



OPEN

Applying taper function models for black locust plantations in Greek post-mining areas

Florian Wilms^{1,2}, Ferréol Berendt², Karol Bronisz³, Ulyana Bashutska^{2,4}, Mariangela Fotelli¹, Kalliopi Radoglou⁵ & Gavriil Spyroglou¹✉

A key process in forest management planning is the estimation of tree volume and, more specifically, merchantable volume. The ability to predict the cumulative stem volume relative to any upper stem diameter on standing trees or stands is essential for forest inventories and the management of forest resources. In the 1980s, the Hellenic Public Power Corporation (HPPC) started the rehabilitation of lignite post-mining areas in Greece by planting mainly black locust (*Robinia pseudoacacia*, L.). Today, these plantations occupy an area of approximately 2570 ha, but the stem volume has not yet been estimated. Therefore, we aimed to estimate the over- and under-bark stem volume using taper function models for 30 destructively sampled trees. Of the nineteen calibrated fixed-effects models, Kozak's (2004) equation performed best for both the over-bark and under-bark datasets, followed by Lee's (2003) and Muhairwe's (1999) equations. Two fixed effect models were compared with fitted coefficients from Poland and the United States confirming that the local model fits were better suited, as the foreign model coefficients caused an increase in root mean square error (RMSE) for stem diameter predictions of 13% and 218%, respectively. The addition of random effects on a single-stem basis for two coefficients of Kozak's (2004) equation improved the model fit significantly at 86% of the over-bark fixed effect RMSE and 69% for the under-bark model. Integrated taper functions were found to slightly outperform three volume equations for predictions of single stem volume over and under bark. Ultimately it was shown that these models can be used to precisely predict stem diameters and total stem volume for the population average as well as for specific trees of the black locust plantations in the study area.

Keywords Forest landscape restoration, Stem taper, Tree volume, Fixed- and mixed-effects modelling, Random coefficients, *Robinia pseudoacacia* L.

Black locust (*Robinia pseudoacacia*, L.) is a tree species native to North America. The medium-sized, deciduous, thorny, fast-growing species grows on a wide range of soil types and is one of the most important and widespread broadleaved alien tree species in Europe¹. The ability of black locust to perform well in various habitats, especially during early succession, is noted in the review of Cierjacks and others². The authors refer to the documentation of black locust invasion on abandoned gravel-sand pits and landfills, urban brownfield sites, secondary forests on farmland, coppice forests and pastures, along disturbed roadsides and on burned sites. Black locust was first introduced to Europe at the beginning of the seventeenth century³ and is now the second most often planted broadleaved tree species worldwide, after eucalyptus⁴. The resistance of black locust to heat and drought, its salt tolerance, and its symbiosis with nitrogen-fixing bacteria are particularly important under extreme environmental conditions and therefore for the rehabilitation of degraded areas^{5–7}. Because of these ecological and biological properties, black locust is often planted in the steppe zone of Ukraine to create forest plantations, rehabilitate forests and reclaim degraded soils^{8–11}. Black locust is widely used to obtain above-ground biomass in short rotation coppices^{12–16}. Moreover, it is a suitable tree species for the restoration of sites that have been significantly degraded through anthropogenic activities, such as former open-cast mining areas¹⁷. However, there is some

¹Forest Research Institute, Hellenic Agricultural Organization DIMITRA, 57006 Thessaloniki, Greece. ²Department of Forest Utilization and Timber Markets, Eberswalde University for Sustainable Development, 16225 Eberswalde, Germany. ³Institute of Forest Sciences, Warsaw University of Life Sciences-SGGW, 02-776 Warsaw, Poland. ⁴Department of Ecology, Ukrainian National Forestry University, Lviv 79057, Ukraine. ⁵Department of Forestry and Management of the Environment and Natural Resources, Democritus University of Thrace, 68200 Orestiada, Greece. ✉email: spyroglou@fri.gr

concern about the future management of black locust in Europe due to its possible ecological disadvantages related to invasiveness, threats to biodiversity and induced changes in microclimate and soil traits^{18–21}.

In Greece, the National Forest Service introduced and planted black locust to stabilize torrents on mountains and prevent soil erosion near rivers, roads and railway banks. It has also been used as fodder in silvopastures²² and as an alternative plantation crop for privately owned marginal agricultural lands, in line with the 2080/92 and 1257/99 European Union (EU) Regulations²³. In the 1980s, the Hellenic Public Power Corporation (HPPC) started to rehabilitate open-cast coal mining fields in the Lignite Center of Northwestern Greece by planting mainly black locust, both alone and in mixture with other species. The establishment of plantations is still ongoing, but already with a planted area of 2570 ha, the HPPC is the largest private forest plantation owner in Greece, owning 26% of the country's black locust plantation area²⁴. To date, there is no inventory or management plan for those plantations.

One of the key processes in forest management planning is the estimation of tree volume and, more specifically, merchantable volume²⁵. There is no accurate and direct method for calculating tree volume due to the unique and complex shape of the stem. Water displacement methods require destructively sampled trees and are both time consuming and expensive. Therefore, tree volume estimation is usually based on empirical—heuristic volume models. Although a tree stem cannot be completely described in mathematical terms, it is convenient to assume that segments of a tree bole approximate various geometric solids. The lower portion of the stem is generally assumed to be a neiloid frustum, the middle portion a paraboloid frustum and the upper portion a cone^{26,27}. The form factor approach is widespread in forest practice and is used to calculate volume²⁸. A limitation of this method is that while it can result in accurate estimates of whole stem volume, it is not useful for estimating tree section or log volumes. This problem can be overcome by describing the stem profile through the development of taper curves and equations. There are two common methodologies for obtaining taper and volume models^{29,30}. One is to develop volume ratio equations that predict merchantable volume as a percentage of total tree volume³¹; the other is to define an equation describing the stem taper.

Taper equations are flexible models that can provide estimates of (1) total stem volume, (2) merchantable volume to any top diameter and from any stump height diameter, (3) log volumes of any length at any height from the ground and (4) biomass³². Moreover, taper equations may be used to produce assortment tables based on log lengths or specific diameter thresholds^{25,33}. Closely related volume models can only estimate the total or merchantable volume of a tree. Tree taper is important for making accurate estimates of standing timber, especially in terms of product potential. Taper equations can be useful in predicting the volume of an individual tree or a stand of trees^{34,35}. Many empirical models have been used to describe tree profiles. Ormerod³⁶ cited Höjer's³⁷ logarithmic equation as the first attempt to model tree taper. Research has shown that stem profile could not be described with simple equations. The idea was that the new taper equations should produce the whole stem volume estimate as existing volume equations. Goulding and Murray³⁸ provided the so-called compatible taper equations, but these equations exhibited significant bias in certain regions of the stem. Grosenbauch³⁹ noted that tree profile is best modeled in sections and that the necessary conditions are that the sub-models produce identical diameters at joint points and that diameters decrease monotonically from the base of tree to the top. Max and Burkhart⁴⁰ used their segmented polynomial model and demonstrated the superiority of three sub-models over two. Thomas and Parresol⁴¹ applied flexible trigonometric taper equations for both coniferous and broadleaved tree forms.

Many forms and types of stem profile models have been published⁴² and evaluated for accuracy and precision^{43–45}. Martin⁴⁵ evaluated five profile models for predicting diameter, height, and volume and found that no single model performed best for all applications. He reported that the Max and Burkhart⁴⁰ model was best overall. Cao and others⁴⁴ evaluated six profile models for predicting diameter and ranked Max and Burkhart's segmented-profile polynomial model first and Cao's segmented-profile model second. Cao's segmented-profile model was ranked first for estimating volumes to various top diameters. A polynomial ratio model was ranked second, Burkhart's³¹ ratio model was third, and Max and Burkhart's segmented-profile model was fourth. The Alberta Forest Service⁴³ evaluated 15 stem-profile models and judged Max and Burkhart's model best because of its accuracy and ease of computation.

While the number of taper equations now available has grown significantly, they can generally be classified into two categories: parametric and non-parametric equations⁴⁶. Parametric approaches include polynomial, trigonometric and variable-exponent equations. Segmented polynomial equations, e.g. Max and Burkhart 1976⁴⁰, contain conditional clauses with different parts of the function applying to the taper depending on relative height along the stem. Elaborate functions with a large number of independent variables and coefficients to be estimated, especially variable-exponent equations^{33,47} are often found to perform better than their less complex counterparts⁴⁸.

The empirical data intended to create taper models are characterized by a hierarchical structure containing information about individual diameters, trees or sites. One possibility of modeling these data in the case of parametric equations is the mixed-effects modelling approach^{49–51}. The fixed component defines the level for a typical basic group (e.g., trees or sites), whereas the random part describes the difference between each group and a typical value. This approach allows both the capture of all dependencies occurring in the analyzed data and the assessment of variability occurring at various levels, as well as the use of the obtained information in the prediction of the modeled parameter (usually the diameter along the tree⁴⁸). On the other hand, machine-learning methods, artificial neural networks especially, have been explored as viable non-parametric approaches to modelling tree taper^{52–54}. Calibrated taper models may in turn be applied to predict the timber assortment to be harvested from each stem.

Because there are no models available for calculating the wood volume of black locust stands and no taper functions to estimate the merchantable volume at any stem diameter along the trunk for Greek conditions, the aims of this study were to (i) fit taper equations for black locust from the restored area of the Lignite Center of

Northwestern Greece, (ii) compare the taper equations with other published taper models for planted black locust in restored coal mines, and (iii) fit volume equations for tree volume using the best performing taper models.

Materials and methods

Site description

In this study, we inventoried the black locust restoration plantations on the former open-cast mining areas of the lignite center of Northwestern Greece (Fig. 1). The plantations are located near Amyntaio (40.56° to 40.61° N and 21.62° to 21.69° E) and Ptolemaida (40.39° to 40.51° N and 21.7° to 21.89° E). Black locust covers more than 95% of the planted area, followed by weaver's broom (*Spartium junceum* L) and Arizona cypress (*Cupressus arizonica* Greene), covering 2.45% and 1.44%, respectively. Other planted species include oaks, maples, pines and various deciduous broadleaves but these make up only very small percentages of the total planted area. The plantations are established on open-cast mining deposits with varying topography (moderate to steep slopes, plains and terraces) and altitudes ranging from 530 to 950 m. a.s.l. The landscape is fragmented by different land uses, which, in addition to forest plantations, also include grasslands, agricultural lands and bare lands used for photovoltaic parks and recycling facilities. The air temperature of the region ranges from 6.1 to 17.4 °C with a mean annual temperature of 12.2 °C. The total annual precipitation is 664 mm (mean values over the last 50 years).

Sampling and measurements

Sampling took place in August–September 2019 and July–September 2020. In total, 30 trees representing the locally present diameter range (3–22 cm) were destructively sampled. Each tree was then cut into 6 sections using a chainsaw or branch cutter, depending on diameter. The sections are located at (1) stump height = 0.3 m, (2) breast height = 1.3 m, (3) mid-bole height, which was defined as the middle point between the ground and crown base height, (4) live crown base height, (5) lower third of the crown, and (6) upper third of the crown. The crown base height was defined as the lowest knot of a live branch.

The stem and branch section discs were sanded and subsequently scanned with an Expression 11000XL scanner (Seiko Epson, Japan). Tree ring increment and section diameters were digitally recorded using WinDENDRO 6.0.4 software⁵⁵ and analyzed via the associated XLStem add-in⁵⁶ for Microsoft Excel⁵⁷.

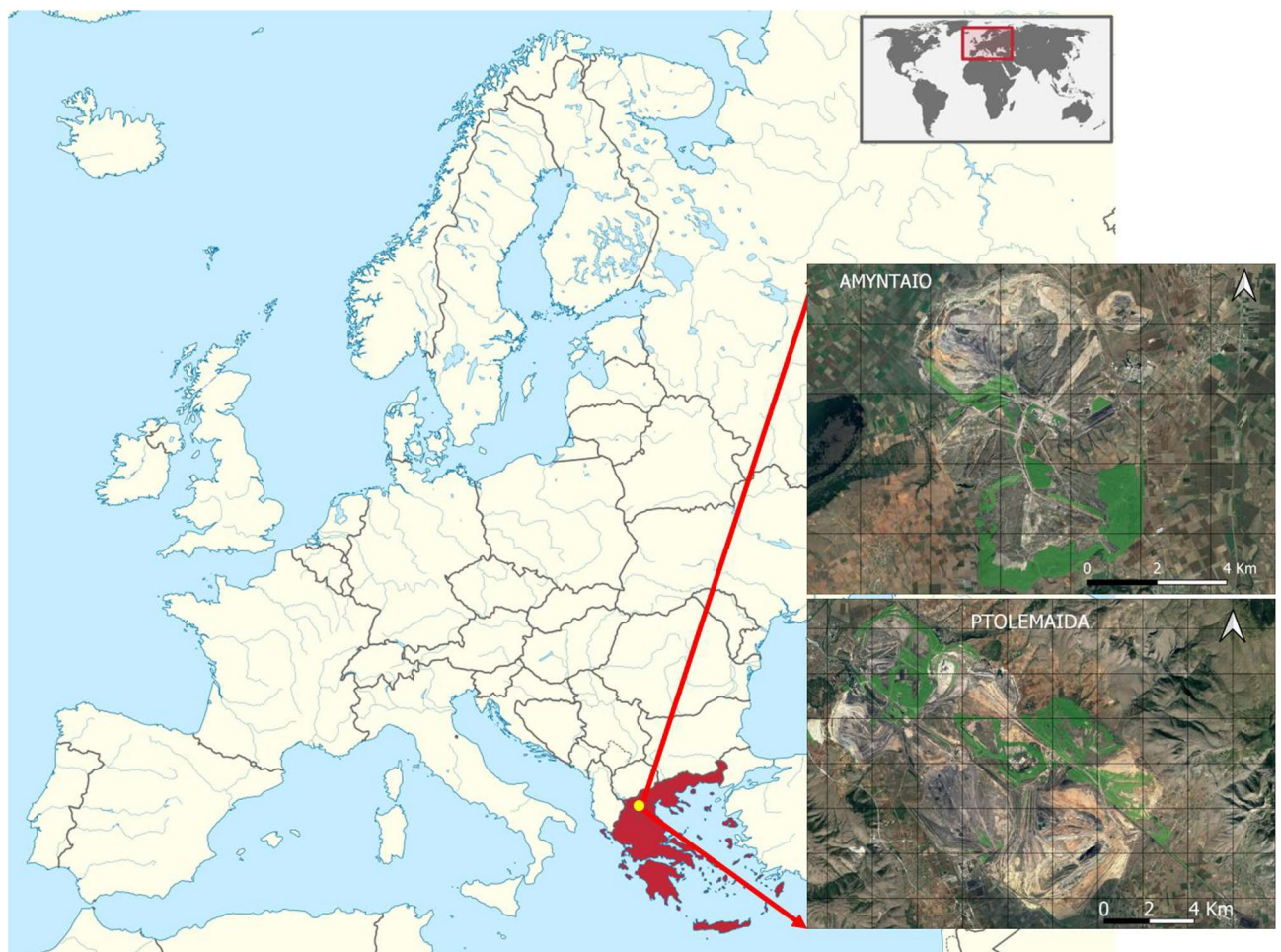


Figure 1. Map of the Amyntaio and Ptolemaida open-cast mine sites. The green colored polygons represent the restored areas planted with black locust.

On each disc, increment paths were identified in four perpendicular directions. For the sections between ground and crown base (sections 1–4), one disc was identified and digitally analyzed. For the crown sections (sections 5 and 6), every branch dissecting the respective imaginary line was cut and a disc was sampled. The three discs with the largest diameters per crown section were subsequently used for further analysis in WinDENDRO⁵⁵. In cases where discs were damaged due to stripped bark or being rotten, the digital measurements were conducted on the next intact disc. The reference stem volume [m³] was calculated via Smalian's equation by the XLStem analysis output.

Taper equations for fixed-effect models

Since the nature of taper datasets violates the assumption of homogenous variance, modelling was not conducted via the basic nonlinear least squares (*nls*) approach. Instead, using the generalized nonlinear least squares function *gnls* of the nlme package⁵⁸, a power type function was applied to model residual variance, as reported by Bronisz and Zasada⁴⁹. Additionally, an autoregressive structure of order 1 was applied on a stem basis.

Some taper equations use only a single explanatory variable such as the segmented model⁴⁰ or the trigonometric model⁴¹, while others use up to four explanatory variables^{59,60}. The selection criteria for taper equations were to include this spectrum of simple to complex functions, while also focusing on those that were shown in the literature to perform well.

In order to predict the same dependent variable 'section diameter squared', all functions were modified according to Bruce et al.⁶¹ and de-Miguel et al.⁴⁸.

Nomenclature:

D	section diameter [cm], over or under bark
DBH	diameter o.b. [cm] at breast height
H	total tree height [m]
HD	section height [m]
RH	$\frac{HD}{H}$ relative height: section height [m] / total tree height [m]
$\beta_1 \dots \beta_{10}$	fixed-effect parameters to be calibrated

Taper equations:

Bi and Long⁶²

$$D^2 \sim \left\{ \left[\frac{\ln \left(\sin \left(\frac{\pi}{2} * RH \right) \right)}{\ln \left(\sin \left(\frac{\pi}{2} * t \right) \right)} \right]^{\beta_1 + \beta_2 * \sin \left(\frac{\pi}{2} * RH \right) + \beta_3 * \cos \left(\frac{3\pi}{2} * RH \right) + \beta_4 * \frac{\sin \left(\frac{\pi}{2} * RH \right)}{RH} + \beta_5 * D + \beta_6 * RH * \sqrt{D} + \beta_7 * RH * \sqrt{H}} \right\}^2 * DBH \quad (1)$$

where $t = \frac{1.3m}{H}$, all other variables are described above.

Biging⁶³

$$D^2 \sim \left\{ \left[\beta_1 + \beta_2 * \ln \left(1 - RH^{\frac{1}{3}} \right) * \left(1 - e^{\left(-\frac{\beta_1}{\beta_2} \right)} \right) \right] * DBH \right\}^2 \quad (2)$$

Bruce et al.⁶¹

$$D^2 \sim \left\{ \left[\beta_1 * k^{0.1} + \beta_2 * [k^{1.5} - k^3] * DBH * 0.01 + \beta_3 * [k^{1.5} - k^3] * H * 0.001 + \beta_4 * [k^{1.5} - k^{32}] * HD * 0.0001 + \beta_5 * [k^{1.5} - k^{32}] * H^{0.5} * 0.001 + \beta_6 * [k^{1.5} - k^{40}] * H^2 * 0.00001 \right] * DBH \right\}^2 \quad (3)$$

Cao⁴⁴ as found in Parresol et al.⁶⁴

$$D^2 \sim \left[\sqrt{\frac{2 * z + \beta_1 * (3 * z^2 - 2 * z) + \beta_2 * (z - b_4)^2 * (z \geq b_5) + \beta_3 * (z - b_4)^2 * (z \geq \beta_5)}{\beta_3 * (z - b_4)^2 * (z \geq \beta_5)}} * DBH \right]^2 \quad (4)$$

where $z = \frac{H - HD}{H}$, all other variables are described above

Demaerschalk⁶⁵

$$D^2 \sim [\beta_1 * DBH^{\beta_2} * (H - HD)^{\beta_3} * H^{\beta_4}]^2 \quad (5)$$

Kozak et al.⁶⁶

$$D^2 \sim \left\{ [\beta_1 * (RH - 1) + \beta_2 * (RH^2 - 1)] * DBH \right\}^2 \quad (6)$$

Kozak⁵⁹

$$D^2 = \left[\beta_1 (DBH^{\beta_2}) + (H^{\beta_3}) + \left(\frac{1 - \sqrt{RH}}{1 - \sqrt{t}} \right)^{\beta_4^{0.1}} + \beta_5 (RH^4) + \beta_6 * \arcsin(1 - \sqrt{RH}) + \beta_7 * \left(\frac{1}{e^{DBH/H}} \right) + \beta_8 * DBH \left(\frac{1 - \sqrt{RH}}{1 - \sqrt{t}} \right) \right]^2 \quad (7)$$

where $t = \frac{1.3m}{H}$, all other variables are described above
Kozak³³

$$D^2 \sim \left[\beta_1 * DBH^{\beta_2} * H^{\beta_3} * C^{\beta_4 * H^4 + \beta_5 * \left(1 - e^{\frac{DBH}{H}} \right) + \beta_6 * C^{0.1} + \beta_7 * \frac{1}{DBH} + \beta_8 * H^Q + \beta_9 * C} \right]^2 \quad (8)$$

where $C = \frac{1 - \left(\frac{HD}{H} \right)^{\frac{1}{3}}}{1 - \left(\frac{1.3}{H} \right)^{\frac{1}{3}}}$, $Q = 1 - \left(\frac{HD}{H} \right)^{\frac{1}{3}}$, all other variables are described above
Laasasenaho⁶⁷ as found in de-Miguel et al.⁴⁸

$$D^2 \sim \{ [\beta_1 * RH + \beta_2 * RH^2 + \beta_3 * RH^8 + \beta_4 * RH^{34}] * DBH \}^2 \quad (9)$$

Lee et al.⁶⁸

$$D^2 \sim \beta_1 * DBH^{\beta_2} * \left[(1 - RH)^{(\beta_3 * RH^2 + \beta_4 * RH + \beta_5)} \right]^2 \quad (10)$$

Max and Burkhardt⁴⁰

$$D^2 \sim [\beta_1 (RH - 1) + \beta_2 (RH^2 - 1) - \beta_3 (\alpha_1 - RH)^2 * I_{\beta_5 - RH} + \beta_4 (\alpha_2 - RH)^2 * I_{\beta_6 - RH}] * DBH^2 \quad (11)$$

where: $I_{\beta_5 - RH} = 1, \beta_5 - RH \geq 0; I_{\beta_5 - RH} = 0, \beta_5 - RH < 0; I_{\beta_6 - RH} = 1, \beta_6 - RH \geq 0; I_{\beta_6 - RH} = 0, \beta_6 - RH < 0$
Muhairwe⁴⁷

$$D^2 \sim \beta_1 * DBH^{\beta_2} * \left[(1 - RH)^{\beta_3 * RH + \beta_4 * RH^2 + \frac{\beta_5}{RH} + \beta_6 * RH^3 + \beta_7 * DBH + \beta_8 * \left(\frac{DBH}{H} \right)} \right]^2 \quad (12)$$

Newberry and Burkhardt⁶⁹

$$D^2 \sim \left[\beta_1 * DBH * \left(\frac{H - HD}{H - 1.3} \right)^{\beta_2} \right]^2 \quad (13)$$

Riemer et al.⁷⁰

$$D^2 \sim \left\{ \frac{\beta_1 * DBH}{1 - e^{\beta_3 * (1.3 - H)}} + \left(\frac{DBH}{2} - \beta_1 * DBH \right) * \left[1 - \frac{1}{1 - e^{\beta_2 * (1.3 - H)}} \right] + e^{-\beta_2 * H} * \left[\frac{\left(\frac{DBH}{2} - \beta_1 * DBH \right) * e^{1.3 * \beta_2}}{1 - e^{\beta_2 * (1.3 - H)}} \right] - e^{\beta_3 * H} * \left[\frac{\beta_1 * DBH * e^{-\beta_3 * H}}{1 - e^{\beta_3 * (1.3 - H)}} \right] \right\}^2 \quad (14)$$

Sharma and Oderwald⁷¹

$$D^2 \sim DBH^2 * \frac{HD^{2 - \beta_1}}{1.3} * \frac{H - HD}{H - 1.3} \quad (15)$$

Sharma and Zhang⁷²

$$D^2 \sim \left\{ DBH * \beta_1 * \left(\frac{HD}{1.3} \right)^{[2 - (\beta_2 + \beta_3 * RH + \beta_4 * RH^2) * \left(\frac{H - HD}{H - 1.3} \right)^{0.5}]} \right\}^2 \quad (16)$$

Sharma and Parton⁷³

$$D^2 \sim \left\{ DBH * \beta_1 * \left(\frac{HD}{1.3} \right)^{[2 - (\beta_2 + \beta_3 * RH + \beta_4 * RH^2) * \left(\frac{H - HD}{H - 1.3} \right)^{0.5}]} \right\}^2 \quad (17)$$

Thomas and Parressol⁴¹

$$D^2 = \left[\beta_1(RH - 1) + \beta_2 \sin(2 * \pi * RH) + \beta_3 \cot \arctan\left(\pi * \frac{RH}{2}\right) \right] * DBH^2 \quad (18)$$

Westfall and Scott⁶⁰

$$D^2 = DBH^2 * \left(\frac{1.3}{\beta_5} \right)^{\beta_{10}} * \left(\frac{1 - \frac{HD}{H}}{1 - \beta_5} \right)^{(\varphi_{1j} + \beta_{1j}) + \left[\frac{\varphi_{2j} + \beta_7}{\left(1 + \frac{HD/H}{\beta_8}\right)^{\beta_9}} \right]} * \left(\frac{1 - \frac{HD}{H}}{1 - \beta_6} \right)^{\beta_2 * \frac{\left(\frac{HD}{H}\right)^{\beta_4} * \frac{DBH}{H}}{1 + \left(\frac{HD}{H}\right)^{\beta_4} * \frac{DBH}{H}}} \quad (19)$$

where $\varphi_{1j}, \varphi_{2j}$ —random effects for tree j .

Comparison with studies on Black locust taper

The predicted values from the present Greek dataset were compared to two sets of calibrated coefficients of closely related publications: the Western Polish model for Kozak's 1995 taper equation⁵⁹ (Eq. 7) for 48 black locust trees⁴⁹ and the North-Eastern US taper equation (Eq. 19) calibrated with a tree species group including black locust⁶⁰. Both models were analyzed alongside the fixed-effect models for diameters over bark. The lack of fit compared to the present model based on Greek data was addressed via residual plots and correlation testing with a linear model.

Mixed-effect taper models

As a result of the fixed-effect (FE) model ranking, the best taper model was selected to be fitted with random effects. Mixed-effect (ME) modelling was conducted using the same application of the power type function to model the residual variance as previously described for FE. Random effects were assigned at the tree grouping-level to improve the individual tree fit along the stem, thus accounting for possible differences in taper behaviour e.g. by age. To include all possible combinations of random effects for every coefficient, a grid search with 1, 2 and 3 random effects was performed. ME model performance was ranked using the same selection of goodness-of-fit metrics as the FE models. In the third step, a likelihood-ratio test was applied to confirm a statistically significant difference of coefficients between the best performing ME model and its FE counterpart. Finally, possible decreases in heteroscedasticity through the addition of random effects to the FE model were assessed via residual analysis. Residuals were plotted against the relative height RH along the stem and section diameter D_{ob} and D_{ub} .

Stem volume equations

The reference stem volume [m^3] was calculated using Smalian's equation applied to the measured section diameters. The predicted stem volume for the FE models was derived using numerical integration provided by the *integrate()*-function in R. Three volume equations from the literature, previously tested for black locust as a function of DBH and H,⁴⁹ were used for comparison without refitting their coefficients. In contrast to taper equations being fitted to repeated measurements (i.e., section diameters along the stem), the three applied volume functions (Eqs. 20, 21 and 22) require only total stem volume in the fitting procedure, eliminating the autocorrelation. The argument of simplicity prevails, as the volume equation terms do not include conditional clauses (Eqs. 4, 11) or complex exponents (Eqs. 7, 19).

Smalian's equation

$$V = \sum_{i=1}^{n-1} \frac{g_i + g_{i+1}}{2} * l + V_n$$

where g_i is the cross-sectional area of the i -th section, l is the section length and V_n is the cone at the tip of the tree.

Formulae for stem volume equations

Sopp and Kolosz⁷⁴

$$V_{stem} = 10^{-8} * DBH^2 * H * \left[\frac{H}{(H - 1.3)} \right]^2 * (0.6326 * DBH * H + 20.23 * DBH + 3034) \quad (20)$$

Moshki and Lamersdorf⁷⁵

$$V_{stem} = 0.4 * H * DBH^2 \quad (21)$$

Lockow and Lockow⁷⁶

$$V_{stem} = 4.848577837 * 10^{-5} * DBH^{1.946545395} * H^{0.96112739} \quad (22)$$

Statistical analysis

All statistical computations were executed using the R programming language, version 4.0.3⁷⁷. The R-packages used for statistical analysis were *minpackLM*⁷⁸ and *nlme*⁵⁸. Graphs were created using *ggplot2*⁷⁹. A significance level of $\alpha = 5\%$ was applied to confirm the statistical significance of model coefficients. Five goodness-of-fit metrics were used to evaluate model performance. While it has been argued that Root Mean Square Error (RMSE) (Eq. 23) is a worse assessor of average errors than Mean Absolute Error (MAE) (Eq. 24)⁸⁰, it is widely used in

performance rankings of taper model calibrations and is included here for comparability with other publications. Mean Bias (*MB*) (Eq. 25) and Mean Percent Bias (*MPB*) (Eq. 26) are both used to quantify a trend in the model's predicted values. *MPB* was chosen in addition to *MB* in order to be able to assess bias relative to the actual value. As taper data naturally covers a wide range of diameters, the same absolute error value is increasingly more influential with decreasing diameter. Lastly, Akaike's Information Criterion (*AIC*) (Eq. 27)⁸¹ measures the model fit by likelihood testing. *AIC* favors parameter parsimonious functions, i.e., the model performance receives a better score for fewer parameters being estimated. For *RMSE*, *MAE* and *AIC*, the lowest values are desired, whereas for *MB* and *MPB*, values closer to 0 indicate the best performance.

$$RMSE = \sqrt{\frac{\sum (y - \hat{y})^2}{n}} \quad (23)$$

$$MAE = \frac{\sum |y - \hat{y}|}{n} \quad (24)$$

$$MB = \frac{\sum (y - \hat{y})}{n} \quad (25)$$

$$MPB = 100 * \frac{\sum (y - \hat{y})}{\sum y} \quad (26)$$

$$AIC = 2k - 2 * \ln(\hat{L}) \quad (27)$$

where y = observed value, $\hat{y}$ = predicted value, n = sample size, k = number of estimated parameters, $\hat{L}$ = likelihood of the estimator

The residual variance of the fixed and mixed effect models was modelled using a power function of DBH (Eq. 28), where ε_{ij} is the residual of the individual section diameter squared, σ being the shape and ∂ being the scale parameter. The function was fitted on an individual tree basis, such that each individual sample tree received its own ∂ , which is reported as the mean $\pm$ standard deviation.

$$var(\varepsilon_{ij}) = \sigma^2 DBH_{ij}^{\partial} \quad (28)$$

A model's relative *Rank_i* (Eq. 28) by a certain goodness-of-fit metric (e.g. *RMSE*), i.e., between the best and worst performing model, was calculated using the ranking method of Poudel and Cao⁸²:

$$Rank_i = 1 + \frac{(n - 1) * (M_i - M_{min})}{M_{max} - M_{min}} \quad (29)$$

where

n —number of compared models.

M_i —respective evaluated value of the model

M_{min} , M_{max} —the lowest (mostly best) and highest (worst) overall values of a metric.

The final model performance was assessed by ranking the mean of each metric performance as *RankAVG*.

Data overview

The dendrochronological analysis showed that the age of the trees ranged between 4 and 24 years. Tree age was determined via ring count at the stump height (HST) section disc (0.3 m). Mean DBH was 10.91 ± 4.52 cm and mean tree height was 11.79 ± 3.28 m (Table 1). Due to the overall small stem dimensions, the volumes of the observed trees were small with a maximum of 0.24 m³. Stem volume also displayed the highest coefficient of variation with almost 100% around the mean.

The section-wise diameter distribution is displayed in Fig. 2. With the exception of one outlier, values recorded for crown diameters were less than 5 cm. There was a sharp absolute decrease in mean section diameter from crown base (DCB = 7.5 cm) to the lower section within the crown branches (D23 = 4.0 cm).

	Mean	±SD	CV%	Min	Max
DBH[cm]	10.91	4.52	41.43	3.20	21.93
H[m]	11.79	3.28	27.79	5.95	17.80
HCB[m]	6.64	2.57	38.69	1.70	11.90
V[m ³]	0.07	0.06	92.34	0.002	0.238

Table 1. Summary statistics for diameter at breast height = DBH, tree height = H, crown base height = HCB and stem volume = V ($n = 30$).

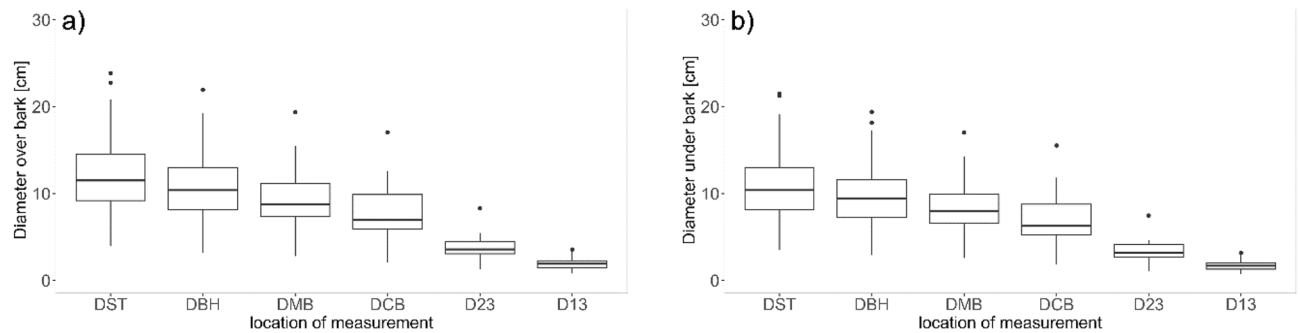


Figure 2. Distribution of over-bark (a) and under-bark (b) diameters at section heights with DST = diameter at stump height, DBH = diameter at breast height, DMB = diameter at mid-bole, DCB = diameter at base of live crown, D23 and D13 = diameter at lower and upper third of live crown.

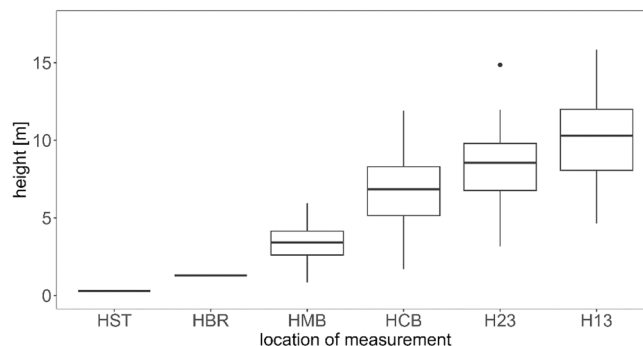


Figure 3. Distribution of each section height; HGR = stump height (0.3 m), HBR = breast height (1.3 m), HMB = mid-bole height, HCB = height at the base of live crown, H23 and H13 = Height of lower and upper third of live crown.

Figure 3 shows the total tree height as well as height distribution for each section. Total tree height H ranged from 6 to 18 m. Mid-bole height HMB was recorded lower than breast height for two small trees.

Ethical approval

The study complies with local and national guidelines and regulations.

Results

Fixed-effects (FE) taper models

Nineteen over-bark (o.b.) taper functions were fitted to model tree taper of black locust. Out of 19 over-bark gnls models, 15 converged and 11 had fully significant coefficient estimates. Table 2 summarizes the statistics for the trees selected for taper functions o.b. and Table 3 for the under-bark (u.b.) dataset. The model with the highest number of calibrated coefficients without non-significant terms was Eq. (11) for both o.b. and u.b. Equations 1, 9, 17 and 19 did not converge for either dataset using the gnls approach, Eq. (19) is therefore only included for the comparison in 3.2.

The subsequent ranking of the 18 models was performed for four error metrics calculated between the observed $D_{o.b.2}$ and the respective model predictions as well as AIC. Next, rankAVG was calculated as the horizontal mean of each individual rank. The top five equations by rankAVG are displayed in Table 4 (o.b. models) and Table 5 (u.b. models). Equation (8) performed best by lowest rankAVG for both the o.b. and u.b. datasets. Equation (10), Eq. (12) Eq. (7) and Eq. (11) were the next four best models. Generally, the first three models performed similarly, with only slight deviations in mean bias.

Overall, the u.b. models showed lower absolute errors, which is due to these diameter values being naturally lower than their o.b. equivalent. In terms of relative errors, regarding mean percent bias (MPB), the u.b. models performed similar to the o.b. models. All models that were biased positively for o.b. were also biased with the u.b. data and vice versa.

When analyzing model performance for relative height by RMSE of $D_{o.b.2}$, Eq. (8) reached the lowest error in four classes (Table 6), including the lowest segment at the stump. Four other models achieved the lowest RMSE in at least a single class. As bark volume per meter decreases to the tip of the stem, the lower stem regions should be prioritized for reducing errors. It is notable that RMSE decreases from at least 25 cm for measurements in the lowest tenth to less than 5 cm in the next, while increasing again to ca. 15 cm in the third.

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
Equation (2) Biging ⁶³	1.278 (0.009)	0.359 (0.008)							
Equation (3) Bruce ⁶¹	0.722 (0.050)	2.436 (6.907) n.s	−152.319 (71.534)	4163.554 (2585.434) n.s	9.696 (8.391) n.s	−222.634 (85.816)			
Equation (4) Cao ^{44,64}	1.18 (0.035)	−7.087 (1.299)	9.681 (1.384)	0.369 (0.027)	0.674 (0.030)				
Equation (5) Demaerschalk ⁶⁵	1.274 (0.059)	0.906 (0.024)	0.752 (0.019)	−0.713 (0.042)					
Equation (6) Kozak ⁶⁶	0.611 (0.050)	0.492 (0.045)							
Equation (7) Kozak ⁵⁹	0.865 (0.104)	0.836 (0.05)	0.187 (0.038)	0.311 (0.052)	−8.285 (1.125)	−0.804 (0.458) n.s	0.556 (0.471) n.s	0.101 (0.026)	
Equation (8) Kozak ³³	0.859 (0.038)	0.924 (0.018)	0.13 (0.032)	1.16 (0.138)	−1.135 (0.181)	0.577 (0.078)	1.708 (0.268)	0.021 (0.017) n.s	−0.024 (0.070) n.s
Equation (10) Lee ⁶⁸	1.471 (0.037)	0.91 (0.010)	4.953 (0.544)	−5.193 (0.538)	2.036 (0.134)				
Equation (11) Max ⁴⁰	−5.929 (0.758)	3.097 (0.450)	−7.062 (1.981)	9.065 (2.063)	0.602 (0.048)	0.346 (0.040)			
Equation (12) Muhairwe ⁴⁷	19.152 (9.836) n.s	0.91 (0.013)	−2.188 (1.514) n.s	−2.818 (3.284) n.s	2.59 (0.499)	5.648 (2.429)	−0.015 (0.006)	0.476 (0.143)	
Equation (13) Newberry ⁶⁹	1.021 (0.004)	0.751 (0.020)							
Equation (14) Riemer ⁷⁰	0.116 (0.047)	−0.333 (0.166)	−1.595 (0.319)						
Equation (15) Sharma ⁷¹	2.121 (0.009)								
Equation (16) Sharma ⁷²	0.976 (0.014)	2.138 (0.017)	−0.907 (0.187)	1.506 (0.255)					
Equation (18) Thomas ⁴¹	−6.429 (0.299)	1.344 (0.241)	0.015 (0.001)						

Table 2. Coefficient estimates of the converged fixed effect taper models, dependent variable: diameter over bark squared D_{ob}^2 ; values in parentheses represent standard errors of the coefficients; n.s. marks non-significant coefficients at $p > 0.05$; values rounded to three decimals.

As for D_{ub}^2 , Eq. (12) was the only equation to achieve the lowest RMSE in three classes (Table 7), with Eq. (8) only achieving two.

Comparison of over-bark fixed effect models with foreign coefficients

The first comparison between a foreign and the current Greek (GR) over-bark fixed effect model was conducted with a west-Polish model (PL) of Eq. (7)⁴⁹. While the fitting of Eq. (7) for the Greek sample yielded two non-significant coefficients β_6 and β_7 , the sample from Poland showed a different selection with β_3 , β_6 and β_7 being non-significant (Table 8).

Secondly, Eq. (19) was fitted in the North-Eastern United States⁶⁰. In their original study, FE coefficients for D_{ob}^2 were provided for a sample of “species group 18” trees (Table 9). No coefficient was reported as non-significant. However, in the present study, the Greek sample produced non-significant estimates for β_1 , β_2 and β_{10} using the o.b. dataset. Furthermore, the addition of both random effects φ_1 and φ_2 , as they are included in the equation of the original study, did not lead to a converging model. The Greek model fit is therefore a FE model rather than a mixed model.

Figure 4a shows the comparison of actual section diameter with both locally fitted coefficients (GR) and coefficients from the Polish model (PL). RMSE was calculated between the GR model predictions and the observed D , as well as with the foreign model's predictions. While a visual inspection only showed slight differences in deviation from the equality line for PL points, RMSE was about 13% higher compared to the GR model fit (GR RMSE = 1.234 cm vs PL RMSE = 1.387 cm).

Observing the values predicted by the North-Eastern US model coefficients however, errors were much higher than for the Greek model (Fig. 4b). The increase in RMSE amounted to 218% over the Greek *nlsLM()* function for fitting nonlinear least squares models equivalent (GR RMSE = 0.912 cm and US RMSE = 1.991 cm).

Mixed-effect (ME) taper models

Kozak's 2004 Eq. (8)³³ proved to be the best performing FE model for both the o.b. and u.b. dataset. In the subsequent grid search for the best allocation of 1, 2 or 3 random effects, the o.b. model assigned with random effects for parameters β_4 and β_8 reached the best goodness-of-fit out of 70 created models and β_6 and β_8 for u.b. (Table 10). There was a substantial increase in model accuracy by all four means of error calculation, while AIC also decreased by 31 for o.b. and 35 for u.b.

While the FE model featured the two non-significant coefficients β_8 and β_9 in both o.b. and u.b. calibrations, the o.b. and u.b. ME model also showed this behavior with the same two non-significant coefficients (Table 11).

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
Equation (2) Biging ⁶³	1.150 (0.008)	0.321 (0.007)							
Equation (3) Bruce ⁶¹	0.586 (0.040)	1.691 (5.640) n.s	− 116.489 (58.191)	2785.956 (2069.559) n.s	9.610 (6.767) n.s	− 195.002 (69.529)			
Equation (4) Cao ^{44,64}	1.230 (0.031)	− 4.829 (0.386)	7.894 (1.229)	0.274 (0.016)	0.762 (0.028)				
Equation (5) Demaerschalk ⁶⁵	1.054 (0.051)	0.890 (0.025)	0.747 (0.019)	− 0.658 (0.044)					
Equation (6) Kozak ⁶⁶	0.533 (0.043)	0.460 (0.038)							
Equation (7) Kozak ⁵⁹	0.596 (0.075)	0.896 (0.046)	0.223 (0.042)	0.304 (0.048)	− 7.872 (0.929)	− 0.219 (0.378) n.s	0.664 (0.387) n.s	0.074 (0.023)	
Equation (8) Kozak ³³	0.753 (0.037)	0.914 (0.019)	0.152 (0.034)	1.057 (0.128)	− 0.732 (0.182)	0.476 (0.078)	1.118 (0.303)	0.006 (0.018) n.s	0.029 (0.080) n.s
Equation (10) Lee ⁶⁸	1.259 (0.031)	0.928 (0.01)	4.770 (0.525)	− 4.918 (0.516)	1.942 (0.128)				
Equation (11) Max ⁴⁰	− 4.970 (0.587)	2.596 (0.347)	− 5.018 (0.910)	7.671 (1.106)	0.623 (0.041)	0.302 (0.029)			
Equation (12) Muhairwe ⁴⁷	26.486 (13.288)	0.930 (0.012)	− 4.485 (1.459)	3.094 (3.085) n.s	3.064 (0.487)	0.927 (2.224) n.s	− 0.018(0.006)	0.540 (0.143)	
Equation (13) Newberry ⁶⁹	0.924 (0.004)	0.734 (0.020)							
Equation (14) Riemer ⁷⁰	0.123 (0.003)	0.110 (0.016)	0.446 (0.594) n.s						
Equation (15) Sharma ⁷¹	2.053 (0.008)								
Equation (16) Sharma ⁷²	0.793 (0.011)	2.138 (0.016)	− 0.948 (0.177)	1.550 (0.243)					
Equation (18) Thomas ⁴¹	− 5.667 (0.275)	1.326 (0.221)	0.013 (0.001)						

Table 3. Coefficient estimates of the converged fixed effect taper models, dependent variable: diameter under bark squared $D_{u.b.}^2$; values in parentheses represent standard errors of the coefficients; ns marks non-significant coefficients at $p > 0.05$; values rounded to three decimals.

	RMSE	MAE	MB	MPB	AIC	
Equation (8) Kozak ³³	17.025	9.655	1.357	− 0.067	1381	
Equation (10) Lee ⁶⁸	17.392	9.623	1.844	− 0.017	1391	
Equation (12) Muhairwe ⁴⁷	17.779	9.700	3.771	0.077	1391	
Equation (7) Kozak ⁵⁹	19.990	11.404	− 0.033	− 0.185	1418	
Equation (11) Max ⁴⁰	18.822	11.127	− 2.033	− 0.094	1435	
	rankRMSE	rankMAE	rankMB	rankMPB	rankAIC	rankAVG
Equation (8) Kozak ³³	1.000	1.010	1.360	1.210	1.000	1.116
Equation (10) Lee ⁶⁸	1.080	1.000	1.490	1.000	1.330	1.180
Equation (12) Muhairwe ⁴⁷	1.160	1.030	2.010	1.250	1.330	1.356
Equation (7) Kozak ⁵⁹	1.620	1.580	1.000	1.710	2.230	1.628
Equation (11) Max ⁴⁰	1.370	1.490	1.540	1.330	2.800	1.706

Table 4. Goodness-of-fit statistics for the five best performing fixed effect models (diameter over bark squared $D_{o.b.}^2$). RMSE=Root Mean Square Error, MAE=Mean Absolute Error, MB=Mean Bias, MPB=Mean Percent Bias, AIC=Akaike’s Information Criterion.

The fixed coefficients of both ME models are displayed in Table 11. Both random effects β_4j and β_8j of the o.b. ME model showed a standard deviation of less than 0.02. The correlation between the two effects reached − 1.

The likelihood ratio test between the FE and ME models of Eq. (8) revealed a statistically significant difference for o.b. as well as for the u.b. models.

The over- and under-bark taper curves from the Eq. (8) mixed model predictions for three sample combinations of DBH, 15 cm, 20 cm and 25 cm at the same tree height $H = 20$ m displayed the largest differences at the center of the stem (Fig. 5 a). A similar behavior was observed for the taper curves at different tree heights of 15, 20 and 25 m but with a fixed DBH = 20 cm (Fig. 5b).

	RMSE	MAE	MB	MPB	AIC	
Equation (8) Kozak ³³	14.925	7.909	0.920	−0.132	1307	
Equation (10) Lee ⁶⁸	15.236	8.022	2.024	−0.006	1314	
Equation (12) Muhairwe ⁴⁷	15.292	7.968	2.874	0.018	1312	
Equation (7) Kozak ⁵⁹	16.812	9.202	−0.072	−0.168	1342	
Equation (11) Max ⁴⁰	16.207	9.137	−2.534	−0.133	1346	
	rankRMSE	rankMAE	rankMB	rankMPB	rankAIC	rankAVG
Equation (8) Kozak ³³	1.000	1.000	1.280	1.520	1.000	1.160
Equation (10) Lee ⁶⁸	1.030	1.030	1.640	1.000	1.230	1.186
Equation (12) Muhairwe ⁴⁷	1.030	1.020	1.910	1.050	1.160	1.234
Equation (7) Kozak ⁵⁹	1.160	1.330	1.010	1.660	2.140	1.460
Equation (11) Max ⁴⁰	1.110	1.320	1.800	1.520	2.270	1.604

Table 5. Goodness-of-fit statistics for the five best fixed effect models (diameter under bark squared $D_{u.b.}^2$). RMSE = Root Mean Square Error, MAE = Mean Absolute Error, MB = Mean Bias, MPB = Mean Percent Bias, AIC = Akaike's Information Criterion.

RH class	0 < RH ≤ 0.1	0.1 < RH ≤ 0.2	0.2 < RH ≤ 0.3	0.3 < RH ≤ 0.4	0.4 < RH ≤ 0.5	0.5 < RH ≤ 0.6	0.6 < RH ≤ 0.7	0.7 < RH ≤ 0.8	0.8 < RH ≤ 0.9	0.9 < RH ≤ 1	N lowest RMSE
N of meas	41	18	21	13	7	11	23	18	27	1	
Equation (4) Cao ^{44,64}	29.269	3.829	14.995	9.593	12.489	16.982	32.194	6.019	4.654	0.000	1
Equation (8) Kozak ³³	25.954	3.276	15.198	7.814	30.287	14.057	19.740	6.368	3.840	5.309	4
Equation (10) Lee ⁶⁸	26.217	2.887	13.249	10.961	14.564	15.181	26.129	5.814	4.320	6.020	3
Equation (11) Max ⁴⁰	28.071	3.660	15.095	10.247	10.796	16.243	29.563	5.920	4.666	5.756	1
Equation (12) Muhairwe ⁴⁷	28.583	3.854	15.671	9.757	25.861	10.068	20.064	6.424	5.770	6.347	1

Table 6. Comparison of Root Mean Square Error (RMSE) over 10 classes of relative height (RH) for FE models of diameter over bark squared ($D_{o.b.}^2$).

RH class	0 < RH ≤ 0.1	0.1 < RH ≤ 0.2	0.2 < RH ≤ 0.3	0.3 < RH ≤ 0.4	0.4 < RH ≤ 0.5	0.5 < RH ≤ 0.6	0.6 < RH ≤ 0.7	0.7 < RH ≤ 0.8	0.8 < RH ≤ 0.9	0.9 < RH ≤ 1	N lowest RMSE
N of meas	41	18	21	13	7	11	23	18	27	1	
Equation (4) Cao ^{44,64}	24.834	3.172	13.382	7.537	7.735	10.650	27.664	5.655	3.745	0.000	2
Equation (8) Kozak ³³	23.565	3.508	7.383	8.842	17.739	12.210	27.354	7.760	11.498	4.127	2
Equation (10) Lee ⁶⁸	24.362	2.794	8.226	7.039	21.553	11.931	18.512	5.583	3.239	3.530	2
Equation (11) Max ⁴⁰	24.872	2.604	7.699	9.980	12.197	12.844	20.917	5.079	3.642	4.343	1
Equation (12) Muhairwe ³³	26.197	3.689	8.150	7.889	21.435	9.903	16.603	5.049	3.814	4.450	3

Table 7. Comparison of Root Mean Square Error (RMSE) over 10 classes of relative height (RH) for FE models of diameter under bark squared ($D_{u.b.}^2$).

Stem volume

All three volume equations showed a tighter volume error spread compared to the numerically integrated FE models in the o.b. category (Fig. 6). In contrast, the taper models were less biased with medians closer to zero. Overall, stem volume errors ranged between −0.07 and 0.15 m³ o.b. and for u.b. from −0.07 to 0.07 m³. The Eq. (8) ME model surpassed all other equations for o.b. but not for u.b. where the FE model had fewer outliers and less bias. Calculating o.b. volume using the PL coefficients, a slightly wider spread was observed compared to the respective Greek equivalent. For the US coefficients, the errors were also higher compared to the Eq. (19) FE model. However, there were no u.b. coefficients available for the said equation.

The Kruskal–Wallis test indicated statistically significant differences in volume error depending on the equation used for prediction for the o.b. dataset, but not the u.b. errors. A subsequent Dunn-test revealed differences between the Eq. 8 ME model with only the Eq. (7) PL and Eq. (19) US errors. The volume errors from the FE models and volume equations were not found to significantly differ among each other.

Coefficient	Estimate	Coefficient	Estimate
β_1	0.962	β_5	-7.643
β_2	0.983	β_6	n.s
β_3	n.s	β_7	n.s
β_4	0.289	β_8	0.061

Table 8. Coefficients for Eq. (7) calibrated by Bronisz and Zasada⁴⁹ (fixed effects, over bark).

Coefficient	Estimate	SE	Coefficient	Estimate	SE
β_1	1.298	0.06	β_6	0.456	0.03
β_2	0.768	0.05	β_7	9.051	1.10
β_3	0.541	0.02	β_8	0.024	0.00
β_4	4.182	0.77	β_9	1.668	0.20
β_5	0.068	0.01	β_{10}	0.177	0.01

Table 9. Coefficients for Eq. (19) calibrated by Westfall and Scott⁶⁰ (fixed effects, over bark diameter squared).

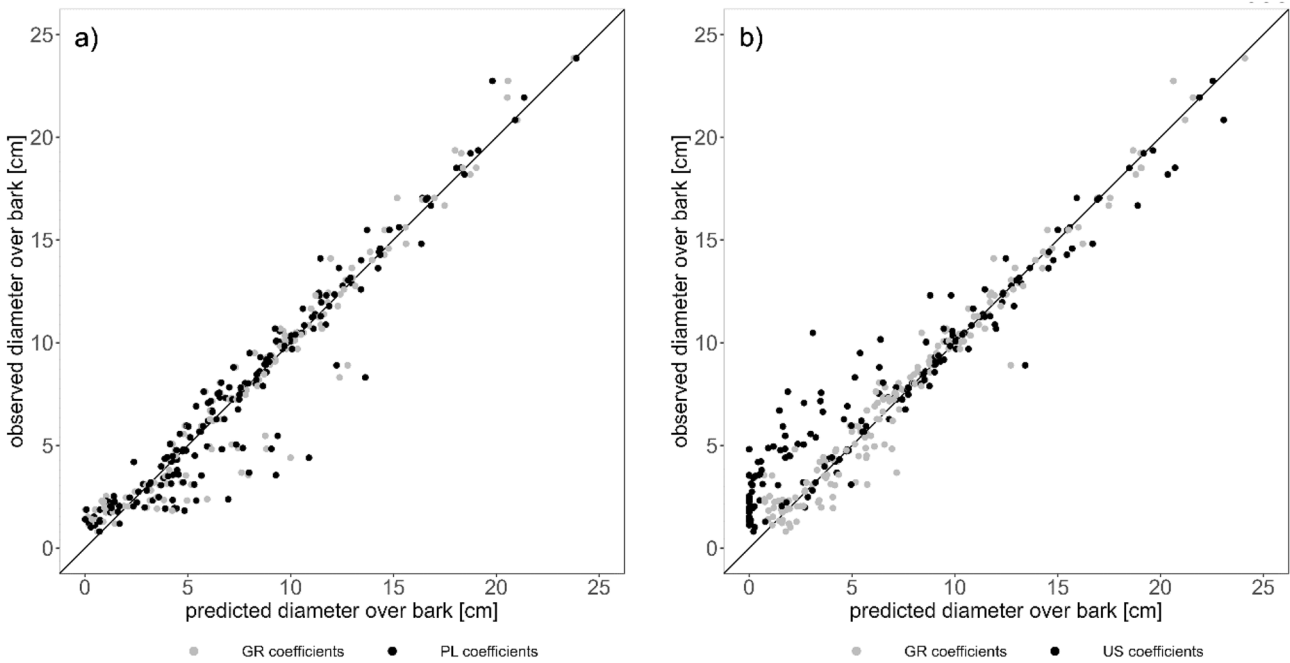


Figure 4. Actual section diameter (D o.b.) vs. GR and PL predictions of Eq. (7) (a) and vs. GR and US predictions of Eq. (19) (b).

Over bark	RMSE	MAE	MB	MPB	AIC
FE	17.025	9.655	1.357	-0.067	1381
ME	14.594	8.580	0.095	-0.111	1350
Under bark	RMSE	MAE	MB	MPB	AIC
FE	14.925	7.909	0.920	-0.132	1307
ME	10.260	5.828	-0.108	-0.075	1272

Table 10. Goodness-of-fit comparison for the Eq. (8) fixed-effect (FE) model and best mixed-effect (ME) model. RMSE=Root mean square error, MAE=Mean Absolute Error, MB=Mean Bias, MPB=Mean percent bias, AIC=Akaike’s information criterion.

Over bark		Under bark	
β1	0.911	β1	0.768
β2	0.966	β2	0.941
β3	0.068	β3	0.118
β4	1.029	β4	1.274
β5	− 1.225	β5	− 1.039
β6	0.618	β6	0.496
β7	1.910	β7	1.530
β8	0.031 n.s	β8	0.027 n.s
β9	− 0.069 n.s	β9	0.029 n.s
Variance power estimate δ mean ± standard deviation	3.361 ± 0.554		4.801 ± 0.839
Correlation structure order 1	− 0.334		− 0.209
Random effects structure			
β4j SD	0.018	β6j SD	0.100
β8j SD	0.019	β8j SD	0.048
β4j-β8j correlation	− 1.000	β6j-β8j correlation	− 1.000
Likelihood ratio test			
fixed effects AIC	1381.192	fixed effects AIC	1306.552
mixed effects AIC	1349.640	mixed effects AIC	1272.148
Likelihood ratio	37.552	Likelihood ratio	40.405
p-value	3.516*10 ^{−08}	p-value	8.745*10 ^{−09}

Table 11. Calibrated fixed effect coefficients for the best ME models of the over and under-bark dataset. Random effects structure of the best ME model; likelihood ratio test results comparing with the respective fixed effect model; Estimate of the power function to model the residual variance as mean $\pm$ standard deviation; Estimate of the first-order autoregressive structure;

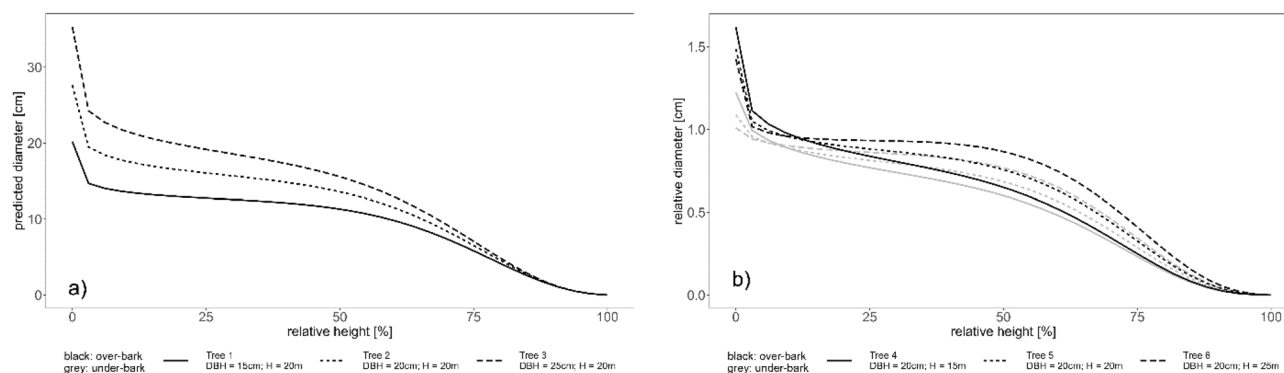


Figure 5. Sample taper curves for over and under-bark of relative diameter along relative height for the fitted Eq. 8. mixed effects model. (a) Varying diameter at breast height DBH = 15, 20, 25 cm and fixed tree height H = 20 m. (b) Varying H = 15, 20, 25 m and fixed DBH = 20 cm.

Discussion

To model both o.b. and u.b. taper of black locust, 18 different equations were analyzed. The models were ranked using a relative ranking of five separate goodness-of-fit metrics: root mean square error (RMSE), mean absolute error (MAE), mean bias (MB), mean percentage bias (MPB) and Akaike's Information Criterion (AIC). The variable exponent equation Eq. (8)³³ ranked best for both o.b. and u.b. datasets. Kozak's⁵⁹ variable exponent model Eq. (7), which had been identified as the best performing model for black locust in Bronisz and Zasada⁴⁹, was found to be the fourth-best model in this study. In their own study, Kozak³³ found Eq. 8 to be superior to Eq. (7) for 38 species groups in British Columbia, Canada. Rojo et al.⁸³ found Eq. (8) to be the best equation among 31 taper functions applied to Maritime pine (*Pinus pinaster* Ait.). Poudel et al.⁸⁴ also observed that Eq. (8) performed best for Douglas fir (*Pseudotsuga menziesii* Franco) compared to three variable-exponent equations.

Variable exponent equations of Eq. (12) (Muhairwe)⁴⁷ and Eq. (10) (Lee et al.)⁶⁸ were the second and third best models respectively, adding to the hypothesis that complex, i.e. non-parameter-parsimonious equations perform better than their less complex counterparts. This has been observed in multiple studies comparing several taper equations, including de-Miguel et al.⁴⁸ and Bronisz and Zasada⁴⁹. These equations aim to achieve good predictive ability for all parts of the stem, which was confirmed by measuring RMSE by 10 classes of relative height. Kozak's Eq. (8)³³ scored best four times for o.b. and twice for u.b., followed by further variable-exponent models.

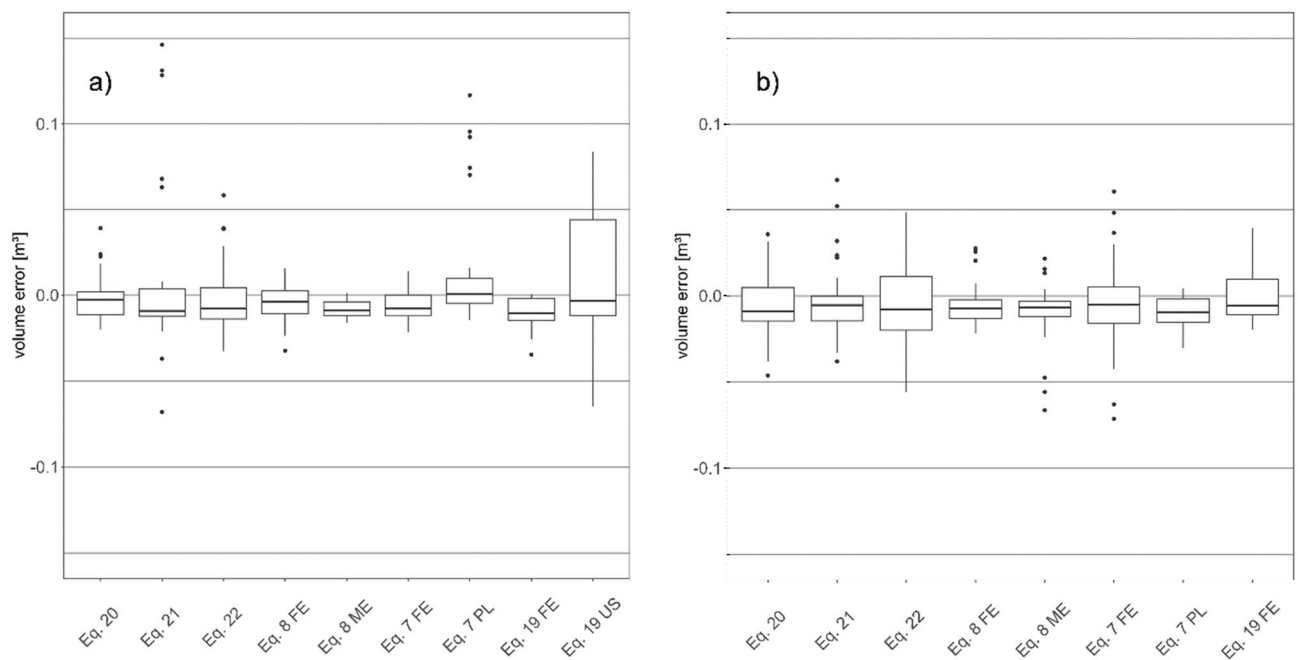


Figure 6. Stem volume error distribution [m^3] for three volume models and integrated Eq. 7, Eq. 8 compared to reference volume; left: over bark– right: under bark.

Recently published taper models for black locust from Poland⁴⁹ with Eq. (7), Kozak's⁵⁹ variable exponent model, and Eq. (19), the Westfall and Scott 2010 model⁶⁰, achieved error values of 13 and 218% of the Greek model fit. Therefore, neither taper model calibrated for foreign datasets proved suitable for predicting stem diameters of the present sample. As these two studies are among, if not the only, applications of taper modelling for black locust, further research is needed to provide adequate predictions for stands within Europe and North America. In addition, the tapering properties of planted black locust in post-mining areas might differ from black locust growing in forest sites due to different growing conditions on disturbed soils. Research showed that even among and within black locust families there is a large amount of genetic variation for growth, form, thorn length and other important traits.⁸⁵

An exhaustive search for all possible combinations of up to three random effects was performed on the basis of the best FE taper equation. Thus, the addition of two random effects to the generalized least-squares model of Eq. (8) provided a substantial decrease in errors for both o.b and u.b. datasets, as the random effects allowed for slight adjustment to the individual growth characteristics of each tree. Variance modelling, i.e., applying weights to the FE and ME models on a single-tree basis depending on a power-type function of DBH was included to account for heteroscedasticity.

In the direct comparison between the three stem volume functions and numerically integrated taper equations, the ME model was able to surpass the other methods slightly. However, as all three volume equations presented a negative bias, predictions could be adjusted. That said, given the small sample size of 30 stems in this study, bias should be investigated on a larger sample that may also expand the diameter and height range of these relatively young trees. Without destructively sampled allometric data and a locally calibrated model, the three volume equations from previous studies^{74–76} serve as a rather precise means of estimating o.b. and u.b. stem volume for the Greek plantation black locust trees. Therefore, diameter at breast height and total tree height would be the only required measurements of the respective sample trees. Under-bark diameter would then need to be sampled via a bark gauge along the stem to verify an additional model for stem wood volume.

Valentine and Gregoire⁸⁶ point out that the butt part of the stem is the most error-prone in taper modelling. In contrast, the number of measurements taken in this area is relatively low, with only stump height and breast height in most taper datasets. Westfall and Scott⁶⁰ also indicate that using few measurements in the lowest part of the stem might result in an underestimation of errors. This issue is present with this study especially, as the small dataset of just 180 measurements may only serve as a minimum sample. In turn, the primarily desired area of application in the Western Greece Lignite Center covers about 2500 ha. However, Yang and Burkhart⁸⁷ found that withholding certain classes of DBH for parametric fitting approaches, i.e. the smallest or largest diameter trees, did not effectively influence the predictive ability of taper functions Eq. (8) and Eq. (11) for the rest of their sample. Therefore, the present dataset of young trees used for calibration might be sufficient to predict larger tree diameters in the growing stands of the Western Greece Lignite Center. Suitable trees should be measured in 10 years to confirm the taper models applied for the 2020/2021 sample.

Nevertheless, with more sample trees from Greece and the broader Balkan region, a larger dataset could be obtained. With plans to phase out the use of fossil resources (especially lignite) for energy in Europe in line with international agreements to limit the effects of human-induced climate change, most open-cast mines will be depleted or abandoned within coming decades. Exotic species can be recommended for the rehabilitation of

former open cast bare coal mine spoils due to their fast growth and establishment⁸⁸. Black locust is an appropriate tree species for the regeneration of post-mining sites. However, the growth and taper of black locust on degraded sites are not well studied. Some studies have focussed on black locust as a suitable tree species for agroforestry systems^{14,85,89–92}. In short-rotation agroforestry systems, the harvested annual yield is between 8.1 and 9.7 t_{odt} ha⁻¹ year⁻¹^{114,90}. In such systems, the aboveground biomass is mostly used to produce bioenergy.

The cited research from several European countries, such as Germany, Greece and Ukraine, shows that black locust is indeed a suitable species for reforesting these disturbed sites. The production of woody biomass in unfavorable forest vegetation conditions of disturbed lands in combination with effective soil erosion protection is increasingly important to respond to the growing demand for bioenergy and conservation^{2,17,21,93,94}. Due to its environment-creating and productive functions, black locust is a very promising tree species in the fight against desertification and climate change as global challenges. However, as a non-native and invasive species, it should be used for land reclamation purposes in a clearly defined area of contaminated sites and for the production of durable and long-lived wood products if it is to be regarded as environmentally justified^{18,21}. A wider network of allometric datasets from reforestation areas, including taper and volume models, would be beneficial for more precise predictions of timber assortments, volume and carbon accounting for post-mining landscapes in the near future.

Conclusions

The following conclusions can be drawn from the present study at the black locust restoration plantations of the Lignite Centre of Northwestern Greece:

- From the eighteen different taper equations that were ranked, Kozak's³³ model performed best for both the over- and under-bark data followed by the Muhairwe's⁴⁷, Lee's et al.⁶⁸, Kozak's⁵⁹ and Max and Burkhardt's⁴⁰ equations.
- ME model regression substantially increased the over and under-bark taper model accuracy by all four means of error calculation, AIC also decreased by over 30 units and presented the largest decrease in residual variance for both responses D^2 and RH over the original FE models.
- A comparison for weighted fixed effects o.b. showed that Kozak's 1995 variable exponent model⁵⁹ exhibited two non-significant coefficients in contrast to three other non-significant coefficients in Bronisz and Zasada's Polish dataset⁴⁹.
- Stem volume equations proved to be performing well and close to the numerically integrated taper equations.

Data availability

The experimental data that support the findings of this study are available in the BonaRes repository site with the identifier: Wilms, F., Berendt, F., & Spyroglou, G. (2024). Taper dataset for planted Black locust (*Robinia pseudoacacia*, L.) in Greek post-mining areas [Data set]. Leibniz Centre for Agricultural Landscape Research (ZALF). <https://doi.org/10.4228/ZALF-KH4C-0B85>.

Received: 8 January 2024; Accepted: 24 May 2024

Published online: 12 June 2024

References

1. Sitzia, T., Cierjacks, A., de Rigo, D., Caudullo, G. *Robinia pseudoacacia* in Europe: distribution, habitat, usage and threats. In: *European Atlas of Forest Tree Species. Publ. Off.* (Eds San-Miguel-Ayán, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A.), e014e79+ (EU, Luxembourg, 2016).
2. Cierjacks, A. et al. Biological Flora of the British Isles: *Robinia pseudoacacia*. *J. Ecol.* **101**, 1623–1640. <https://doi.org/10.1111/1365-2745.12162> (2013).
3. Peabody, F. J. A 350-Year-Old American Legume in Paris. *Castanea*, **47**, 99–104 (1982).
4. Niclescu, V.-N. et al. Ecology, growth and management of black locust (*Robinia pseudoacacia* L.), a non-native species integrated into European forests. *J. For. Res.* **31**, 1081–1101. <https://doi.org/10.1007/s11676-020-01116-8> (2020).
5. Monk, R. W. & Wiebe, H. H. Salt tolerance & protoplasmic salt hardness of various woody & herbaceous ornamental plants. *Plant Physiol.* **36**, 478. <https://doi.org/10.1104/pp.36.4.478> (1961).
6. Rahmonov, O. The chemical composition of plant litter of black locust (*Robinia pseudoacacia* L.) and its ecological role in sandy ecosystems. *Acta Ecol. Sin.* **29**, 237–243. <https://doi.org/10.1016/j.chnaes.2009.08.006> (2009).
7. Umarov, M. M. Assotsiativnaya azotifikatsiya [Associative Nitrogen Fixation]. *Mosk. Gos. Univ* **132**, 22–23 (1986).
8. Masiuk, O. M., Kharytonov, M. M., Stankevich, S. A. & Stankevich, S. A. Remote and ground-based observations of land cover restoration after forest reclamation within a brown coal basin. *J. Geol. Geogr. Geocool.* **29**, 135–145 (2020).
9. Sytnyk, S. A. Моделювання морфометричних показників крони робінії несправжньоакації в умовах Північного Степу України. *Науковий вісник НЛТУ України* **28**, 34–37 (2018).
10. Sytnyk, S. A. Energy potential of the black locust stands within Ukraine's North Steppe. *Науковий вісник НЛТУ України* **27**, 79–82 (2017).
11. Zverkovsky, V. M. & Zubkova, O. S. The assessment sectional area of stem wood and stocks in experimental trees reclamation Western Donbass plot No 1. *Питання степового лісознавства та лісової рекультивації земель* **45**, 76–81 (2019).
12. Danilović, M., Stojnić, D., Vasiljević, V. & Gačić, D. Biomass from short rotation energy plantations of black locust on tailing dump of “Field B” open pit in “Kolubara” mining basin. *Nova mehanizacija šumarstva* **34**, 11–19 (2013).
13. Engel, J., Knoche, D. Die Robinie: Eine »kurzumtriebig« Baumart mit vielfältigen Nutzungsoptionen. *LWF Wissen* **84**, 67–75 (2020).
14. Grünwald, H. et al. *Robinia pseudoacacia* L.: a lesser known tree species for biomass production. *BioEnergy Res.* **2**, 123–133. <https://doi.org/10.1007/s12155-009-9038-x> (2009).
15. Malvolti, M. E. et al. Black locust (*Robinia pseudoacacia* L.) root cuttings: diversity and identity revealed by SSR genotyping: A case study. *SEEFOR* **6**, 201–217. <https://doi.org/10.15177/seefor.15-19> (2015).

16. Mantovani, D., Veste, M. & Freese, D. Biomass production and water use of Black Locust (*Robinia pseudoacacia* L.) for short-rotation plantation. *Geophysical Research Abstracts*, **14** (2012).
17. Mantovani, D., Veste, M., Böhm, C., Vignudelli, M. & Freese, D. Spatial and temporal variation of drought impact on black locust (*Robinia pseudoacacia* L.) water status and growth. *iFor. Biogeosci. For.* **8**, 743–747. <https://doi.org/10.3832/for1299-008> (2015).
18. Brundu, G. *et al.* Global guidelines for the sustainable use of non-native trees to prevent tree invasions and mitigate their negative impacts. *NeoBiota* **61**, 65–116. <https://doi.org/10.3897/neobiota.61.58380> (2020).
19. Keskin, T. & Makineci, E. Some soil properties on coal mine spoils reclaimed with black locust (*Robinia pseudoacacia* L.) and umbrella pine (*Pinus pinea* L.) in Agaçlı-Istanbul. *Environ. Monit. Assess.* **159**, 407–414. <https://doi.org/10.1007/s10661-008-0638-2> (2009).
20. Usuga, J. C. L., Toro, J. A. R., Alzate, M. V. R., de Jesús, Á. & Tapias, L. Estimation of biomass and carbon stocks in plants, soil and forest floor in different tropical forests. *For. Ecol. Manag.* **260**(10), 1906–1913. <https://doi.org/10.1016/j.foreco.2010.08.040> (2010).
21. Vítková, M., Müllerová, J., Sádlo, J., Pergl, J. & Pyšek, P. Black locust (*Robinia pseudoacacia*) beloved and despised: A story of an invasive tree in Central Europe. *For. Ecol. Manag.* **384**, 287–302. <https://doi.org/10.1016/j.foreco.2016.10.057> (2017).
22. Papachristou, T. G. *et al.* How the structure and form of vegetation in a black locust (*Robinia pseudoacacia* L.) silvopastoral system influences tree growth, forage mass and its nutrient content. *Agrofor. Syst.* **94**, 2317–2330. <https://doi.org/10.1007/s10457-020-00552-z> (2020).
23. Dini-Papanastasi, O. Contribution to the selection of productive progenies of *Robinia pseudoacacia* var. *monophylla* Carr. from young plantations in Northern Greece. *For. Genet.* **11**, 113–123 (2004).
24. YPEKA. *Annual Report of Forest Service Activities during 2011* (Special Secretariat of Forest Environment, 2014).
25. McTague, J. P. & Weiskittel, A. Evolution, history, and use of stem taper equations: A review of their development, application, and implementation. *Can. J. For. Res.* **51**, 210–235. <https://doi.org/10.1139/cjfr-2020-0326> (2021).
26. Burkhart, H. E. & Tomé, M. *Modeling Forest Trees and Stands* (Springer, 2012). <https://doi.org/10.1007/978-90-481-3170-9>.
27. Husch, B., Miller, C. I. & Beers, T. W. *Forest Mensuration* (Wiley, 1972).
28. Bruchwald, A. New empirical formulae for determination of volume of scots pine stands. *Folia Forestalia Polonica. Series A. For.* **38**, 5–10 (1996).
29. Byrne, J. C. & Reed, D. D. Complex compatible taper and volume estimation systems for red and loblolly pine. *For. Sci.* **32**, 423–443. <https://doi.org/10.1093/forestscience/32.2.423> (1986).
30. Rustagi, K. P. & Loveless, R. S. Jr. Compatible variable-form volume and stem-profile equations for Douglas-fir. *Can. J. For. Res.* **21**, 143–151. <https://doi.org/10.1139/x91-018> (1991).
31. Burkhart, H. E. Cubic-foot volume of loblolly pine to any merchantable top limit. *Southern J. Appl. For.* **1**, 7–9. <https://doi.org/10.1093/sjaf/1.2.7> (1977).
32. Calegario, N., Gregoire, T. G., Da Silva, T. A., Tomazello Filho, M. & Alves, J. A. Integrated system of equations for estimating stem volume, density, and biomass for Australian redcedar (*Toona ciliata*) plantations. *Can. J. For. Res.* **47**, 681–689. <https://doi.org/10.1139/cjfr-2016-0135> (2017).
33. Kozak, A. My last words on taper equations. *For. Chron.* **80**, 507–515. <https://doi.org/10.5558/tfc80507-4> (2004).
34. Cormier, K. L., Reich, R. M., Czaplewski, R. L. & Bechtold, W. A. Evaluation of weighted regression and sample size in developing a taper model for loblolly pine. *For. Ecol. Manag.* **53**, 65–76. [https://doi.org/10.1016/0378-1127\(92\)90034-7](https://doi.org/10.1016/0378-1127(92)90034-7) (1992).
35. Spellmann, H., Nagel, J. Auswertung des Nelder-Pflanzverbandversuches mit Kiefer im Forstamt Walsrode. (Evaluating the Nelder-Spacing Experiment with Scots Pine in Walsrode). *Niedersächsische Forstliche Versuchsanstalt*, S. 149–161 (1992).
36. Ormerod, D. W. The diameter-point method for tree taper description. *Can. J. For. Res.* **16**, 484–490. <https://doi.org/10.1139/x86-086> (1986).
37. Höjer, A. G. *Tallens Och Granens Tillväxt Bihang till Fr. Lovén Om Våra Barrskogar* (Department of Agriculture, Forest Service, Stockholm, 1903).
38. Goulding, C. J. & Murray, J. C. Polynomial taper equations that are compatible with tree volume equations. *N. Z. J. For. Sci.* **5**, 313–322 (1976).
39. Grosenbaugh, L. R. Tree form: Definition, interpolation, extrapolation. *For. Chron.* **42**, 444–457. <https://doi.org/10.5558/tfc42444-4> (1966).
40. Max, T. A. & Burkhart, H. E. Segmented polynomial regression applied to taper equations. *For. Sci.* **22**, 283–289. <https://doi.org/10.1093/forestscience/22.3.283> (1976).
41. Thomas, C. R. & Parressol, B. R. Simple, flexible, trigonometric taper equations. *Can. J. For. Res.* **21**, 1132–1137. <https://doi.org/10.1139/x91-157> (1991).
42. Sterba, H. Stem curves—a review of the literature. *Forestry Abstracts* **41**, 141–145 (1980).
43. Alberta Forest Service. *Evaluation and Fitting of Tree Taper Functions for Alberta* (Edmonton, 1987).
44. Cao, Q. V., Burkhart, H. E. & Max, T. A. Evaluation of two methods for cubic-volume prediction of loblolly pine to any merchantable limit. *For. Sci.* **26**, 71–80. <https://doi.org/10.1093/forestscience/26.1.71> (1980).
45. Martin, A. J. *Taper and Volume Equations for Selected Appalachian Hardwood Species* (U.S. Department of Agriculture, Forest Service, Berlin, 1981).
46. Salekin, S. *et al.* Global tree taper modelling: A review of applications, methods, functions, and their parameters. *Forests* **12**, 913. <https://doi.org/10.3390/f12070913> (2021).
47. Muhairwe, C. K. Taper equations for *Eucalyptus pilularis* and *Eucalyptus grandis* for the north coast in New South Wales, Australia. *For. Ecol. Manag.* **113**, 251–269. [https://doi.org/10.1016/S0378-1127\(98\)00431-9](https://doi.org/10.1016/S0378-1127(98)00431-9) (1999).
48. De-Miguel, S., Mehtätalo, L., Shater, Z., Kraid, B. & Pukkala, T. Evaluating marginal and conditional predictions of taper models in the absence of calibration data. *Can. J. For. Res.* **42**, 1383–1394. <https://doi.org/10.1139/x2012-090> (2012).
49. Bronisz, K. & Zasada, M. Taper models for black locust in west Poland. *Silva Fennica* <https://doi.org/10.14214/sf.10351> (2020).
50. Mehtätalo, L. & Lappi, J. *Forest Biometrics with Examples in R* (Chapman & Hall/CRC, 2020).
51. Pinheiro, J. & Bates, D. *Mixed-Effects Models in S and S-PLUS* (Springer, 2006).
52. Diamantopoulou, M. J. Predicting fir trees stem diameters using Artificial Neural Network models. *South. Afr. For. J.* **205**, 39–44. <https://doi.org/10.2989/10295920509505236> (2005).
53. Sakici, O. E. & Ozdemir, G. Stem taper estimations with Artificial Neural Networks for mixed Oriental Beech and Kazdağı Fir stands in Karabük Region, Turkey. *CERNE* **24**, 439–451 (2018).
54. Socha, J., Netzel, P. & Cywicka, D. Stem taper approximation by artificial neural network and a regression set models. *Forests* **11**, 79 (2020).
55. Regent Instruments Inc. WinDENDRO. (2000).
56. Regent Instruments Inc. XLStem. (2000).
57. Microsoft Corporation. Microsoft Excel 97. (1995).
58. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. (2021).
59. Kozak, A. *A Comparison of the 1994 and 1995 Taper Equations: Report to the Resources Inventory Branch* (B.C., Ministry of Forests, 1995).
60. Westfall, J. A. & Scott, C. T. Taper models for commercial tree species in the Northeastern United States. *For. Sci.* **56**, 515–528. <https://doi.org/10.1093/forestscience/56.6.515> (2010).

61. Bruce, D., Curtis, R. O. & VanCoeveering, C. *Volume and Taper Tables for Red Alder* (Dept. of Agriculture, Portland, Pacific Northwest Forest and Range Experiment Station, Forest Service, 1968).
62. Bi, H. & Long, Y. Flexible taper equation for site-specific management of *Pinus radiata* in New South Wales, Australia. *For. Ecol. Manag.* **148**, 79–91. [https://doi.org/10.1016/S0378-1127\(00\)00526-0](https://doi.org/10.1016/S0378-1127(00)00526-0) (2001).
63. Biging, G. S. Taper equations for second-growth mixed conifers of Northern California. *For. Sci.* **30**, 1103–1117. <https://doi.org/10.1093/forestscience/30.4.1103> (1984).
64. Parresol, B. R., Hotvedt, J. E. & Cao, Q. V. A volume and taper prediction system for bald cypress. *Can. J. For. Res.* **17**, 250–259. <https://doi.org/10.1139/x87-042> (1987).
65. Demaerschalk, J. P. Converting volume equations to compatible taper equations. *For. Sci.* **18**, 241–245. <https://doi.org/10.1093/forestscience/18.3.241> (1972).
66. Kozak, A., Munro, D. D. & Smith, J. H. G. Taper functions and their application in forest inventory. *For. Chron.* **45**, 278–283. <https://doi.org/10.5558/tfc45278-4> (1969).
67. Laasasenaho, J. Taper curve and volume functions for pine, spruce and birch [*Pinus sylvestris*, *Picea abies*, *Betula pendula*, *Betula pubescens*]. in (1982).
68. Lee, W.-K., Seo, J.-H., Son, Y.-M., Lee, K.-H. & von Gadow, K. Modeling stem profiles for *Pinus densiflora* in Korea. *For. Ecol. Manag.* **172**, 69–77. [https://doi.org/10.1016/S0378-1127\(02\)00139-1](https://doi.org/10.1016/S0378-1127(02)00139-1) (2003).
69. Newberry, J. D. & Burkhart, H. E. Variable-form stem profile models for loblolly pine. *Can. J. For. Res.* **16**, 109–114. <https://doi.org/10.1139/x86-018> (1986).
70. Riemer, T., von Gadow, K. & Sloboda, B. Ein Modell zur Beschreibung von Baumschäften. *Allgemeine Forst-und Jagdzeitung* **166**, 144–147 (1995).
71. Sharma, M. & Oderwald, R. G. Dimensionally compatible volume and taper equations. *Can. J. For. Res.* **31**, 797–803. <https://doi.org/10.1139/x01-005> (2001).
72. Sharma, M. & Zhang, S. Y. Variable-exponent taper equations for jack pine, black spruce, and balsam fir in eastern Canada. *For. Ecol. Manag.* **198**, 39–53. <https://doi.org/10.1016/j.foreco.2004.03.035> (2004).
73. Sharma, M. & Parton, J. Modeling stand density effects on taper for jack pine and black spruce plantations using dimensional analysis. *For. Sci.* **55**, 268–282. <https://doi.org/10.1093/forestscience/55.3.268> (2009).
74. Sopp, L., Kolozs, L. Volume tables (in Hungarian). Forest Service, Budapest, Hungary 58–66 (2000).
75. Moshki, A. & Lamersdorf, N. P. Growth and nutrient status of introduced black locust (*Robinia pseudoacacia* L.) afforestation in arid and semi arid areas of Iran. *Res. J. Environ. Sci.* **5**, 259–2686. <https://doi.org/10.3923/rjes.2011.259.268> (2011).
76. Lockow, K. W. & Lockow, J. Ertragstafel Robinie (*Robinia Pseudoacacia* L.), Gesellschaft zur Förderung schnellwachsender Baumarten in Norddeutschland e.V. (2015).
77. R Core Team. R: A Language and Environment for Statistical Computing. (2020).
78. Baty, F. et al. A toolbox for nonlinear regression in R: The package nlstools. *Journal of Statistical Software* **66**, 1–21. <https://doi.org/10.18637/jss.v066.i05> (2015).
79. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2016).
80. Willmott, C. & Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **30**, 79–82 (2005).
81. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**, 716–723. <https://doi.org/10.1109/TAC.1974.1100705> (1974).
82. Poudel, K. P. & Cao, Q. V. Evaluation of methods to predict weibull parameters for characterizing diameter distributions. *For. Sci.* **59**, 243–252. <https://doi.org/10.5849/forsci.12-001> (2013).
83. Rojo, A., Perales, X., Sánchez-Rodríguez, F., Álvarez-González, J. G. & von Gadow, K. Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (Northwestern Spain). *Eur. J. For. Res.* **124**, 177–186. <https://doi.org/10.1007/s10342-005-0066-6> (2005).
84. Poudel, K. P., Temesgen, H. & Gray, A. N. Estimating upper stem diameters and volume of Douglas-fir and Western hemlock trees in the Pacific northwest. *For. Ecosyst.* **5**, 1–12. <https://doi.org/10.1186/s40663-018-0134-2> (2018).
85. Hanover, J. W., Mebrathu, T. & Bloese, P. Genetic improvement of black locust: A prime agroforestry species. *For. Chron.* **67**, 227–231. <https://doi.org/10.5558/tfc67227-3> (1991).
86. Valentine, H. T. & Gregoire, T. G. A switching model of bole taper. *Can. J. For. Res.* **31**, 1400–1409. <https://doi.org/10.1139/x01-061> (2001).
87. Yang, S.-I. & Burkhart, H. E. Robustness of parametric and nonparametric fitting procedures of tree-stem taper with alternative definitions for validation data. *J. For.* **118**, 576–583. <https://doi.org/10.1093/jofore/fvaa036> (2020).
88. Dutta, R. K. & Agrawal, M. Restoration of opencast coal mine spoil by planting exotic tree species: A case study in dry tropical region. *Ecol. Eng.* **21**, 143–151. <https://doi.org/10.1016/j.ecoleng.2003.10.002> (2003).
89. Grünwald, H. et al. Agroforestry systems for the production of woody biomass for energy transformation purposes. *Ecol. Eng.* **29**, 319–328. <https://doi.org/10.1016/j.ecoleng.2006.09.012> (2007).
90. Huber, J. A. et al. Yield potential of tree species in organic and conventional short-rotation agroforestry systems in southern Germany. *BioEnergy Res.* **9**, 955–968. <https://doi.org/10.1007/s12155-016-9750-2> (2016).
91. Küppers, M. et al. Photosynthetic characteristics and simulation of annual leaf carbon gains of hybrid poplar (*Populus nigra* L. × *P. maximowiczii* Henry) and black locust (*Robinia pseudoacacia* L.) in a temperate agroforestry system. *Agrofor. Syst.* **92**, 1267–1286. <https://doi.org/10.1007/s10457-017-0071-z> (2018).
92. Ntayombya, P. & Gordon, A. M. Effects of black locust on productivity and nitrogen nutrition of intercropped barley. *Agrofor. Syst.* **29**, 239–254. <https://doi.org/10.1007/BF00704871> (1995).
93. DeGomez, T. & Wagner, M. R. Culture and use of black locust. *HortTechnology* **11**, 279–288 (2001).
94. Liu, Z., Hu, B., Bell, T. L., Flemetakis, E. & Rennenberg, H. Significance of mycorrhizal associations for the performance of N₂-fixing Black Locust (*Robinia pseudoacacia* L.). *Soil Biol. Biochem.* **145**, 107776. <https://doi.org/10.1016/j.soilbio.2020.107776> (2020).

Acknowledgements

This research was funded by Single RTDI state Aid Action Research–Create–Innovation with co-financing from Greece and the European Union (European Regional Development Fund) as part of the Operational Program Competitiveness, Entrepreneurship and Innovation (EPANEK) of the NSRF 2014–2020 (project COFORMIT—Contribution of tree plantations in the Northwest Greece lignite center to environmental protection and climate change mitigation T1EDK-02521). The authors would like to acknowledge the Hellenic Public Power Corporation (HPPC S.A.) for its substantial contribution of the necessary staff and machinery to carry out the fieldwork and destructive sampling data collection. Special thanks are due to Marina Tentsoglidou, Aris Azas, Christos Papadopoulos and their technical team. Special thanks are also due to Stamatis Tziaferidis, for his valuable help in inventory and tree logging. Special thanks are due to Dr. Dimitris Zianis for his guidance with the formulation

of mixed effect models and prediction procedures. We also thank the anonymous reviewers for their helpful comments and suggestions to improve this manuscript.

Author contributions

GS conceived and designed the study, GS, MF performed field and laboratory measurements. KR was responsible for funding acquisition and project administration. FW, GS organized the database and involved in first draft preparation. FW performed the statistical analyses and wrote the manuscript. FB, KB, UB provided comparative data from Poland reviewed the manuscript and contributed to the discussion section, MF, KR and GS reviewed and revised the manuscript. All authors contributed to the article and approved the submitted version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024