STEM TAPER EQUATIONS FOR *BETULA ALNOIDES* IN SOUTH CHINA

C Tang[1, 2], CS Wang[1], SJ Pang[3], ZG Zhao[1], JJ Guo[1], YC Lei[4] & J Zeng[1, *]

[1]*Research Institute of Tropical Forestry, Chinese Academy of Forestry, Longdong, Guangzhou 510520, China*
[2]*Agricultural College, Shihezi University, Shihezi, Xinjiang 832003, China*
[3]*Experimental Center of Tropical Forestry, Chinese Academy of Forestry, Pingxiang, Guangxi 536000, China*
[4]*Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing, China*

**zengdasan@hotmail.com*

Submitted December 2015; accepted July 2016

Accurate stem taper functions are necessary for estimating stem diameter, form and tree volume, and are important for wood estimation and timber utilisation. *Betula alnoides* is a valuable plantation hardwood species under large-scale development in South-East Asia and south China, but no study on its stem taper has been reported yet. Here, 28 well-known taper functions from three groups of models (single, segmented and variable-form taper functions) were fitted and validated separately with diameter–height datasets from 90 and 29 trees of *B. alnoides*. All trees were sampled from its 8–36-year-old even-aged plantations in south China. Non-linear regression techniques were applied to estimate the parameters in the taper equations using the fitting data set. Then bias, absolute bias, mean squared error and per cent explained variation were applied to examine model accuracy. Student's paired *t*-tests for observed and predicted diameters were used to examine the validity of these functions based on the validation data. Of the 28 functions, the Bruce, Max-Burkhart and Muhairwe models were the best single, segmented and variable-form taper equations respectively. In general, variable-form taper equations performed better than simple and segmented functions. The Muhairwe model best estimated stem taper without obvious multicollinearity for this species.

Keywords: Non-linear regression analysis, taper functions, variable-form model

INTRODUCTION

Stem form differs among tree species and is influenced by factors such as site condition and stand density (Sharma & Parton 2009, Jiang & Liu 2011). Stem form varies along the length of the tree from ground to top with bole shapes that include neiloid, paraboloid and conic forms (Kozak 1997). Stem taper functions can provide forest managers with estimates of diameter variation at any height on the stem (Kozak 2004, De-Miguel et al. 2012, Özçelik & Crecente-Campo 2016). Accurate taper functions are needed in most inventory systems for estimating upper stem diameter, form and tree volume, and are important for wood harvesting and timber utilisation. Forest researchers have thus been aware of the variability and importance of individual tree stem form and have been modelling it for over 100 years (Fang & Bailey 1999).

Many taper models have been reported over the past several decades and are usually grouped into three types, namely, single, segmented and variable-exponent. Single taper models describe the whole stem profile using a polynomial, trigonometric or power function. The general consensus among researchers is that single taper functions sometimes fail to represent the entire stem profile, particularly near the butt and top of trees (Jiang et al. 2005). Max and Burkhart (1976) introduced the first segmented taper function which modelled stems divided into three segments, and increased taper prediction accuracy. Variable-exponent models were also developed to reduce local bias in taper predictions for the entire stem (Kozak 1988, Newnham 1988, Muhairwe 1999).

Most taper systems have focused on softwoods species, particularly pine, spruce and fir. Fewer taper systems have been published for hardwood species. The segmented polynomial model developed by Max and Burkhart (1976) performed consistently and was good for lower bole diameter predictions for 18 commercial hardwood species (Martin 1981). Jiang et al. (2005) also suggested that for yellow poplar, segmented polynomial models developed by

Max and Burkhart (1976) and Clark et al. (1991) performed well in describing tree form along the entire stem. For stem taper of tropical and subtropical hardwood species only preliminary reports exist, e.g. for teak (Shuaibu 2015), some eucalypts (Gomat et al. 2011), 23 tropical hardwood species in Hainan Island (Fang & Bailey 1999) and four tropical tree species in Mount Makiling (Lumbres et al. 2016).

Betula alnoides is a fast-growing valuable hardwood species naturally distributed in South-East Asia and south China (Zeng et al. 2003). Its wood is commonly used in floorboard and furniture manufacture as well as decorative veneers, and is very popular with the salariat (Zeng et al. 2006). Driven by huge demands for its wood and wood products, *B. alnoides* plantations have increased rapidly in the past 15 years and exceeded 150,000 ha in south China in 2014. Since the rotation of *B. alnoides* is about 20 years (Zeng et al. 2010), a number of these plantations will reach harvestable age in the next 10 years. Thus a practicable stem taper equation for this species is urgently needed by forest managers to estimate wood volume in *B. alnoides* plantations more accurately. In this context, no proper stem taper equation currently exists for this tree species. The present study was therefore carried out to evaluate selected existing taper functions and to choose the equation best suited for stem diameter prediction of *B. alnoides*.

MATERIALS AND METHODS

Data collection

A total of 119 trees were used for stem taper analysis in the present study, of which 77 were sampled from the Experimental Center of Tropical Forestry in Pingxiang City, and 42 from the Laoshan Forest Farm in Baise City. Both cities are located in the western part of the Guangxi Zhuang Autonomous Region, which is one of the most important areas of *B. alnoides* plantations in south China. The sampled trees were dominant or co-dominant in 8–36-year-old even-aged stands of *B. alnoides*. Trees with multiple stems, obvious insect damage, broken tops or crooked boles were not sampled.

Diameters at 0, 0.3, 0.6 and 1.3 (breast height) m above ground level were measured to the nearest 0.1 cm for each sampled tree. The trees were then felled leaving stumps about 0.2 m high, and whole stem length was measured to the nearest 0.01 m using a 50-m measuring tape. Disks 2–3 cm thick were obtained at 0.3, 1.0, 1.3 and 3.0 m, then at 2 m intervals along the whole stem to the top. If the length of the last section was less than 2 m, one disk was extracted at a distance of 1 m from that section. All disks were labelled, placed in plastic bags and brought back to the laboratory where outside and inside bark diameters of each disk were measured at two perpendicular directions to the nearest 0.01 cm. Diameter–height data from the sampled trees were randomly divided into two subsets, namely, fitting data from 90 trees and validation data from 29 trees. Descriptive statistics for these trees are given in Table 1. Relative diameter was calculated as the quotient between diameter at any height and diameter at breast height, and relative height as the ratio between height at any sampled point and total height. Relative diameter against relative height was then plotted for the fitting and validation datasets (Figure 1).

Model fitting and validation

Twenty-eight well-known stem taper equations comprising 18 single taper functions, 4 segmented taper models and 6 variable-exponent taper models (Table 2) were fitted to our data. Non-linear regression analysis was carried out and generalised least squares were used for fitting and estimating the parameters in the 28 models. Models for which all parameters tested significant

Table 1 Descriptive statistics for fitting and validation data sets of *Betula alnoides* in Guangxi, China

Data	No. of trees	No. of data points	Diameter at breast height (cm)			Height (m)		
			Range	Mean	Standard deviation	Range	Mean	Standard deviation
Fitting	90	1147	11.0–42.2	20.3	6.0	9.9–28.7	18.8	4.0
Validation	29	348	10.4–31.0	19.3	5.6	11.5–23.2	18.3	3.1

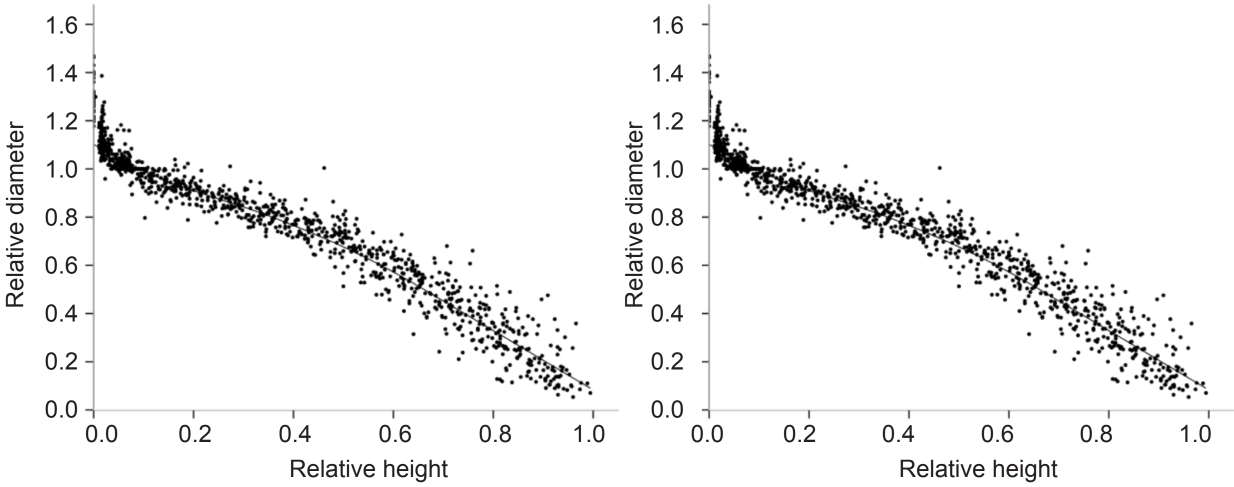

Figure 1 Plots of relative diameter against relative height for (a) fitting and (b) validation datasets of *Betula alnoides* in Guangxi, China

at the 5% level were chosen for further analysis. The remaining models were evaluated with a ranking system of four statistical criteria: per cent unexplained variation (1–PVE) in which PVE means per cent explained variation, bias (B), absolute bias (AB) and mean squared error (MSE) using the following equations (Sakici et al. 2008):

$$1-\mathrm{PVE}=\frac{\sum_{i=1}^{n}\left(d_i-\hat{d}_i\right)^2}{\sum_{i=1}^{n}\left(d_i-\bar{d}\right)^2}$$

$$B=\sum_{i=1}^{n}\frac{\left(d_i-\hat{d}_i\right)}{n}$$

$$AB=\sum_{i=1}^{n}\frac{\left|d_i-\hat{d}_i\right|}{n}$$

$$\mathrm{MSE}=\frac{\sum_{i=1}^{n}\left(d_i-\hat{d}_i\right)^2}{n-m}$$

where d_i, $\hat{d}_i$ and $\bar{d}$ = observed, predicted and mean diameters respectively, n = number of observations in the fitting dataset and m = number of model parameters.

We used the method proposed by Poudel and Cao (2013) to obtain the relative rank of each model for each of the four statistical criteria. The relative rank of model i is defined as:

$$R_i = 1 + \frac{(w-1)(S_i - S_{min})}{S_{max} - S_{min}}$$

where R_i = relative rank of model i (i = 1, 2, …, w), w = total number of models, S_i = value of each statistic criterium produced by model i, S_{min} = minimum value of S_i and S_{max} = maximum value of S_i. The average relative rank of these four statistical criteria for a given model was calculated and sorted, and the best to worst models were determined.

The Student's paired *t*-test was applied to the validation data to test the validity of these functions. The function was excluded if its predicted values differed significantly from observed values at the 5% level. Selected taper models were further assessed using box plots for diameter residuals by relative height along the stem (5, 15, 25 and so on up to 95%). Graphs were used to determine which stem sections could obtain good predictions of taper. Assumptions of residual homoscedasticity were also analysed by means of graphical analysis (Cellini et al. 2012, Özçelik & Crecente-Campo 2016).

To determine whether or not severe multicollinearity was present among independent variables in each selected model, condition number was further calculated as the square root of the ratio between the largest and smallest eigenvalues for their correlation matrix. Condition number higher than $1000^{0.5}$ indicated potentially severe multicollinearity (Myers 1990). Autocorrelation was assessed for these models using the Durbin-Watson test. Positive (or negative) autocorrelation was present in a model if its Durbin-Watson test value was close to zero (or four), while test values close to two indicated no autocorrelation (Durbin & Watson,

Table 2 Taper functions evaluated in the present study

Model	Expression	Reference
Single taper equation		
1	$(\frac{d}{D})^2 = b_1 - b_2 \cdot (\frac{h}{H-1.3})$	Munro (1966)
2	$(\frac{d}{D})^2 = (\frac{b_1 \cdot X^{1.5}}{10}) + (\frac{b_2 \cdot (X^{1.5} - X^3) \cdot D}{10^2}) + (\frac{b_3 \cdot (X^{1.5} - X^3) \cdot H}{10^3}) + (\frac{b_4 \cdot (X^{1.5} - X^{32}) \cdot H \cdot D}{10^5}) + (\frac{b_5 \cdot (X^{1.5} - X^{32}) \cdot \sqrt{H}}{10^3}) + (\frac{b_6 \cdot (X^{1.5} - X^{40}) \cdot H^2}{10^6})$	Bruce et al. (1968)
3	$(\frac{d}{D})^2 = b_1 \cdot (T-1) + b_2 \cdot (T^2 - 1)$	Kozak et al. (1969)
4	$(\frac{d}{D})^2 = b_1 \cdot (1 - 2 \cdot T + T^2)$	Kozak et al. (1969)
5	$\frac{d}{D} = b_1 \cdot X + b_2 \cdot \frac{(H-h)(h-1.3)}{D} + b_3 \cdot \frac{(H-h)(h-1.3) \cdot H}{D} + b_4 \cdot \frac{(H-h)(h-1.3) \cdot (H+h+1.3)}{D}$	Bennett and Swindel (1972)
6	$\frac{d}{D} = b_1 + b_2 \cdot X + b_3 \cdot X^2 + b_4 \cdot X^3 + b_5 \cdot X^4$	Cervera (1973)
7	$(\frac{d}{D})^2 = b_1 \cdot X + b_2 \cdot X^2 + b_3 \cdot X^3$	Coffre (1982)
8	$\frac{d^2}{D^2} = X^2 + b_1 \cdot (X^3 - X^2) + b_2 \cdot (X^8 - X^2) + b_3 \cdot (X^{40} - X^2)$	Real and Moore (1986)
9	$(\frac{d}{D})^2 = b_1 + b_2 \cdot T + b_3 \cdot T^2 + b_4 \cdot T^3 + b_5 \cdot T^4 + b_6 \cdot T^5$	Jiménez et al. (1994)
10	$d = b_1 \cdot D^{b_2} \cdot (H-h)^{b_3} H^{b_4}$	Demaerschalk (1972)
11	$(\frac{d}{D})^2 = b_1 \cdot \left[\frac{(H-h)^{b_2}}{b_3 \cdot H^{b_2+1} + b_4 \cdot H^{b_2}}\right]$	Demaerschalk (1973)
12	$(\frac{d}{D})^2 = b_1 \cdot (\frac{1}{D^2 \cdot H}) \cdot (\frac{H-h}{H})^{b_2} + b_3 \cdot (\frac{H-h}{H})^{b_4}$	Demaerschalk (1973)
13	$\frac{d}{D} = (\frac{H-h}{H-1.3})^{b_1}$	Ormerod (1973)
14	$d = b_1 \cdot D \cdot (H-h)^{b_2}$	Newberry and Burkhart (1986)
15	$d = b_1 \cdot D \cdot (\frac{H-h}{H-1.3})^{b_2}$	Newberry and Burkhart (1986)
16	$(\frac{d}{D})^2 = b_1 \cdot (1 - \frac{h}{H})^{b_2}$	Reed and Green (1984)
17	$\frac{d}{D} = (1 - T^{b_1})^{\frac{1}{b_2}}$	Forslund (1990)

(continued)

Table 2 (continued)

Model	Expression	Reference
18	$d = D\cdot\left[b_1 + b_2\cdot\log\left[1 - T^{\frac{1}{3}}\cdot\left[1 - e^{(-\frac{b_1}{b_2})}\right]\right]\right]$	Biging (1984)
Segmented taper equation		
19	$(\frac{d}{D})^2 = b_1\cdot(T-1) + b_2\cdot(T^2-1) + b_3\cdot(b_5 - T)^2\cdot I_1 + b_4\cdot(b_6 - T)^2\cdot I_2 \quad I_1 = \begin{cases}1 \text{ if } T \geq b_5\\ 0 \text{ if } T < b_5\end{cases} \quad I_2 = \begin{cases}1 \text{ if } T \geq b_6\\ 0 \text{ if } T < b_6\end{cases}$	Max and Burkhart (1976)
20	$(\frac{d}{D})^2 = b_1\cdot Z + b_2\cdot Z^2 + b_3\cdot(Z - b_5)\cdot I_1 + b_4\cdot(Z - b_6)\cdot I_2 \quad I_1 = \begin{cases}1 \text{ if } Z \geq b_5\\ 0 \text{ if } Z < b_5\end{cases} \quad I_2 = \begin{cases}1 \text{ if } Z \geq b_6\\ 0 \text{ if } Z < b_6\end{cases}$	Valenti and Cao (1986)
21	$(\frac{d}{D})^2 = Z^2\cdot(b_1 + b_2\cdot Z) + (Z - b_5)^2\cdot\left[b_3 + b_4\cdot(Z + 2\cdot b_5)\right]\cdot I \quad I = \begin{cases}1 \text{ if } Z \geq b_5\\ 0 \text{ if } Z < b_5\end{cases}$	Parresol et al. (1987)
22	$d = D.\left(\frac{h}{1.3}\right)^{b5}$ if $h < 1.30$	Farrar (1987)
	$d = D\cdot X + b_1\cdot\frac{(H-h)\cdot(h-1.3)}{H^2} + b_2\cdot\frac{D\cdot(H-h)\cdot(h-1.3)}{H^2} + b_3\cdot\frac{D^2\cdot(H-h)\cdot(h-1.3)}{H^2} + b_4\cdot\frac{(H-h)\cdot(h-1.3)\cdot(2\cdot H - h - 1.3)}{H^3}$ if $1.30 \leq h \leq H$	
Variable-form taper equation		
23	$d = b_1\cdot D^{b_2}\cdot b_3^D\cdot\left[\frac{1-\sqrt{T}}{1-\sqrt{p}}\right]^{\left[b_4\cdot T^2 + b_5\cdot \mathrm{Log}(T+0.001) + b_6\cdot\sqrt{T} + b_7\cdot e^T + b_8\cdot(\frac{D}{H})\right]}$	Kozak (1988)
24	$d = b_1\cdot D^{b_2}\cdot b_3^D\cdot\left[1-\sqrt{T}\right]^{\left[b_4\cdot T^2 + (\frac{b_5}{T}) + b_6\cdot D + b_7\cdot H + b_8\cdot(\frac{D}{H})\right]}$	Muhairwe (1999)
25	$d = b_1\cdot D^{b_2}\cdot\left[1-\sqrt{T}\right]^{\left[b_3\cdot T + b_4\cdot T^2 + (\frac{b_5}{T}) + b_6\cdot T^3 + b_7\cdot D + b_8\cdot(\frac{D}{H})\right]}$	Muhairwe (1999)
26	$d = b_1\cdot D^{b_2}\cdot(1-T)^{b_3\cdot T^2 + b_4\cdot T + b_5}$	Lee et al. (2003)
27	$d = b_1\cdot D^{b_2}\cdot H^{b_3}\cdot\left[\frac{1-T^{\frac{1}{3}}}{1-p^{\frac{1}{3}}}\right]^{\left[b_4\cdot T^4 + b_5\cdot\left(\frac{1}{e^{\frac{D}{H}}}\right) + b_6\cdot\left(\frac{1-T^{\frac{1}{3}}}{1-p^{\frac{1}{3}}}\right)^{0.1} + b_7\cdot\left(\frac{1}{D}\right) + b_8\cdot H^{1-\left(\frac{h}{H}\right)^{\frac{1}{3}}} + b_9\left(\frac{1-T^{\frac{1}{3}}}{1-p^{\frac{1}{3}}}\right)\right]}$	Kozak (2004)
28	$\frac{d}{D} = b_1\cdot\left(\frac{H-h}{H-1.37}\right)\left(\frac{H}{1.37}\right)^{b_2 + b_3\cdot T + b_4\cdot T^2}$	Sharma and Parton (2009)

D = diameter at breast height (1.3 m) outside bark (cm), H = tree total height (m), h = height above ground level(m), d = diameter outside bark at height h (cm), $b_{i=}$ coefficients to be estimated, p = relative height of inflection point where the taper curve changes from neiloid to paraboloid X; X = (H – h)/(H – 1.3), T = h/H and Z = (H – h)/H

1951). The best taper model was determined for *B. alnoides* based on statistical and graphical analysis. The diameter residuals obtained from the best model were plotted against observed diameters, then the fitting and validation datasets were pooled together to refit the best equation. All statistical analyses were conducted using R software version 3.2.1 (2015).

RESULTS

Among the 28 stem taper functions tested, 8 functions, i.e. 6, 9, 11, 21, 22, 23, 27 and 28 (Table 2), had at least one parameter without significant difference from zero at the 0.05 level (Table 3). These functions were therefore excluded from further analysis. Models 23 and 27, which were excluded in the present study, have performed well for a wide range of species in other studies (e.g. Yang et al. 2009, Rojo et al. 2005, Sakici et al. 2008). The four statistical criteria and their ranking were calculated for each of the remaining 20 models (Table 4). Markedly different performances were seen between models within each model group. Models 2, 19 and 24 performed the best among simple, segmented and variable-form models respectively. Models 2 and 24 differed little in mean rank value, indicating that both models were the most suitable for stem taper stimulation of *B. alnoides*.

Student's paired *t*-test showed significant differences between observed and estimated diameters for equations 3, 4, 13, 14, 17 and 18. All six equations were single models with largely high-ranking values and were excluded from further analysis (Table 5). Subsequently, the remaining 14 models were evaluated for their accuracy in estimating diameter (Figure 2). The box-plots showed a homogeneous distribution of the diameter residuals for all the models, i.e. no heteroscedasticity was observed. Among the nine remaining single taper functions, model 2 gave the most precise diameter predictions along the whole stem. Other functions either overestimated stem diameters at the lower sections (5–25%) or underestimated stem diameters at the middle sections (35–75%). As for the two remaining segmented models, model 19 performed better than model 20. Among the three remaining variable-form models, models 24 and 25 had similar diameter residual distributions, i.e. medians largely distributed near zero. Diameter residuals fluctuated noticeably near the stem butt and top for model 26. In general, variable-form taper functions appeared more accurate than the single and segmented functions. As a whole, the inter-quartile ranges of all 14 models were smaller at the lower sections than those at higher sections, indicating that the models performed better for the lower than upper sections.

The condition numbers of models 2, 7, 25 and 26 were much higher than $1000^{0.5}$, indicating the potential presence of severe multicollinearity in these models (Table 6). Condition numbers were much lower than this criterion in other 10 models, i.e. no obvious multicollinearity was indicated in these models, including model 24. The Durbin-Watson test indicated severe autocorrelation in model 10, while weak autocorrelation was detected in the other models according to Durbin and Watson (1951). Nonetheless, autocorrelation may be disregarded in these models because the parameter estimates and the predicted values remain unbiased in the presence of autocorrelation (Kozak 1997). Based on statistical and graphical analysis, model 24 was the optimum model for *B. alnoides*. The scatterplot of residuals obtained from this function against observed diameters showed equal distribution across the center of the ordinate axis with no trend of increasing error variance (Figure 3). When fitting and validation datasets were pooled to refit the function, the parameter estimates of the refitted model 24 all differed significantly from zero and their values fell within the preliminary confidence limits (Table 7).

DISCUSSION

Of the three groups of tested functions, there were 14 equations in which parameter estimates were all significant, with no evident difference between observed and predicted diameters for *B. alnoides*. Models 2, 19 and 24 were the best single, segmented and variable-form taper equations for this species respectively. Four equations, including model 2, showed severe multicollinearity and were not considered further for *B. alnoides* because severe multicollinearity may inflate the variance of the predicted values (Kozak 1997). Ten taper functions were found appropriate for application in describing stem taper of this species. In particular, model 24, a variable-form equation, had the lowest ranking value on the basis of the four fitting statistics with

Table 3 Parameter estimates for 28 taper functions based on the fitting dataset for *Betula alnoides*

Model	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	p
Single taper equation										
1	1.1372	1.1936								
2	10.0535	-3.2611	45.1656	6.2009	20.3105	-270.5559				
3	-2.1228	0.8703								
4	1.3489									
5	1.0545	0.0689	0.0057	-0.0032						
6	0.1400	-0.3218 ns	6.9538	-11.1336	5.3938					
7	1.0000	-0.9355	0.8643							
8	-2.1063	0.3669	0.0191							
9	1.4316	-7.7025 ns	32.9947 ns	-69.4933 ns	64.2827 ns	-21.5160 ns				
10	1.2752	0.9534	0.7693	-0.7668						
11	37.8995 ns	1.5378	0.1627 ns	27.3704 ns						
12	1.7478	0.6219	1.1697	1.6690						
13	0.7083									
14	0.1752	0.5878								
15	1.0541	0.7691								
16	1.2316	1.5372								
17	1.5068	1.0889								
18	1.2411	0.4297								
Segmented taper equation										
19	-3.5710	1.7139	-1.7665	345.5065	0.6634	0.0420				
20	0.3822	0.8703	5.0000	-0.3000	2.0000	2.0000				
21	2.2164	-1.0081	-56.4961 ns	27.1328	4.0697					
22	17.2237 ns	0.6523 ns	-0.0096	-9.0697	-0.0594 ns					
Variable-form taper equation										
23	1.3164	0.8712	1.0028	0.6288 ns	-0.2403 ns	0.2962	-0.1187	0.1648		0.1356 ns
24	1.1847	0.9636	1.0018	0.3517	-0.0038	-0.0020	0.0091	0.1681		
25	1.0069	1.0214	0.6025	-0.8723	-0.0043	0.8080	0.0056	0.0691		
26	1.3870	0.9448	2.7322	-3.4532	1.8332					
27	1.1829 ns	0.9820 ns	-0.0265 ns	-2.8621 ns	-0.9903 ns	9.1135 ns	-5.9256 ns	0.8491 ns	-9.5670 ns	0.0489 ns
28	1.0667	0.2333 ns	-0.5197	0.2849						

ns = not significant at 5% level

Table 4 Fit statistics and their ranks of the taper functions

Model	B	RK_B	AB	RK_{AB}	MSE	RK_{MSE}	PVE	RK_{PVE}	Mean RK	Rank
Single taper equation										
1	-0.0399	1.0000	1.1763	5.0716	3.5854	6.1492	0.0554	6.1629	4.5959	13
2	0.0064	2.2395	0.8573	1.0000	1.5781	1.1754	0.0243	1.1764	1.3978	2
3	0.0902	4.4830	1.2229	5.6664	2.8635	4.3604	0.0442	4.3671	4.7192	15
4	0.6014	18.1688	1.6641	11.2977	4.3203	7.9701	0.0668	7.9907	11.3568	19
5	0.0086	2.2984	1.1436	4.6542	2.5455	3.5725	0.0392	3.5654	3.5226	7
7	0.0525	3.4737	1.2081	5.4775	2.8173	4.2460	0.0435	4.2549	4.3630	12
8	0.1746	6.7426	0.9625	2.3427	2.2059	2.7310	0.034	2.7316	3.6370	8
10	0.0483	3.3613	1.1843	5.1737	2.7677	4.1231	0.0427	4.1266	4.1962	10
12	0.0056	2.2181	1.1753	5.0588	2.6958	3.9449	0.0416	3.9502	3.7930	9
13	0.5943	17.9787	1.1666	4.9478	3.4301	5.7644	0.053	5.7781	8.6172	17
14	0.5037	15.5532	2.3459	20.0000	9.1753	20.0000	0.1417	20.0000	18.8883	20
15	0.0631	3.7575	1.1858	5.1929	2.7770	4.1461	0.0429	4.1586	4.3138	11
16	0.1096	5.0024	1.1899	5.2452	2.8167	4.2445	0.0435	4.2549	4.6867	14
17	0.6698	20.0000	1.3114	6.7960	4.2802	7.8708	0.0661	7.8785	10.6363	18
18	0.1123	5.0747	1.0294	3.1966	1.9996	2.2198	0.0309	2.2346	3.1814	5
Segmented taper equation										
19	0.1333	5.6369	0.9035	1.5897	1.6727	1.4098	0.0257	1.4008	2.5093	4
20	0.0902	4.4830	1.2229	5.6664	2.8735	4.3852	0.0442	4.3671	4.7254	16
Variable-form taper equation										
24	0.0074	2.2663	0.8659	1.1098	1.5073	1.0000	0.0232	1.0000	1.3440	1
25	0.0192	2.5822	0.8775	1.2578	1.5665	1.1467	0.0241	1.1443	1.5328	3
26	0.0868	4.3920	1.0713	3.7314	2.2909	2.9416	0.0353	2.9401	3.5013	6

B = bias, AB = absolute bias, MSE = mean squared error, PVE = per cent explained variation, RK = relative rank of model

Table 5 Student's paired *t*-test on the validation dataset for *Betula alnoides*

Model	Paired differences				Significance level
	Mean	Standard deviation	Standard error mean	t	
Single taper equation					
1	-0.067	1.669	0.091	-0.740	0.460
2	0.032	1.092	0.059	0.553	0.580
3	14.861	7.464	0.400	37.140	0.000 *
4	0.594	1.869	0.100	5.929	0.000 *
5	-0.027	1.382	0.074	-0.363	0.717
7	0.025	1.473	0.079	0.315	0.753
8	0.080	1.218	0.065	1.230	0.220
10	0.035	1.461	0.078	0.448	0.654
12	0.021	1.461	0.078	0.272	0.786
13	0.460	1.533	0.082	5.597	0.000 *
14	14.890	7.416	0.398	37.457	0.000 *
15	0.0352	1.461	0.078	0.450	0.653
16	0.063	1.468	0.079	0.803	0.422
17	0.535	1.735	0.093	5.754	0.000 *
18	0.466	3.836	0.206	2.264	0.024*
Segmented taper equation					
19	0.085	1.103	0.059	1.430	0.154
20	0.049	1.481	0.079	0.618	0.537
Variable-form taper equation					
24	0.004	1.057	0.057	0.071	0.943
25	0.006	1.088	0.058	0.096	0.923
26	0.057	1.310	0.070	0.812	0.417

*significant at 0.05 level

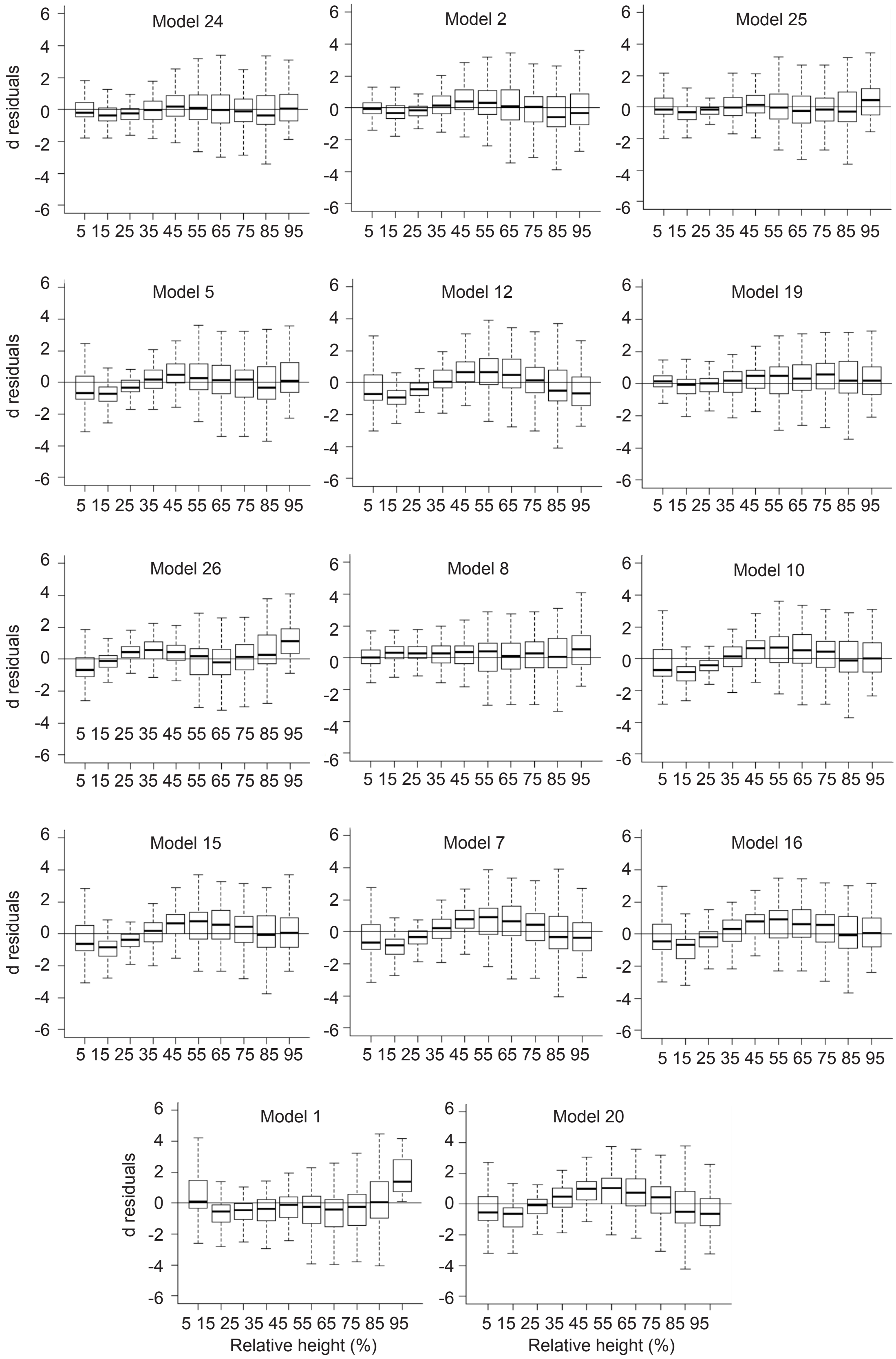

Figure 2 Box plots of diameter (d) residuals against relative stem height classes for the 14 models; the boxes represent the interquartile ranges with their edges being 25th and 75th percentiles, maximum and minimum errors are represented as the upper and lower horizontal lines crossing the vertical bars respectively

Table 6 Multicollinearity and autocorrelation of the 14 models

Model	Condition index	Durbin-Watson test
Single taper equation		
1	-	-
2	67.17	0.991
5	18.00	1.396
7	204.10	0.999
8	9.31	0.999
10	3.77	0.793
12	1.07	1.208
15	1.01	1.293
16	-	-
Segmented taper equations		
19	7.36	1.478
20	9.87	1.478
Variable-form equations		
24	21.51	1.125
25	157.87	1.08
26	76.72	1.287

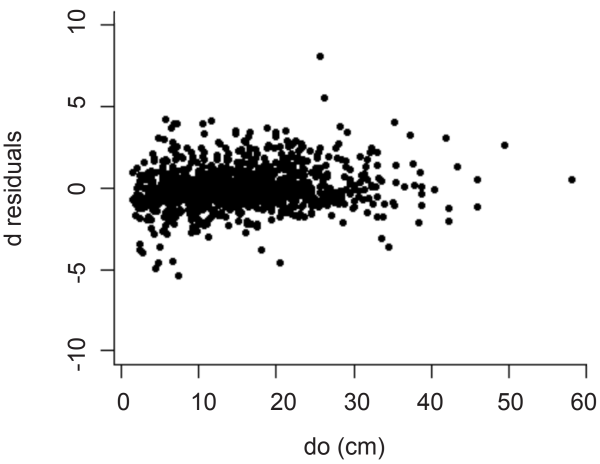

Figure 3 Plot of diameter (d) residuals against observed values of outside bark diameter (do) for stem taper model 24

a condition number below $1000^{0.5}$ (Myers 1990), and thus performed well in fitting and validation. This model estimated diameters along the stem most accurately and was recommended for use in *B. alnoides* plantations.

However, there were disadvantages for variable-form taper equations such as model 24. For example, they cannot be integrated analytically for calculation of merchantable and total stem volumes. For this purpose, numerical integration methods and iterative procedures should be conducted for estimating merchantable height, i.e. the height to any specific diameter (Gómez-García et al. 2013). Model 19, a segmented equation, could provide a closed-form solution for estimating merchantable and total volume in *B. alnoides.* The model had two inflection points at 4.2 and 66.3% on *B. alnoides* trees. Similar inflection points have been reported for oak (5.5 and 59.9%, Pompa-García et al. 2009), *Eucalyptus pilularis* (6.8 and 70.0%, Muhairwe 1999) and *E. grandis* (6.5 and 79.0%, Muhairwe 1999). The heights of both inflection points were located approximately below breast height and at crown base for *B. alnoides,* indicating that the stem profile could be separated into three parts (neiloid, paraboloid and conoid) from the base to the top of stem which was consistent with many previous studies (e.g. Leites & Robinson 2004, Brooks et al. 2008). The model could thus satisfactorily describe the stem profile of *B. alnoides.*

Many studies have shown that model 23 performs well across a range of species (Rojo et al. 2005, Sakici et al. 2008, Yang et al. 2009), and different values have been proposed for inflection points in this model, e.g. 7 and 57% of total stem height for both loblolly and slash pine

Table 7 Parameter estimates and standard errors (SE) of the refitted stem taper model 24 for *Betula alnoides*

Parameter	Estimate	Standard error
b_1	1.0886	0.0869
b_2	1.0433	0.0365
b_3	0.9998	0.0015
b_4	0.1875	0.0147
b_5	0.0087	0.0008
b_6	-0.0051	0.0012
b_7	0.0154	0.0011
b_8	0.2338	0.0176

(Fang et al. 2000), 25% for *Pinus oocarpa* (Perez et al. 1990) and 37 and 40% for *Picea glauca* (Yang et al. 2009). In the present study, model 23 was not suitable for *B. alnoides* because at least one parameter was not significantly different from zero when starting values of the inflection points were attempted as 5 to 95% at intervals of 10%. Model 24, which is a modification of model 23 with p removed (Muhairwe 1999), was the most suitable model for *B. alnoides*. The improved performance following exclusion of p for model 24 could perhaps be explained by the degree of variability inherent in stem form for broadleaf versus coniferous species. The coniferous species for which model 23 was reported suitable, vary less in stem form than *B. alnoides,* which is a broadleaf species. Future research should evaluate and develop specific stem taper equations for broadleaf species such as *B. alnoides*. Some models, e.g. model 2, are single taper equations that performed well in goodness of fit and box-plot analysis of residuals but showed severe multicollinearity in the present study. They could be modified to eliminate multicollinearity and thus be applicable in practice.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science and Technology of China (2012BAD21B0102). The authors are grateful to RY Huang, BG Yang and SK Deng of the Experimental Center of Tropical Forestry, Chinese Academy of Forestry and to TL Lin and GM Liang of the Laoshan Forest Farm, Baise City for their assistance in the field.

REFERENCES

Bennett FA & Swindel BF. 1972. *Taper Curves for Planted Slash Pine*. Research Note No. 179. USDA Forest Service, Asheville.

Biging GS. 1984. Taper equations for second-growth mixed conifers of Northern California. *Forest Science* 30: 1103–1117.

Brooks JR, Jiang L & Ozcelik R. 2008. Compatible stem volume and taper equations for Brutian pine, Cedar of Lebanon, and Cilicica fir in Turkey. *Forest Ecology and Management* 256: 147–151.

Bruce R, Curtiss L & Van Coveering C. 1968. Development of a system of taper and volume tables for red alder. *Forest Science* 14: 339–350.

Cellini JM, Galarza M, Burns SL, Martinez-Pastur GJ, Lencinas MV. 2012. Equations of bark thickness and volume profiles at different heights with easy-measurement variables. *Forest Systems* 21: 23–30.

Cervera JM. 1973. El área basimétrica reducida, el volumen reducido y el perfil. *Montes* 174: 415–418.

Clark A, Souter RA & Schlaegel BE. 1991. *Stem Profile Equations for Southern Tree Species*. Research Paper SE-282. USDA Forest Service, Asheville.

Coffre M. 1982. *Modelos fustales*. Tesis de Ingeniería Forestal, Universidad Austral de Chile, Valdivia.

Demaerschalk J. 1972. Converting volume equations to compatible taper equations. *Forest Science* 18: 241–245.

Demaerschalk J. 1973. Integrated systems for the estimation of tree taper and volume. *Canadian Journal of Forest Research* 3: 90–94.

De-Miguel S, Mehtätalo L, Shater Z, Kraid B, Pukkala T. 2012. Evaluating marginal and conditional predictions of taper models in the absence of calibration data. *Canadian Journal of Forest Research* 42: 1383–1394.

Durbin J & Watson GS. 1951. Testing for serial correlation in least squares regression. II. *Biometrika* 38: 159–177.

Fang ZX & Bailey RL. 1999. Compatible volume and taper models with coefficients for tropical species on Hainan Island in Southern China. *Forest Science* 46: 1–12.

Fang ZX, Borfers 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: 1–12.

Farrar RM. 1987. Stem profile functions for predicting multiple product volumes in natural longleaf pines. *Southern Journal of Applied Forestry* 11: 161–167.

Forslund R. 1990. The power function as a simple stem profile examination tool. *Canadian Journal of Forest Research* 21: 193–198.

Gomat HY, Deleporte P, Moukini R. et al. 2011. What factors influence the stem taper of *Eucalyptus*: growth, environmental conditions, or genetics? *Annals of Forest Science* 68: 109–120.

Gómez-García E, Crecnte-Campo F, DiÉGuez-Aranda U. 2013. Selection of mixed-effects parameters in a variable–exponent taper equation for birch trees in northwestern Spain. *Annals of Forest Science* 70: 707–715.

Jiang LC, Brooks JR, Wang JX. 2005. Compatible taper and volume equations for yellow-poplar in West Virginia. *Forest Ecology and Management* 213: 399–409.

Jiang LC & Liu RL. 2011. Segmented taper equations with crown ratio and stand density for Dahurian larch (*Larix gmelinii*) in northeastern China. *Journal of Forestry Research* 22: 347–352.

Jiménez J, Aguirre O, Niembro R, Navar J & Dominguez A. 1994. Determinación de la forma externa de *Pinus hartwegii* Lindl. en el noreste de México. *Forest Systems* 3: 175–182.

Kozak A, Munro D & Smith J. 1969. Taper functions and their application in forest inventory. *Forestry Chronicle* 45: 278–283.

Kozak A. 1997. Effects of multicollinearity and autocorrelation on the variable-exponent taper functions. *Canadian Journal of Forest Research* 27: 619–629.

Kozak A. 1988. A variable-exponent taper equation. *Canadian Journal of Forest Research* 18: 1363–1368.

Kozak A. 2004. My last words on taper equations. *Forestry Chronicle* 80: 507–515.

Lee WK, Seo JH, Son YM, Lee KH & Von Gadow K. 2003. Modeling stem profiles for *Pinus densiflora* in Korea. *Forest Ecology and Management* 172: 69–77.

Leites, LP, Robinson AP. 2004. Improving taper equations of loblolly pine with crown dimensions in a mixed-effects modeling framework. *Forest Science* 50: 204–212.

Lumbres RIC, Abino AC, Pamolina NM, Clora FG, Lee YJ. 2016. Comparison of stem taper models for the four tropical tree species in Mount Makiling, Philippines. *Journal of Mountain Science* 13: 536–545.

Martin AJ. 1981. *Taper and Volume Equations for Selected Appalachian Hardwood Species.* Research Paper NE-490. USDA Forest Service, Princeton.

Max T & Burkhart H. 1976. Segmented polynomial regression applied to taper equations. *Forest Science* 22: 283–289.

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: 251–269.

Munro DD. 1966. *The Distribution of Log Size and Volume Within Trees: A Preliminary Investigation.* University of British Columbia, Vancouver.

Myers RH. 1990. *Classical and Modern Regression with Applications.* Second edition. Duxbury Press, Belmont.

Newberry J & Burkhart H. 1986. Variable form stem profile models for loblolly pine. *Canadian Journal of Forest Research* 16: 109–114.

Newnham RM. 1988. *A Variable-Form Taper Function.* Information Report PI-X-83. Petawawa National Forestry Institute, Chalk River.

Ormerod D. 1973. A simple bole model. *Forestry Chronicle* 49: 136–138.

Özçelik R & Crecente-Campo F. 2016. Stem taper equations for estimating merchantable volume of Lebanon cedar trees in the Taurus Mountains, southern Turkey. *Forest Science* 62: 78–91.

Parresol B, Hotvedt J & Cao Q. 1987. A volume and taper prediction system for bald cypress. *Canadian Journal of Forest Research* 17: 250–259.

Perez DN, Burkhart HE & Stuff CT. 1990. A variable-form taper function for *Pinus oocarpa* Schiede in Central Honduras. *Forest Science* 36: 186–191.

Pompa-García M, Corral-Rivas JJ, Hernández-Díaz JC, Alvarez González JG. 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: 293–300.

Poudel KP & Cao QV. 2013. Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. *Forest Science* 59: 243–252.

Real PL & Moore JA. 1986. *An individual tree system for Douglas fir in the inland north-west.* Pp 1037–1044 in Ek AR et al. (eds) *Forest Growth Modeling and Prediction.* General Technical Report NC-120. USDA Forest Service, St. Paul.

Reed D & Green E. 1984. Compatible stem taper and volume ratio equations. *Forest Science* 30: 977–990.

Rojo A, Perales X, Sánchez-Rodríguez F, ÁLvarez-González JG & Von Gadow K. 2005. Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (Northwestern Spain). *European Journal of Forest Research* 124: 177–186.

Sakici OE, Misir N, Yavuz H & Misir M. 2008. Stem taper functions for *Abies nordmanniana* subsp. *Bornmulleriana* in Turkey. *Scandinavian Journal of Forest Research* 23: 522–533.

Sharma M & Parton J. 2009. Modeling stand density effects on taper for jack pine and black spruce plantations using dimensional analysis. *Forest Science* 55: 268–282.

Shuaibu RB. 2015. Developing stem taper equation for *Tectona grandis* (teak) plantation in Agudu Forest Reserve, Nasarawa State, Nigeria. *NSUK Journal of Science and Technology* 5: 199–206

Valenti M & Cao QV. 1986. Use of crown ratio to improve loblolly pine taper equations. *Canadian Journal of Forest Research* 16: 1141–1145.

Yang Y, Huang S & Meng SX. 2009. Development of a tree-specific stem profile model for white spruce: a nonlinear mixed model approach with a generalized covariance structure. *Forestry* 82: 541–555.

Zeng J, Guo WF, Zhao ZG, Weng QJ, Yin GT & Zheng HS. 2006. Domestication of *Betula alnoides* in China: current status and perspectives. *Forest Research* 12: 379–384. (In Chinese)

Zeng J, Wang ZR, Zhou SL, Bai JY & Zheng HS. 2003. Allozyme variation and population genetic structure of *Betula alnoides* from Guangxi, China. *Biochemical Genetics* 41: 61–75.

Zeng J, Zhao ZG, Weng QJ, Guo JJ. 2010. *Questions and Answers on High-Yield Cultivation Techniques for Betula alnoides*. China Forestry Publishing House, Beijing. (In Chinese)