A quantile regression approach to model stand survival in Chinese fir plantations

Hanyue Chen[a,b] , Quang V. Cao[c] , Yihang Jiang[a] , Jianguo Zhang[a] , and Xiongqing Zhang

[a] Key Laboratory of Tree Breeding and Cultivation of the National Forestry and Grassland Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China; [b] Collaborative Innovation Center of Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, P.R. China; [c] School of Renewable Natural Resources, Louisiana State University Agricultural Center, Baton Rouge, LA 70803, USA

Corresponding author: Xiongqing Zhang (email: xqzhang85@caf.ac.cn)

Abstract

The development of stand survival models can provide an important basis for the sustainable management of forest re- sources. In a new approach developed in this study, parameters of four survival quantile regression models were predicted from a quantile associated with a current stand density. The curves from these quantile regression models were then used to project future stand density for that stand. A three-fold cross-validation revealed that the quantile regression approach outper- formed the least squares method based on three evaluation statistics, especially for longer projection lengths. These results were consistent for all four survival models evaluated. The best survival model is Clutter–Jones model, without constraints, but its ln(N)–ln(Dq) trajectories (N = stand density and Dq = quadratic mean diameter) from the quantile regression showed the linear self-thinning trend.

Key words: Chinese fir, stand survival, quantile regression, three-fold cross-validation, self-thinning

Introduction

A major goal to achieve sustainable management of forest is to develop and evaluate a model system for stand growth and yield prediction (Özçelik et al. 2018). Mortality in this sys- tem can be classified as regular (or noncatastrophic) and ir- regular (or catastrophic). The former mortality results from competition within a stand and the latter results from ex- ternal factors such as fire, wind, snow, and insect outbreaks (VanClay 1995; Thapa and Burkhart 2015). Mortality is one of the most important functions in dynamic growth and yield model systems (Álvarez-González et al. 2010), and reg- ular mortality is included in growth models of even-aged stands (Monserud 1976;Monserud and Sterba 1999; Zhang et al. 2017). However, forest mortality or survival models were overlooked due to insufficient long-term databased on per- manent plots to model or predict trees in the early period (Thapa and Burkhart 2015).

Whole-stand survival models predict the number of sur- viving trees per unit area for a given initial number of trees and the corresponding age. Stand survival has been employed as a driver for growth and yield predictions for plantations (Dieguéz-Aranda et al. 2005;Lonsdale et al. 2015; Scolforo et al. 2019). The ordinary least squares (OLS) procedure is the most commonly used method in forestry to obtain regression coefficients (Zhang et al. 1993; Diéguez-Aranda et al. 2005). The OLS method addresses only the conditional mean or the central effects of the covariates (Özçelik et al. 2018) and might not perform well for asymmetric and abnormally distributed data (Lanssanova et al. 2021). In the presence of outliers, the use of this method may lead to inadequate, low precision, and biased estimates (Koenker and Bassett 1978; Lanssanova et al. 2021).

The quantile regression (QR) method was developed by Koenker and Bassett (1978)to examine how covariates affect the location,size, and shape of the entire response distribu- tion (Tian 2019). QR is an alternative to solve problems caused by the presence of outliers or asymmetries in the data distri- bution (Lanssanova et al. 2021). In ecology, QR is advocated as an effective method to detect relationships between variables when there is either no relationship or a weak relationship between means (Cade and Noon 2003).

Recently, there has been a growing interest in applying QR to solve forestry regression problems, such as the esti- mation of maximum values (Weiskittel et al. 2011). QR has been used in research dealing with maximum size-density and self-thinning boundary lines (Zhang et al. 2005; Ducey and Knapp 2010), potential diameter increments (Pretzsch and Biber 2010),diameter percentiles (Mehtätalo et al. 2008), branch growth (Miao et al. 2021), maximum crown width of individual trees (Russell and Weiskittel 2011), and crown modeling (Sun et al. 2017). In addition, QR has been used in studies of tree taper (Cao and Wang 2015), tree height– diameter relationships (Zang et al. 2016; Özçelik et al. 2018; Lanssanova et al. 2021), insect and disease transmission rates

Table 1. Means (and standard deviations) of stand variables, by site.

Site	Coordinates	*n*	*A*	*Dq*	*N*
Fujian	27◦05\N 117◦43\E	192	16 (8.05)	11.1 (4.39)	5103 (2427.05)
Jiangxi	27◦30\N 114◦33\E	120	12 (4.60)	9.4 (2.35)	5866 (2176.39)
Guangxi	22◦06\N 106◦43\E	132	19 (6.34)	11.9 (2.86)	3612 (1825.51)
Sichuan	28◦47\N 105◦23\E	192	16 (7.98)	9.1 (2.77)	4798 (2343.81)

Note: Four density levels (i.e., 2 m × 1.5 m (3333 trees/ha), 2 m × 1 m (5000 trees/ha), 1 m × 1.5 m (6667 trees/ha), and 1 m × 1 m (10 000 trees/ha)) were replicated three times at each site. *n*, total number of mea- surements across all plots; *A*, stand age (year); *Dq*, stand quadratic mean of diameter (cm);*N*, surviving number of trees per hectare.

(Evans and Gregoire 2007;Evans and Finkral 2010), and con- structing growth curves (Muggeo et al. 2013). QR can also be used to fitlongitudinal data (Geraci and Bottai 2007; Karlsson 2008;Bohora and Cao 2014).

Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook.) is a fast-growing evergreen coniferous tree species and is a major plantation timber species in southern China (Zhang et al. 2019). As a native tree species, Chinese fir occupies an im- portant position in terms of area, production, and resource utilization, and the current plantation area of this species is about 9.215 million ha, accounting for about 28.54% of the national afforestation area (Zhang et al. 2013). Its wood has good raw material quality, straight board shape, and high resistance to decay (Zhang et al. 2020). Thus, establishing an accurate stand survival model for Chinese fir planta- tions can provide a scientific basis for subsequent effective management of forests. Although there have been studies to establish stand survival models for plantation stands, such as Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco) whole stand survival model by Zhang et al. (1993) and the systematic estimation of longleaf pine (*Pinus palustris* Mill.) survival equation in the United States by Gonzalez-Benecke et al.

(2012), there are few studies on modeling and projecting stand-level survival of Chinese fir plantations by using the QR approach.

The objective of this study is to evaluate the use of least squares (LS) and QR methods for estimating parameters of different stand-level survival models to predict self-thinning trajectories for Chinese fir plantations.

Data

In this study, four experimental sites of Chinese fir were located in Jiangxi, Fujian, Guangxi, and Sichuan provinces in southern China, with bare-root seedlings as experi- mental materials. Four planting densities were designed by random block design: 2 m × 1.5 m (3333 trees/ha), 2 m × 1 m (5000 trees/ha), 1 m × 1.5 m (6667 trees/ha), and 1 m × 1 m (10 000 trees/ha). Three replications were set for each level, resulting in a total of 12 plots. Each plot includes an area of 0.06 ha (20 m × 30 m), sur- rounded by two rows of similarly treated trees with the same spacing, forming a buffer zone. The diameters of 1.3 m (Diameter at Breast Height, DBH) of all trees were mea- sured in each plot.Table 1 describes the statistical data of stand factors of each site, and Fig. 1 shows the change of stand density and quadratic mean diameter with increasing age.

Out of the four sites, Jiangxi, Fujian, and Sichuan belong to the middle-subtropical climate zone, and Guangxi belongs to the southern-subtropical climate zone. In Jiangxi, the yellow–brown earth was developed by sand shale, and the ex- perimental site has been set up since 1981. Because the snow- storm in 1998 in Jiangxi led to the deaths of a large number of trees, the data after 1999 were not used. Fujian, Guangxi, and Sichuan are mainly red earth developed on granite and other parent materials, all of which were established in 1982. The data of 1985–2010, 1990–2009, and 1985–2013 were measured, respectively. Samples were measured every 1–3 years.

Methods

Survival models

All of the following four survival models satisfy the prop- erties stated by Clutter et al. (1983) for stand survival models.

Model 1: Clutter–Jones

The stand survival model by Clutter and Jones (1980) has the following form:

$1/b_1$

(1) $N_i = {}_{0i}1000$

where N_i is the predicted surviving number of trees/hain plot i at age A_i, N_{0i} is the number of trees/hain plot i at age A_{0i}, and b_j is the regression coefficients, where j = 1, 2, 3.

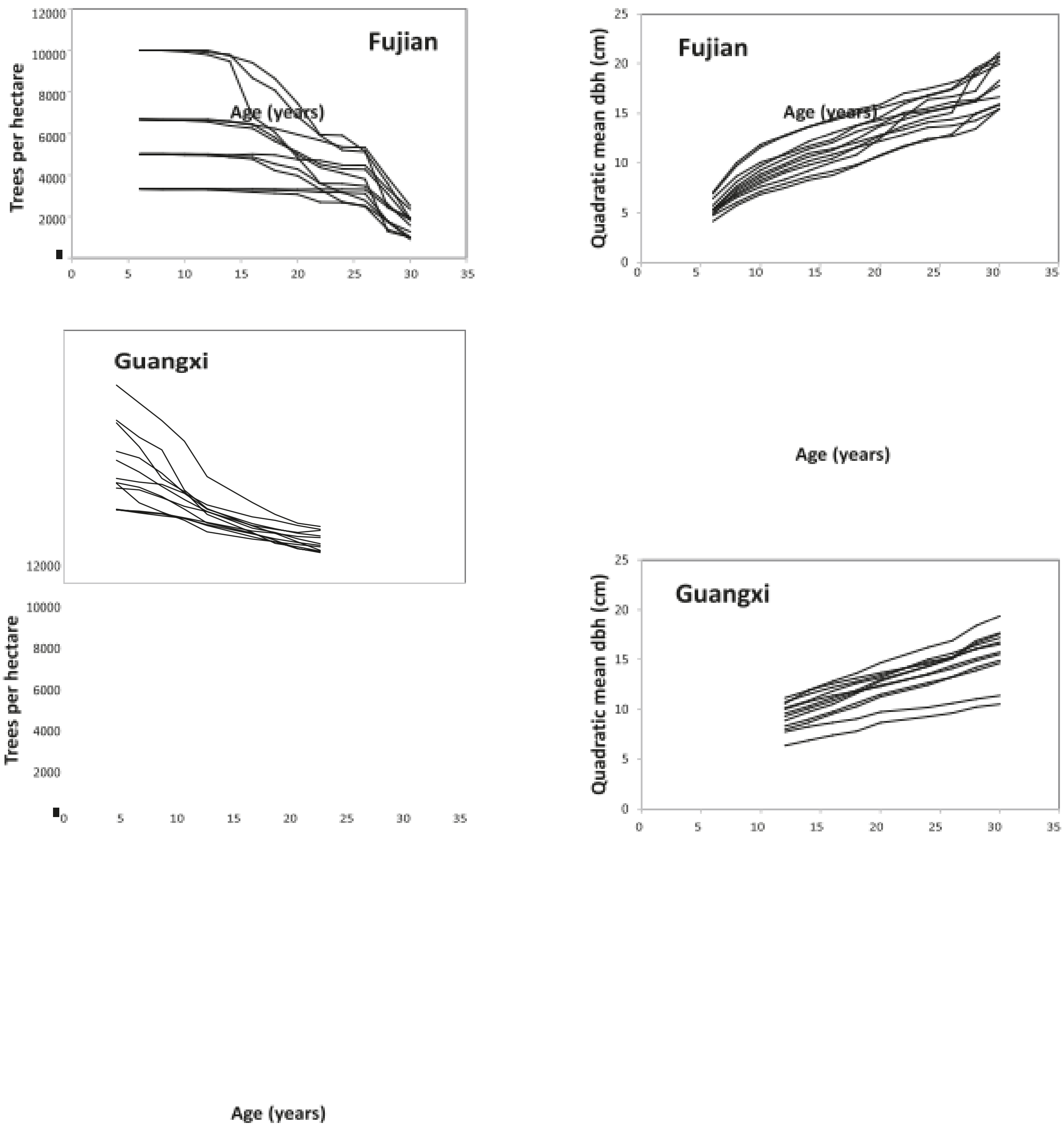
Note that stand density and age were scaled by 1000 and 100, respectively, to ensure that the parameter b_2 was significant. Because no mortality occurred before age 6, we set A_{0i} = 6 years, and N_{0i} becomes the planting den- sity at plot i. Future survival (N_{2i}) at age A_{2i} was predicted from:

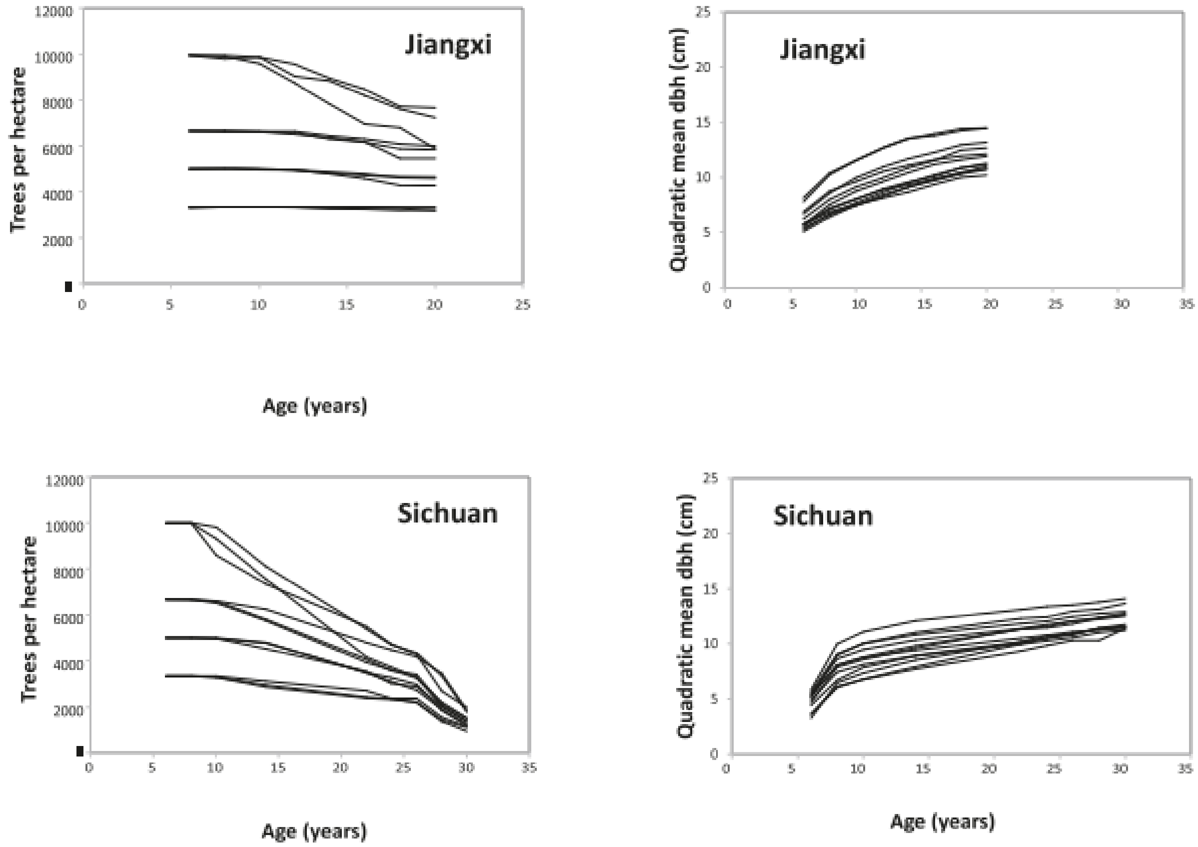
$1/b_1$ ⌋

N_{2i}

where N_{1i} is the observed stand survival (number of trees/ha) of plot i at age A_{1i}.

Fig. 1. Graphs of stand density (left) and quadratic mean diameter (right) versus age, by site.





Model 2: modified Clutter–J

Equation 1was modified as follows:

Age (years)

(3) $$\bar{N}_i = 1000\left[\left(\frac{N_{0i}}{1000}\right)^{b_1} + b_2\left\{\left(\exp\left(\frac{A_i}{100}\right)-1\right)^{b_3} - \left(\exp\left(\frac{A_{0i}}{100}\right)-1\right)^{b_3}\right\}\right]^{1/b_1}$$

Future survival ($\tilde{N}_{2i}$) was predicted from:

(4) $$\tilde{N}_{2i} = N_{1i} \times \frac{\left[\left(\frac{N_{0i}}{1000}\right)^{b_1} + b_2\left\{\left(\exp\left(\frac{A_{2i}}{100}\right)-1\right)^{b_3} - \left(\exp\left(\frac{A_{0i}}{100}\right)-1\right)^{b_3}\right\}\right]^{1/b_1}}{\left[\left(\frac{N_{0i}}{1000}\right)^{b_1} + b_2\left\{\left(\exp\left(\frac{A_{1i}}{100}\right)-1\right)^{b_3} - \left(\exp\left(\frac{A_{0i}}{100}\right)-1\right)^{b_3}\right\}\right]}$$

Model 3: Weibull

The Weibull cumulative distribution function (cdf) (Bailey and Dell 1973) is as follows:

(5) $$F(x) = 1 - \exp\left[-\left(\frac{x}{b_1}\right)^{b_2}\right],\ x \geq 0$$

Because $1 - F(x)$ varies from 1 to 0 as x goes from 0 to ∞, a survival model can be constructed as:

(6) $$\tilde{N}_i = N_{0i} \{\exp [-\left(\frac{A_i}{b_1}\right)^{b_2}]]$$

By using eq. 6 to express future survival (N_{2i}) and current survival (N_{1i}), future survival was computed as follows:

(7) $$\tilde{N}_{2i} = N_{1i} - N_{0i} \{\exp [-\left(\frac{A_{1i}}{b_1}\right)^{b_2}] - \exp\left[-\left(\frac{A_{2i}}{b_1}\right)^{b_2}\right]\}$$

Model 4: Dagum

Dagum (1977)proposed the following cdf:

$$F \quad \frac{1}{1 + \left(\frac{x}{b_1}\right)^{b_2}}, x \geq 0.$$

Similar to model 3, model 4 based on the Dagum cdf is as follows:

(9) $$N_i = N_{0i} \left\{ 1 + \left(\frac{A_i}{}\right)^{b_2} \right\}$$

Future survival was computed in the same manner as in eq. 7:

(10) $$N_{2i} = N_{1i} - N_{0i} \left\{ 1 + \left(\frac{A_{1i}}{}\right)^{b_2} - 1 + \left(\frac{A_{2i}}{}\right)^{b_2} \right\}$$

Estimation methods

Least squares (LS) approach

Parameters for eqs. 2, 4, 7, and 10 (models 1–4, respectively) were obtained by using the OLS method.

Quantile regression (QR) approach

The same forms in eqs. 1,3,6, and 9 were used to predict the qth quantile, with $y_q(A_i)$ replacing N_{2i}, where $y_q(A_i)$ is the

predicted value of the qth quantile of stand density of plot i at age A_i.

Parameters from the above equations varied with q and were estimated from:

(11) $$b_j = c_{0j} + c_{1j}q + c_{2j}q^2, j = 1, 2, 3$$

where j has a value of 1, 2, or 3 for models 1 and 2, or 1 or 2 for models 3 and 4, and the coefficients c were obtained by minimizing

$$(12)\qquad S = \sum_{N_i \geq \hat{y}_q(A_i)} \left\{ \sum_q q \left(N_i - \hat{y}_q(A_i) \right) \right\} + \sum_{N_i < \hat{y}_q(A_i)} \left\{ \sum_q (1-q) \left(\hat{y}_q(A_i) - N_i \right) \right\}$$

where N_i is the observed stand survival of plot i at age A_i, and $\sum_q(\cdot)$ denotes the sum for values of q varying from 0.1 to 0.9 by 0.1.

The following steps were taken to compute $\hat{N}_{2i}$, projected future stand survival at age A_{2i}, from current stand density (N_{1i} at age A_{1i}):

1. Numerically solve the quantile q such that curve for the qth QR passes through the point (A_{1i}, N_{1i}).
2. Compute parameters b_1, b_2, and b_3 from this q value by using eq. 11.
3. Using these parameter estimates, compute N_{2i} from eqs. 2, 4, 7, or 10.

When the quantile q computed from step 1 was less than 0.1, q was set to 0.1. On the other hand, if $q > 0.9$, then it was set to 0.9.

Evaluation

A three-fold cross-validation was carried out to evaluate the two approaches. This process required the construction of three groups of data, each containing a fit data set (consist- ing of two replications from each treatment) and a validation data set (the remaining replication) (Table 2). Stand survival of the validation data was predicted by using parameter es- timates obtained from the fit data. The following evaluation statistics were computed:

$$(13)\qquad \text{Mean difference:}\quad \mathrm{MD} = \sum_i \left(N_{2i} - \hat{N}_{2i} \right) / n$$

$$(14)\qquad \text{Mean absolute difference:}\quad \mathrm{MAD} = \sum_i \left| N_{2i} - \hat{N}_{2i} \right| / n$$

$$(15)\qquad \text{Fit index:}\quad \mathrm{FI} = 1 - \sum_i \left(N_{2i} - \hat{N}_{2i} \right)^2 / \sum_i \left(N_{2i} - \bar{N}_2 \right)^2$$

Mean difference (MD) and mean absolute difference (MAD) are measures of accuracy and precision, respectively, whereas

Table 2. Distribution of replications, by group.

Group	Fit data	Validation data
	Replication	
1	2, 3	1
2	1, 3	2
3	1, 2	3

fit index (FI) is a measure of both (Cao 2019). The better ap- proach should have a smaller absolute value of MD, smaller MAD, and higher FI.

The relative position of each method was displayed in terms of a relative rank, introduced by Poudel and Cao (2013). The relative rank of method i is defined as:

(16) $R_i = 1 + \frac{(k-1)(S_i - S_{\min})}{S_{\max} - S_{\min}}$ for minimization objective, and (17) $R_i = k - \frac{(k-1)(S_i - S_{\min})}{S_{\max} - S_{\min}}$ for maximization objective

where R_i is the relative rank of method i (i = 1, 2, …, k),k is the number of methods evaluated, S_i is the evaluation statis- tic produced by method i, $S_{\min}$ is the minimum value of S_i, and $S_{\max}$ is the maximum value of S_i. Note that R_i is a real number rather than an integer, and for either minimization or maximization objective, the best method receives a rank of 1, whereas the worst method receives a rank of k (k = 8 in this study).

Results

Table 3 presents estimates of the parameters of the four survival models for each group from the LS method. The regression coefficients for predicting parameters b_j of these models from the quantile q are shown in Table 4 or each group.

Table 5 shows values of the evaluation statistics computed on the validation data for the eight methods (4 models × 2es- timation methods), based on various projection lengths rang- ing from 2 to 10 years. As expected, the predictions were worse as projection length increased. On average, the MAD values increased by 144% and the FI values decreased by 20% as projection length increased from 2 to 10 years.

Generally speaking, all methods resulted in overprediction (negative MD) for short-term projection (six years or less) and underprediction for longer projection lengths. The best MD value corresponded to a different method for each projec- tion length. On the other hand, the QR Clutter–Jones method scored best in terms of MAD for all projection lengths, and best in terms of FI for projection lengths of six years or more. For each projection rank, a relative rank was computed for each of the eight methods, based on the sum of ranks from the three evaluation statistics (Table 6). The “best” method was the QR Weibull for short projections (less than six years) and the QR Clutter–Jones for long projections (more than six years). The overall winner was the QR Clutter–Jones, with the QR modified Clutter–Jones being a close second (rank of 1.34). The QR modified Clutter–Jones was also the best method for six-year projections.

Discussion

Survival models

The Clutter–Jones model was implemented in the past to predict stand survival of loblolly pine plantations. The FI value ranged from 0.7423 (Cao 2019) to 0.7439 (Cao 2022) for five-year projections. It has generally been taken for granted

that the inclusion of additional stand variables into the base models automatically results in better predictions (Wang et al. 2021). In this study, planting density was incorporated into the survival model and appeared to significantly improve the

performance of the model. The FI value varied from 0.8099 (10-year projection) to 0.9711 (two-year projection) for the LS method, as compared with 0.8280–0.9730 for the QR method (Table 5).

The difference between the Clutter–Jones and modified Clutter–Jones models is that $\left(\frac{A_{2i}}{100}\right)^{b_3}$ in eq. 1 is replaced with $\left(\exp\left(\frac{A_i}{100}\right)-1\right)^{b_3}$ in eq. 3. Graphs drawn from these two survival models were similar, leading to their equiva- lent performances. The modified Clutter–Jones was slightly better when their parameters were estimated by using the LS method (an overall rank of 3.58 versus 3.76, Table 6). The reverse was true in the case of the QR estimation method; the Clutter–Jones was slightly better with an over- all rank of 1.00 versus 1.34 for the modified Clutter–Jones (Table 6).

For short projection periods (less than six years), the QR Weibull and Dagum models performed well. The two mod- els did not perform that well for longer projection lengths (Table 6). The reason might be that they contain only two pa- rameters as compared to three parameters for the Clutter– Jones and modified Clutter–Jones models. Attempts to add one more parameter to either the Weibull or Dagum model was not successful because the extra parameter was not sig- nificant.

LS versus QR

The QR method outperformed the LS method, regardless of the model used or the projection length. The sum of the overall ranks (last column ofTable 6) across models is 7.12 for the QR method, which is much lower as compared to 22.01 for the LS method.

Survival curves from the LS method (also called the mean regression) are similar to those from the QR method when $q = 0.5$ (or the median regression), and the model with the quantile adjusted to 0.50 covers most ofthe observations pre- dicted by the LS method. On the other hand, the QR method allows for variable curve shapes that fitvarious trajectories for each planting density (see Fig. 2for the QR Clutter–Jones). The LS method was based on a single set ofparameters ofthe survival model, whereas the parameters from the QR method varied as functions of the quantile q, determined by the location of the point (A_{1i}, N_{1i}) in Fig. 2. The survival curves from the QR method were therefore much more flexible than those from the LS method, resulting in better evaluation statistics, especially for longer projections.

Table 3. Regression coefficients for the four survival models from the least squares method, by group.

Model	Parameter	Group 1	Group 2	Group 3	All
1. Clutter–Jones	b_1	− 0.6364	− 0.4275	− 0.6808	− 0.5648
	b_2	37.0738	30.4101	33.9567	32.3947
	b_3	3.8082	3.8397	3.8284	3.7835

2. Modified Clutter–Jones	b_1	− 0.6710	− 0.4512	− 0.7165	− 0.5917
	b_2	14.3111	11.3394	12.9624	12.4130
	b_3	3.4369	3.4305	3.4547	3.4007
3. Weibull	b_1	27.4074	29.1114	28.3844	28.2974
	b_2	2.7713	2.8808	2.7533	2.7911
4. Dagum	b_1	24.6220	26.2789	25.3685	25.4016
	b_2	− 3.6065	− 3.6129	− 3.5109	− 3.5644

Table 4. Regression coefficients for predicting parameters b_j of the four survival models from the quantile q, by group. The general model is $b_j = c_{0j} + c_{1j}q + c_{2j}q^2$, where j = 1, 2, 3.

Model	Parameter	Group 1	Group 2	Group 3	All
1. Clutter–Jones	b_1 c_{01}	− 1.4089	− 1.9435	− 1.6948	− 1.6832
	c_{11}	− 0.3920	2.5249	2.1978	1.5344
	c_{21}	0.2354	− 2.7005	− 3.2847	− 1.8325
	b_2 c_{02}	100.8100	101.3290	105.6280	94.3408
	c_{12}	− 150.4340	− 152.0570	− 151.0040	− 149.7530
	c_{22}	350.6200	351.1310	349.4650	356.9400
	b_3 c_{03}	4.0469	4.1059	4.0797	3.9763
	c_{13}	1.0120	0.1854	0.1165	0.6484
	c_{23}	1.2992	2.7043	3.2252	2.1281
2. Modified Clutter-Jones	b_1 c_{01}	− 0.5518	− 2.4934	− 2.0767	− 2.0661
	c_{11}	− 4.8073	3.9313	2.1400	2.0159
	c_{21}	4.6663	− 3.7522	− 2.5991	− 1.8748
	b_2 c_{02}	21.7300	54.2570	68.1292	45.5848
	c_{12}	5.8364	3.5001	− 23.8825	21.0324
	c_{22}	− 7.6519	− 49.3041	− 13.6195	− 35.1151
	b_3 c_{03}	2.9203	4.3213	4.2841	4.0468
	c_{13}	4.3184	0.2118	0.7385	1.2226
	c_{23}	− 3.0403	0.8922	1.0522	0.2606
3. Weibull	b_1 c_{01}	21.5094	21.1328	21.1588	21.3477
	c_{11}	12.9061	15.1900	14.3568	13.5862
	c_{21}	− 3.9027	− 5.4310	− 4.2561	− 3.8060
	b_2 c_{02}	2.3302	2.0655	2.4185	2.2699
	c_{12}	2.8030	3.8109	2.5964	3.1133
	c_{22}	0.1234	− 0.1514	0.7587	0.1742
4. Dagum	b_1 c_{01}	17.4938	15.9263	17.3238	17.1273
	c_{11}	14.4373	21.9137	15.0621	16.2643
	c_{21}	− 3.1355	− 9.0503	− 2.3061	− 4.1309
	b_2 c_{02}	− 3.5389	− 3.3235	− 3.7329	− 3.6317
	c_{12}	− 2.0285	− 2.2987	− 1.6157	− 1.6151
	c_{22}	− 0.4405	− 1.0006	− 1.0678	− 1.1593

Improvement of the QR method

QR was successfully used to calibrate tree diameters (Bohora and Cao 2014) and tree taper (Cao and Wang 2015). This method involves two steps: (1) computing a quantile, say 0.27, for a given observation, and (2) predicting for other observations by interpolating between outputs from the QR equations based on 0.2 and 0.3 quantiles (if nine quantiles were used)

Table 5. Evaluation statistics for each method, by projection length. A bold–italic number denotes the best statistic among the eight methods.

Evaluation statistic*	Survival model	Estimation method†	Projection length (years) 2	4	6	8	10
MD	1. Clutter–Jones	LS	−5.72	−3.43	19.57	85.99	200.19
		QR	−20.95	−25.71	−43.65	31.47	102.40
	2. Modified Clutter–Jones	LS	***−4.06***	***−1.29***	21.14	84.64	197.73
		QR	−16.83	−18.73	−33.05	41.06	110.06
	3. Weibull	LS	−17.76	−34.27	***−11.68***	27.01	139.60
		QR	5.76	−2.70	−39.86	−19.02	***35.76***
	4. Dagum	LS	−39.31	−69.92	−55.25	***−7.05***	101.25
		QR	−5.38	−3.36	−15.10	59.92	142.56
MAD	1. Clutter–Jones	LS	258.67	405.73	476.69	524.19	559.99
		QR	***224.45***	***336.28***	***394.39***	***447.85***	***502.29***
	2. Modified Clutter–Jones	LS	258.18	404.47	474.55	520.67	556.51
		QR	228.10	344.63	394.74	459.29	530.28
	3. Weibull	LS	270.04	446.99	539.05	610.54	685.16
		QR	225.10	353.56	431.57	489.59	610.29
	4. Dagum	LS	264.87	445.26	541.64	619.18	681.96
		QR	229.42	357.15	437.72	523.00	657.95

FI	1. Clutter–Jones	LS	0.9711	0.9322	0.9013	0.8545	0.8099
		QR	0.9730	0.9418	***0.9172***	***0.8792***	***0.8280***
	2. Modified Clutter–Jones	LS	0.9714	0.9330	0.9024	0.8561	0.8121
		QR	0.9728	0.9400	0.9151	0.8701	0.8126
	3. Weibull	LS	0.9688	0.9185	0.8810	0.8165	0.7208
		QR	***0.9744***	***0.9422***	0.8985	0.8412	0.7407
	4. Dagum	LS	0.9678	0.9146	0.8772	0.8105	0.7224
		QR	0.9735	0.9401	0.9030	0.8413	0.7370

*MD, mean difference; MAD, mean absolute difference; FI, fit index.

†LS, least squares; QR, quantile regression.

Table 6. Relative rank for each method, by projection length. Each rank was based on the sum of the ranks from the three evaluation statistics. The overall rank was based on the sum of the ranks over all projection lengths. A bold–italic number denotes the best relative rank among the eight methods.

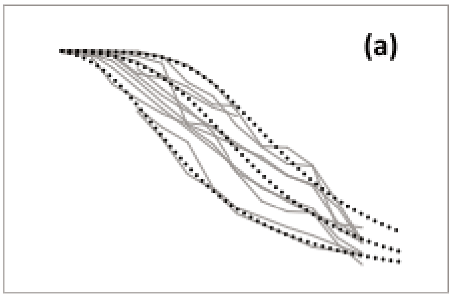
Survival model	Estimation method*	Projection length (years) 2	4	6	8	10	Overall rank
1. Clutter–Jones	LS	4.05	3.10	2.69	6.80	4.39	3.76
	QR	2.56	1.48	1.53	***1.00***	***1.00***	***1.00***
2. Modified Clutter–Jones	LS	3.82	2.93	2.67	6.56	4.22	3.58
	QR	2.56	1.57	***1.00***	2.24	2.08	1.34
3. Weibull	LS	6.37	6.39	4.82	8.00	8.00	6.67
	QR	***1.00***	***1.00***	3.34	3.48	4.14	1.95
4. Dagum	LS	8.00	8.00	8.00	7.55	7.16	8.00
	QR	1.56	1.29	1.52	6.24	7.11	2.84

*LS, least squares; QR, quantile regression.

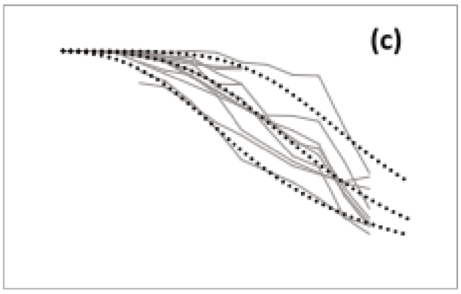
or on 0.25 and 0.50 quantiles (if five quantiles were used).

The QR method developed in this study is an improvement over the method previously implemented by Bohora and Cao (2014) andCao and Wang (2015). After the quantile q is deter-mined for the current stand density, parameters of the sur- vival model are computed from q by usingeq. 11and then em- ployed to provide predictions for future stand densities. The new method is therefore more flexible than the old method and requires no interpolation.

Fig. 2. Graphs of observed survival (solid line) and the 10%, 50%, and 90% quantile regression estimates of the Clutter–Jones model (dotted line) versus stand age for planting densities of (*a*) 10 000, (*b*) 6667, (*c*) 5000, and (*d*) 3333 trees/ha.



Age (years)



Age (years)

Two-step process

Woollons (1998)advocated the use of a two-step process for modeling stand survival. In the first step, a logistic regres- sion is used to predict the probability that the stand has suf- fered from mortality. The second step involves constructing a survival model using only data where mortality has occurred over the growth period. The final model is obtained by com- bining results from these two steps.

In a preliminary analysis, the two-step process was applied to both LS and QR methods. Results indicated that the LS and QR methods actually performed better than their two-step counterparts for all projection lengths, based on the three evaluation statistics. The two-step LS and QR are therefore not included in the evaluation of survival models and their estimation methods.

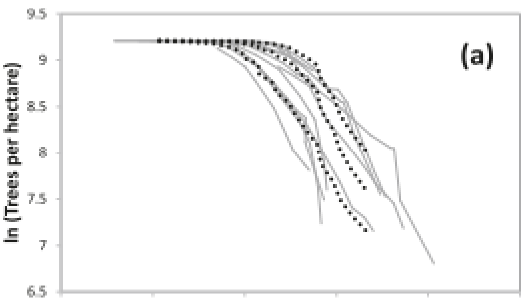
Self-thinning

The self-thinning rule describes mortality related to com- petition within even-aged stands (Hynynen 1993). The rule has been a popular topic of research and discussion in forestry (Reineke 1933; Burkhart 2013). Zhang et al. (2016) have used segmented regression techniques to fit self- thinning trajectories in Chinese fir plantations, and the QR method used in this study is an alternative way to fit the self-thinning boundary line and the maximum density line. The linear relationship for older ages between observed stand density and tree size (in terms of quadratic mean diameter) on a log–log scale indicates that the stands had already en- tered the self-thinning stage (Fig. 3). Also displayed in this figure are trajectories computed from the 10%, 50%, and 90%QR for the Clutter–Jones model. The quadratic mean diam- eter used for each curve is the average value for each age– planting density combination. It can be seen that different self-thinning boundary lines were obtained by choosing dif- ferent quartiles, indicating that different slopes and intercept ranges of self-thinning boundary lines can be calculated by QR, which can provide reference for the subsequent construc- tion of maximum density lines for the density management diagram of Chinese fir plantations. It is notable that these trajectories also exhibit the linear trend found in the curves based on observed values, even though the survival model contains no constraints to accommodate self-thinning.

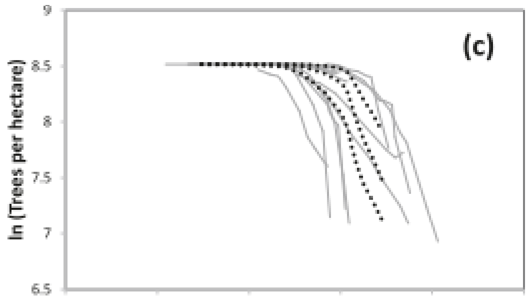
Summary and conclusions

In a new approach developed in this study, parameters of four survival QR models were predicted from a quantile asso- ciated with a current stand density. The curves from these QR models were then used to project future stand density for that stand. A three-fold cross-validation revealed that the QR ap- proach outperformed the LS method based on three evalua- tion statistics, especially for longer projection lengths. These results were consistent for all four survival models evaluated. Even though the best survival model, the Clutter–Jones model, contains no constraints, its ln(N)–ln(Dq) trajectories (N = stand density and Dq = quadratic mean diameter) from the QR still showed a linear trend, which was observed in the self-thinning data. These findings suggest that the Clutter– Jones model might be capable of modeling survival of self- thinned stands of other species as well. Furthermore, the new technique to implement the QR approach proved to perform

Fig. 3. Graphs of stand density versus tree size. The y- and x-axes are natural logarithm of a number of trees/ha and quadratic mean diameter (in cm), respectively. Solid lines represent the observed data, while dotted lines represent the 10%, 50%, and 90% quantile regression estimates from the Clutter–Jones model for planting densities of (a) 10 000, (b) 6667, (c) 5000, and (d) 3333 trees/ha.



ln (Quadratic mean diameter)



ln (Quadratic mean diameter)

better than the method previously applied by Bohora and Cao (2014) andCao and Wang (2015), and should be considered in the future for prediction of stand mortality/survival.

Acknowledgements

The authors appreciate the support of the National Natural Science Foundation of China (No. 31670634), and the National Institute of Food and Agriculture, US Department of Agri- culture, McIntire–Stennis project LAB94379. They would also like to thank Dr. Aiguo Duan for the field work.

Article information

History dates

Received: 5 August 2022

Accepted: 23 September 2022

Version of record online: 8 February 2023

Copyright

Data availability

Data generated or analyzed during this study are not avail- able due to the nature of this research.

Author information

Author ORCIDs

Xiongqing Zhanghttps://orcid.org/0000-0001-5592-6730

Author Contributions

Formal analysis: QVC

Investigation: XZ
Methodology: QVC, XZ

Supervision: XZ

Writing – original draft: HC

Writing – review & editing: QVC,JZ, XZ

Competing interests

The authors declare that they have no known competing fi- nancial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Álvarez-González, J.G., Zingg, A., and Gadow, K.V. 2010. Estimating growth in beech forests: a study based on long term experiments in Switzerland. Ann. For. Sci. **67**: 307–307. doi:10.1051/forest/2009113.

Bailey, R.L., and Dell, T.R. 1973. Quantifying diameter distributions with Weibull function. For. Sci. **19**: 97–104.

Bohora,B.S., and Cao, Q.V. 2014. Prediction of tree diameter growth using quantile regression and mixed-effects models. For. Ecol. Manage. **319**: 62–66. doi:10.1016/j.foreco.2014.02.006.

Burkhart, H.E. 2013. Comparison of maximum size-density relationships based on alternate stand attributes for predicting tree numbers andstand growth. For. Ecol. Manage. **289**: 404–408. doi:10.1016/j.foreco. 2012.10.04.

Cade, B.S., and Noon, B.R. 2003. A gentle introduction to quantile re- gression for ecologists. Front. Ecol. Environ. **1**: 412–420. doi:10.1890/ 1540-9295(2003)001[0412:AGITQR]2.0.CO2.

Cao, Q.V. 2019. A method to derive a tree survival model from any exist- ingstand survival model. Can.J. For. Res. **49**: 1598–1603. doi:10.1139/ cjfr-2019-0171.

Cao, Q.V. 2022. Deriving a tree growth model from any existing stand growth model. Can. J. For. Res. **52**: 169–174. doi:10.1139/ cjfr-2021-0106.

Cao,Q.V., and Wang,J. 2015. Evaluation of methods for calibrating a tree taper equation. For. Sci. **61**: 213–219. doi:10.5849/forsci.14-008.

Clutter, J.L., and Jones, E.P. 1980. Predicting of growth after thinning in oldfield slash pine plantations. USDA For. Serv. Res. Pap. SE-217.

Clutter, J.L., Fortson,J.C., Pienaar, L.V., Brister, G.H., and Bailey, R.L. 1983. Timber management: a quantitative approach. Wiley, New York, NY. 333 p.

Dagum, C. 1977. A new model of personal income distribution: specifi- cation and estimation. Economie Appliquée, **30**: 413–437.

Diéguez-Aranda, U., Castedo-Dorado, F., Álvarez-González, J.G., and Rodríguez-Soalleiro, R. 2005. Modelling mortality of Scots pine (*Pi- nus sylvestris* L.) plantations in the northwest of Spain. Eur. J. For. Res. **124**: 143–153. doi:10.1007/s10342-004-0043-5.

Ducey, M.J., and Knapp, R.A. 2010. A stand density index for complex mixed-species forests in the northeastern United States. For. Ecol. Manage. **260**: 1613–1622. doi:10.1016/j.foreco.2010.08.014.

Evans, A.M., and Finkral, A.J. 2010. A new look at spread rates of exotic diseases in North American forests. For. Sci. **56**: 453–459. doi:10.1093/ forestscience/56.5.453.

Evans,A.M., and Gregoire, T.G. 2007. A geographically variable model of hemlock woolly adelgid spread. Biol. Invasions, **9**: 369–382. doi:10. 1007/s10530-006-9039-z.

Geraci, M., and Bottai, M. 2007. Quantile regression for longitudinal data using the asymmetric Laplace distribution. Biostatistics, **8**(1): 140– 154. doi:10.1093/biostatistics/kxj039. PMID: 16636139.

Gonzalez-Benecke, C.A., Gazan, S.A., Leduc, D.J., Martin, T.A., ,Cropper, W.P., Jr., and Samuelson, L.J. 2012. Modeling survival, yield, volume partitioning and their response to thinning for longleaf pine planta- tions. Forests, **3**: 1104–1132. doi:10.3390/f3041104.

Hynynen, J. 1993. Self-thinning models for even-aged stands of *Pinus sylvestris*, *Picea abies*, and *Betula pendula*. Scand. J. For. Res. **8**: 326–336. doi:10.1080/02827589309382781.

Karlsson, A. 2008. Nonlinear quantile regression estimation of longitu- dinal data. Commun. Stat. Sim. Computat. **37**: 114–131. doi:10.1080/ 03610910701723963.

Koenker, R., and Bassett,G., Jr. 1978. Regression quantiles. Econometrica, **46**: 211–244. doi:10.2307/1913643.

Lanssanova, L.R., Silva, F.A., Machado, S.A., Pelissari, A.L., Figueiredo Filho, A., Fiorentin, L., and Cerqueira, C.L. 2021. Hypsometric rela- tionship in *Tectona grandis* L. F. stands using quantile regression. Sci. For. **49**(131): e3559. doi:10.18671/scifor.v49n131.07.

Lonsdale,J., Xenakis, G., Mencuccini, M., and Perks, M. 2015. A compar- ison of models for quantifying growth and standing carbon in UK Scots pine forests. iForest, **8**: 596–605. doi:10.3832/ifor1403-008.

Mehtätalo, L., Gregoire, T.G., and Burkhart, H.E. 2008. Comparing strate- gies for modeling tree diameter percentiles from remeasured plots. Environmetrics, **19**: 529–548. doi:10.1002/env.896.

Miao, Z., Widagdo, F. R. A., Dong, L., and Li, F. 2021. Prediction ofbranch growth using quantile regression and mixed-effects models: an ex- ample with planted *Larix olgensis* Henry trees in Northeast China. For. Ecol. Manage. **496**: 119407. doi:10.1016/j.foreco.2021.119407.

Monserud, R.A. 1976. Simulation of forest tree mortality. For. Sci. **22**: 438–444. doi:10.1016/0378-1127(76)90006-2.

Monserud, R.A., and Sterba, H. 1999. Modeling individual tree mortality for Austrian forest species. For. Ecol. Manage. **113**: 109–123. doi:10. 1016/S0378-1127(98)00419-8.

Muggeo,V.M.R., Sciandra, M., Tomasello,A., and Calvo, S. 2013. Estimat- ing growth charts via nonparametric quantile regression: a practical framework with application in ecology. Environ. Ecol. Stat. **20**: 519– 531. doi:10.1007/s10651-012-0232-1.

Özçelik, R., Cao, Q.V., Trincado, G., and Göçer, N. 2018. Predicting tree height from tree diameter and dominant height using mixed-effects and quantile regression models for two species in Turkey. For. Ecol. Manage. **419–420**: 240–248. doi:10.1016/j.foreco.2018.03.051.

Poudel, K.P., and Cao, Q.V. 2013. Evaluation of methods to predict Weibull parameters to characterize diameter distributions. For. Sci. **59**: 243–252. doi:10.5849/forsci.12-001.

Pretzsch, H., and Biber, P. 2010. Size-symmetric versus size-asymmetric competition and growth partitioning among trees in forest stands along an ecological gradient in Central Europe. Can. J. For. Res. **40**: 370–384. doi:10.1139/X09-195.

Reineke,L.H. 1933. Perfecting a stand-density index for even-aged forests. J. Agric. Res. **46**: 627–638.

Russell, M.B., and Weiskittel,A.R. 2011. Maximum and stand-level crown width equations for 15 tree species in Maine. North.J. Appl. For. **28**(2): 84–91. doi:10.1093/njaf/28.2.84.

Scolforo., H.F., Mctague, J.P., Burkhart, H., Roise, J., Alvares, C.A., and Stape, J.L. 2019. Modeling whole-stand survival in clonal eucalypt stands in Brazil as a function of water availability. For. Ecol. Manage. **432**: 1002–1012. doi:10.1016/j.foreco.2018.10.044.

Sun, Y., Gao, H., and Li, F. 2017. Using linear mixed-effects models with quantile regression to simulate the crown profile of planted *Pinus sylvestris var. Mongolica* trees. Forests, **8**(11): 446. doi:10.3390/f8110446. Thapa, R., and Burkhart, H.E. 2015. Modeling stand-level mortality of loblolly pine (*Pinus taeda* L.) using stand, climate, and soil variables.

For. Sci. **61**: 834–846. doi:10.5849/forsci.14-125.

Tian, D. 2019. Study on the self-thinning boundary surface for *Larix olgen- sis* plantation. Northeast Forestry University, Harbin, China. [In Chi- nese.]doi:10.27009/d.cnki.gdblu.2019.000819.

Vanclay, J.K. 1995. Synthesis: growth models for tropical forests: A synthesis of models and methods. For. Sci. **41**: 7–42. doi:10.1093/ forestscience/41.1.7.

Wang,T., Xie,L., Miao,Z., Widagdo,F.R.A., Dong,L., and Li,F. 2021. Stand volume growth modeling with mixed-effects models and quantile re- gressions for major forest types in the eastern Daxing'an Mountains, Northeast China. Forests, **12**: 1111. doi:10.3390/f12081111.

Weiskittel, A.R., Hann, D.W., Kershaw, J.A., Jr., and Vanclay, J.K. 2011. Forest growth and yield modeling. Wiley-Blackwell, Hoboken, NJ. 415p.

Woollons, R.C. 1998. Even-aged stand mortality estimation through a two-step regression process. For. Ecol. Manage. **105**: 189–195. doi:10. 1016/S0378-1127(97)00279-X.

Zang, H., Lei,X., and Zeng,W. 2016. Height–diameter equations for larch plantations in northern and northeastern China: a comparison of the mixed-effects, quantile regression and generalized additive models. Forestry, **89**: 434–445. doi:10.1093/forestry/cpw022.

Zhang, L., Moore, J.A., and Newberry, J.D. 1993. A whole-stand growth and yield model for interior Douglas-fir. West. J. Appl. For. **8**: 120– 125. doi:10.1093/wjaf/8.4.120.

Zhang, L., Bi, H., Gove, J.H., and Heath, L.S. 2005.A comparison of alter- native methods for estimating the self-thinning boundary line. Can. J. For. Res. **35**: 1507–1514. doi:10.1139/x05-070.

Zhang, X., Duan, A., and Zhang,J. 2013. Tree biomass estimation of Chi- nese fir (*Cunninghamia lanceolata*) based on Bayesian method. PLoS One, **8**(11): e79868. doi:10.1371/journal.pone.0079868.

Zhang, X., Cao, Q.V., Duan, A., and Zhang,J. 2016. Self-thinning trajec- tories of Chinese fir plantations in southern China. For. Sci. **62**: 594– 599. doi:10.5849/forsci.16-004.

Zhang,X., Cao, Q.V., Duan,A., and Zhang,J. 2017. Modeling tree mortality in relation to climate, initial planting density, and competition in Chinese fir plantations using a Bayesian logistic multilevel method. Can. J. For. Res. **47**: 1278–1285. doi:10.1139/cjfr-2017-0215.

Zhang, X., Cao, V. Q., Lu, L., Wang, H., Duan, A., and Zhang,J. 2019. Use of modified Reineke's stand density index in predicting growth and survival of Chinese fir plantations. For. Sci. **65**: 776–783. doi:10.1093/ forsci/fxz033.

Zhang, X., Wang, Z., Chhin, S., Wang, H., Duan, A., and Zhang,J. 2020. Relative contributions of competition, stand structure, age, and cli- mate factors to tree mortality of Chinese fir plantations: long-term spacing trials in southern China. For. Ecol. Manage. **465**: 118103. doi:10.1016/j.foreco.2020.118103.