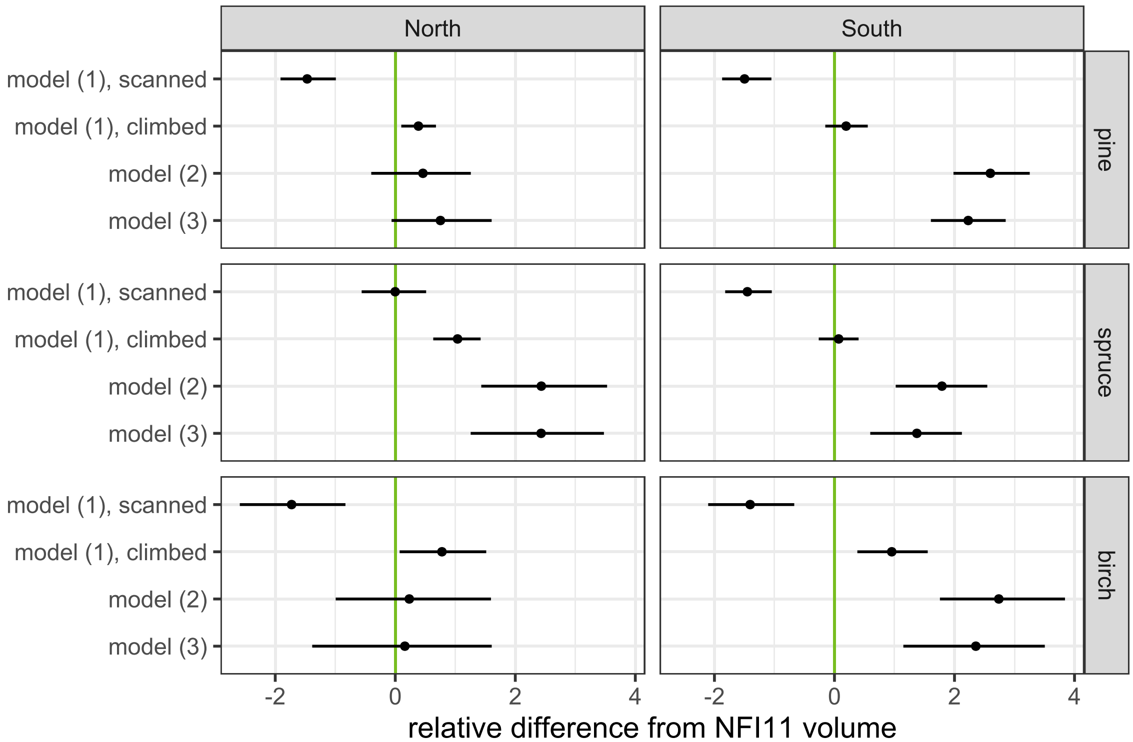

Figure 6 Relative differences, %, of mean volume estimates based on sample tree predictions with the models of this study to those based on sample tree volumes used in Korhonen et al. (2017). Horizontal bars indicate uncertainty due to parameter estimation. Further details are available in the Methods section and the caption of Table 7.

the shape of stems has really changed consistently in time, with the trees on average getting slenderer in time. From the average taper curves, it is also evident that the largest change has happened in the butt area (Figure 3).

Second, the error sources were minimized as much as possible by using field-measured stump height and stump height diameter as well as total height in the volume calculations (Pitkänen *et al.*, 2021). Using the field-measured values of stump-height diameter should markedly reduce the possibility that the thinner butts of the scanned trees is a phenomena related to the scanning rather than the tree shape. Moreover, the explanatory variables included in the datasets were all measured in the field so that any error-in-variables bias would have been similar across datasets.

At regional level, i.e. in the NFI sample trees, the differences between the models were fairly large and also significant. Especially, the large difference between the three-predictor models and two-predictor models in the regional test was surprising. Even though the volumes conditioned to dbh and h are smaller in the new datasets, the three-predictor models (both old and new) provided even smaller volumes on average. The difference is due to the NFI11 trees being slenderer on average than the scanned data for all tree species. This, in turn, can partly be explained by the representativeness of the data and partly by the measurement method. Particularly, the $d6$ measurements with a calliper and a 5-m rod have been noted to be on average underestimated (by 0.1 cm, Päivinen *et al.*, 1992). This highlights the importance of updating the approach used to calculate volume.

In this study, we linearized the models (1) and (2) but fitted model (3) as a non-linear model. Model (1) was fitted as a form factor model and transformed into a volume model afterwards. Model (2) was linearized using the logarithmic transformation, a bias correction being then needed to make predictions in the linear scale. As a result of this transformation, heteroscedasticity was not apparent in the residuals and thus an explicit model for the variance was not needed. In model (3), a specific variance function was needed to account for the heteroscedasticity. While there was no need for backtransformation, there is still a bias correction to be considered as a result of having plot and cluster random effects within the logit link function: if the model function is non-linear with respect to the random effects, the 'plug-in' predictions where random effect are given their expected value of zero are biased. However, that bias component is often small compared with linear models for transformed variables, as the residual error is additive and not affected by the non-linear logit link function. Also with our model (3), the effect of this bias correction was negligible.

We tested different structures to the random effects. However, only random intercepts were finally used. More complicated models often had convergence problems and, when successfully estimated, did not significantly improve the model fit.

We also developed tools to address the prediction error due to the parameter uncertainty. The effect of parameter uncertainty to the confidence interval of the predictions was analysed through a simulation exercise that used the distribution of the parameters. A computationally less intensive approach based

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Table 6 Estimates of mean volume, $m^3\ ha^{-1}$, from the sample trees of NFI11 in South Finland (excluding Åland islands) and North Finland (excluding northernmost Lapland); column 'NFI11' reports results based on sample tree volumes that were used in Korhonen et al. (2017), mainly based on the three-predictor models of Laasasenaho (1982); the other columns report results based on sample tree volumes predicted with the models presented in this study for dataset scanned; columns 's.e.' report standard errors in the estimates due to parameter uncertainty

Region	Species	Nfi11	Model (1)		Model (2)		Model (3)	
		$m^3\ ha^{-1}$	$m^3\ ha^{-1}$	s.e., %	$m^3\ ha^{-1}$	s.e., %	$m^3\ ha^{-1}$	s.e., %
South	Pine	63.2	62.3	0.2	64.9	0.3	64.7	0.3
	Spruce	50.4	49.7	0.2	51.3	0.4	51.1	0.4
	Birch	23.5	23.2	0.4	24.2	0.5	24.1	0.6
	Total	137.1	135.1	0.1	140.3	0.2	139.8	0.2
North	Pine	44.6	44.0	0.2	44.8	0.4	45.0	0.4
	Spruce	15.7	15.7	0.3	16.1	0.5	16.1	0.6
	Birch	12.7	12.5	0.5	12.7	0.7	12.7	0.8
	Total	73.1	72.2	0.2	73.7	0.3	73.8	0.3
Whole region	Pine	54.7	53.9	0.2	55.7	0.3	55.6	0.3
	Spruce	34.4	34.0	0.2	35.1	0.4	35.0	0.4
	Birch	18.5	18.3	0.4	18.9	0.5	18.8	0.6
	Total	107.7	106.2	0.1	109.7	0.2	109.4	0.2

on first-order Taylor approximation for linearized models was additionally implemented for model (3). The results from both approaches were very similar. In general, the standard errors related to the parameter uncertainty were approximately some 0.4 per cent for each species separately, which implies that there is a systematic error in all volume estimates based on these models. This likely amounts to <1 per cent for each of the three species. While it is not really possible to correct that systematic error due to the prediction errors (otherwise we could have estimated an unbiased model in the first place), it is possible to address its importance in all analyses. It is important to account this effect, especially with large-scale calculations, such as NFI results, and long-term calculations.

Conclusion

The strong effect of dataset (representing a time-period of 50 years in this study) in the volume models identified in this study serves as evidence of a changed tree stem shape in the studies time period. Introduction of the independent felled dataset into the analysis corroborates the observed change in stem shape over time, even though it is not possible to fully exclude the effect of the measurement technique. This is due to the measurement effect being confounded with the time. This change can be attributed either to a changed management or to a changed environment. The significant effect of temperature sum, on one hand, points to environmental change, and the role of upper diameter in the three-predictor model points to the changed management, on the other. The old stem taper models were based on thorough theoretical analysis of the stem form, and therefore, no linear or linearizable model form could outperform the forms selected in 1980s. With non-linear modelling, it was possible to make a more rigorous theoretical analysis and find a model that outperforms the old two-parameter models with small trees and also fits to larger trees. That required using a variable form factor model with weighted mean of two approximations of a stump diameter, one suitable for smallest trees and the other for larger trees. The change in the models is significant when compared with the model used in NFI calculations. We furthermore found out that including environmental variables, in our case temperature sum, into the stem taper models can improve the model performance on national scale. This suggests that volume models need to be updated from time to time since tree shapes can change due to changes in management or environmental condition and that the integration of environmental variables into the stem taper models can help to improve their performance in countries with pronounced environmental gradients.

Supplementary data

Supplementary data are available at *Forestry* online.

Data availability

The data underlying this article are available in the article and in its online supplementary material.

Conflict of interest statement

None declared.

Funding

Academy of Finland flagship programme; Forest-Human-Machine Interplay (UNITE) flagship (grant number 337655).

References

Bates, D., Mächler, M., Bolker, B. and Walker, S. 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48. https://doi.org/10.18637/jss.v067.i01.

Burkhart, H.E. and Tomé, M. 2012 Tree-stem volume equations. In *Modeling Forest Trees and Stands*. Springer. doi: https://doi.org/10.1007/978-90-481-3170-9_3.

Chave, J., Rejou-Mechain, M., Burquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C., *et al.* 2014 Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang. Biol.* **20**, 3177–3190. https://doi.org/10.1111/gcb.12629

Cysneiros, C.V., de Souza, F.C., Gaui, T.D., Pelissari, A.L., Orso, G.A., Machado, S.A., *et al.* 2021 Integrating climate, soil and stand structure into allometric models: an approach of site-effects on tree allometry in Atlantic forest. *Ecol. Indic.* **127**, 107794. https://doi.org/10.1016/j.ecolind.2021.107794.

Davison, A. and Hinkley, D. 1997 *Bootstrap Methods and Their Application*. Cambridge University Press, p. 582.

Fischler, M.A. and Bolles, R.C. 1981 Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **24**, 381–395. https://doi.org/10.1145/358669.358692.

Fortin, M., Van Couwenberghe, R., Perez, V. and Piedallu, C. 2019 Evidence of climate effects on the height-diameter relationships of tree species. *Ann. For. Sci.* **76**, 1. https://doi.org/10.1007/s13595-018-0784-9.

Fu, L., Lei, X., Hu, Z., Zeng, W., Tang, S., Marshall, P., *et al.* 2017 Integrating regional climate change into allometric equations for estimating tree aboveground biomass of Masson pine in China. *Ann. For. Sci.* **74**, 42. https://doi.org/10.1007/s13595-017-0636-z.

Goussanou, C.A., Guendehou, S., Assogbadjo, A.E., Kaire, M., Sinsin, B. and Cuni-Sanchez, A. 2016 Specific and generic stem biomass and volume models of tree species in a West African tropical semi-deciduous forest. *Silva Fennica* **50**, 1474. https://doi.org/10.14214/sf.1474.

Henttonen, H.M., Nöjd, P. and Mäkinen, H. 2017 Environment-induced growth changes in the Finnish forests during 1971-2010—an analysis based on National Forest Inventory. *For. Ecol. Manag.* **386**, 22–36. https://doi.org/10.1016/j.foreco.2016.11.044.

Henttonen, H., Nöjd, P. and Mäkinen, H. 2020 Size-class structure of the forests of Finland during 1921-2013: a recovery from centuries of of explotation, guided by forest policies. *Eur. J. For. Res.* **139**, 279–293. https://doi.org/10.1007/s10342-019-01241-y.

Husch, B., Miller, C.I., and Beers, T.W. 1972 Forest Mensuration. Ronald Press.

Kangas, A., Henttonen, H., Pitkänen, T., Sarkkola, S. and Heikkinen, J. 2020 Re-calibrating stem volume models—Is there change in the tree trunk form from the 1970s to the 2010s? *Silva Fennica* **54**, 10269.

Kershaw, J.A.J., Ducey, M.J., Beers, T.W. and Husch, B. 2016. *Forest Mensuration*. Wiley. https://doi.org/10.1002/9781118902028.

Korhonen, K.T., Ihalainen, A., Ahola, A., Heikkinen, J., Henttonen, H.M., Hotanen, J.-P., *et al.* 2017 *Suomen metsät 2009–2013 ja niiden kehitys 1921–2013*. Luonnonvara- ja biotalouden tutkimus 59/2017, Luonnonvarakeskus, [In Finnish].

Korhonen, K. and Maltamo, M. 1990 *Männyn maanpäällisten osien kuivamassat Etelä-Suomessa*. Metsäntutkimuslaitoksen tiedonantoja 371. Metsäntutkimuslaitos. [In Finnish].

Laasasenaho, J. 1982 Taper curve and volume functions for pine, spruce and birch. *Communicationes Instituti Forestalia Fennica* **108**, 1–74.

Lahtinen, A. and Laasasenaho, J. 1979 On the construction of taper curves by using spline functions. *Communications Instituti Forestalia Fennica* **95**, 1–63.

Lappi, J. 1986 Mixed linear models for analyzing and predicting stem form variation of scots pine. *Communications Instituti Forestalia Fennica* **134**, 1–69.

Lappi, J. 2006 A multivariate, nonparametric stem-curve prediction method. *Can. J. For. Res.* **36**, 1017–1027. https://doi.org/10.1139/X05-305.

Li, D., Guo, H., Jia, W. and Wang, F. 2021 Analysis of taper functions for Larix olgensis using mixed models and TLS. *Forests* **12**, 196. https://doi.org/10.3390/f12020196.

Mehtätalo, L., de-Miguel, S. and Gregoire, T.G. 2015 Modeling height-diameter curves for prediction. *Can. J. For. Res.* **45**, 826–837. https://doi.org/10.1139/cjfr-2015-0054.

Mehtatalo L., and Kansanen K. 2020 lmfor: functions for forest biometrics. R package version 1.5. https://CRAN.R-project.org/package=lmfor.

Mehtätalo, L. and Lappi, J. 2020 *Biometry for Forestry and Environmental Data*. CRC Press, p. 411.

Nelder, J.A. 1977 A reformulation of linear models. *J. R. Stat. Soc.* **140**, 48–77. https://doi.org/10.2307/2344517.

Päivinen, R., Nousiainen, M. and Korhonen, K.T. 1992 Puutunnusten mittaamisen luotettavuus. Summary: accuracy of certain tree measurements. *Folia For.* **787**, 1–20. [In Finnish with English summary].

Pinheiro, J.C. and Bates, D.M. 2000. Mixed-effects Models in S and Splus. Springer.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D and the R Development Core Team 2013. nlme: linear and nonlinear mixed effects models. R package version 3.1-108.

Pitkänen, T.P., Raumonen, P. and Kangas, A. 2019 Measuring stem diameters with TLS in boreal forests by complementary fitting procedure. *ISPRS J. Photogramm. Remote Sens.* **147**, 294–306. https://doi.org/10.1016/j.isprsjprs.2018.11.027.

Pitkänen, T.P., Raumonen, P., Liang, X., Lehtomäki, M. and Kangas, A.S. 2021 Improving TLS-based stem volume estimates by field measurements. *Comput. Electron. Agric.* **180**, 105882. https://doi.org/10.1016/j.compag.2020.105882.

Qiu, H., Liu, S., Zhang, Y. and Li, J. 2021 Variation in height-diameter allometry of ponderosa pine along competition, climate, and species diversity gradients in the western United States. *For. Ecol. Manag.* **497**, 119477. https://doi.org/10.1016/j.foreco.2021.119477.

R Core Team 2021 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, https://www.R-project.org/.

Robert, C.P. and Casella, G. 2004 *Monte Carlo Statistical Methods*. Springer, p. 649.

Schneider, R., Franceschini, T., Fortin, M. and Saucier, J.-P. 2018 Climate-induced changes in the stem form of 5 North American tree species. *For. Ecol. Manag.* **427**, 446–455. https://doi.org/10.1016/j.foreco.2017.12.026.

Schumacher, F.X. and Hall, F.S. 1933 Logarithmic expression of timber tree volume. *J. Agric. Res.* **47**, 719–734.

Sileshi, G.W. 2014 A critical review of forest biomass estimation models, common mistakes and corrective measures. *For. Ecol. Manag.* **329**, 237–254. https://doi.org/10.1016/j.foreco.2014.06.026.

Tomppo, E., Heikkinen, J., Henttonen, H.M., Ihalainen, A., Katila, M., Mäkelä, H., Tuomainen, T. and Vainikainen, N. 2011 *Designing and Conducting a Forest Inventory—Case: 9th National Forest Inventory of Finland*. Springer. https://doi.org/10.1007/978-94-007-1652-0.

Vallet, P., Dhôte, J.-F., Le Moguédec, G., Ravart, M. and Pignard, G. 2006 Development of total aboveground volume equations for seven

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

important forest tree species in France. *For. Ecol. Manag.* **229**, 98–110. https://doi.org/10.1016/j.foreco.2006.03.013.

Vibrans, A.C., Moser, P., Oliveira, L.Z. and de Maçaneiro, J.P. 2015 Generic and specific stem volume models for three subtropical forest types in southern Brazil. *Ann. For. Sci.* **72**, 865–874. https://doi.org/10.1007/s13595-015-0481-x.

Xu, Q., Lei, X., Zang, H. and Zeng, W. 2022 Climate change effects on height–diameter allometric relationship vary with tree species and size for larch plantations in northern and northeastern China. *Forests* **13**, 468. https://doi.org/10.3390/f13030468.

Xu, K., Jiang, J. and He, F. 2021. Climate-based allometric biomass equations for five major Canadian timber species. *Can. J. For. Res.* **51**, 1633–1642. doi:https://doi.org/10.1139/cjfr-2020-0485

Zianis, D., Muukkonen, P., Mäkipää, R. and Mencuccini, M. 2005 Biomass and stem volume equations for tree species in Europe. *Silva Fennica Monographs* **2005**, 1–63. https://doi.org/10.14214/sf.sfm4.

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Appendix A Parameter estimates for models defined by equations (1) and (6). The right-hand side of equation (1) with $\varepsilon_{\text{cluster}} = \varepsilon_{\text{plot}} = \varepsilon_{\text{tree}} = 0$ gives the global prediction of form factor and needs to be multiplied with $\pi \cdot \text{dbh}^2 \cdot h/40$ to obtain the volume prediction. Estimates of parameters appearing in equations (1) and (6) are accompanied by standard errors and *P*-values from the Wald's test of difference from 0. β parameters describe the temperature sum effects. α_a, α_b and α_c are the ***a***, ***b*** and ***c*** parameters for the mineral soils of the climbed dataset and γ parameters quantify their difference from those of the other datasets, e.g. $a_{\text{felled}} = \alpha_a + \gamma_{a,\text{felled}}$. Note: Estimates are provided here only for illustration. For proper numerical precision, R model objects provided as electronic supplements should be used when applying the models.

Parameter	Pine			Spruce			Birch		
	Estimate	Standard error	$P>0$	Estimate	Standard error	$P>0$	Estimate	Standard error	$P>0$
α_a	−0.08654	0.01558	0.0000	−0.10467	0.01496	0.0000	−0.07734	0.02396	0.0013
α_b	0.00070	0.00033	0.0360	−0.00175	0.00012	0.0000	−0.00198	0.00019	0.0000
α_c	−0.00005	0.00093	0.9548	0.00306	0.00079	0.0001	0.00635	0.00083	0.0000
d	2.88178	0.04476	0.0000	2.53886	0.04313	0.0000	2.37385	0.07796	0.0000
e	1.28343	0.03379	0.0000	1.59561	0.03533	0.0000	1.32783	0.07508	0.0000
f	0.49211	0.00927	0.0000	0.46154	0.00908	0.0000	0.37397	0.01548	0.0000
β_a	−0.00034	0.00007	0.0000	−0.00038	0.00007	0.0000	−0.00014	0.00005	0.0085
β_b	−0.00002	0.00000	0.0000						
β_c	0.00004	0.00001	0.0000	0.00002	0.00000	0.0002			
$\gamma_{a,\text{felled}}$	−0.00566	0.00325	0.0817	0.00656	0.00271	0.0161	−0.00182	0.00225	0.4193
$\gamma_{b,\text{felled}}$	−0.00091	0.00025	0.0004	−0.00031	0.00013	0.0146			
$\gamma_{c,\text{felled}}$	0.00142	0.00041	0.0005						
$\gamma_{a,\text{scanned}}$	−0.00316	0.00321	0.3247	0.00771	0.00245	0.0019	−0.00981	0.00196	0.0000
$\gamma_{b,\text{scanned}}$	−0.00110	0.00017	0.0000	−0.00060	0.00011	0.0000			
$\gamma_{c,\text{scanned}}$	0.00118	0.00030	0.0001						
σ_{cluster}	0.00403			0.00097			0.00508		
σ_{plot}	0.00680			0.00498			0.00620		
$\sigma_{\text{tree,climbed}}$	0.01712			0.01651			0.02276		
$\sigma_{\text{tree,felled}}$	0.02279			0.01943			0.02565		
$\sigma_{\text{tree,scanned}}$	0.01644			0.01692			0.02321		

Appendix B Parameter estimates for models defined by equations (2) and (6). The right-hand side of equation (2) with $\varepsilon_{\text{cluster}} = \varepsilon_{\text{plot}} = \varepsilon_{\text{tree}} = 0$ gives the global prediction of log-volume, which needs to be exponentiated to obtain a (biased) volume prediction; the bias correction is presented in Volume predictions and model diagnostics section. Estimates of parameters appearing in equations (2) and (6) are accompanied by standard errors and *P*-values from the Wald's test of difference from 0. β parameters describe the temperature sum effects, γ parameters describe the dataset effects, δ parameters the soil effects and α parameters describe the intercepts (see caption of Appendix A for further details). Note: Estimates are provided here only for illustration. For proper numerical precision, R model objects provided as electronic supplements should be used when applying the models.

Parameter	Pine			Spruce			Birch		
	Estimate	Standard error	$P > 0$	Estimate	Standard error	$P > 0$	Estimate	Standard error	$P > 0$
α_a	−3.68354	0.05275	0.0000	−3.42955	0.07059	0.0000	−4.65382	0.11833	0.0000
α_b	2.18652	0.02664	0.0000	1.78430	0.02584	0.0000	2.12488	0.03852	0.0000
α_c	2.21490	0.04849	0.0000	2.78307	0.04388	0.0000	3.79665	0.13451	0.0000
d	−1.23539	0.03046	0.0000	−1.47244	0.03807	0.0000	−2.41251	0.11858	0.0000
e	−0.00381	0.00050	0.0000	−0.00690	0.00063	0.0000	−0.00771	0.00117	0.0000
β_a	0.00216	0.00036	0.0000	−0.00235	0.00048	0.0000	0.00264	0.00070	0.0002
β_b	−0.00195	0.00022	0.0000	0.00065	0.00017	0.0001	−0.00139	0.00028	0.0000
β_c	0.00100	0.00028	0.0004						
$\gamma_{a,\text{felled}}$	−0.00376	0.01939	0.8462	0.07736	0.02169	0.0004	−0.01438	0.00692	0.0385
$\gamma_{b,\text{felled}}$	−0.08112	0.01362	0.0000	−0.02892	0.00763	0.0002			
$\gamma_{c,\text{felled}}$	0.08922	0.01538	0.0000						
$\gamma_{a,\text{scanned}}$	0.02241	0.02105	0.2870	0.08346	0.02220	0.0002	−0.03312	0.00651	0.0000
$\gamma_{b,\text{scanned}}$	−0.07316	0.01120	0.0000	−0.03325	0.00743	0.0000			
$\gamma_{c,\text{scanned}}$	0.06688	0.01400	0.0000						
$\delta_{a,\text{peatland}}$				0.00345	0.01941	0.8591			
$\delta_{b,\text{peatland}}$				0.06530	0.01854	0.0004			
$\delta_{c,\text{peatland}}$				−0.07016	0.01985	0.0004			
σ_{cluster}	0.01514			0.01135			0.01982		
σ_{plot}	0.02497			0.02760			0.02264		
$\sigma_{\text{tree,climbed}}$	0.06058			0.06496			0.07490		
$\sigma_{\text{tree,felled}}$	0.06435			0.06297			0.07130		
$\sigma_{\text{tree,scanned}}$	0.04802			0.04999			0.07058		

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022

Appendix C Parameter estimates for models defined by equations (3–6). The right-hand side of equation (3) with $\varepsilon_{cluster} = \varepsilon_{plot} = \varepsilon_{tree} = 0$ gives a slightly biased global volume prediction; the bias correction is presented in Volume predictions and model diagnostics section. Estimates of parameters appearing in equations (3), (5) and (6) are accompanied by standard errors, and *P*-values from the Wald's test of difference from 0. β parameters describe the temperature sum effects, γ parameters describe the dataset effects and α parameters describe the intercepts (see caption of Appendix A for further details). Note: Estimates are provided here only for illustration. For proper numerical precision, R model objects provided as electronic supplements should be used when applying the models.

Parameter	Pine			Spruce			Birch		
	Estimate	Standard error	$P > 0$	Estimate	Standard error	$P > 0$	Estimate	Standard error	$P > 0$
α_a	0.01368	0.04493	0.7609	0.23709	0.06839	0.0005	−0.67021	0.10719	0.0000
α_b	0.00704	0.00186	0.0002	−0.02423	0.00119	0.0000	0.00910	0.00453	0.0447
α_c	−0.01102	0.00425	0.0095	0.00652	0.00415	0.1160	0.03081	0.00355	0.0000
d	−3.04833	0.13171	0.0000	−3.68858	0.22357	0.0000	−2.15717	0.41955	0.0000
e				0.00015	0.00005	0.0020	−0.00036	0.00012	0.0029
f	1.00773	0.29721	0.0007	2.99481	0.53185	0.0000	2.79948	0.70398	0.0001
λ	−1.73570	0.13012	0.0000	−1.32830	0.12557	0.0000	−0.78318	0.27983	0.0053
β_a	−0.00043	0.00032	0.1745	−0.00223	0.00040	0.0000	0.00076	0.00053	0.1564
β_b	−0.00014	0.00002	0.0000				−0.00014	0.00003	0.0001
β_c	0.00013	0.00003	0.0003	0.00011	0.00003	0.0001			
$\gamma_{a,felled}$	−0.02526	0.01513	0.0951	0.05019	0.01750	0.0042	−0.02081	0.01094	0.0576
$\gamma_{b,felled}$	−0.00589	0.00123	0.0000						
$\gamma_{c,felled}$	0.00902	0.00200	0.0000	−0.00390	0.00106	0.0003			
$\gamma_{a,scanned}$	0.00035	0.01706	0.9838	0.05793	0.01764	0.0010	−0.04935	0.01043	0.0000
$\gamma_{b,scanned}$	−0.00407	0.00092	0.0000						
$\gamma_{c,scanned}$	0.00396	0.00160	0.0136	−0.00515	0.00097	0.0000			
$\sigma_{cluster}$	0.02694			0.02016			0.03033		
σ_{plot}	0.04282			0.04573			0.03964		
$\sigma_{tree,climbed}$	0.07526			0.08248			0.07571		
$\sigma_{tree,felled}$	0.07977			0.08323			0.07195		
$\sigma_{tree,scanned}$	0.06301			0.06895			0.07381		
φ	0.95396			0.94456			0.99095		

Appendix D Parameter estimates for taper curve models (7).

Dataset		Pine			Spruce			Birch		
		Estimate	Std. Error	*t* value	Estimate	Std. Error	*t* value	Estimate	Std. Error	*t* value
Climbed	b_1	2.209654	0.02610472	84.64576	2.332404	0.03310898	70.44626	0.9836599	0.05281394	18.625
	b_2	−1.253057	0.2136512	−5.864966	−3.266161	0.2709988	−12.05231	3.845336	0.4322544	8.896002
	b_3	−0.4930451	0.4713317	−1.046068	3.605485	0.597877	6.030479	−7.758683	0.9535947	−8.136248
	b_4	1.448394	0.665182	2.177441	−1.963125	0.843823	−2.326466	8.780191	1.345801	6.524136
	b_5	−1.732559	0.7179847	−2.413086	−0.7298747	0.9108443	−0.8013167	−9.483963	1.45264	−6.528778
	b_6	2.128949	0.5529165	3.850398	3.102304	0.7014588	4.422646	8.923729	1.118676	7.977048
	b_7	−1.747868	0.2971082	−5.882936	−3.392779	0.3769354	−9.000956	−6.416295	0.601119	−10.67392
	b_8	1.090554	0.08409131	12.96868	2.063306	0.1066867	19.33986	2.937096	0.1701366	17.26316
Felled	b_1	2.558634	0.03376175	75.78499	2.658635	0.04560419	58.29803	1.218756	0.07872799	15.48059
	b_2	−4.119581	0.2945941	−13.98392	−5.274632	0.3980434	−13.2514	1.723098	0.6806427	2.531575
	b_3	5.30609	0.6755294	7.854713	7.359041	0.9129659	8.060586	−3.365435	1.55176	−2.168786
	b_4	−5.373937	0.9940529	−5.406088	−6.660855	1.34408	−4.955699	3.101181	2.267446	1.367698
	b_5	4.398924	1.10618	3.97668	4.411605	1.496449	2.948048	−3.209363	2.507854	−1.279725
	b_6	−1.977288	0.8739062	−2.262586	−1.3565	1.182941	−1.146719	3.443886	1.970122	1.748057
	b_7	0.3657713	0.4801638	0.7617636	−0.4385974	0.6508456	−0.6738886	−2.807817	1.077735	−2.605296
	b_8	0.3571666	0.1390495	2.568629	0.8862659	0.1891905	4.684516	1.563363	0.3111132	5.025061
Scanned	b_1	2.303274	0.02511136	91.72237	2.009117	0.03436261	58.46811	1.7017	0.0607347	28.01858
	b_2	−1.543539	0.2188965	−7.051457	−1.344592	0.2995304	−4.489	−0.5352576	0.5294261	−1.011015
	b_3	−0.02144857	0.5014357	−0.0427743	0.6853101	0.6861284	0.9988074	0.5671405	1.212779	0.467637
	b_4	0.358095	0.7364097	0.4862714	−0.7139362	1.007613	−0.7085423	−2.239303	1.781091	−1.257265
	b_5	0.6176541	0.8171306	0.7558817	1.45179	1.118027	1.298529	3.649125	1.976324	1.846421
	b_6	−1.405169	0.6426328	−2.186582	−2.275319	0.8792501	−2.587794	−3.863442	1.554281	−2.485679
	b_7	1.51449	0.3506372	4.31925	2.137261	0.4797323	4.455113	2.988462	0.8480561	3.523897
	b_8	−0.4192849	0.1003352	−4.178841	−0.5428525	0.1372739	−3.954521	−0.7675704	0.2426722	−3.162993

Downloaded from https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpac038/6711417 by Natural Resources Institute Finland (Luke) user on 26 September 2022