RESEARCH PAPER

Predicting individual tree growth using stand-level simulation, diameter distribution, and Bayesian calibration

Xianglin Tian[1,2] · Shuaichao Sun[1] · Blas Mola-Yudego[3] · Tianjian Cao[1]

Abstract

• *Key message* **We propose a methodology to develop a preliminary version of a growth model when tree-level growth data are unavailable. This modelling approach predicts individual tree growth using only one-time observations from temporary plots. A virtual dataset was generated by linking the whole stand and diameter distribution models. The individual tree model was parameterized using Bayesian calibration and a likelihood of diameter distributions.**

• *Context* A key component of tree-level growth and yield prediction is the diameter increment model that requires at least two different points in time with individual tree measurements. In some cases, however, sufficient inventory data from remeasured permanent or semitemporary plots are unavailable or difficult to access.

• *Aims* The purpose of this study was to propose a three-stage approach for modelling individual tree diameter growth based on temporary plots.

• *Methods* The first stage is to predict stand dynamics at 5-year intervals based on stand-level resource inventory data. The second stage is to simulate diameter distribution at 5-year intervals using a Weibull function based on tree-level research inventory data. The final stage is to improve the reliability of individual tree diameter estimates by updating parameters with Bayesian calibration based on a likelihood of diameter distributions.

• *Results* The virtual-data-based diameter increment model provided logical patterns and reliable performances in both tree- and stand-level predictions. Although it underestimated the growth of suppressed trees compared with tree cores and remeasurements, this bias was negligible when aggregating tree-level simulations into stand-level growth predictions.

• *Conclusion* Our virtual-data-based modelling approach only requires one-time observations from temporary plots, but provide reliable predictions of stand- and tree-level growth when validated with tree cores and whole-stand models. This preliminary model can be updated in a Bayesian framework when growth data from tree cores or remeasurements were obtained.

Keywords Diameter increment · Growth and yield · Sparse data · Data aggregation · Parametric uncertainty

✉ Tianjian Cao
   cao@nwafu.edu.cn

[1] Simulation Optimization Lab, Northwest A&F University, Taicheng Road 3, Yangling 712100, Shaanxi, China

[2] Department of Forest Sciences, University of Helsinki, P.O. Box 27, 00014 Helsinki, FI, Finland

[3] School of Forest Sciences, University of Eastern Finland, PO Box 111, 80101 Joensuu, FI, Finland

1 Introduction

Efficient and accurate models for growth and yield are a fundamental tool in forest sciences, playing a key role in forest management, forest planning, ecological studies, or in fact any discipline within the field. Traditionally, large efforts have been placed to design and retrieve the necessary data, often from large databases or long-term observations in permanent plots. Several places have had these data available since long ago: there are long-term experimental plots being still surveyed in Germany that date back to the 1860s (Prodan 1965; Pretzsch 2010). As a second source, plots from national forest inventory (NFI) are a good alternative for data retrieval, although, may lack accuracy and detail as are not necessarily designed for growth and yield models. In this case, the large amount of data available may compensate these limitations.

INRAE ✦ Springer

However, what are the alternatives when the purpose of analysis is a new species or a new region with limitations in longitudinal data for single trees? For instance, despite its large role in forest resources, Brazil has only recently started a country-level NFI based on a standardized and systematized methodology, which for some time will provide a single measurement (Freitas et al. 2016; da Luz et al. 2018). The same would apply to large countries in Asia, such as Mongolia, with 11.3 million hectares of boreal forest, which has completed its first NFI recently (Altrell 2019). In Europe, despite the large experience in NFI, several countries lack permanent plots as a main source of data (Tomppo et al. 2010). In fact, according to the global forest resource assessment, 40 of the 99 tropical countries in the report did not have at least two comparable NFI measurements on permanent plots that could be the basis for individual tree growth models (FAO 2015; Romijn et al. 2015).

When sufficient data from permanent plots are not available, semitemporary or interval plots are useful to develop forest growth models. Those plots are established in stands of different development stages and measured at least in two points in time. However, local forest regulations often limit the establishment, collection, and availability of extensive plot data along time; but, while remeasurement data could be more difficult, one-time observations in resource inventory data from temporary plots are more commonly available. This type of inventory relies on temporary plot sampling that may provide databases for modelling sample whole-stand models (e.g., Vanclay 2010) and diameter distribution models (e.g., Cao 2004). The basic data units of whole-stand models are stand and site parameters. Consequently, their construction and simulation do not require tree-level information of stand dynamics. In contrast, size-class models use an entire tree class as the basic unit, and individual tree models rely on the information from each tree. This approach, although useful, presents a more restricted suitability for individual tree models.

The individual tree approach, although being more prescriptive in data needs, performs largely better to characterize growth under variable stand conditions and management practices (Weiskittel et al. 2011). For instance, concerning thinnings, a whole-stand model is limited since thinning treatments are designed to selectively remove individual trees, which leads to alteration of the residual stand structure. To address these limitations, an alternative use of temporary plots is to estimate previous stand conditions (e.g., stand basal area 5 years earlier) based on tree cores; thus, a timeline of stand attributes, along with the increment measurement from tree cores, can be used to fit individual tree models. This approach is applied, for example by several versions of the ORGANON (ORegon Growth ANalysis and projectiON) growth model (Hann et al. 2011) and many variants of the western US' Forest Vegetation Simulator (FVS) growth model (Crookston and Dixon 2005).

Another method is to incrementally modify the explanatory variables of tree vigor and competition (Schröder et al. 2002).

If explanatory variables change slowly over time, errors introduced by not having quantitative backdated measures are negligible over a short interval. However, large datasets of tree cores to construct these models are often not available in most forest inventories, and although alternative approaches have been developed to model diameter increments, these are much less accurate (Vanclay 1994).
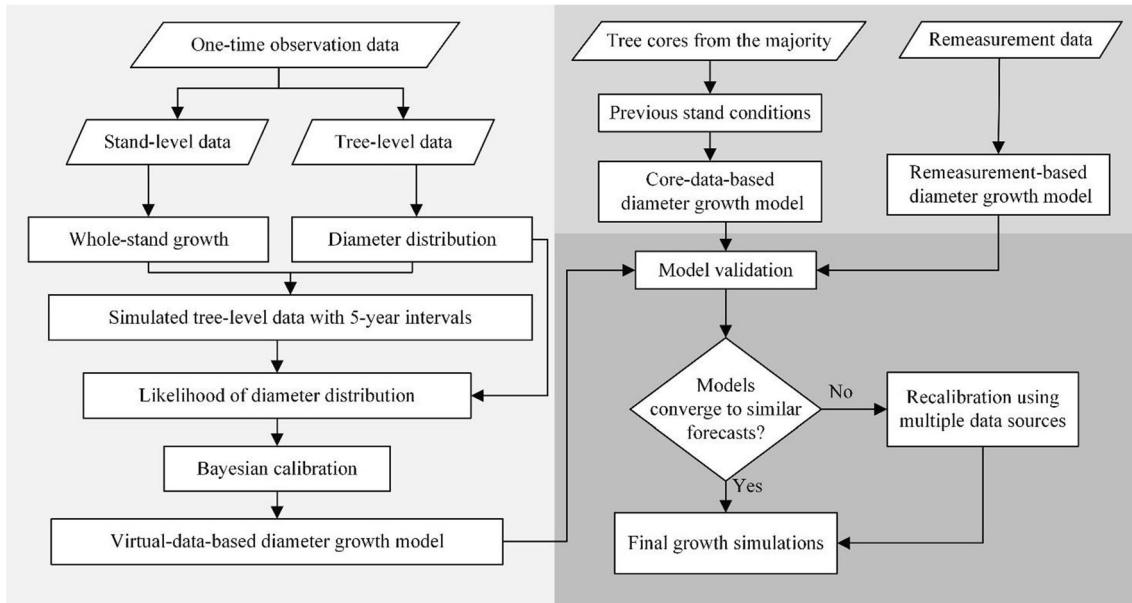
Limited by the availability of data, most modelling efforts are quite opportunistic and often commence with any data available (Weiskittel et al. 2011). In our case, we investigate the growth and yield of *Pinus tabuliformis*, a dominant conifer species native to China which plays an important ecological role in soil erosion and water protection and on the regional socioeconomic development (Wu and Feng 1994), with a wide distribution area and a large altitudinal gradient (Mao and Wang 2011). In recent years, adequate designs of forest resource management, in terms of tending, thinning, timber stand management, and bush management, have been needed for restoring the degraded forest and reducing the risk of forest fires. The lack of proper individual tree level growth models due to data limitations, and the ecological and economic importance of the species was the rationale for our research efforts to provide suitable modelling approaches that could be used in forest planning related to the species. Thus, we aim to establish a modelling approach to predict individual tree diameter growth based on one-time observation data from temporary plots, with special focus on the assessment of the effects on model performance of different data types and the resulting parametric uncertainty.

2 Materials and methods

We propose a three-stage approach to generate virtual data for fitting individual tree growth models by linking stand-level and size-class simulations (Fig. 1). First, we construct a whole-stand model, and then predict stand-level dynamics. Second, we simulate diameter distributions using the Weibull function that is parameterized with observed size-class structure data. Finally, we use the simulated data to fit individual tree growth models with Bayesian calibration using a diameter distribution likelihood. All data used to parameterize models in our study were collected from temporary plots. We collected one reference dataset of tree cores and one small dataset of diameter remeasurements to validate our simulated data modelling approach.

2.1 Study area

Temporary plot data were collected in the Huoditang Forest (108° 21′–108° 39′ E, 33° 18′–33° 28′ N) and the Xunyangba Forest (103° 58′–109° 48′ E, 32° 29′–33° 13′ N), in Ningshan County, Shaanxi, China (Fig. 2). The altitude of all sampling

Fig. 1 Flowchart of the main processes in the individual tree diameter growth modelling

sites varied from 1200 to 1800 m, with a slope gradient ranging from 15 to 42°. The dominant soil is a brown forest soil with an average thickness of about 50 cm. This region has a subtropical humid mountain climate. The mean annual temperature is 8–12 °C, and the mean annual precipitation is 1100 mm. The primeval forests in this region were harvested during the 1960s and 1970s. Now most of the area is covered with secondary forests or plantations.
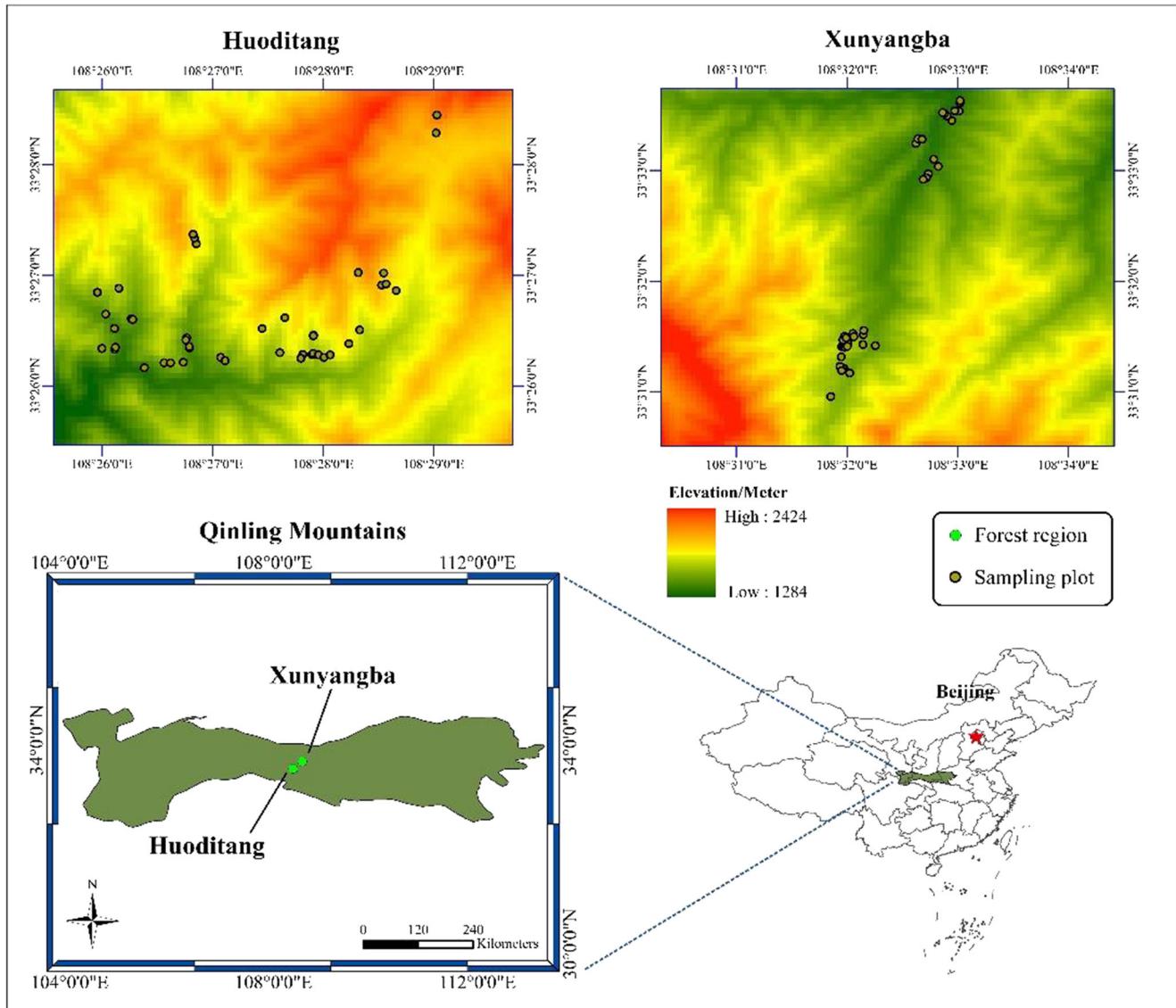
2.2 Modelling data

The data for model parameterization were collected from 215, variable-radius plots using angle gauge sampling and 22 square plots with a fixed area of 0.04 ha (Table 1). The basal area factor of angle gauge sampling is 1 $m^2\,ha^{-1}$. This dataset was provided by the forest resource inventory (1990, 2005, and 2015); most of which comes from level II forest inventory in China (Lei et al. 2009). The level II inventory is designed for forest management planning, while most plots from 1990 and 2005 lack exact coordinates. All data are from temporary plots established in *Pinus tabuliformis* plantations from which only one-time observations were collected (Table 2). One-quarter of plots were randomly withheld for benchmarking, while the remaining three quarters were used for model fitting. Additional data of stem analysis from 11 medium trees across various site types and conditions were combined with plot sample data to calibrate the height-age equation in the whole stand model and the height-diameter equation in the individual tree model, taking into consideration that stem analysis data and plot data perform differently in height curve estimation (Perin et al. 2013).

For each size class (4 cm) in each square plot, increment measures from one or two sample trees were selected randomly, taking two cores in opposite directions at the breast height. This provided 338 tree cores from 169 trees in the 22 square plots. Increment core data were processed specifically to validate the individual tree diameter increment model (Fig. 1). Therefore, individual tree diameters and stand total basal areas from 5 years earlier were back-dated based on tree cores and the successional stages of snags. Besides the tree cores, another dataset was also obtained for model validation (Fig. 1). Trees from 18 of the 22 square plots were remeasured during 2019. The other four plots were unable to locate because of inaccurate coordinate records.

2.3 Model simulations

We generated a virtual dataset for fitting an individual tree model using the following six steps: (1) constructed a whole-stand model based entirely on temporary plot data, (2) using this model, we simulated the development of stand-level attributes for 30, virtual, 1-ha plots from ages 20 to 80 years that include six site fertility classes and five initial densities (6 × 5 = 30 plots), (3) we constructed a diameter distribution model based on temporary plot data, (4) using the diameter distribution model, we simulated the diameter distributions of virtual plots at 5-year intervals, (5) based on diameter distributions and stand density, we generated individual tree diameters, and finally (6), we simulated 30 virtual 1-ha plots among sites and competition conditions, with 13 ((80–20)/5 + 1) measurement times per plot.

INRAE Springer

Fig. 2 Locations of the sampling plots. The 90m-DEM (Digital Elevation Model) was downloaded from the CGIAR-CSI GeoPortal (http://srtm.csi. cgiar.org). The locations only include plots being established after 2005.

Plots of other datasets were from the same compartments but lacking exact coordinates

2.3.1 Whole-stand model

A whole-stand model that estimates stand average height, per hectare basal area, and quadratic mean diameter was built and fitted using temporary plots data. The following basic functions were hypothesized and fit:

$$\overline{h}_{gc} = \exp(\alpha_1 + \alpha_2 t^{\alpha_3}) \tag{1}$$

$$\overline{h}_{30} = \exp\left(\alpha_1 + \left(\ln\overline{h}-\alpha_1\right) \cdot \left(\frac{30}{t}\right)^{\alpha_3}\right) \tag{2}$$

$$\overline{h} = \exp\left(\alpha_1 + \left(\ln\overline{h}_{30}-\alpha_1\right) \cdot \left(\frac{t}{30}\right)^{\alpha_3}\right) \tag{3}$$

$$SDI = N \times \left(\frac{\overline{d}_g}{20}\right)^{\alpha_4} \tag{4}$$

$$G = \alpha_5 \times \overline{h}_{30}^{\ \alpha_6} \times \exp\left(-\frac{\alpha_7 \times \left(\frac{SDI}{1000}\right)^{-\alpha_8}}{t}\right) \tag{5}$$

$$\overline{d}_g = \alpha_9 \times \overline{h}_{30}^{\ \alpha_{10}} \times \exp\left(-\frac{\alpha_{11} \times \left(\frac{SDI}{1000}\right)^{\alpha_{12}}}{t}\right) \tag{6}$$

$$N = \frac{40000 \times G}{\pi \times \overline{d}_g^{\ 2}} \tag{7}$$

where t is the stand age (years), $\overline{h}_{gc}$ is the average stand height in guide curve (m), and $\overline{h}_{30}$ is average height at the reference age (m), of which the reference age is 30 (years) for *Pinus*

Table 1 Characteristics of the two types of plot data used in modelling

Plot type		Variable-radius plot	Square plot
Investigation year		1989-1990, 2004-2005, 2014-2015	2014-2015
Sample size		215	22
Stand age (year)	Min.	15	25
	Mean	46	45
	Max.	87	73
Stand height (m)	Min.	5.0	8.6
	Mean	13.4	12.5
	Max.	25.0	20.4
Mean diameter (cm)	Min.	8.0	12.9
	Mean	22.0	17.4
	Max.	44.0	24.8
Number of trees per hectare	Min.	151	600
	Mean	628	1595
	Max.	2355	3123

tabuliformis. $\overline{h}$ is stand height (m), SDI is stand density index (trees ha^{-1}), N is stand density (trees ha^{-1}), G is stand basal area (m^2 ha^{-1}), and $\overline{d}_g$ is quadratic mean diameter (cm).

There was a lack of records of dominant trees in previous inventory datasets. Therefore, the expected stand average

Table 2 Posterior probability distribution of the parameters in the whole-stand model. The posterior distribution is not analytical and it is characterized here by the mean values of the parameters, the standard deviation (SD), and the maximum a posteriori estimates (MAP). The prior was set as independent uniform distributions, of which the maxima and minima are listed

Parameters	Prior	Mean	SD	MAP
α_1	(− 100, 100)	4.1543	0.3513	4.2016
α_2	(− 100, 0)	− 6.1392	0.3391	− 5.8887
α_3	(− 100, 100)	− 0.3743	0.0659	− 0.3488
α_4	(0, 10)	1.5928	0.1221	1.6080
α_5	(0, 100)	21.1022	5.7949	21.5691
α_6	(− 10, 100)	0.2248	0.1236	0.2086
α_7	(0, 100)	6.9665	2.0722	6.6053
α_8	(0, 10)	2.0750	0.3157	2.0297
α_9	(0, 100)	30.3612	3.6852	32.0450
α_{10}	(− 10, 100)	0.0593	0.0518	0.0107
α_{11}	(0, 100)	20.0725	1.5733	20.1571
α_{12}	(0, 10)	0.0371	0.0602	0.0305
γ_1	(− 100, 100)	− 1.0195169	0.2944	− 1.2737
γ_2	(− 10, 10)	1.0307536	0.2443	0.9113
γ_3	(− 10, 10)	0.2815911	0.0796	0.3956
γ_4	(− 100, 100)	− 0.3399967	0.1129	− 0.3833
γ_5	(− 10, 10)	1.6498416	0.0356	1.6916
γ_6	(− 10, 10)	− 0.7033391	0.03290	− 0.7150

height at the reference age, $\overline{h}_{30}$, was used to track site quality. We assumed that, similar to the widely used site index, the relationship between stand age and stand average height correlates closely with total stand production and is less dependent on thinning intensity (Eichhorn 1902). Site fertility classification can also be implemented based on average height and age relations (von Baur 1876; de Perthuis Laillevault 1803), and this site evaluation method has been widely applied in Chinese forest inventory since the 1950s. Tree height growth has been found to be impacted by stand density (Lynch 1958; Cieszewski and Bella 1993; MacFarlane et al. 2000; Flewelling et al. 2001), and the severity of this impact is related to species tolerance and site conditions (Weiskittel et al. 2011). Furthermore, the intensification of thinning from below can significantly affect the calculation of the stand average height (Assmann 1970; Pretzsch 2010). Based on this, the limitation of using average height was not considered as a severe issue in our study, because thinning was forbidden after 1998 in this region due to the National Forest Conservation Programme (NFCP) for soil and water protection (Li 2004). A guide curve of stand average height (Eq. 1) was fitted with the Bailey–Clutter model (Bailey and Clutter 1974). The equations of $\overline{h}_{30}$ and stand height (Eqs. 2 and 3) were then derived.

Reineke's (1933) stand density index (SDI), which is fitted by regression of number of stems (N) and quadratic mean diameter ($\overline{d}_g$) using fully stocked stand observations, was used to describe competition (Eq. 4). Stand total basal area (G) model (Eq. 5) and quadratic mean diameter model (Eq. 6) were modifications of the Schumacher model (Schumacher 1939), and number of stems was derived (Eq. 7).

Full-stocked stands were classified using the following three steps: (1) fit the Reineke's (1933) equation ln$N$ =

INRAE 🔍 Springer

$\lambda_1 + \lambda_2 \ln \overline{d}_g$ with all stand observations, (2) drop all points that fall below this regression line and repeat the regression, and (3) drop all points falling below the new regression line. The remaining was considered fully stocked. We also validated this method with local stock tables of "normal" stands.

This whole stand model was chosen because it can be fitted with temporary plots, although, it may be biased depending on the data distribution: the range of site fertility classes, and stand densities must be covered in all age classes. This model system was used for generating variable-density yield tables with given SDI values but not sufficient for simulating a specific real stand. Therefore, when using this whole-stand model, we set the SDI and $\overline{h}_{30}$ manually and then ran the $\overline{d}_g$ and G equations (Eqs. 5 and 6) to estimate the dynamics of the virtual stands.

2.3.2 Diameter distribution model

The Weibull distribution was used to characterize the diameter distribution. This probability density function is calculated as:

$$f(x; \beta_1, \beta_2, \beta_3) = \left(\frac{\beta_3}{\beta_2}\right)\left(\frac{x-\beta_1}{\beta_2}\right)^{\beta_3-1} \exp\left[-\left(\frac{x-\beta_1}{\beta_2}\right)^{\beta_3}\right] \quad (8)$$

where β_1, β_2, and β_3 are the location, scale, and shape parameters of Weibull distribution, respectively, and x is tree diameter at breast height. We used a moment-based parameter recovery method to calculate scale and shape parameters from quadratic mean diameter and diameter variance. The regression equation of diameter variance is:

$$D_{\mathrm{var}} = \exp\left(\gamma_1 + \gamma_2 \ln \overline{d}_g + \gamma_3 \ln G\right) \quad (9)$$

where D_{var} is diameter variance. The location parameter β_1, lower limit of x, was similarly set as:

$$\beta_1 = \exp\left(\gamma_4 + \gamma_5 \ln \overline{d}_g + \gamma_5 \ln G\right) \quad (10)$$

Parameter γ_is was obtained by maximum likelihood estimator regression (Cao 2004). The β_2 and β_3 were recovered by:

$$\beta_2 = -\beta_1 G_1/G_2 + \left[(\beta_1/G_2)^2\left(G_1{}^2 - G_2\right) + D^2/G_2\right]^{0.5} \quad (11)$$

$$\beta_2{}^2\left(G_2 - G_1{}^2\right) - D_{\mathrm{var}} = 0 \quad (12)$$

where G_i is $\Gamma(1 + i/\beta_3)$; $\Gamma(\cdot)$ is the complete gamma function. The diameter distribution model was developed using only square plot data because the size-class information of variable radius plots was inadequate.

2.4 Individual tree diameter model

An individual tree model simplified from Pukkala et al. (2009) was fitted using Bayesian calibration and simulated virtual data. This model system includes a distance-independent diameter growth, survival, and height prediction models:

$$i_d = \exp\left(c_1 + c_2 G_{>d} + c_3 \ln G + c_4 \sqrt{d} + c_5 d^2 + c_6 \overline{h}_{30}\right) \quad (13)$$

$$p = \frac{1}{1 + \exp\left[-\left(c_7 + c_8 \sqrt{d} + c_9 G_{>d}\right)\right]} \quad (14)$$

$$h = \frac{c_{10} + c_{11}\overline{h}_{30}}{1 + (c_{12}/d) + \left(c_{13}/d^2\right)} \quad (15)$$

where i_d is 5-year diameter increment (cm), $G_{>d}$ is basal area in larger trees (m² ha⁻¹), G is stand total basal area (m² ha⁻¹), d is individual tree diameter at breast height (cm), $\overline{h}_{30}$ represents site fertility class (m), p is the probability to survive for the coming 5 years, h is individual tree height (m).

The virtual data simulated in this study were designed for the diameter increment and survival models. The ingrowth model was deliberately omitted because of ingrowth stochasticity in pure, even-aged *Pinus tabuliformis* plantations. Therefore, recruitment was assumed to be null. The height projection model is static and can be fitted separately from the diameter increment and survival equations.

2.5 Parameter estimation and uncertainty analysis

The Bayes' theorem offers an approach for updating distributions, and can be applied repeatedly when new data are available. In its simplified notion, this theorem states:

$$p(\theta|D) \propto p(D|\theta) \times p(\theta) \quad (16)$$

where $p(\theta|D)$ and $p(\theta)$ are the posterior and prior distribution for the parameters θ, respectively, and $p(D|\theta)$ is the likelihood function, which is the probability of the data D for a given θ. For the first calibration of physiological process-based model (e.g., Van Oijen et al. 2005), prior distribution information is usually based on best guesses or published values. In this study, empirical model parameters have no physiological meaning. Thus, vague priors were used for the first calibration. Parameters were assumed as independent uniform distributions, of which the minima and maxima are shown in Tables 2, 3, and 4.

Because multiple sources of data were used in the Bayesian calibration, the likelihood function was decomposed based on the assumptions of independent errors. For example, the calibration frame of height model (Eqs. 1 and 15) was:

$$p(\theta|D) \propto p(D_{VR}|\theta) \times p(D_{SP}|\theta) \times p(D_{SA}|\theta) \times p(\theta) \quad (17)$$

Table 3 Posterior probability distribution of the parameters in the individual tree diameter increment model. The posterior distribution is not analytical and it is characterized here by the mean values of the parameters, the standard deviation (SD) and the maximum a posteriori estimates (MAP). The prior was set as independent uniform distributions, of which the maxima and minima are listed

Parameters	Predictor	Prior	VB model			CB model			RB model			Recalibrated model		
			Mean	SD	MAP	Mean	SD	MAP	Mean	SD	MAP	Mean	SD	MAP
c_1	Constant	(−100,100)	−0.3014	0.0712	−0.3129	−1.5807	0.2388	−1.5154	−0.2250	0.8122	−0.3144	−1.3257	0.3396	−1.1309
c_2	$G_{>d}$	(−100,10)	−0.1364	0.0011	−0.1368	−0.0579	0.0040	−0.0578	−0.0355	0.0096	−0.0335	−0.0905	0.0034	−0.0901
c_3	$\ln G$	(−100,10)	−0.2878	0.0213	−0.2887	−0.0480	0.0405	−0.0027	−0.5429	0.1739	−0.5211	−0.2751	0.0372	−0.2796
c_4	$\sqrt{d}$	(−10,100)	0.8385	0.0162	0.8366	0.6758	0.0623	0.6576	0.48914	0.1089	0.4764	0.7822	0.0858	0.7406
c_5	d^2	(−10,1)	−0.0039	0.0001	−0.0039	−0.0024	0.0001	−0.0023	−0.0015	0.0003	−0.0014	−0.0030	0.0002	−0.0029
c_6	$\bar{h}_{30}$	(−10,100)	0.0636	0.0040	0.0662	0.0915	0.0166	0.0792	0.1408	0.0681	0.1420	0.1360	0.0136	0.1323

where D_{VR}, D_{SP}, and D_{SA} are observations from variable-radius plots, square plots, and stem analysis respectively. A heavy tailed likelihood probability density function (Sivia and Skilling 2006) was used:

$$p(D|\theta) = \prod_{i=1}^{n} \frac{1}{\sigma_i\sqrt{2\pi}} \left[\frac{1-\exp\left(-R_i^2/2\right)}{R_i^2}\right] \quad (18)$$

where σ_i is the noise-scaling factor; a measure of the uncertainty of random error of the ith observation. R_i is the residual for the ith datum and prediction, divided by σ_i.

Data from variable-radius plots, square plots, and stem analyses were not mixed since the noise-scaling factor, σ, is different for each dataset. For each data type, a linear relation between error and prediction was established as $\sigma_{mn} = a_m \cdot \widehat{y}_{mn} + b_m$, where $\widehat{y}_{mn}$ is the nth prediction of dataset m. Parameters a and b were calibrated along with parameters θ.

A special likelihood function modified from Pukkala et al. (2011) was used to simultaneously fit individual tree diameter growth model and survival model:

$$p(D|\theta) = p(E = D-M(\theta)) =$$
$$\prod_{k=1}^{K}\prod_{j=1}^{J_k}\prod_{i=1}^{I_j} \left[\varphi\left(G_{jk}^D(d_{ijk})-G_{jk}^M(d_{ijk});0,\sigma_{ijk}^{G\,2}\right) \cdot \varphi\left(F_{jk}^D(d_{ijk})-F_{jk}^M(d_{ijk});0,\sigma_{ijk}^{F\,2}\right)\right] \quad (19)$$

where θ is the set of coefficients of the individual tree models, K is number of plots, J_k is the number of measurement intervals of plot k, I_j is number of 4-cm diameter classes in measurement interval j of plot k, φ denotes a Gaussian probability density function with a given mean and variance, $G_{jk}^D(d_{ijk})$ and $G_{jk}^M(d_{ijk})$ are, respectively, cumulative basal area (m^2 ha^{-1}) of data and prediction at diameter d_{ijk} (upper limit of diameter class i) at the end of measurement interval j of plot k, and $F_{jk}^D(d_{ijk})$ and $F_{jk}^M(d_{ijk})$ are, respectively, cumulative number of trees per hectare of data and prediction at diameter d_{ijk} at the end of measurement interval j of plot k.

Detailed information on diameter increments and survival were unavailable from the simulated data because of an identification problem: we did not specifically know which trees die or individual tree increments during the 5-year period because trees in virtual plots were not identified. Thus, diameter increment and survival models (Eqs. 13 and 14) were fitted together using the likelihood function (Eq. 19) based on diameter distribution changes. The height prediction model (Eq. 15) was parameterized independently.

Pukkala et al. (2011) used least-squares to balance two variables at different scales (G and F in Eq. 19) by multiplying the sum of squared errors of F by 0.0001. However, this choice could be entirely arbitrary and should be justified. We converted the objective function of the optimization-based modelling into a likelihood function. Linear relations between the standard

Table 4 Posterior probability distribution of the parameters in the individual tree survival and height prediction models. The prior was set as independent uniform distributions, of which the maxima and minima are listed

Parameters	Predictor	Prior	Mean	SD	MAP
c_7	Constant	$(-100,100)$	-4.6433	0.1467	-4.7345
c_8	$\sqrt{d}$	$(-10,100)$	2.0424	0.0387	2.0652
c_9	$G_{>d}$	$(-100,10)$	-0.0333	0.0008	-0.0327
c_{10}	Constant	$(-100,100)$	48.3936	1.5133	49.9526
c_{11}	$\overline{h}_{30}$	$(0,100)$	2.8640	0.1191	2.9929
c_{12}	$1/d$	$(-10,100)$	80.9377	4.0777	85.7924
c_{13}	$1/d^2$	$(-100,1000)$	159.8110	58.8020	152.1656

deviation and the model predictions were calculated as $\sigma_{ijk}^G = a_G \cdot G_{jk}^M(d_{ijk}) + b_G$ and $\sigma_{ijk}^F = a_F \cdot F_{jk}^M(d_{ijk}) + b_F$, where parameters a_G, b_G, a_F, b_F were calibrated together with model parameters. Here, two variables, G and F, will have two different linear relations. In this way, variable error increased with its magnitude and the likelihood automatically balanced the scales of the two variables.

Unlike the likelihood function for fitting whole stand model or the individual tree model, the likelihood for diameter distribution model was not based on the above assumptions of error distributions, but on the probability density function of Weibull distribution:

$$p(D|\theta) = \prod_{i=1}^{p}\prod_{j=1}^{n_i} \left(\frac{\beta_3}{\beta_2}\right)\left(\frac{x_{ij}-\beta_1}{\beta_2}\right)^{\beta_3-1} \exp\left[-\left(\frac{x_{ij}-\beta_1}{\beta_2}\right)^{\beta_3}\right]$$
(20)

where p is the number of plots, n_i is the number of trees in the ith plot, x_{ij} is the diameter at breast height of tree j in the ith plot. The regression equations for D_{var} and β_1 (Eqs. 9 and 10) were simultaneously fitted while β_2 and β_3 were recovered (Eqs. 11 and 12).

The MCMC was simulated, using differential evolution adaptive Metropolis with snooker updating (DREAMzs), which runs a few chains in parallel and explores the parameter space in an efficient way (Laloy and Vrugt 2012; Vrugt et al. 2009). We used the DREAMzs algorithm implemented in the R package BayesianTools (Hartig et al. 2017; R Development Core Team 2017). When using unanalytical joint posterior distribution as the prior for a recalibration, a multivariate sample was extracted from the previous MCMC. With this sample, a multivariate normal density was fitted and used as the prior for the new calibration.
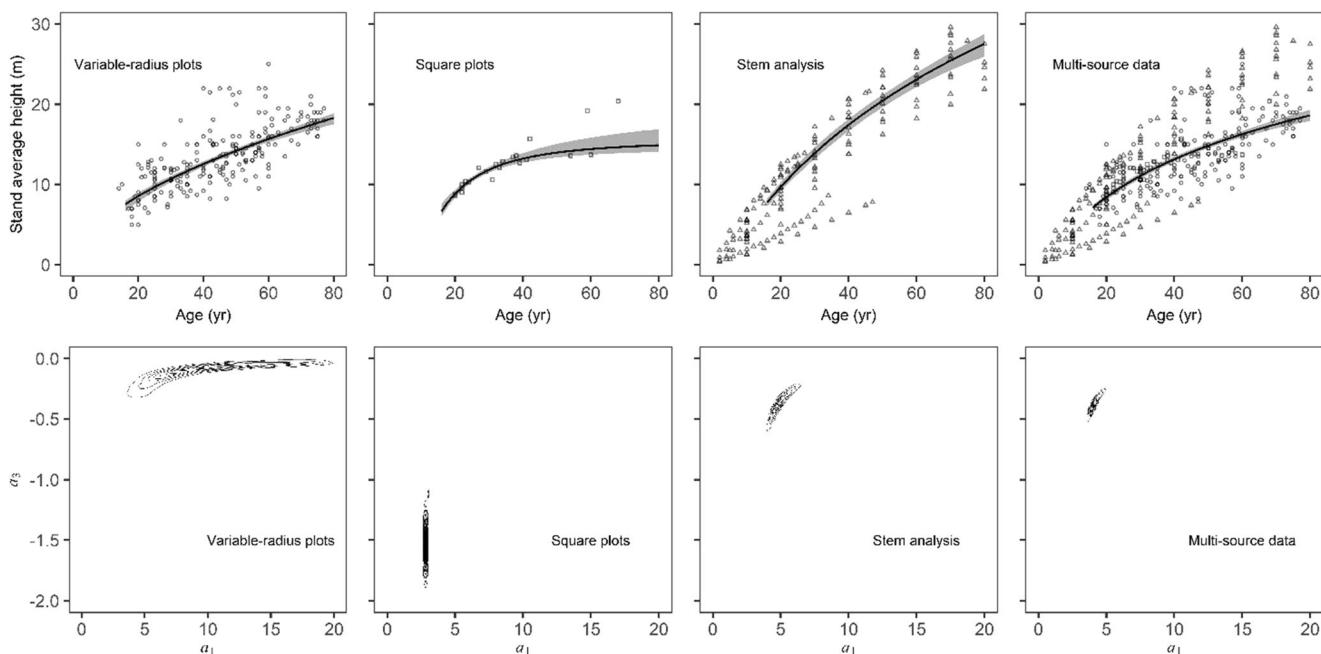
The MCMC convergence diagnostic (Brooks and Gelman 1998; Gelman and Rubin 1992) was used to monitor the convergence in the MCMC output. The multivariate potential scale reduction factor (MPSRF) was calculated, based on two MCMC runs, each of which has three internal chains. The number of iterations was set based on the complexity of the model and data, ranging from 10^5 to 10^7. A large MPSRF means that the output from all chains is distinguishable, and a notable difference exists between variance and intrachain variance. In our study, convergence was diagnosed when the MPSRF was below 1.05, which is a relatively strict criterion (Brooks and Gelman 1998).

2.6 Model validation

One-quarter of plot observations were randomly selected for benchmarking, including the whole-stand, diameter distribution, and individual tree height prediction models. The increment core observations and remeasurements were respectively processed to validate the individual tree diameter increment equation. For the tree core data, we back-dated tree and stand attributes to the start of previous 5-year growth period, to create explanatory variables. The inside bark diameters from 5-year earlier were calculated for each subject tree by subtracting the 5-year inside bark increment. Bark thicknesses were estimated based on the assumption that ratios between outside and inside bark diameters were constant over a relatively short, 5-year period (e.g., Pukkala 1989). Diameter growth rate of each size class was calculated using increment cores. Tree diameters from 5-years previous were roughly estimated using estimated growth rates. From this we obtained the 5-year earlier stand basal area and basal area in larger trees. For the remeasurements, the interval years were calculated based on the number of growing seasons, considering that diameter growth mainly occurs in the early summer. Growth data of 4-, 6- or 7-year interval were linearly converted to the growth of 5 years.

We used a regression-based equivalence test developed by Robinson et al. (2005) for model benchmarking. Unlike usual parametric methods, equivalence tests rely on a null hypothesis that a model is unacceptable rather than acceptable. The
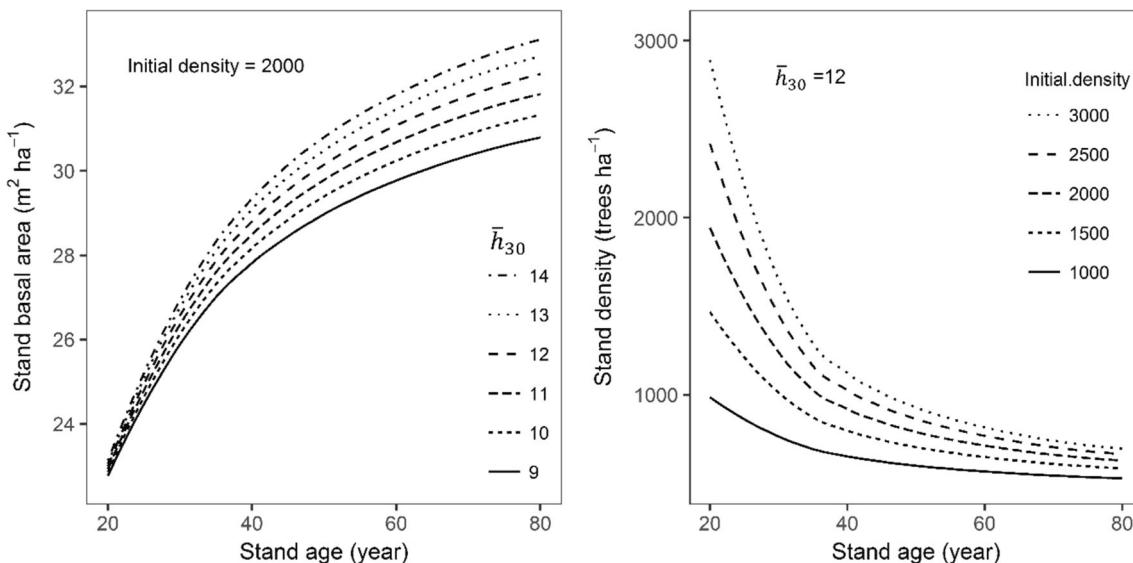
Fig. 3 Guide curves of stand average height and joint distributions of parameters in the guide curve equation (Eq. 1) when based on different types of data. The grey area is the 95% Bayesian credible interval based on parameter uncertainty
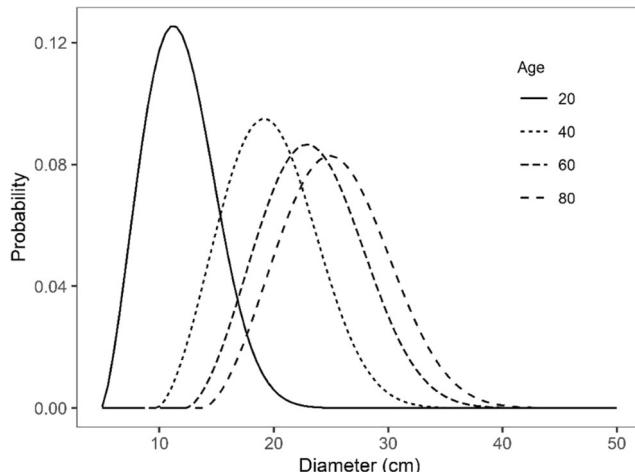
equivalence test not only compares means but also for checks for similarity between individual predictions and observations. The validation strategy suggested by Robinson et al. (2005) followed the following six steps: (1) subtract mean prediction from component predictions, (2) establish regions of equivalence for the shifted intercept and slope (e.g.,±25%), (3) fit a linear regression between observations and adjusted predictions, (4) test the intercept for equality by calculating two, one-sided confidence intervals for the intercept and comparing these to the estimated equivalence regions, (5) test the slope for equality by calculating the two, one-sided confidence intervals for the slope and comparing

to its equivalence regions, (6) accept or reject the hypothesis of dissimilarity based on whether the confidence interval falls within the equivalence region. The R module "equivalence" was used to perform the regression-based equivalence test (Robinson 2016). We calculated the minimum percentage that would result in rejection of the null hypothesis (Froese and Robinson 2007).

The individual tree model and the diameter increment equation (Eq. 13) were fitted separately by simulated virtual data, increment core data, and remeasurement data (Fig. 1). The performance of the core-data-based (CB) model, remeasurement-based model (RB), and virtual-

Fig. 4 **a** Development of stand total basal area on different site fertility classes. **b** Development of stand density with different initial density

Fig. 5 Development of diameter distribution at different ages. The probability density functions were for stands with 2000 trees ha^{-1} initial density at age 20 and $\overline{h}_{30}$ =12 m

data-based (VB) model was compared at both tree- and stand-levels. In addition to the equivalence test, mismatches between real growth observations and VB model predictions were quantified by decomposed mean squared deviation (MSD), which identified causes of possible large deviations between observations and predictions (Kobayashi and Salam 2000). As suggested by Gauch et al. (2003) the MSD was partitioned into three parts: squared bias (SB), nonunity slope (NU), and lack of correlation (LC). When comparing observations and predictions, squared bias means the deviation of means; nonunity slope is relevant to the slope of linear regression; lack of correlation means a scatter. In this case, we expected the mean of model predictions and the mean of observations to be the same, and the slope of observations on predictions to be 1. Thus, the data model mismatch was decomposed as:

$$\text{MSD} = \frac{\sum_{i=1}^{n}(X_n - Y_n)^2}{N} = \text{SB} + \text{NU} + \text{LC} \qquad (21)$$

$$\text{SB} = \left(\overline{X} - \overline{Y}\right)^2 \qquad (22)$$

$$\text{NU} = (1-\beta)^2 \left(\sum \frac{\left(X_n - \overline{X}\right)^2}{N}\right) \qquad (23)$$

$$\text{LC} = (1-r^2) \left(\sum \frac{\left(Y_n - \overline{Y}\right)^2}{N}\right) \qquad (24)$$

where $\overline{X}$ and $\overline{Y}$ are the means of model predictions (X) and observations (Y), respectively, β is the slope of the least-squares regression of Y on X, r^2 is the square of the correlation coefficient, and N is the number of observations.

3 Results
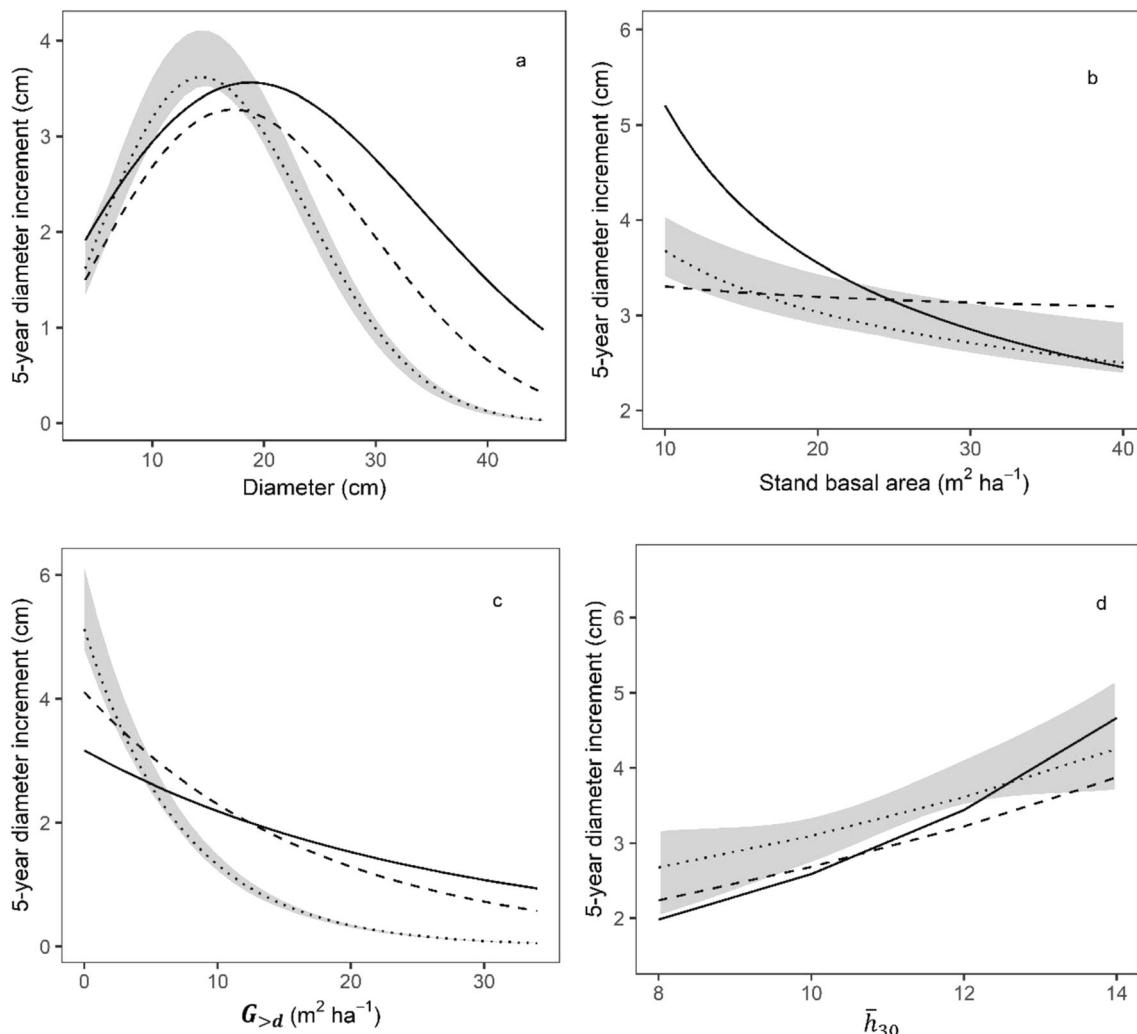
3.1 Whole-stand and diameter distribution models

In the Bayesian calibration process, the joint posterior distribution of parameters is approximated by Markov Chain Monte Carlo simulation. The mean value, standard deviation, and MAP (maximum a posteriori probability estimates) of parameters in the whole-stand model are given in Table 2. As shown in the model of guide curves, different types of inventory data resulted in contrasting predictions and joint parametric distributions (Fig. 3). The performance of the guide curve estimated from multisource data was similar to that of the curve based on various-radius plots. However, the joint posterior distribution of a_1 and a_3 from multisource distribution was more similar to the distribution from equation based on stem analysis, because the likelihood weighted each dataset automatically based on the estimated error distributions. Based on the whole-stand model, we simulated the growth and mortality of 30, 1-ha virtual plots (Fig. 4). These plantations possessed the characteristics of both relatively high initial stand basal area (Fig. 4a) and mortality (Fig. 4b), since no thinning schemes were designed in the virtual plots.

The parameters that predict diameter variance (Eq. 9) and the minimum diameter (Eq. 10) are given in Table 2. The diameter distribution changes due to growth and self-thinning in a medium site and density condition (Fig. 5). The most prominent changes were tree frequencies in small size classes, especially at young ages. The minimum diameter increased over time.

3.2 Individual tree diameter model

The calibrated parameters of the individual tree model are given in Table 3. Parameter vectors based on simulated virtual data, increment core method, and remeasurements were logical and general similar. Compared with the CB and RB models, the diameter increment predictions of the VB model were more sensitive to their own diameters (Fig. 6). The diameter growth rate reached its maximum at about 15–20 cm for *P. tabuliformis*. Simulated VB model was most sensitive to competition variable *BAL*. The RB model was most sensitive to the total stand basal area and site fertility. The standard deviations in Table 3 represented the parametric uncertainty; the standard deviations of parameters for ln*TBA* and $\overline{h}_{30}$ were larger than those of *BAL*, which suggests more uncertainty in predictions. Using the parameters' joint posterior distribution of the VB model as the prior, we recalibrated the model using tree cores and remeasurements, which combined information of both virtual data and real growth observations in a Bayesian framework (Eq. 16). The parameter

Fig. 6 Comparison of individual tree diameter increment predictions based on simulated virtual data, increment cores, and remeasurements. (**a** The relationships of 5-year diameter increment over diameter, **b** stand basal area, **c** basal area of larger trees ($G_{>d}$), **d** and site fertility ($\overline{h}_{30}$, average height on the reference age). The dotted, dashed, and solid lines are, respectively, virtual-data-based predictions, increment core method predictions, and remeasurement-based predictions. Lines were generated by MAP (the maximum a posterior parameter vector). The grey area represents the 95% Bayesian credible interval based on parametric uncertainty of virtual data method. For Fig. 5a, the stand total basal area is 25 m^2 ha^{-1}, the basal area of larger trees is 5 m^2 ha^{-1}, and the $\overline{h}_{30}$ is 12 m. For Fig. 5b, the diameter is 20 cm, the basal area of larger trees is 5 m^2 ha^{-1}, and the $\overline{h}_{30}$ is 12 m. For Fig. 5c, the diameter is 20 cm, the stand basal area is 35 m^2 ha^{-1}, and the $\overline{h}_{30}$ is 12 m. For Fig. 5d, the diameter is 15 cm, the stand basal area is 20 m^2 ha^{-1}, and the basal area of larger trees is 5 m^2 ha^{-1})
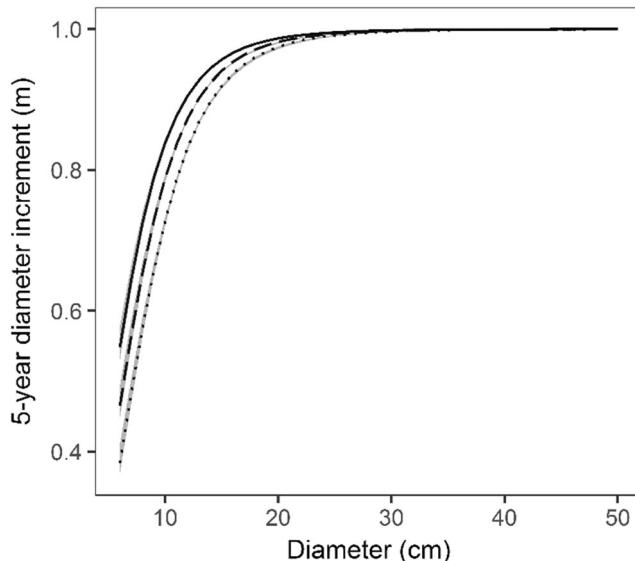
vector of the recalibrated model was somewhat a compromise among parameter vectors of the VB, CB, and RB models (Table 3).

Survival model coefficients (Table 4) show that the survival probability increased with increasing diameter or decreasing competition (Fig. 7). When the diameter of a tree reached 30 cm, it was unlikely to die in subsequent measurement intervals, because the mortality from senescence and disturbances were not considered in the survival model. The height prediction model indicated trees of equivalent diameter are higher in more fertile sites (Fig. 8), and height uncertainty increases as the diameter increases.
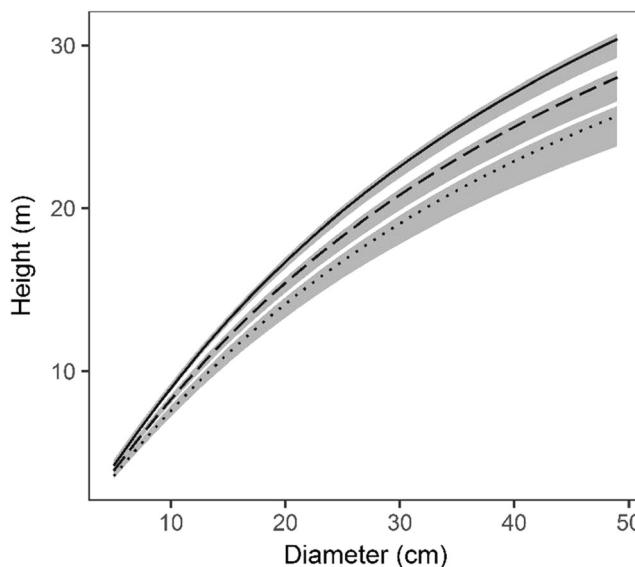
3.3 Validation of different resolution models

Tests for the intercept only failed to reject the null hypothesis of dissimilarity in the individual tree diameter increment model (Table 5). However, tests for the slope failed to reject the null of quadratic mean diameter prediction function, diameter distribution function, and individual tree diameter increment model. The stand-level basal area model showed the best performance in both intercept and slope tests. Since the area of the square plot was only 0.04 ha, random errors in Weibull distributions could be large when compared with observations. The minimum rejection interval of the slope in dbh frequencies was as large as 80.25% (Table 5), which was

Fig. 7 Dependence of the probability that a tree to survives for 5 years relative to diameter at breast height. The solid line, dashed line and dotted line represent trees on $G_{>d}$ of 5, 15, and 25 m^2 ha^{-1}, respectively. The stand total basal area is 30 m^2 ha^{-1}, and the $\overline{h}_{30}$ is 12 m. The grey area is the 95% Bayesian credible interval based on parameter uncertainty

caused by the randomness of observed diameter distribution on a limited population. As the difference between virtual data and increment CB models shows (Fig. 6), equivalence tests for individual tree diameter increment model also largely failed to reject null hypotheses of dissimilarity (Table 5).

The virtual data-based individual tree diameter increment model underestimated the growth of the suppressed trees compared with the tree core observations (Fig. 9a) or

Fig. 8 Dependence of the height of *Pinus tabuliformis* on diameter at breast height. The solid line, dashed line and dotted line represent trees on good quality site ($\overline{h}_{30}$ =14 m), median site ($\overline{h}_{30}$ =12 m), and poor site ($\overline{h}_{30}$ =10 m), respectively. The grey area is the 95% Bayesian credible interval based on parameter uncertainty

remeasurements (Fig. 9b). The VB model was more sensitive to the competition variable $G_{>d}$ than the CB and RB models (Fig. 6c and Table 3). When validated with tree core data, the MSD of the VB model was as high as 1.92 times of MSD of the CB model (Fig. 10). When validated with remeasurements, the MSD of the VB model was as high as 1.38 times of MSD of RB model. The distinction between the CB and RB model was smaller.

On the individual tree scale, this model-data mismatch means biased predictions for the highly suppressed trees (Fig. 9). However, this bias was negligible on the stand-level growth predictions (Fig. 11). The VB model shows the highest consistency with the whole-stand model.

4 Discussion

The principles of stand development have long been applied to model diameter growth, competition, stand structure, and diameter distribution in forests (Vanclay 1994). With remeasured data available, more complicated competition variables for modelling detailed individual tree models, such as FVS (Crookston and Dixon 2005) and ORGANON (Hann et al. 2011), are possible. Vanclay (2010) reported the feasibility of robust relationships to predict forest growth with sparse data at the stand level. At the tree level, we have applied a simple but robust competition index, $G_{>d}$, to demonstrate the sensitivity of competition effects on individual tree diameter growth. The empirical whole-stand and diameter distribution models were used as constraints to control the bias between the stand and the tree level. This makes the individual tree diameter model more reliable to track stand diameter structure and basal area development.

We used both tree core observations and tree-level remeasurements to validate the predicted individual tree diameter growth. The VB, CB, and RB models presented somewhat different predictions of diameter increment. The growth prediction from CB model was not sensitive to stand basal area, and, on the contrary, the RB model was largely sensitive to basal area, possibly due to the fact that both tree cores and remeasurements were from a limited number of plots. The large uncertainty resulting in the predictions of the RB model (Fig. 12) indicates that the remeasurement data were unable to sufficiently constrain the parameter distributions. These divergent results reflected the fact that these three methods sampled and simulated the stands in different ways. Nevertheless, the accuracy or biases were primarily dependent on database size, measurement error, and inventory sampling design rather than modelling techniques.

One of the greatest problems in using data from temporary plots is that biases could be easily introduced into the model. The guide curve method assumes that average site productivity and site fertility distribution is the same in all age classes. If

Table 5 Summary of equivalence-based regression results. Sample size is denoted by n, and the approximate joint two one-sided 95% confidence intervals for the intercept β_0 and the slope β_1 are (C_β^-, C_β^+), the intercept interval of equivalence (I_β^-, I_β^+) for intercept and slope are, respectively, $\bar{y}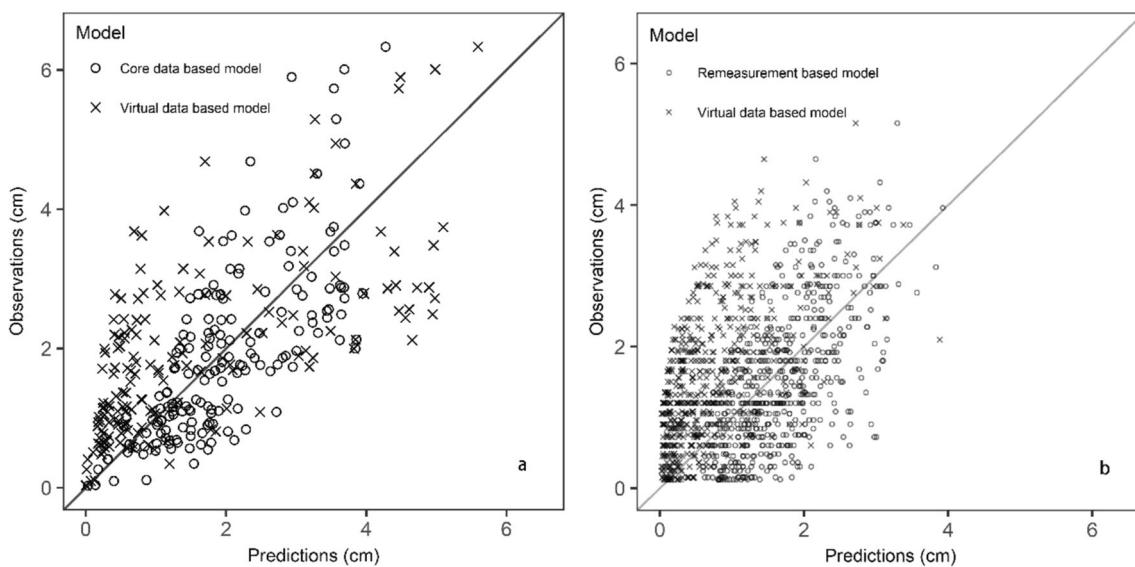 \pm 25\%$ and 1 ± 0.25. $F(8,9,10)$ means frequency in specific size class, which were predicted using diameter distribution model (Eqs. 8, 9, and 10)

	Variable (Equation)	n	Confidence interval		Indifference region		Reject dissimilarity?	Min. rejection interval
			C_β^-	C_β^+	I_β^-	I_β^+		
β_0	$\bar{h}_{gc}$ (1)	59	12.61	14.47	9.73	16.21	Yes	11.57%
	G (5)	59	21.61	23.97	17.13	28.54	Yes	5.36%
	$\bar{d}_g$ (6)	59	19.85	21.73	14.68	24.46	Yes	11.04%
	F(8,9,10)	53	2.25	2.96	2.11	3.51	Yes	19.93%
	i_d(13)	169	1.64	1.81	1.05	1.75	No	29.29%
	h (15)	130	15.71	16.93	13.76	22.94	Yes	15.32%
β_1	$\bar{h}_{gc}$ (1)	59	0.92	1.22	0.75	1.25	Yes	20.07%
	G (5)	59	0.82	1.10	0.75	1.25	Yes	19.24%
	$\bar{d}_g$ (6)	59	0.96	1.28	0.75	1.25	No	29.89%
	F(8,9,10)	53	0.20	0.67	0.75	1.25	No	80.25%
	i_d(13)	169	0.52	0.65	0.75	1.25	No	46.54%
	h (15)	130	0.76	0.97	0.75	1.25	Yes	22.56%

this is not the case, the model for stand height (Eq. 1) will be biased, affecting all other models. In practice, a temporary plot dataset with the same site fertility distributions in all age classes is almost impossible. Thus, we combined a small dataset of stem analysis in the parameterization of height prediction models by using multiplied likelihood (Eq. 17). Our dataset contains 237 temporary plots and only 11 trees with stem-analysis records. The temporary plot data were more representative, while the stem analysis data were more accurate by avoiding the biases of changing site distributions. By
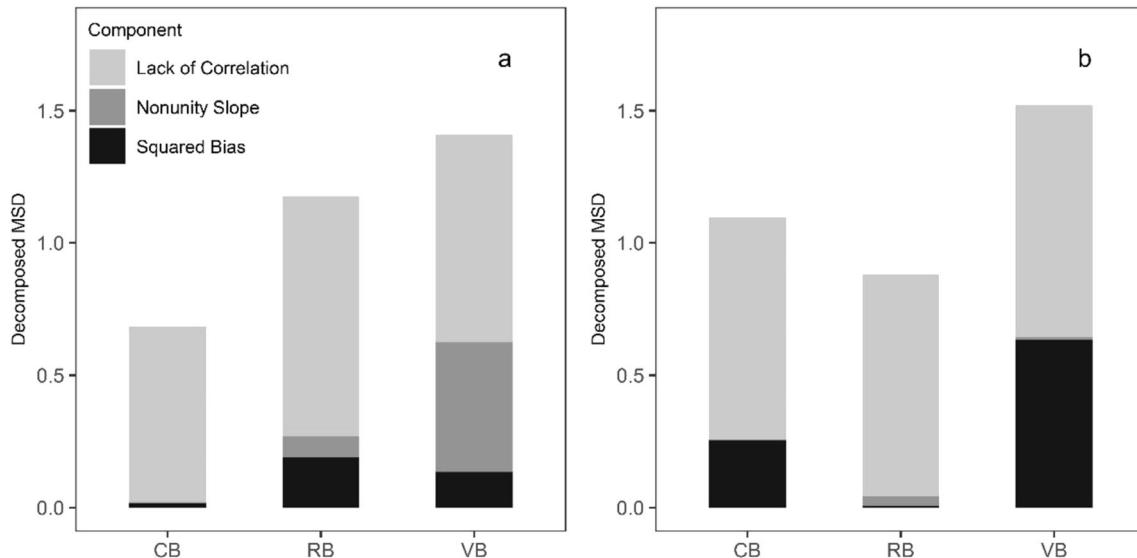
calibrating the values of noise-scaling factor σ, the weights of different datasets were balanced automatically.

The use of the parameter recovery estimator in a Weibull distribution linked the whole-stand and size-class models. The estimation of diameter distributions relies on stand-level variables, so variability of individuals will change consistently with the stand-level growth and mortality. The individual tree model based on virtual data will also provide consistent projections with stand-level attributes. This is because model data were simulated directly from the whole-stand model. This

Fig. 9 Simulated and observed 5-year individual tree diameter increments. Comparisons between the virtual-data-based model and **a** tree cores or **b** remeasurements
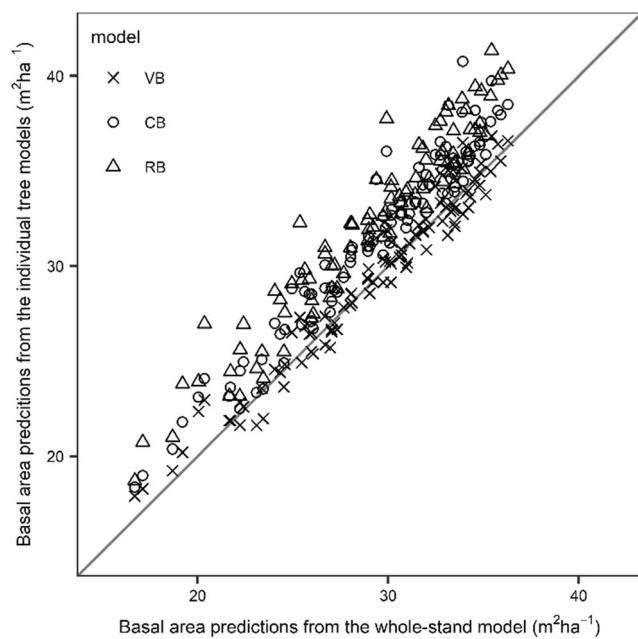
Fig. 10 Decomposed mean squared deviation (MSD) between predictions and the observations of **a** tree cores or **b** remeasurements. VB = virtual-data-based individual tree diameter growth model; CB = core-
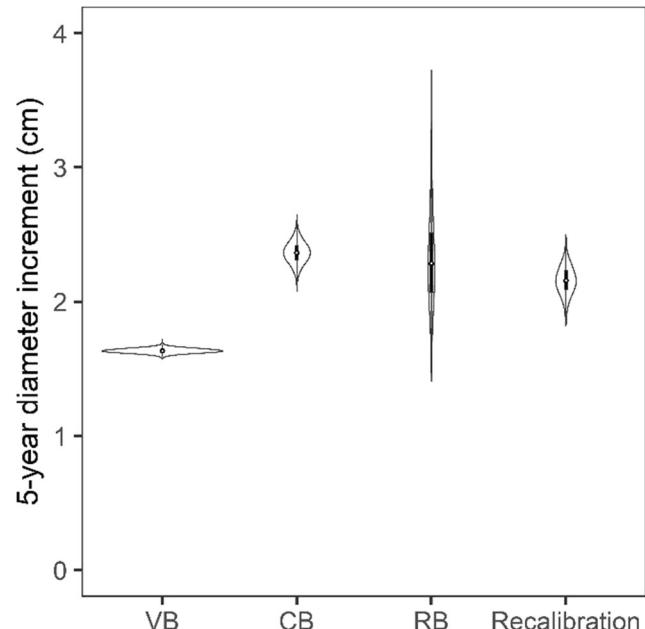
data-based individual tree diameter growth model; RB = remeasurement-based individual tree diameter growth model

could be a distinct advantage when model users are concerned about accumulated error from aggregations of tree-level projections. However, systematic errors of whole-stand and diameter distribution models are also passed on to the individual tree model. In contrast, tree core–based individual tree models were developed independently of other models' resolution.

Thus, the CB model provided more reliable, tree-level projections that were well matched with tree core observations. However, the models that rely on tree cores or remeasurement may also provide unrealistic projections of stand-level attributes because changes in stand-level variables are not

Fig. 11 Comparison of basal area predictions between the whole-stand model and individual tree models. Initial states of 100 plots were randomly selected as input, and the stand total basal area after five years were respectively predicted by the whole stand model and three individual tree models. VB = virtual-data-based individual tree diameter growth model; CB = core-data-based individual tree diameter growth model; RB = remeasurement-based individual tree diameter growth model

Fig. 12 Posterior uncertainty of predictions from three diameter increment models, by 3000 runs of each model. The diameter of the objective tree is 15 cm, the stand basal area is 30 m^2 ha^{-1}, basal area in larger trees is 10 m^2 ha^{-1}, and the $\overline{h}_{30}$ is 12 m. Violin plots represent the density estimates of predictions from minimum to maximum values. The boxplots and white points at the median represent the quartiles of distributions. VB = virtual-data-based individual tree diameter growth model; CB = core-data-based individual tree diameter growth model; RB = remeasurement-based individual tree diameter growth model

simulated directly but aggregated from tree-level projections (Ritchie and Hann 1997). In our case, the VB and CB models performed similarly in stand level simulations and were consistent with the whole-stand model predictions. However, the risk in the aggregation was revealed in the RB model. Several approaches can be used to modify stand-level projections by linking stand- and tree-level models. These include disaggregating and allocating stand-level growth and mortality into individual trees (e.g., Clutter and Allison 1974; Qin and Cao 2006), estimating multiresponse parameters to optimize tree-level predictions at multilevels (e.g., Zhang et al. 1997; Cao 2006) and using a composite estimator to joint estimates of stand- and tree-level models (e.g., Yue et al. 2008; Zhang et al. 2010).

With the optimization-based modelling method of Pukkala et al. (2011), we fit the individual tree model by minimizing the difference between measured and predicted diameter distributions. de Miguel et al. (2014) also found that this optimization-based method can solve tree misidentification and uneven measurement interval problems in individual tree models. Instead of the optimization algorithm, we chose Bayesian calibration for fitting models for two reasons; to quantify uncertainty of parameters and predictions, and to update models when new data became available. Considering that the predictions of the VB model were poorly matched with tree core observations, we recalibrated the model using the Bayesian calibration framework (Eq. 16). The posterior distributions of parameters in the VB model were taken as the prior information. Increment cores and remeasurements were used to calculate likelihood. The posterior distributions of parameters contained the information from both the simulated virtual stand and real growth measurements. At the end of this process, the predictions of the recalibrated model can be thought as a dynamic combination of the three datasets (Fig. 12). The basal area parameter ($\ln G$) in the recalibrated model was greater than that in the CB model (Table 3), which means that the recalibrated model was more sensitive to the change in stand basal area. It is important to note that the posterior information of VB model was not completely converted into the new, prior information during recalibration. The posterior parameter distributions cannot be determined analytically. Therefore, we adopted the multivariate normal distribution, which was only a general description of the joint posterior distributions, to fit the density function.

We fit the individual tree diameter growth model and the individual tree survival model by linking stand-level and size-class simulations. Another important component, the in-growth model, was deliberately omitted. This was because all data were collected from pure even-aged stands, and information of ingrowth from temporary plots was very limited. To validate our method, we used tree cores to backdate the stand condition from 5 years earlier and then fit another individual tree diameter growth model. Theoretically, an individual tree survival model could be constructed based on the backdated information. However, the exact time of tree death could not be estimated because of small age differences among trees and unclear deterioration characteristics.

More than two hundred plots were used to calibrate the whole-stand model. For other more limited cases or areas, other forms of whole-stand models or diameter distribution models could possibly be adopted. Vanclay (2010) developed a set of sample stand–level models that only contained three parameters and could be fitted with fewer data (even a single observation) for plantations. A variety of parameter recovery or parameter prediction methods are also available for simulating diameter distributions (e.g., Poudel and Cao 2013).

5 Conclusions

In this study, we proposed a three-stage modelling approach to forecast individual tree diameter growth using limited single observation data. The empirical whole stand model provides an accurate forecast of basal area development at the stand level. Links between the whole-stand model and diameter distribution model across forest sites and densities effectively enlarged the database of reliable stand structure information. This permits Bayesian calibration to model individual tree diameter without remeasured data. Our results demonstrated the individual tree diameter growth model effectively predicts competition at individual tree level; compared with empirical predictions based on incremental tree core measures, although, the simulated results seem to underestimate the diameter growth of suppressed trees.

The ultimate goal of constructing stand level models and diameter distributions is to provide adequate simulated data and appropriate likelihood functions for the calibration of individual tree models. Under the Bayesian framework, the proposed individual tree growth model can be updated when data from increment cores, interval plots or permanent plots are available, providing more accurate and robust predictions along time.

Authors' contributions XT implemented most of the calculations and writing the original draft. SS produced data of tree cores for model validation and constructed the diameter distribution model. BM contributed significantly to the design of linking models of different resolutions and the writing. TC conceived the original idea and also participated in the design of experiments, the analysis of results, and the writing.

Data availability The datasets generated or analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Altrell D (2019) Multipurpose National Forest Inventory in Mongolia, 2014-2017-a tool to support sustainable forest management. Geogr Environ Sustain 12:167–183. https://doi.org/10.24057/2071-9388-2019-36

Assmann E (1970) The principles of forest yield study. Pergamon, Oxford

Bailey RL, Clutter JL (1974) Base-age invariant polymorphic site curves. For Sci 20:155–159

von Baur F (1876) Die Fichte in Bezug auf Ertrag, Zuwachs und Form. Springer, Berlin https://doi.org/10.1007/978-3-642-91377-8

Brooks SP, Gelman A (1998) General methods for monitoring convergence of iterative simulations. J Comput Graph Stat 7:434–455. https://doi.org/10.1080/10618600.1998.10474787

Cao QV (2004) Predicting parameters of a Weibull function for modeling diameter distribution. For Sci 50:682–685

Cao QV (2006) Predictions of individual-tree and whole-stand attributes for loblolly pine plantations. Forest Ecol Manag 236:342–347. https://doi.org/10.1016/j.foreco.2006.09.019

Cieszewski CJ, Bella IE (1993) Modeling density-related lodgepole pine height growth, using Czarnowski's stand dynamics theory. Can J For Res 23:2499–2506

Clutter J, Allison B (1974) A growth and yield model for Pinus radiata in New Zealand. Growth Models for Tree and Stand Simulation. Fries J (Ed), Royal College of Forestry, Stockholm, Sweden. Department of Forest Yield Research Note 30, pp 136–160.

Crookston NL, Dixon GE (2005) The forest vegetation simulator: a review of its structure, content, and applications. Comput Electron Agric 49:60–80. https://doi.org/10.1016/j.compag.2005.02.003

de Miguel S, Pukkala T, Morales M (2014) Using optimization to solve tree misidentification and uneven measurement interval problems in individual-tree modeling of Balsa stand dynamics. Ecol Eng 69:232–236. https://doi.org/10.1016/j.ecoleng.2014.04.008

da Luz NB, Garrastazu MC, Doetzer Rosot MA, Maran J C, de Oliveira YMM, Franciscon L, et al. (2018) Brazilian National Forest Inventory-a landscape scale approach to monitoring and assessing forested landscapes. Brazilian Journal of Forest Research/Pesquisa Florestal Brasileira, 38. https://doi.org/10.4336/2018.pfb.38e201701493

Eichhorn F (1902) Ertragstafeln für die Weißtanne. Verlag Julius Springer, Berlin

FAO (2015) Global Forest Resources Assessment 2015. Food and Agriculture Organization, Rome

Flewelling J, Collier R, Gonyea B, Marshall D, Turnblom E (2001) Height-age curves for planted stands of Douglas-fir, with adjustments for density. Stand Management Cooperative Working Paper 1. University of Washington, College of Forest Resources, Seattle.

Freitas JV, Malheriros de Oliveira Y M, Mello Rosa CA, Póvoa de Mattos P, Doetzer Rosot CM, Brena DA, et al. Chapter 10 Brazil.

In: Vidal C, Alberdi I, Hernandez L, Redmon JJ Eds. (2016). National Forest Inventories-Assessment of Wood Availability and Use. Springer Int. Publishing. 845 pp. https://doi.org/10.1007/978-3-319-44015-6

Froese RE, Robinson AP (2007) A validation and evaluation of the prognosis individual-tree basal area increment model. Can J For Res 37:1438–1449. https://doi.org/10.1139/x07-002

Gauch HG, Hwang JT, Fick GW (2003) Model evaluation by comparison of model-based predictions and measured values. Agron J 95:1442–1446. https://doi.org/10.2134/agronj2003.1442

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–472. https://doi.org/10.1214/ss/1177011136

Hann DW, Bluhm A, Hibbs D (2011) Development and evaluation of the tree-level equations and their combined stand-level behavior in the red alder plantation version of ORGANON. Forest Biometrics Research Note 1. Forest Biometrics Research Paper 1. Oregon State University, Department of Forest Engineering, Resources, and Management, Corvallis, Oregon, USA

Hartig F, Minunno F, Paul S (2017) BayesianTools: general-purpose MCMC and SMC samplers and tools for Bayesian statistics. R package version 0.1.3. https://CRAN.R-project.org/package = BayesianTools

Kobayashi K, Salam MU (2000) Comparing simulated and measured values using mean squared deviation and its components. Agron J 92:345–352. https://doi.org/10.1007/s100870050043

Laloy E, Vrugt JA (2012) High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing. Water Resour Res 48. 10.1029/2011wr010608

Lei X, Tang M, Lu Y, Hong L, Tian D (2009) Forest inventory in China: status and challenges. Int Forest Rev 11:52–63. https://doi.org/10.1505/ifor.11.1.52

Li W (2004) Degradation and restoration of forest ecosystems in China. Forest Ecol Manag 201:33–41

Lynch DW (1958) Effects of stocking on site measurement and yield of second-growth ponderosa pine in the Inland Empire. Research Paper 56. USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden, UT.

MacFarlane DW, Green EJ, Burkhart HE (2000) Population density influences assessment and application of site index. Can J For Res 30:1472–1475

Mao J, Wang X (2011) Distinct niche divergence characterizes the homoploid hybrid speciation of Pinus densata on the Tibetan Plateau. Am Nat 177:424–439. https://doi.org/10.1086/658905

Perin J, Hébert J, Brostaux Y, Lejeune P, Claessens H (2013) Modelling the top-height growth and site index of Norway spruce in Southern Belgium. Forest Ecol Manag 298:62–70. https://doi.org/10.1016/j.foreco.2013.03.009

de Perthuis Laillevault de R (1803) Traité de l'aménagement et de la restauration des bois et forêts de la France. Madame Huzard, Paris, p 384

Poudel KP, Cao QV (2013) Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. For Sci 59:243–252. https://doi.org/10.5849/forsci.12-001

Pretzsch H (2010) Forest Dynamics, Growth and Yield. Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-540-88307-4

Prodan M (1965) Holzmeßlehre. JD Sauerländer's Verlag, Frankfurt am Main, 644 pp

Pukkala T (1989) Predicting diameter growth in even-aged scots pine stands with a spatial and non-spatial model. Silva Fenn 23:101–116. https://doi.org/10.14214/sf.a15533

Pukkala T, Lähde E, Laiho O (2009) Growth and yield models for uneven-sized forest stands in Finland. Forest Ecol Manag 258:207–216. https://doi.org/10.1016/j.foreco.2009.03.052

Pukkala T, Lähde E, Laiho O (2011) Using optimization for fitting individual-tree growth models for uneven-aged stands. Eur J Forest Res 130:829–839. https://doi.org/10.1007/s10342-010-0475-z

Qin J, Cao QV (2006) Using disaggregation to link individual-tree and whole-stand growth models. Can J For Res 36:953–960. https://doi.org/10.1139/x05-284

Reineke LH (1933) Perfecting a stand-density index for even-aged forests. J Agric Res 46:627–638

Ritchie MW, Hann DW (1997) Implications of disaggregation in forest growth and yield modeling. For Sci 43:223–233

Robinson A (2016) equivalence: Provides tests and graphics for assessing tests of equivalence. R package version 0.7.2. https://CRAN.R-project.org/package = equivalence

Robinson AP, Duursma RA, Marshall JD (2005) A regression-based equivalence test for model validation: shifting the burden of proof. Tree Physiol 25:903–913. https://doi.org/10.1093/treephys/25.7.903

Romijn E, Lantican CB, Herold M, Lindquist E, Ochieng R, Wijaya A, Murdiyarso D, Verchot L (2015) Assessing change in national forest monitoring capacities of 99 tropical countries. For Ecol Manag 352:109–123. https://doi.org/10.1016/j.foreco.2015.06.003

R Development Core Team (2017) R: A language and environment for statistical computing. R foundation for statistical computing

Schröder J, Soalleiro RR, Alonso GV (2002) An age-independent basal area increment model for maritime pine trees in northwestern Spain. Forest Ecol Manag 157:55–64. https://doi.org/10.1016/s0378-1127(00)00657-5

Schumacher F (1939) A new growth curve and its application to timber yield studies. J For 37:33

Sivia D, Skilling J (2006) Data analysis: a Bayesian tutorial. Oxford University Press, Oxford

Tomppo E, Gschwantner T, Lawrence M, McRoberts RE, Gabler K, Schadauer K, Vidal C, Lanz A, Ståhl G, Cienciala E. (2010) National forest inventories. Pathways for Common Reporting. European Science Foundation: 541-553.

Van Oijen M, Rougier J, Smith R (2005) Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiol 25:915–927. https://doi.org/10.1093/treephys/25.7.915

Vanclay JK (1994) Modelling forest growth and yield: applications to mixed tropical forests. CAB International, Wallingford

Vanclay JK (2010) Robust relationships for simple plantation growth models based on sparse data. Forest Ecol Manag 259:1050–1054. https://doi.org/10.1016/j.foreco.2009.12.026

Vrugt JA, ter Braak CJF, Diks CGH, Robinson BA, Hyman JM, Higdon D (2009) Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. In: Int J Nonlin Sci Num 10: 273-290. https://doi.org/10.1515/ijnsns.2009.10.3.273

Weiskittel AR, Hann DW, Kershaw JA Jr, Vanclay JK (2011) Forest Growth and Yield Modeling. John Wiley & Sons, Oxford. https://doi.org/10.1002/9781119998518

Wu G, Feng Z (1994) Study on the social characteristics and biomass of the Pinus tabulaeformis forest systems in China. Acta Ecol Sin 14(4):415–422

Yue C, Kohnle U, Hein S (2008) Combining tree-and stand-level models: a new approach to growth prediction. For Sci 54:553–566

Zhang S, Amateis RL, Burkhart HE (1997) Constraining individual tree diameter increment and survival models for loblolly pine plantations. For Sci 43:414–423

Zhang X, Lei Y, Cao QV (2010) Compatibility of stand basal area predictions based on forecast combination. For Sci 56:552–557