iForest
Biogeosciences and Forestry

Research Article
doi: 10.3832/ifor2114-010
vol. 10, pp. 567-575

Compatible taper-volume models of *Quercus variabilis* Blume forests in north China

Conghui Zheng [1-2], Yuzhong Wang [1], Liming Jia [2], Euan Gordon Mason [3], Songpo We [2], Caowen Sun [2], Jie Duan [2]

Compatible taper and volume models were created for *Quercus variabilis* Blume (cork oak) forests in North China. 174 trees were felled to obtain stem analysis data. Linear mixed effects analyses were used in modelling. Firstly, a bark thickness model was built. Then diameter at breast height over bark (DBH_{ob}) for the inner layers of the 174 trees could be calculated, based on which a total volume model was built. The estimated volume and a specific parameter restriction were then substituted into a polynomial taper model, finally the taper model was fitted and compatible taper and volume models were obtained. Four sets of models based on different data sets were separately built and compared through coefficients of determination (R^2), root mean square error (*RMSE*), value of Akaike's information criterion (*AIC*), residuals plots and histograms of residuals. Models based on data of the analyzed stems without ramicorns and simultaneously with relative diameter under 1.5 were chosen as the most precise. Further testing of the chosen models using the jackknife method for the bark thickness and total volume models and a validation data set for the taper model verified that those models can be used to predict bark thickness, diameter at a specific point along the stem, merchantable volume and total stem volume of cork oak forests in North China within specific tree diameter at breast height and height ranges.

Keywords: *Quercus variabilis* Blume, Dummy Variable, Box-Cox Transformation, Linear Mixed Effects Models, Compatible Taper-Volume Model

Introduction

Compatible taper-volume models are flexible tools for estimating total and merchantable tree volume that can meet the demands of market trends as product specifications change. A compatible taper-volume estimation system contains a taper equation and a total volume equation. The taper equation can provide estimations of diameter at a given height up a tree and merchantable tree volume (Diéguez-Aranda et al. 2006), and the total volume equation can easily estimate the total volume of a tree. Both models require diameter at breast height over bark (DBH_{ob}) and height as inputs. Compatible taper-volume estimation systems allow the volume computed by integration of the taper equation from the ground to the top of the tree to equal that calculated by a total volume equation.

Taper and volume estimation systems can be divided into two types: Type (1), the total volume model is directly derived through integration of the taper equation; Type (2), equation form of the total volume model is independent from the taper equation. For type (1), two methods can be used to estimate parameters of the two models: Method (1), firstly fit the taper equation, then the volume model with its parameters can be directly obtained by integration (Martin 1981); Method (2), after obtaining the total volume equation through integration of the taper equation, the two models are fitted simultaneously to get their parameters by seemingly unrelated regression (SUR) or full information maximum likelihood (FIML) procedures (Jiang et al. 2005, Brooks et al. 2008, Pompa-García et al. 2009, Ozçelik & Brooks 2012). For type (2), there are also two methods to estimate the parameters: Method (1), firstly estimate the total volume equation using the total volume observations, then substitute the estimated total volume from the volume model and a specific parameter restriction into a taper model, so that a compatible taper model can be estimated (Goulding & Murray 1976, Malimbwi & Philip 1989, Fang & Bailey 1999, Muhairwe 1999); Method (2), simultaneous estimation of parameters of taper and volume models using SUR or FIML (Diéguez-Aranda et al. 2006). Type (2) can provide an easily applied total volume model which can rapidly estimate tree volume (Diéguez-Aranda et al. 2006), so this type is often preferred. Method (1) for Type (2) was especially useful when simultaneous estimation caused difficulty in achieving convergence, while Method (2) for Type (2) could make a reasonable compromise among the components in the system in the process of minimizing the sum of square errors (Fang et al. 2000, Diéguez-Aranda et al. 2006). Method (2) for Type (2) is more difficult when equations in the system have different numbers of observations. In this case, weights may be needed (Fang & Bailey 1999, Diéguez-Aranda et al. 2006). Some researchers compared the

(1) Hebei Engineering and Technology Center of Forest Improved Variety, Hebei Academy of Forestry, Shijiazhuang 050000 (China); (2) Ministry of Education Key Laboratory of Silviculture and Conservation, Beijing Forestry University, Beijing 100083 (China); (3) School of Forestry, University of Canterbury, Christchurch 8140 (New Zealand)

@ Euan Gordon Mason (euan.mason@canterbury.ac.nz)

Received: May 16, 2016 - Accepted: Feb 22, 2017

Citation: Zheng C, Wang Y, Jia L, Mason EG, We S, Sun C, Duan J (2017). Compatible taper-volume models of *Quercus variabilis* Blume forests in north China. iForest 10: 567-575. - doi: 10.3832/ifor2114-010 [online 2017-05-08]

Communicated by: Rupert Seidl

two methods and found that similar results were obtained (Fang et al. 2000). Method (1) for Type (2) also could make the system more flexible in application, *i.e.*, for users who would like to use an existing volume table or volume equation to estimate volume; they can just use the taper model to obtain the diameter predictions (Fang & Bailey 1999).

A large number of compatible taper-volume systems based on type (1) have been created for oak species in Greece, America, Denmark, Spain and Mexico (Hilt 1980, Thomas & Parresol 1991, Tarp-Johansen et al. 1997, Pompa-García et al. 2009, Kitikidou 2010). Simple equations (Hilt 1980, Thomas & Parresol 1991, Kitikidou 2010), variable exponent equations (Tarp-Johansen et al. 1997) or segmented-polynomial equations (Pompa-García et al. 2009) were chosen for taper modelling. There has been no similar research for *Quercus variabilis* Blume, an important broadleaf tree species in North China, prior to the study reported here. The objective of our study was to develop a suitable compatible taper and volume estimation system including a polynomial taper equation and a total volume model basing on type (2) using Method (1), which can describe the stem profile well and provide accurate estimates of the stem volume of *Quercus variabilis* Blume (cork oak) forests in Northern China.

Material and methods

Measurements

174 trees from 104 plots with an area of 20 × 20 m for cork oak natural forests and plantations in North China were used in this study, 57 of those trees were from plantations (including 31 average trees and 26 dominant trees), and the other 117 trees were from natural forests (including 60 average trees and 57 dominant trees). These plots were created in the following locations with different site conditions and age distributions: Gao-Luo forestry station, Qi-Jiahe forestry station, Bei-Tan forestry station and Heng-He forestry station of Zhong-tiaoshan region in Shanxi province, collective forests of Da-Geliao village in Xingtai city of Hebei province, Si-Zuolou forestry station and Xi-Shan forestry station in Beijing. Measured and computed variables were as follows: (1) single tree variables including diameter at breast height over bark (DBH_{ob}), total tree height (H); (2) two perpendicular diameters inside-bark (dib) of every five rings of each disc, starting with the outermost ring, at 0.0, 0.5, 1.3 and 1.5 m above the ground and then every 1.0 m along the remainder of the stem which were measured and averaged. For the outermost layer of those stem analyzed trees, diameters outside-bark at those same heights were also measured; (3) log volumes were calculated using the Huber's formula (Figueiredo-Filho et al. 2000) where the top section was treated as a cone. Inside-bark total stem volumes (vib) were obtained by summing the log volumes and the volume of the top of the tree. Each tree contributed to the data set with as many height-diameter measurements from the stem analysis data as possible. Total stem volume (vib), diameter at breast height over bark (DBH_{ob}) and height (H) were repeated for each analyzed stem defined by 5 ring measurements (Nunes et al. 2010). The data with DBH_{ob} equals to 0 (or H < 1.3 m) were deleted from the data set because the total volume under bark model (vib) would rely on DBH_{ob} as an independent variable. Finally, a total of 2358 bark thickness observations, 12814 diameter-height observations and 1299 volume observations from the 174 trees were obtained.

Model building

Four alternative modelling strategies were tried, because (1) a few analyzed stems ("trees") had ramicorns and (2) some of the small analyzed stems had very high values of relative diameter (Rd, which is equal to dib/DBH_{ib} where dib is the diameter under bark at height h in cm, DBH_{ib} is the diameter at breast height under bark in cm, and h is the height from ground in m). Therefore, four sets of compatible bark thickness-taper-volume model systems were built, one for each dataset type: (i) System 1, using all the data of the analyzed stems; (ii) System 2, using data of analyzed stems without ramicorns; (iii) System 3, using data of analyzed stems with a Rd less than 1.5; (iv) System 4, using data of the analyzed stems without ramicorns and simultaneously with a Rd less than 1.5. Descriptive statistics for those data sets are shown in Tab. 1. Each model system included three models, *i.e.*, a bark thickness model, a volume model and a taper model. Two dummy variables were used in System 1: *Rddummy* (*Rddummy* = 1 when Rd < 1.5, and 0 otherwise) and *branch* (*branch* = 1 when the tree has ramicorns, and *branch* = 0 otherwise). The dummy variable *Rddummy* was used in System 2, and dummy variable *branch* was used in System 3.

For bark thickness models and volume models from the above-mentioned four systems, all data were used for model fitting. Before modelling, some explanatory and response variables (V) were transformed to new variables (tV) by Box-Cox transformation to make frequency distributions of those variables (tV) as close to normal distributions as possible. The following equation expresses a Box-Cox transformation (Sakia 1992 – eqn. 1):

$$tV=\begin{cases}\dfrac{V^{\lambda}-1}{\lambda} & \text{if } \lambda\neq 0\\ \log_e V & \text{if } \lambda=0\end{cases}$$

where V was the original response or explanation variable, tV was the response or explanatory variable after Box-Cox transformation, λ was the parameter in the Box-Cox transformation.

To establish bark thickness models and volume models, linear mixed effects analyses were used on transformed variables with the tree number (*tree.no*) as random effect. Model form was expressed as in eqn.2 (see below – Peng & Lu 2012, Su et al. 2012). The autocorrelation was addressed using three residual autocorrelation structures: a first-order autoregressive structure [AR(1)], a moving average structure [MA(1)] and combination of first-order autoregressive and moving average structures [ARMA(1.1)]. For bark thickness models of the four systems, $tdib^n$, Rh^n, *natural, dominant, branch* (for System 1 and System 3) and their interaction terms were chosen as possible explanatory variables and *tbt* is

Tab. 1 - Summary statistics of four data sets used for modelling. (bt): bark thickness; (vib): stem total volume under bark; (dib): diameter under bark at height h; (h): height from ground; (H): total tree height; (DBH_{ob}): diameter at breast height over bark, breast height is 1.3 m height above the ground; (Rd): relative diameter, equal to dib/DBH_{ib}.

Model type	Model system	Sample number	Range of age (year)	Range of Rd	Range of DBH_{ob} (cm)	Range of H (m)	Range of response variable
bt (cm)	1, 3	2358	16-84	0.01-1.50	3.8-39.9	5.0-21.0	0.0-3.5
	2, 4	2059	16-84	0.02-1.50	3.8-22.6	5.0-18.2	0.0-3.0
vib (m^3)	1	1299	5-84	1.00	0.3-39.9	1.4-21.0	0.00001-0.649
	2	1201	5-84	1.00	0.3-23.1	1.4-18.2	0.00001-0.224
	3	1035	5-84	1.00	1.6-39.9	1.4-21.0	0.0002-0.649
	4	937	5-84	1.00	1.6-23.1	1.4-18.2	0.0002-0.224
dib (cm)	1	12814	5-84	0.01-12.00	0.3-39.9	1.4-21.0	0.1-41.0
	2	11336	5-84	0.01-12.00	0.3-23.1	1.4-18.2	0.1-25.5
	3	11419	5-84	0.01-1.50	1.6-39.9	1.4-21.0	0.1-41.0
	4	9942	5-84	0.01-1.50	1.6-23.1	1.4-18.2	0.1-25.5

the response variable, where *tdib*^{n} is equal to $(tdib)^n$ (n=1, 2, …, 5), *tdib* is the transformed diameter under bark at height *h* by the Box-Cox method (cm), *Rh*^{n} is equal to $(Rh)^n$ (n=1, 2, …, 5), *Rh* is the relative height and equal to *h*/*H*, where *h* is height from ground (m), *H* is total tree height (m); *natural* and *dominant* are dummy variables to define the forest origin and the tree size, respectively: *natural* = 1 when the origin was a natural forest, *natural* = 0 in the case of plantations; *dominant* = 1 when the tree is dominant, *dominant* = 0 otherwise); *tbt* is the transformed bark thickness by the Box-Cox method (cm). The dummy variable *Rddummy* was not used in System 1 and System 2 because the data used for fitting the *tbt* model were all with a *Rd*<1.5 (Tab. 1). For volume models of the four systems, $tDBH_{ob}$, *tH, td*^{2}*h, natural, dominant, branch* (for System 1 and System 3), *Rddummy* (for System 1 and System 2) and their interaction terms were chosen as possible explanatory variables, while *tvib* was the response variable, where $tDBH_{ob}$, *tH*, and *tvib* are transformed diameter at breast height over bark, transformed total tree height and transformed stem total volume under bark by the Box-Cox method, respectively (cm, m, m^3), td^2h is the transformed d^2h by the Box-Cox method, d^2h is equal to $DBH_{ob}\times DBH_{ob}\times H$; the other variables have the same specifications as in the bark thickness models. DBH_{ob} of inner layers of the 174 trees were calculated by eqn. 3 (see below), where bark thickness (*bt*) could be obtained from the bark thickness model. Model parameters were estimated by the ordinary least squares method (OLS – eqn. 2):

$$Y = X\beta + Z\mu + \varepsilon$$

where *Y* is the vector of the response variable; *X* is the vector of fixed-effect regressors; *Z* is the vector of random-effect regressors; β is the vector of fixed-effect coefficients; μ is the vector of the random-effect coefficients; ε is the vector of errors. DBH_{ob} (diameter at breast height over bark, cm) was obtained as follows (eqn. 3):

$$DBH_{ob} = DBH_{ib} + 2\cdot bt$$

where DBH_{ib} is the diameter at breast height under bark (cm), breast height is 1.3m height from ground, and *bt* is the bark thickness (cm).

For each model system, an overall merit-based method was used to select model explanatory variables. Regression equations for bark thickness models and volume models with different variable combinations were compared. Four sets of optimal base equations were obtained by examining the coefficients of determination (R^2) and root mean square errors (*RMSE*); then the Akaike's information criterion (*AIC*) was used to successively determine the best random-effects combination and the best residual autocorrelation structure for each optimal base model to obtain four sets of optimal *tbt* models and *tvib* models. For those optimal models, residual distribution homogeneity and model bias were visually checked by residual plots with loess regression lines overlaid on the plots. For an unbiased model, a loess line should be flat and located at the zero value on the vertical axis in the residual plot (Jacoby 2000). Normality of residuals was checked with histograms of residuals and by using a Shapiro-Wilk test (probabilities of type I error *or p-values* < 0.05 indicate a departure from a normal distribution). Finally, *bt* models and *vib* models were obtained by back-transforming *tbt* models and *tvib* models (see eqn. 4, where *exp* is the natural exponential function, other notations have the same meanings with those in eqn. 1) and the residuals were also examined (eqn. 4):

$$V = \begin{cases} (\lambda \cdot tV + 1)^{1/\lambda} & \text{if } \lambda \neq 0 \\ \exp(tV) & \text{if } \lambda = 0 \end{cases}$$

For taper modelling of the four systems, a subset of data (80%) from the analyzed stems were randomly selected for the fitting phase, while the remaining data were used for model validation (Muhairwe 1999, Brooks et al. 2008, Ozçelik & Brooks 2012). Using the fitting data, the estimated total volume by the previously built total volume model (*vib* model) and a specific parameter restriction (see eqn. 7 below) could be substituted in a polynomial taper model (see eqn. 5 and eqn. 6 below – Goulding & Murray 1976), then the parameters of the taper models could be fitted by ordinary least squares (OLS – eqn. 5, eqn. 6, eqn. 7):

$$dib^2 = \frac{Vv}{(\pi/40000)\cdot H}\cdot \left(c_1 Z_1 + c_2 Z_2 + c_3 Z_3 + c_4 Z_4 + c_5 Z_5 + c_6 Z_6\right)$$

$$Z_n = \left(\frac{H-h}{H}\right)^n$$

$$\frac{c_1}{2} + \frac{c_2}{3} + \ldots + \frac{c_n}{n+1} = 1$$

where *dib* is the diameter under bark at height *h* (cm), *h* is the height above the ground (m), *H* is the total tree height (m), *Vv* is the estimated stem total volume under bark from a total volume model (m^3), c_i terms are model parameters (i=1, 2, 3, …, 6).

To make taper modelling simpler, statistic *Y* can be calculated according to eqn. 8 (see below) and then the *Y* models were fitted firstly, *i.e.*, the estimated values of total volume obtained by the *vib* model and a specific parameter restriction (see eqn. 7) were substituted into eqn. 8 instead of into eqn. 5. Linear mixed effects models (*Y*) with variables Z_n (n=1, 2, ..., 6), dummy variable *Rddummy* (for System 1 and System 2) and dummy variable *branch* (for System 1 and System 3) as fixed factors and with tree number (*tree.no*) and dummy variable *natural* as random factors were fitted. Whether or not to include any particular estimated parameter was decided by the significance of a *t*-test. The autocorrelation was also addressed using the above-mentioned three residual autocorrelation structures. Then *dib* models can be obtained by eqn. 9 and the residuals were also examined (eqn. 8, eqn. 9):

$$Y = \frac{(\pi/40000)\, dib^2 H}{Vv} = c_1 Z_1 + c_2 Z_2 + c_3 Z_3 + c_4 Z_4 + c_5 Z_5 + c_6 Z_6$$

$$dib = \sqrt{\frac{40000\, Vv\, Y}{\pi H}}$$

Similarly, the R^2, *RMSE* , *AIC* value, residual plots with loess regression lines overlaid on the plots and histograms of residuals of the *Y* models and *dib* models were tested. Then four sets of optimal regression equations (*Y* models and *dib* models) were selected.

To sum up, four sets of data (with a bark thickness model, a volume model and a taper model in each of them) were used for modelling and the most suitable set was then selected using the above-mentioned statistics and residual plots.

Optimal model system evaluation

After the selection of the optimal model system, representing essentially the best dataset, the transformed bark thickness model (*tbt*) and the transformed volume model (*tvib*) in it were tested by the leave-one-out Jackknife method (Sánchez-González et al. 2005, 2007). The residual ranges and prediction ranges of the models and their corresponding jackknife tests were compared. Mean biases (*Bias*) and mean absolute biases (*MAD*) of the back-transformed *vib* model were also assessed for each diameter classes.

For the taper model in the optimal model system, the predictive performance of *dib* model was evaluated using the validation data set. The residual ranges and prediction ranges based on the validation data set were compared with those based on the fitting data set. Mean biases (*Bias*) and mean absolute biases (*MAD*) which were computed respectively using the fit data and validation data were assessed by position (percent relative height, *i.e.*, 5%, 10%, 15%, …, 95%). Finally, the fitting and validation datasets were combined, the taper model (*Y* and *dib*) was refitted and the corresponding statistics and plots were examined again (Muhairwe 1999).

Results

Four sets of models

The *p*-value, R^2 and *RMSE* of eight models (*tbt, tvib, Y, bt vib* and *dib*) in each model system are shown in Tab. 2. R^2 values were higher than 0.85 for all the models and values of *RMSE* of all the models were low compared to the ranges of response variables (Tab. 1, Tab. 2). Probabilities of type I errors of all the models were lower than 0.05 (Tab. 2). Overall, models in System 4

Tab. 2 - Values of fitting statistics for eight models in four modeling systems. (*bt*): bark thickness; (*vib*): stem total volume under bark; (*dib*): diameter under bark at height *h*; (*tbt*) and (*tvib*): transformed values of *bt* and *vib* by the Box-Cox method; (*Y*): calculated according to eqn. 8; (*p*-value): probability of type I error in Shapiro Wilks test (for the *tbt* and *tvib* models) and probability of type I error in Kolmogorov-Smirnov test (for the *Y* models); (R^2): coefficient of determination; (*RMSE*): root mean square errors.

Models type	Model system	Transformed models ***tbt*** **(cm),** ***tvib*** **(m^3) or** ***Y***			Back-transformed models ***bt*** **(cm),** ***vib*** **(m^3) or** ***dib*** **(cm)**	
		p-value	R^2	*RMSE*	R^2	*RMSE*
bt (cm)	1, 3	< 2.2e-16	0.90	0.22	0.86	0.24
	2, 4	9.646e-13	0.94	0.15	0.92	0.15
vib (m^3)	1	3.343e-12	0.99	0.061	0.99	0.006
	2	9.195e-08	0.99	0.043	0.99	0.003
	3	6.921e-09	0.99	0.066	0.99	0.006
	4	2.14e-06	0.99	0.044	0.99	0.003
dib (cm)	1	< 2.2e-16	0.91	0.32	0.97	0.84
	2	< 2.2e-16	0.91	0.32	0.97	0.67
	3	< 2.2e-16	0.95	0.21	0.97	0.80
	4	< 2.2e-16	0.96	0.20	0.98	0.61

Tab. 3 - Summaries for the *tbt* and *bt* models in System 4. The type III sum of squares was used in those models. (*bt*): bark thickness (cm); (*tbt*) and (*tdib*): transformed values of *bt* and *dib* by the Box-Cox method, *i.e.*, $tbt=(bt^{0.538}-1)/0.538$, $tdib=(dib^{0.716}-1)/0.716$; (*dib*): diameter under bark at height *h* (cm); (*Rh*): relative height and equal to *h*/*H*; (*h*): height from ground (m); (*H*): total tree height (m); (*natural*): dummy variable (*natural* = 1 for natural forests; *natural* = 0 for plantations); (a_i): model parameters (*i*=1, 2, 3, ..., *n*); (*SE*): standard errors of coefficients; (ρ): the parameter for first-order autoregressive structure [AR(1)]; (σ^2): the residual variance; (σ_{ai}^2): the variance for the random effects. (***): p<0.001.

Model	Equation form	Coefficients (± *SE*)		Predicted value range/ Residual range Model (Entire data)	Jackknife
tbt	$tbt = a_0 + a_1 \cdot tdib^5/10^4 + a_2 \cdot dib^4/10^3 + a_3 \cdot tdib^3/10^2 + a_4 \cdot tdib + a_5 \cdot Rh + a_6 \cdot Rh^2 + a_7 \cdot Rh^3 + a_8 \cdot Rh^4 + a_9 \cdot natural$	a_0	-0.881 ± 0.058***	(-1.58, 1.22)/ (-0.61, 0.56)	(-1.60, 1.24)/ (-0.64, 0.60)
		a_1	-0.277 ± 0.071***		
		a_2	0.669 ± 0.160***		
		a_3	-0.495 ± 0.101***		
		a_4	0.278 ± 0.018***		
		a_5	-3.182 ± 0.201***		
		a_6	8.094 ± 0.882***		
		a_7	-8.695 ± 1.426***		
		a_8	3.327 ± 0.745***		
		a_9	0.089 ± 0.022***		
		ρ	0.403		
		σ^2	0.025		
		σ_{a0}^2	0.013		
bt	$bt = (0.538 \cdot tbt + 1)^{(1/0.538)}$	-	-	(0.03, 2.56)/ (-0.60, 0.74)	-

Tab. 4 - Summaries for the *tvib* and *vib* models in System 4. The type III sum of squares was used in those models; (*tvib*), (*tH*) and (td^2h): transformed values of *vib*, *H* and d^2h by the Box-Cox method, *i.e.*, $tvib=(vib^{0.18}-1)/0.18$, $tH=(H^{0.81}-1)/0.81$, $td^2h=[(d^2h)^{0.21}-1]/0.21$; ($d^2h$): equal to $DBH_{ob} \cdot DBH_{ob} \cdot H$; ($DBH_{ob}$): diameter at breast height over bark (cm); (*H*): total tree height (m); (*vib*): stem total volume under bark (m^3); (b_i): model parameters (*i*=1, 2, 3, ..., *n*); (*SE*): standard errors of coefficients; (θ): the parameter for moving average structure [MA(1)]; (σ^2): the residual variance; (σ_{bi}^2): the variance for the random effects. (***): p<0.001.

Model	Equation form	Coefficients (± *SE*)		Predicted value range/ Residual range Model	Jackknife
tvib	$tvib = b_0 + b_1 \cdot td^2h + b_2 \cdot tH$	b_0	-4.488 ± 0.007***	(-4.296,-1.244)/ (-0.218, 0.180)	(-4.290,-1.233)/ (-0.243.0.216)
		b_1	0.125 ± 0.002***		
		b_2	-0.021 ± 0.005***		
		θ	0.467		
		σ^2	0.002		
		σ_{b0}^2	0.002		
vib	$vib = (0.18 \cdot tvib+1)^{(1/0.18)}$	-	-	(0.0003.0.245)/ (-0.020, 0.018)	-

had higher R^2 and lower *RMSE* (Tab. 2). The loess curves of models in System 4 were closest to the *x*-axis, followed by those of System 2, then those of System 3, and models in System 1 with the biggest deviation from the *x*-axis; this means that models in System 4 had much lower bias than those in the other three systems. Therefore, System 4 was the optimal model system (only some of the residuals plots are shown in this paper, see below).

Equation forms, coefficients and standard errors of coefficients of the models in System 4 are shown in Tab. 3, Tab. 4 and Tab. 5. All the coefficients were significant at α=0.001 confidence level and standard errors of coefficients were small compared to the coefficients (Tab. 3, Tab. 4 and Tab. 5). A first-order autoregressive structure [AR(1)], a moving average structure [MA(1)] and a combination of first-order autoregressive and moving average struc-

Tab. 5 - Summaries for *Y* and *dib* models in System 4. The type III sum of squares was used in those models. (*dib*): diameter under bark at height *h* (cm); (*h*): height from ground (m); (*H*): total tree height (m); (*Y*): calculated according to eqn.8; (*Zn*) (*n* =1, 2, ..., 6): calculated according to eqn.6; (*Vv*): estimated stem total volume under bark from stem total volume model (m^3); (c_i): model parameters (*i*=1, 2, 3, ..., *n*); (*SE*): standard errors of coefficients; (*ρ*): the parameter for first-order autoregressive structure [AR(1)]; (ϑ): the parameter for moving average structure [MA(1)]; (σ^2): residual variance; (σ_{ci}^2): variance for the random effects (including *natural* and *tree.no.*). (***): p<0.001.

Model	Equation form	Coefficients (± *SE*)		Predicted value range / Residual range		
				Entire data	Fit data	Validation data
Y	$Y = c_3 \cdot Z_3 + c_4 \cdot Z_4 + c_5 \cdot Z_5 + c_6 \cdot Z_6 + [8 \cdot (1-(c_3/4)-(c_4/5)-(c_5/6)-(c_6/7))] \cdot Z_7$	c_3	33.972 ± 1.545***	(0.00, 4.04)/ (-1.18, 1.48)	(0.00, 4.05)/ (-1.19, 1.24)	(0.00, 3.12)/ (-1.81, 1.45)
		c_4	-125.749 ± 8.907***			
		c_5	229.692 ± 18.913***			
		c_6	-211.380 ± 17.418***			
		ρ	0.738			
		θ	-0.359			
		σ^2	0.043			
		σ_{c3}^2	0.050; 8.117			
		σ_{c4}^2	0.210; 24.145			
dib	$dib = [(40000 \cdot Vv \cdot Y)/(\pi \cdot H)]^{(1/2)}$	-	-	(0.00, 24.39)/ (-2.94, 3.03)	(0.00, 24.40)/ (-2.91, 2.70)	(0.00, 22.94)/ (-2.60, 3.05)

tures [ARMA(1.1)] were respectively elected as the best to address the autocorrelation of the *tbt* model, *tvib* model and *Y* model, respectively (Tab. 3, Tab. 4 and Tab. 5). Predicted value ranges and residual ranges of those models are also shown in Tab. 3, Tab. 4 and Tab. 5. Residual plots of the bark thickness model (*bt*), volume model (*vib*) and taper model (*dib*) in System 4 fitted to the combined datasets are shown in Fig. 1, Fig. 2 and Fig. 3, respectively. In general, the loess lines are represented by flat lines located at the baselines, except for the trees with a relative height (*Rh*) higher than 0.9 for the residual plot of the taper model. Heteroscedasticity was not obvious in the transformed bark thickness model (*tbt*), transformed volume model (*tvib*) and the taper models (*Y* and *dib*), while a weak heteroscedasticity could be detected in the back-transformed thickness model (*bt*) and in the back-transformed volume model (*vib*). Low *p*-values of the Shapiro Wilks tests or Kolmogorov-Smirnov tests (Tab. 2) and histograms of residuals (data not shown) suggested that residuals of the models in System 4 did not have a normal distribution. Skewness was not detected in these distributions, but kurtosis was.

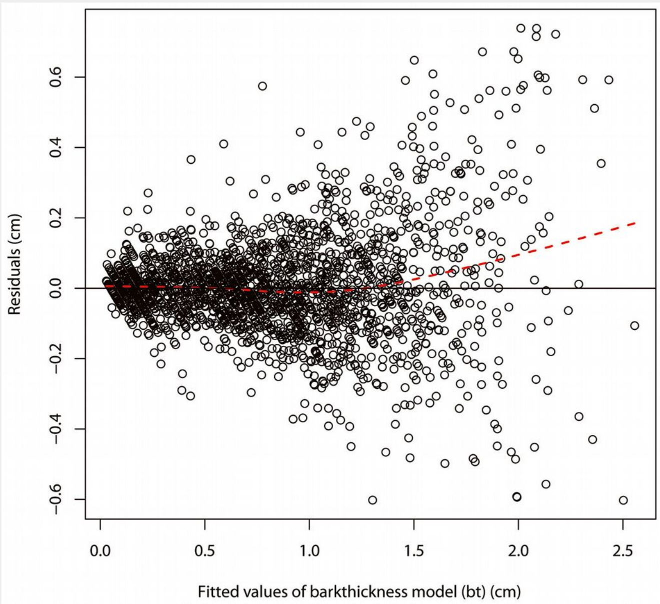

Fig. 1 - Loess residual plot of the back-transformed bark thickness model in System 4. The solid horizontal line indicates the baseline, while the red dotted line represents the loess curve.

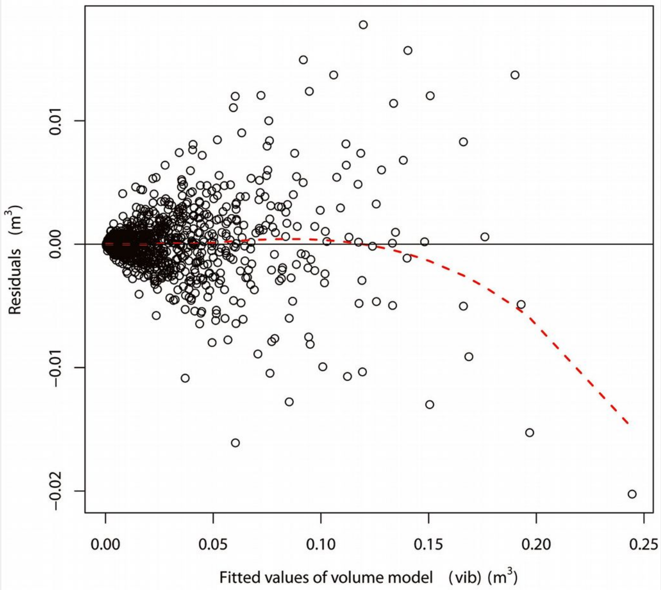

Fig. 2 - Loess residual plot of the back-transformed total stem volume inside bark model in System 4. The solid horizontal line indicates the baseline, while the red dotted line represents the loess curve.

Evaluation of the models in System 4

Predicted value ranges and residual ranges from the Jackknife validations for the transformed bark thickness model (*tbt*) and the transformed volume model (*tvib*) in System 4 were similar with those obtained from fittings and are shown in Tab. 3 and Tab. 4. Scatter distribution in residual plots from fittings and that from the jackknife validations were similar, as well as the frequency distributions in histograms of residuals (data not shown). Moreover, frequencies, mean bias (*Bias*) and mean absolute deviation (*MAD*) of the back-transformed volume model (*vib*)

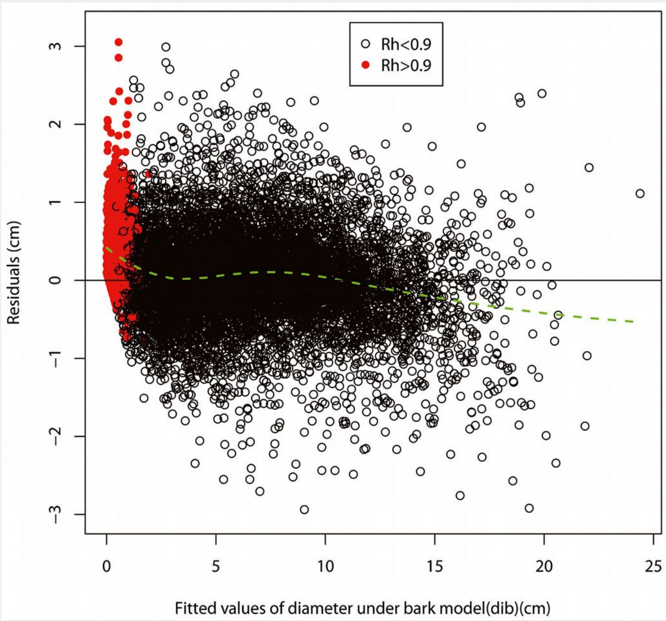

Fig. 3 - Loess residual plot of the back-transformed diameter inside bark model in System 4. The red closed circles represent the residuals under the condition of relative height > 0.9 and the black open circles represent the residuals under the condition of relative height <0.9. The solid horizontal line indicates baseline and the green dotted line indicates loess curve.

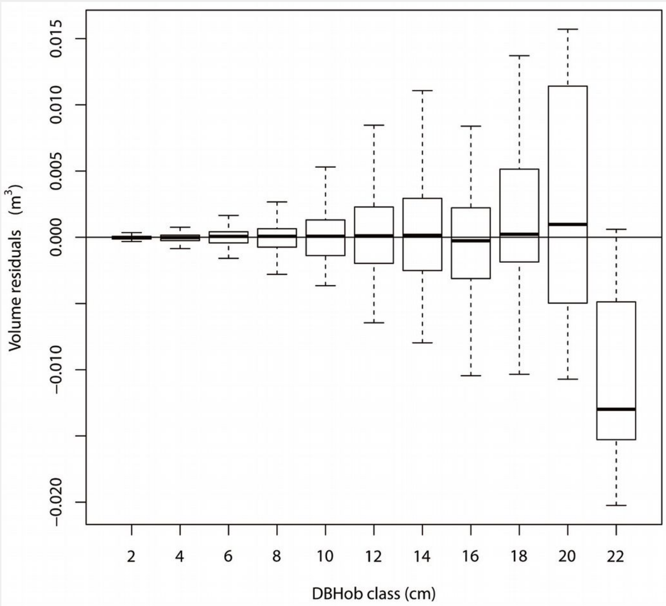

Fig. 4 - Box plot of residuals *versus* diameter at breast height over bark class for the volume inside bark model in System 4.

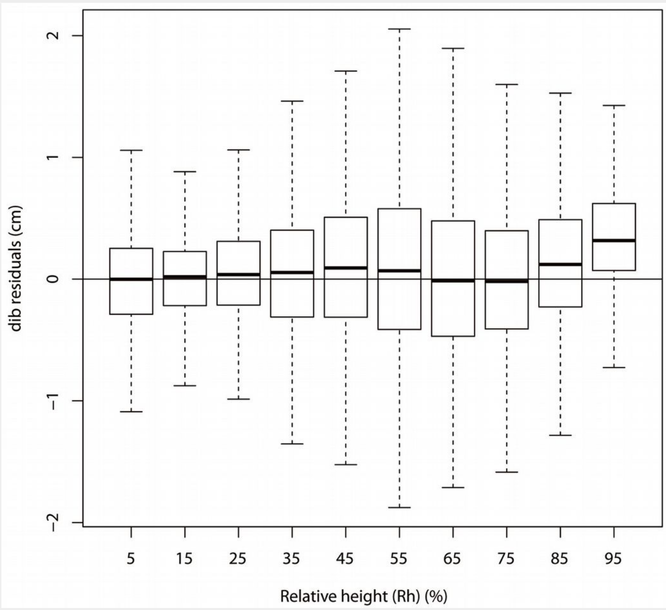

Fig. 5 - Box plot of residuals *versus* relative height for the taper model in System 4.

were tested and the corresponding results are listed in Tab. S7 (Supplementary material). The overall values of *Bias* and *MAD* were $0.09 \cdot 10^{-3}$ m^3 and $1.76 \cdot 10^{-3}$ m^3 (Tab. S7 in Supplementary material). Both *Bias* and *MAD* in different DBH_{ob} classes had an increasing trend with increasing DBH_{ob}. However, all the values of *Bias* and *MAD* were very small compared to the magnitudes of predicted values (see Tab. S4 and Tab. S7 in Supplementary material), so the *vib* model could provide accurate prediction within the DBH_{ob} range of 1.6 to 23.1 cm and within the *H* range of 1.4 to 18.2 m. The box plot of residuals in each diameter class for the *vib* model is shown in Fig. 4.

Residual plots, predicted value ranges, residual ranges, *Bias* and *MAD* of the taper model fitted with the entire data set and the fitting data set were very close to each other, while the corresponding statistics of the taper model fitted with the validation data set were a little different (Tab. S5 and Tab. S8 – Supplementary material). All the values of the *Bias* and *MAD* were small compared to the magnitude of predicted values (see Tab. S5 and Tab. S8 in Supplementary material). For the *dib* model built using the entire data, the overall values of *Bias* and *MAD* were 0.06 cm and 0.45 cm, and the *Bias* and *MAD* in different relative height classes (*Rhg*) respectively ranged from -0.04 to 0.40 cm and from 0.29 to 0.60 cm (Tab. S8 in Supplementary material). This *dib* model could provide accurate predictions within a DBH_{ob} range of 1.6 to 23.1 cm and within an *H* range of 1.4 to 18.2 m. Fig. 5 shows a box plot of residuals versus relative height (*Rh*) for the *dib* model.

Discussion

Equation form of total volume model

Various forms of volume models have been reported in the literature, such as the model represented by eqn. 10 (see below) proposed by Schumacher & Hall in 1933 (Bailey 1994, Fang & Bailey 1999), eqn. 11 (Malimbwi & Philip 1989), eqn. 12 (Diéguez-Aranda et al. 2006), eq. 13 (Honer 1965) and eqn. 14 (Muhairwe 1999). These equations are shown below (eqn. 10 to eqn. 14):

$$V = a_1 \cdot DBH^{a_2} \cdot H^{a_3}$$

$$V = a_1 + a_2 \cdot DBH^2 \cdot H$$

$$V = a_1 \cdot DBH^2 \cdot H$$

$$V = DBH^2 / (a_1 + a_2 / H)$$

$$V = e^{a_1} DBH^{a_2} \left[H^2 / (H - 1.3) \right]^{a_3}$$

where *V* is the volume, *DBH* is the diameter at breast height, *H* is the total height, a_i terms are model parameters, *e* is the base of the natural logarithm. Some of the volume models have a conceptual basis in the geometry of solids of revolution and have a constant form factor (Bailey 1994). In fact, the form of a tree depends upon the actual tree size, *e.g.*, there was a downward trend of form factor along with increased tree height in our study. Some volume models in some studies explicitly represented change of a form factor, which was defined as a function of diameter at breast height, total height, stem height at a predetermined fraction of diameter at breast height outside bark, or the ratio of this height to total height (Bailey 1994, Rustagi & Loveless 1991). However, there is always a problem of heterogeneity in those nonlinear volume models. For removing heterogeneity of variance, weighted non-linear regression was usually used to estimate parameters (Muhairwe 1999). However, computing a suitable weighting variable was awkward. Another simple and common way of removing heterogeneity of residual variance is performing a transformation to stabilize variance. In our study, the Box-Cox transformation was used, then the linear mixed effect equation of the transformed volume was built and no heterogeneity was detected. Meanwhile, bias of the back-transformed volume model was found to be small, and so no correction factor was used in this study. Additionally, an overall merit-based method was used to select model explanatory variables for the volume model, so the volume model did not have a conceptual basis in the geometry of solids of revolution and did not explicitly represent change of a form factor.

In the application of a mixed effect model, when a sub-sample of the dataset is available to calculate the random effects, users can calibrate the coefficients of the linear mixed effect model ("lme" - Temesgen et al. 2008) and then obtain unbiased predictions. However, predicting the random effects is hard for users. Actually, in our study the bias was found small enough, even though just the fixed effect was considered in prediction; thus, there was no need to calibrate the random effect before using the "lme" volume model. Population predictions of volume for a new tree can be obtained using fixed effect coefficients.

Similar features can be found in the bark thickness model, which was also a linear mixed effect equation using variables transformed by the Box-Cox method.

The simple taper model

According to several studies in the literature (Diéguez-Aranda et al. 2006, Oytunemre et al. 2008, Hjelm 2013), taper equations can be grouped in three types: (1) simple taper equations (Bruce et al. 1968, Kozak et al. 1969, Demaerschalk 1972, Demaerschalk 1973, Ormerod 1973, Goulding & Murray 1976, Fang & Bailey 1999, Sharma & Oderwald 2001); (2) segmented taper equations (Max & Burkhart 1976, Cao et al. 1980, Parresol et al. 1987, Fang et al. 2000, Jiang et al. 2005); (3) variable exponent taper equations (Kozak 1988, Newnham 1992, Bi 2000, Lee et al. 2003, Kozak 2004). Some researchers have pointed out that segmented taper equations and variable exponent taper equations can sometimes provide more flexible descriptions of tree profiles than simple taper equations; variable exponent taper equations usually have the least bias and best predictive abilities among the three kinds of models (Kozak 1988, Newnham 1992, Muhairwe 1999, Oytunemre et al. 2008). For simpler equations, the presence of larger residuals located in the lower bole (the stump region) is pronounced in some studies (Hjelm 2013). However, a shortcoming of variable exponent taper equations is that they cannot be analytically integrated to calculate total stem or log volumes (Diéguez-Aranda et al. 2006, Ozçelik & Brooks 2012). Additionally, segmented taper equations and variable exponent taper equations suffer from statistical complexity, difficulties in parameter estimation and difficulties of being understood and correctly used by forest managers. Therefore, when simplicity of use is an objective, the simple taper model would be a good choice, despite its lower accuracy in the lower bole (Martin 1981).

A polynomial taper equation (Goulding & Murray 1976) was used in this study. Larger residuals were only found at about 90% of stem height (Fig. 3, Fig. 5). The poorer performance observed in predictions at the stem top is negligible from a practical point of view (Figueiredo-Filho et al. 1996, Hjelm 2013), as the top part of cork oak is usually collected for bio-fuel. As we are interested only in the middle part of the bole, a simple taper model can be used for practical purposes (Oytunemre et al. 2008).

Ramicorns and relative diameter

Branches are an important aspect of tree form because they affect stem utilization. A ramicorn branch is a steep-angled branch diverging less than 30° from the main stem and significantly smaller than the main stem (Xiong et al. 2014). In this study, the number of trees with ramicorns was very small. Additionally, the values of relative diameter for most of the computed trees in the data set were under 1.5, which is consistent with those in other studies (Cao et al. 1980, Bailey 1994, Diéguez-Aranda et al. 2006, Oytunemre et al. 2008, Li & Weiskittel 2010, Ozçelik & Brooks 2012, Hjelm 2013). A few small trees estimated from inner rings had very high values of relative diameter (up to 12.0). These were very tiny "trees", with ages of 5 years and a diameter at breast height under 1.0 cm.

We built four sets of compatible taper-volume model systems using all the data (Model system 1), using data of stems without ramicorns (Model system 2), data of stems with a *Rd* less than 1.5 (Model system 3) and data of stems without ramicorns and simultaneously with a *Rd* less than 1.5 (Model system 4). Performances of models using those four datasets were compared though the R^2, *RMSE*, residual plots and histograms of residuals. It turned out that models in System 4 had the best performances, followed by System 2 and System 3, while models in System 1 had the worst performances (Tab. 2). Although a dummy variable (*branch*) to define ramnicorn trees was introduced in System 1 and System 3, performances of those models were still not good enough, partly because of the small number of sample trees with ramicorns. More data of trees with ramicorns need to be collected in order to get more integrated and accurate models. Performances of models in System 2 were much better than those of System 1 and System 3, and just a little poorer than those of System 4. However, System 2 contained the data of stems with a *Rd* bigger than 1.5, which were not common in practical application. Therefore, models in System 4 were selected as the most appropriate in terms of precision, lack of bias and practical application. They can be used to predict bark thickness, diameter at a specific point along the stem, merchantable volume and total stem volume of cork oak forests in North China within the specific ranges of *DBH* (1.6-23.1 cm) or *H* (1.4-18.2 m).

In System 4, data from four big trees were removed because they had ramicorns. Due to the small sample size for big trees, more big trees should be measured in the future to obtain a compatible taper-volume model system with a larger useable diameter span. It should be noted that if models created using System 4 are used for predictions of stems with ramicorn branches, then errors would be likely greater than those reported here. Therefore, we suggest that models created with System 4 can be used for predictions of stems without ramicorn branches and simultaneously with a relative diameter less than 1.5.

Conclusion

Linear mixed effect equations with tree number as random factor were used for bark thickness and volume modelling using variables transformed by the Box-Cox method to minimise heteroscedasticity. Using the polynomial equation reported by Goulding & Murray (1976), linear mixed effect equations with tree number and *natural* (a dummy variable specifying the stand origin) as random factors were fitted during taper modelling.

Four sets of compatible taper-volume models systems using different data sets were established and compared. The models in System 4 had superior coefficients of determination (R^2), root mean square errors (*RMSE*) and lack of bias than models from the other three systems and thus were selected as the most suitable in this study. Furthermore, the models in System 4 had good performances in jackknife validation or independent data set validation. Heteroscedasticity was not obvious in the transformed bark thickness model, transformed volume model and the taper

iForest – Biogeosciences and Forestry

model, while a weak heteroscedasticity could be detected in the back-transformed bark thickness model and back-transformed volume model. Residuals of the models in System 4 did not follow normal distribution. Skewness was not detected in these distributions, but they were slightly kurtotic.

Within the specified ranges of *DBH* (1.6-23.1 cm) or *H* (1.4-18.2 m) tested in this study, the compatible taper-volume models system can be used for predicting diameter at a specific point along the stem, merchantable volume and total stem volume of cork oak forests in North China.

Acknowledgements

This research was jointly supported by scientific and research base construction projects of Beijing Municipal Education Commission (SYSBL2009), forestry science promotion project of the State Forestry Bureau (2011-44), open fund project of Beijing Forestry University "985" advantage subject innovation platform (000-1108003), special fund project for forestry public service industry and research (201004021) and China Scholarship Council. We acknowledge the strong support from Zhong Tiaoshan National Forest Authority, Xingtai County Forestry Bureau, Si Zuolou forestry station and Xi Shan forestry station in Beijing.

Conghui Zheng and Yuzhong Wang have equally contributed to this work and should be regarded as co-first authors.

References

Bailey RL (1994). A compatible volume-taper model based on the Schumacher and Hall generalized constant form factor volume equation. Forest Science 40 (2): 303-313. [online] URL: http://www.ingentaconnect.com/content/saf/fs/1994/00000040/00000002/art00008

Bi HQ (2000). Trigonometric variable-form taper equations for Australian eucalypts. Forest Science 46 (3): 397-409. [online] URL: http://www.cabdirect.org/cabdirect/abstract/20003001623

Brooks JR, Jiang L, Ozçelik R (2008). Compatible stem volume and taper equations for Brutian pine, Cedar of Lebanon, and Cilicica fir in Turkey. Forest Ecology and Management 256 (1): 147-151. - doi: 10.1016/j.foreco.2008.04.018

Bruce D, Curtis RO, Vancoevering C (1968). Development of a system of taper and volume tables for red alder. Forest Science 14 (3): 339-350. [online] URL: http://www.ingentaconnect.com/content/saf/fs/1968/00000014/00000003/art00021

Cao QV, Burkhart HE, Max TA (1980). Evaluation of two methods for cubic-volume prediction of loblolly pine to any merchantable limit. Forest Science 26 (1): 71-80. [online] URL: http://www.ingentaconnect.com/content/saf/fs/1980/00000026/00000001/art00012

Demaerschalk JP (1972). Converting volume equations to compatible taper equations. Forest Science 18 (3): 241-245. [online] URL: http://www.ingentaconnect.com/content/saf/fs/1972/00000018/00000003/art00018

Demaerschalk JP (1973). Integrated systems for the estimation of tree taper and volume. Canadian Journal of Forest Research 3 (1): 90-94. - doi: 10.1139/x73-013

Diéguez-Aranda U, Castedo-Dorado F, Alvarez-González JG, Rojo A (2006). Compatible taper function for Scots pine plantations in northwestern Spain. Canadian Journal of Forest Research 36 (5): 1190-1205. - doi: 10.1139/x06-008

Fang ZX, Bailey RL (1999). Compatible volume and taper models with coefficients for tropical species on Hainan Island in Southern China. Forest Science 45 (1): 85-100. [online] URL: http://www.ingentaconnect.com/content/saf/fs/1999/00000045/00000001/art00011

Fang ZX, Borders BE, Bailey RL (2000). Compatible volume-taper models for loblolly and slash pine based on a system with segmented-stem form factors. Forest Science 46 (1): 1-12.

Figueiredo-Filho A, Borders BE, Hitch KL (1996). Taper equations for *Pinus taeda* plantations in Southern Brazil. Forest Ecology and Management 83 (1-2): 39-46. - doi: 10.1016/0378-1127(96)03706-1

Figueiredo-Filho A, Machado SA, Carneiro MRA (2000). Testing accuracy of log volume calculation procedures against water displacement techniques (xylometer). Canadian Journal of Forest Research 30 (6): 990-997. - doi: 10.1139/x00-006

Goulding CJ, Murray JC (1976). Polynomial taper equations that are compatible with tree volume equations. New Zealand Journal of Forestry Science 5 (3): 313-322. [online] URL: http://www.scionresearch.com/__data/assets/pdf_file/0005/31001/NZJFS5197GOULDING313_322.pdf

Hilt DE (1980). Taper-based system for estimating stem volumes of upland oaks. Research Paper NE-458, Northeast Forest Experiment Station, USDA Forest Service, Broomall, PA, USA, pp. 12. [online] URL: http://www.treesearch.fs.fed.us/pubs/15007

Hjelm B (2013). Stem taper equations for poplars growing on farmland in Sweden. Journal of Forestry Research 24 (1): 15-22. - doi: 10.1007/s11676-012-0270-4

Honer TG (1965). A new total cubic foot volume function. The Forestry Chronicle 41 (4): 476-493. - doi: 10.5558/tfc41476-4

Jacoby WG (2000). Loess: a nonparametric, graphical tool for depicting relationships between variables. Electoral Studies 19 (4): 577-613. - doi: 10.1016/S0261-3794(99)00028-1

Jiang L, Brooks JR, Wang J (2005). Compatible taper and volume equations for yellow-poplar in West Virginia. Forest Ecology and Management 213 (1-3): 399-409. - doi: 10.1016/j.foreco.2005.04.006

Kitikidou K (2010). Taper equation compatible with volume equation for the Hungarian oak stands under restoration at Northern Greece. Journal of Food Agriculture and Environment 8 (2): 866-869. [online] URL: http://world-food.net/download/journals/2010-issue_2/e4.pdf

Kozak A (1988). A variable-exponent taper equation. Canadian Journal of Forest Research 18 (11): 1363-1368. - doi: 10.1139/x88-213

Kozak A (2004). My last words on taper equations. The Forestry Chronicle 80 (4): 507-515. - doi: 10.5558/tfc80507-4

Kozak A, Munro DD, Smith JHG (1969). Taper functions and their application in forest inventory. The Forestry Chronicle 45 (4): 278-283. - doi: 10.5558/tfc45278-4

Lee WK, Seo JH, Son YM, Lee KH, Gadow KV (2003). Modeling stem profiles for Pinus densiflora in Korea. Forest Ecology and Management 172 (1): 69-77. - doi: 10.1016/S0378-1127(02)00139-1

Li R, Weiskittel AR (2010). Comparison of model forms for estimating stem taper and volume in the primary conifer species of the North American Acadian Region. Annals of Forest Science 67 (3): 302. - doi: 10.1051/forest/2009109

Malimbwi RE, Philip MS (1989). A compatible taper/volume estimation system for *Pinus patula* at Sao Hill forest project, Southern Tanzania. Forest Ecology and Management 27 (2): 109-115. - doi: 10.1016/0378-1127(89)90033-9

Martin AJ (1981). Taper and volume equations for selected Appalachian hardwood species. Research Paper NE-490, Northeast Forest Experiment Station, USDA Forest Service, Broomall, PA, USA, pp. 22. [online] URL: http://www.fs.fed.us/ne/newtown_square/publications/research_papers/pdfs/scanned/OCR/ne_rp490.pdf

Max TA, Burkhart HE (1976). Segmented polynomial regression applied to taper equations. Forest Science 22 (3): 283-289. [online] URL: http://www.ingentaconnect.com/content/saf/fs/1976/00000022/00000003/art00011

Muhairwe CK (1999). Taper equations for *Eucalyptus pilularis* and *Eucalyptus grandis* for the north coast in New South Wales, Australia. Forest Ecology and Management 113 (2-3): 251-269. - doi: 10.1016/S0378-1127(98)00431-9

Newnham RM (1992). Variable-form taper functions for four Alberta tree species. Canadian Journal of Forest Research 22 (2): 210-223. - doi: 10.1139/x92-028

Nunes L, Tomé J, Tomé M (2010). A system for compatible prediction of total and merchantable volumes allowing for different definitions of tree volume. Canadian Journal of Forest Research 40 (4): 747-760. - doi: 10.1139/X10-030

Ormerod DW (1973). A simple bole model. The Forestry Chronicle 49 (3): 136-138. - doi: 10.5558/tfc49136-3

Oytunemre S, Nuray M, Hakki Y, Mehmet M (2008). Stem taper functions for *Abies nordmanniana* subsp. *bornmulleriana* in Turkey. Scandinavian Journal of Forest Research 23 (6): 522-533. - doi: 10.1080/02827580802552453

Ozçelik R, Brooks JR (2012). Compatible volume and taper models for economically important tree species of Turkey. Annals of Forest Science 69 (1): 105-118. - doi: 10.1007/s13595-011-0137-4

Parresol BR, Hotvedt JE, Cao QV (1987). A volume and taper prediction system for bald cypress. Canadian Journal of Forest Research 17 (3): 250-259. - doi: 10.1139/x87-042

Peng H, Lu Y (2012). Model selection in linear mixed effect models. Journal of Multivariate Analysis 109: 109-129. - doi: 10.1016/j.jmva.2012.02.005

Pompa-García M, Corral-Rivas JJ, Hernández-Díaz JC (2009). A system for calculating the merchantable volume of oak trees in the northwest of the state of Chihuahua, Mexico. Journal of Forestry Research 20 (4): 293-300. - doi:

10.1007/s11676-009-0051-x

Rustagi KP, Loveless RS (1991). Improved cubic volume prediction using a new measure of form factor. Forest Ecology and Management 40 (1-2): 1-11. - doi: 10.1016/0378-1127(91)90087-C

Sánchez-González M, Cañellas I, Montero G (2007). Generalized height-diameter and crown diameter prediction models for cork oak forests in Spain. Forest Systems 16 (1): 76-88. [online] URL: http://revistas.inia.es/index.php/fs/article/view/999

Sánchez-González M, Tomé M, Montero G (2005). Modelling height and diameter growth of dominant cork oak trees in Spain. Annals of Forest Science 62 (7): 633-643. - doi: 10.1051/forest:2005065

Sakia RM (1992). The Box-Cox transformation technique: a review. Journal of the Royal Statistical Society 41 (2): 169-178. [online] URL: http://www.jstor.org/stable/2348250

Sharma M, Oderwald RG (2001). Dimensionally compatible volume and taper equations. Canadian Journal of Forest Research 31 (5): 797-803. - doi: 10.1139/x01-005

Su X, Yan X, Tsai CL (2012). Linear regression. Wiley Interdisciplinary Reviews Computational Statistics 4 (3): 275-294. - doi: 10.1002/wics.1198

Tarp-Johansen MJ, Skovsgaard JP, Madsen SF, Johannsen VK, Skovgaard I (1997). Compatible stem taper and stem volume functions for oak (*Quercus robur* L and *Q. petraea* (Matt) Liebl) in Denmark. Annales Des Sciences Forestières 54 (7): 577-595. - doi: 10.1051/forest:19970701

Temesgen HT, Monleon VJM, Hann DWH (2008). Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. Canadian Journal of Forest Research 38 (3): 553-565. - doi: 10.1139/X07-104

Thomas CE, Parresol BR (1991). Simple, flexible, trigonometric taper equations. Canadian Journal of Forest Research 21 (7): 1132-1137. - doi: 10.1139/x91-157

Xiong JS, Mckeand SE, Whetten RW, Isik FT (2014). Genetics of stem forking and ramicorn branches in a cloned loblolly pine family. Forest Science 60 (2): 360-366. - doi: 10.5849/forsci.12-018

Supplementary Material

Tab. S1 – Summary statistics of four data sets used for modelling.

Tab. S2 – Values of fitting statistics for eight models (*tbt, tvib, Y, bt, vib* and *dib*).

Tab. S3 – Results of *tbt* and *bt* models.

Tab. S4 – Results of *tvib* and *vib* models.

Tab. S5 – Results of *Y* and *dib* models.

Tab. S6 – The estimated taper function (red dotted curve) and the basic taper data for each system.

Tab. S7 – Frequencies, *Bias* and *MAD* of the *vib* model in system 4.

Tab. S8 – Frequencies, *Bias* and *MAD* of *dib* model in system 4.

Link: Zheng_2114@suppl001.pdf