Forestry Advance Access published March 17, 2015

Forestry An International Journal of Forest Research

Institute of Chartered Foresters

Forestry 2015; **00**, 1–10, doi:10.1093/forestry/cpv003

Bayesian calibration and Bayesian model comparison of a stand level dynamic growth model for Sitka spruce and Scots pine

J. Lonsdale[1]*, F. Minunno[2], M. Mencuccini[1,3] and M. Perks[4]

[1]*School of Geosciences, University of Edinburgh, Edinburgh EH9 3JN, UK*
[2]*Department of Forest Sciences, University of Helsinki, Yliopistonkatu 4, 00100 Helsinki, Finland*
[3]*ICREA at CREAF, Cerdanyola del Valles, Barcelona, Spain*
[4]*Forest Research, NRS, Roslin, Midlothian EH25 9SY, UK*

*Corresponding author. E-mail: jack.lonsdale@ed.ac.uk

Received 13 June 2014

Growth models continue to be of importance in modern multi-functional forestry to provide forecasts. Bayesian calibrations allow both model structure and parameters to be assessed simultaneously in a probabilistic framework, providing a model with which forecasts and their uncertainty can be better understood and quantified using posterior probability distributions. A Bayesian calibration of a stand-level dynamic growth (SLeDG) model is carried out for both Sitka spruce and Scots pine in the UK for the first time. The calibration used the differential evolution Markov-Chain method to reduce the required number of iterations for inference. Two different model structures were considered for estimating local stand productivity: one using the measured height–age relationship, and one using estimated site yield class. The height–age relationship was shown to be more probable for both species in a Bayesian model comparison (total model probability = 0.64 and 0.58 for Sitka spruce and Scots pine, respectively), although metrics of model performance were similar for both model structures ($R^2 \geq 0.88$ in all variables). A complete calibration (using all data) of the more probable model structure was then completed, and excellent model fit was observed ($R^2 > 0.95$ for all variables in both species). Example forecasts using the output from the calibration were demonstrated, and are compatible with existing yield tables for both species. This method could be applied to other species or other model structures in the future.

Introduction

Due to the time scales involved in a forest rotation, modern multi-purpose forestry continues to require robust forest growth models to forecast timber, carbon storage and even time to maturity for recreational uses. Current computing power and availability has resulted in a marked increase in the use of Bayesian approaches when calibrating models across a range of fields. Calibration involves the fitting of models to data by adjusting parameters, and Bayesian calibration allows for both the identification of optimal model parameters and the potential to quantify the uncertainty in future predictions (Kennedy and O'Hagan, 2001).

For the parameterization of forestry models, Bayesian approaches have been shown to produce outputs comparable with maximum likelihood approaches (Li *et al.*, 2011; Laloy and Vrugt, 2012), but with the added benefit of a full posterior distribution of parameters, thereby facilitating a more thorough understanding of areas of uncertainty within a model. It is also relatively simple to use the Bayesian method to investigate uncertainty in model structure through Bayesian model comparison (BMC) (van Oijen *et al.*, 2013). While Bayesian approaches are becoming more prevalent for forest ecosystem modelling, particularly for calibration of process-based models (Minunno *et al.*, 2013; van Oijen *et al.*, 2013), they have not seen wide applications in forest growth and yield modelling. Site index (Li *et al.*, 2011), tree mortality (Metcalf *et al.*, 2009), diameter distribution (Bullock and Boone, 2007) and yield models (Green and Strawderman, 1996) have all used Bayesian approaches; however, a complete calibration of an empirical stand-level growth model has not yet been done using a Bayesian approach.

Stand-level dynamic growth (SLeDG) models based on the initial work of García (1979) have been developed for a variety of tree species around the world: most recently trembling aspen (*Populus tremuloides* Michx.) in British Columbia (García, 2013) and Scots pine in the UK (Lonsdale *et al.*, 2014). SLeDG models use differential equations to forecast a state vector composed of common stand-level forest inventory measures (top height, number of stems and basal area). Their strength lies in that they are able to provide robust predictions from any point in state space after calibration with even small datasets (Lonsdale *et al.*, 2014). It should be noted that while some process-based models may also be stand level, dynamic growth models, here SLeDG

Table 1 Summary statistics for permanent sample plot (PSP) stands used

	Mean	SD	Minimum	Quartile 1	Median	Quartile 3	Maximum
Sitka spruce ($n = 171$)							
YC (m^3 ha^{-1} $year^{-1}$)	18.10	4.58	5.00	15.00	17.00	21.00	32.00
Age (years)	34.34	12.94	12.00	23.00	34.00	42.00	87.00
Top height (m)	18.92	6.71	5.20	13.6	18.3	22.8	41.3
Trees ha^{-1}	1660.2	1135.6	119.0	737.5	1413.5	2320.5	7769.0
Basal area (m^2)	43.56	15.95	4.36	31.66	41.03	55.01	93.41
Scots pine ($n = 134$)							
YC (m^3 ha^{-1} $year^{-1}$)	9.61	2.08	5.00	8.00	9.00	11.00	18.00
Age (years)	45.59	19.90	14.00	30.0	43.0	57.0	123.0
Top height (m)	16.27	4.81	5.80	12.60	15.90	20.0	28.90
Trees ha^{-1}	1460.8	1192.6	78.0	618.8	1073.5	1939.0	6415.0
Basal area (m^2)	32.71	11.50	3.23	24.54	30.50	38.96	75.26

Number of available permanent sample plots are given in brackets after each species' name.
SD, standard deviation; YC, yield class.

models specifically refer to models of the form suggested by García (1979) as a convenient acronym.

The prevalent methodology for parameterizing SLeDG models is to use a combination of the values of maximum likelihood to estimate the top height growth parameters (Hu and García, 2010), non-linear least-squares to estimate mortality parameters (García, 2009) and basal area growth parameters (García *et al.*, 2011). While using these methods of parameter estimation to successively parameterize a model undoubtedly produce acceptable results, it is often difficult to identify suitable starting points for parameter calibrations when using these methods of model fitting, and the differential equations strong non-linearity can lead to algorithms reaching false optimization maxima in the state space. It is also suggested that simultaneous parameter estimation may reduce model errors (Vanclay and Skovsgaard, 1997). Over 30 years ago Garcia (1983) suggested Bayesian methods could be used in the place of maximum likelihood for parameter estimation in SLeDG models; however, to date, there appears to have been no published attempts to calibrate this model using Bayesian methods.

Thus here we demonstrate the feasibility of using the Bayesian methods to parameterize dynamic forest growth models. We have completed two Bayesian calibrations of an SLeDG model for the two most important conifer species in the UK: Sitka spruce (*Picea sitchensis*) and Scots pine (*Pinus sylvestris*). Together they make up 68% of British conifer forests (Forestry Commission, 2011). Markov-Chain Monte Carlo (MCMC) methods are often used in Bayesian calibrations for posterior inference (e.g. van Oijen *et al.* 2013) and here we use a recent development of this method: differential evolution Markov-Chain (DE-MC) (ter Braak, 2006). The DE-MC is detailed in the Methods section, but it most importantly provides a method of accelerating chain convergence.

Here we provide a brief description of the SLeDG model, as well as the theory of the DE-MC. Then two different ways of including an estimate of local site productivity are compared using a BMC. Finally we provide parameterizations, validation and uncertainty measures for the optimal models selected by BMC for both species, and provide an example prediction forecast based on the parameterization and forecast. We conclude by discussing the most important implications of this work for future uses of Bayesian calibration of forest growth models.

Methods

Overview

The data for the calibrations were provided by the Forest Research Forest Mensuration, Modelling and Forecasting Group and comprised permanent sample plot (PSP) measurements from 171 Sitka spruce plots and 134 Scots pine plots. The main characteristics of the plots are summarized in Table 1. Each plot included measurements at two or more points in time with intervals between measurements ranging between 1 and 45 years for Sitka spruce and 1 and 49 years for Scots Pine plots. To capture the shape of growth trajectories, stand conditions at planting were added to the measurements based on the recorded initial spacings (ranging from 0.9 m to 2.4 m square spacing for both species), and assumed planted trees were breast height (1.3 m).

The data for each species were initially split into half ($n/2$), fully randomly to provide a cross-validation for the model comparison with separate calibration and validation datasets. Bayesian calibrations using DE-MC were performed for each species using the calibration data. Two different model structures for estimating local site productivity were calibrated (detailed below). A BMC was then performed using the validation dataset to identify the preferred model structure. Finally, a model calibration of the preferred structure was run using the whole dataset to provide posterior parameter distribution estimates which also indicate uncertainty.

SLeDG model

SLeDG models are based on the modelling approach first suggested by García (1979). This family of models is biologically

based and uses a state-space approach to forecast the stand development. Stands are described using a state vector of common forest inventory measures, and forecasts through time can be made from any point in state space. A more detailed explanation of SLeDG models can be obtained in the Supplementary information of Lonsdale *et al.* (2014) or García *et al.* (2011). In this instance three variables make up the state vector: top height (H), number of stems per hectare (N) and the product of basal area and top height ($B \times H = W$), for convenience henceforth referred to as tubular stem. This variable is highly correlated with merchantable volume and biomass, and behaves more simply than the basal area alone.

Changes in the state vector are forecast as a function of itself ($\mathrm{d}X/\mathrm{d}t = f(X)$) with an equation for each part of the state vector:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = b_1[(b_2/H)^{b_3} - H] \tag{1a}$$

$$\frac{\mathrm{d}N}{\mathrm{d}H} = -b_4 H^{b_5} N^{b_6} \tag{1b}$$

$$\frac{\mathrm{d}W}{\mathrm{d}H} = b_7 H^{b_8} N^{b_9} - b_9 \frac{W}{N}\frac{\mathrm{d}N}{\mathrm{d}H} \tag{1c}$$

where b_i are parameters to be estimated. Parameters may be estimated either based on growth intervals as is the case in this paper, or change from planting may be used. García (2010a) suggested that SLeDG model calibrations may be insensitive to the choice of time interval. Equation (1a) is a form of the von Bertalanffy model commonly used in growth models for both plants and animals. A measure of site-specific productivity was included by setting b_1 as a site-specific parameter, which results in anamorphic height growth curves of consistent shape with height scaled by the site-specific parameter. Anamorphic height growth curves have been previously used for both Sitka spruce and Scots pine in the UK (Rennolls, 1995; Lonsdale *et al.*, 2014).

Both changes in stem number and tubular stem (Equations (1b) and (1c)) are for given changes in height rather than time, as such changes are likely to be better predicted by a physiological change (such as height) rather than temporal change (García, 2010b).

Bayesian calibration with differential evolution Markov-Chain

Bayes' formula states that the probability of the parameters (θ) given data (D) in a model is proportional to the current uncertainty in parameters ($P(\theta)$) multiplied by a likelihood function ($\mathrm{P}(D|\theta)$): a measure of fit between the model and the data given the parameters. Thus,

$$P(\theta|D) \propto P(\theta)P(D|\theta) \tag{2}$$

These three terms are usually referred to as the posterior, the prior and the likelihood, respectively. Bayesian calibration begins with the specification of a prior uncertainty in the set of parameters as a multivariate probability distribution. The likelihood function is then determined using the model's ability to match the data (D) with given candidate parameters θ (van Oijen *et al.*, 2005).

To perform the calibration an iterative process which explores parameter space is often used to simulate the distributions of Equation (2). The most common and general of these is the Metropolis algorithm (Metropolis *et al.*, 1953). In this method, a proposal of new parameters (θ_x) is suggested by adding a random shift (γ) to the existing parameters (θ_{x-1}):

$$\theta_x = \theta_{x-1} + \epsilon \tag{3}$$

The ratio between the proposal posterior and the existing posterior is calculated (known as the Metropolis ratio β):

$$\beta = \frac{\mathrm{p}(\theta_x|D)}{\mathrm{p}(\theta_{x-1}|D)} \tag{4}$$

The proposed parameters may be accepted with a probability equal to the Metropolis ratio and the chain of iterations moves forward from that point in parameter space, otherwise returning to θ_x for another random shift (van Oijen *et al.*, 2013). The difficulty in this method lies in trying to balance between making sufficiently large shifts to progress through the parameter space, while still accepting enough proposed parameters.

More recent developments have suggested improvements in this random-walk method, allowing for more directed exploration of parameter space. In this study the DE-MC with snooker updater is used. The standard DE-MC method was developed by ter Braak (2006) and further refined by ter Braak and Vrugt (2008). It aims to improve the scale and orientation of jumps in the MCMC method. This is achieved by using a number of iterative chains of parameters, n. Each parameter vector θ_i (where $i = 1 : n$), or chain state, is updated using a proposal based on the difference between the parameter position of two other chains' states. Thus a given chain's proposal (θ_i^*) is given by

$$\theta_i^* = \theta_i + \gamma(\theta_{R1} - \theta_{R2} + \varepsilon) \tag{5}$$

where γ is a user-defined scalar, ε is a random vector with a small variance compared with the posterior and θ_{Ri} are drawn from other chains' states θ_{-i}. Earlier versions of the DE-MC required a number of chains greater than the dimensionality of the parameters (ter Braak, 2006), however by sampling difference vectors from previous points in the chains the number of chains that must be run is reduced (ter Braak and Vrugt, 2008). Thus θ_{Ri} is sampled from the present state of other chains as well as the previous states back to a user-defined previous number of states (M_0). A further enhancement used in this analysis is the snooker updater which instead of a simple vector difference (Equation (5)) uses the difference in the orthogonal projection of two other chain states (from past or present) onto a line between the chain state being updated and another random chain state, $\mathbf{z}$. The proposed update is given by

$$\theta_i^* = \theta_i + \gamma(\mathbf{z_{P1}} - \mathbf{z_{P2}}) \tag{6}$$

where $\mathbf{z_{Pi}}$ is the orthogonal projected position of other chains onto the line $\theta_i - \mathbf{z}$, and γ is again a user-defined scalar. The proposed update is accepted with a probability equal to to the Metropolis ratio (ter Braak and Vrugt, 2008).

Relatively uninformative uniform priors were used for each of the parameters (b_i) (Table 3). The parameters do not represent any physical value (such as is the case in process-based models), and thus theoretically do not have limits. However, to ensure convergence in a reasonable number of iterations all other parameters were loosely constrained with uniform priors based on previous estimates of parameters for SLeDG models with similar structure (e.g Broad and Lynch, 2006; García *et al.*, 2011; Lonsdale *et al.*, 2014). Parameters were not allowed to go below zero, as values below zero would result in negative growth (e.g. b_1, b_2, b_7), or increases in stem numbers (e.g. b_4). (See Equations (1a)–(1c).) Additionally, no previous parameterizations of an SLeDG style model have found negative parameters. Measurement uncertainty was not provided in the PSP dataset and so it was assumed that top height had 10% measurement uncertainty, number of stems 20% uncertainty and tubular volume 30% uncertainty (as it also includes the uncertainty in top height). Only unthinned stands were used in the calibration of the mortality part of the model (Equation (1b)), and it was assumed that basal area should always be increasing as negative basal area is indicative of disturbance events (García, 2013) such as windthrow which are not accounted for in this model. Alternative forms for predicting basal area growth may accommodate negative growth in the future; however, it is also possible that observed negative growth is a result of the difficulty in measuring diameter (Clark *et al.*, 2007), which is required to estimate the basal area. Thus stands where the basal area decreased were excluded from the calibrations. For each model considered in this study, three million iterations split between three chains were run to determine the probability density functions for parameters b_i. The first half of these iterations were discarded as burn-in (iterations where the chain is yet to reach the area of high posterior probability). Chain convergence was assessed visually and the Gelman–Rubin diagnostic (Gelman and Rubin, 1992) was calculated and checked to have a value less than 1.1 for all parameters.

Bayesian model comparison

BMC follows logically from Bayesian calibration, but instead of model parameters being informed by data it is the model selection. Thus the probability of a model given data $P(M \mid D)$ is proportional to the product of the prior probability of a model $P(M)$ and an integrated likelihood, $P(D \mid M)$. By separating the data (D) in half to give calibration data (D_c) (which has been used to estimate the prior parameter distribution $P(\theta \mid D_c)$) and validation data (D_v) the integrated likelihood can be calculated using the law of total probability as

$$P(D_v|M, D_c) = \int P(D_v \mid M(\theta))P(\theta \mid D_c)\,\mathrm{d}\theta \tag{7}$$

The integrated likelihood was estimated using the validation dataset for each species, with the prior distributions estimated by a calibration of the model with the calibration dataset.

The advantage of BMC versus frequentist model comparison indices such as Akaike information criterion (AIC) is that parameter uncertainty determines a models' probability rather than a single parameter vector (Tuomi *et al.*, 2008).

Two approaches to estimating site-specific productivity were compared using the Bayesian model comparison. The first method (Model 1) used an integrated form of Equation (1a) to estimate the site productivity given a fixed origin $[t_1, H_1] = [0, 0]$, the height at time t_x being given as

$$H_x = b_1(1 - e^{-b_2 t_x})^{1/b_3} \tag{8}$$

This equation can then be solved for b_1 giving a site-specific productivity. The second method (Model 2) of including an estimate of site productivity used the Forestry Commission determined YC. A simple relationship between YC and a SI at age 50 (H_{50}) is given as

$$H_{50} = \alpha_1 + \alpha_2 \mathrm{YC} \tag{9}$$

The estimated H_{50} can then be substituted into Equation (8) with $t_x = 50$, and the local parameter calculated as for the first method. The parameters α_i are estimated as part of the calibration, constrained by uniform priors of −50–50.

The model comparison used the half of the dataset not used for the initial Bayesian calibration. Model probabilities for determining each of the state variables (H, N and W) as well as for the model as a whole were compared to identify the most suitable model structure.

Results

Bayesian model comparison

The results from the comparison of the two alternative model structures employed to include site quality when predicting the validation dataset are given in Table 2. The fit of the validation half of the data to the highest probability parameter vector predictions was assessed using the root mean standard error (RMSE) and a psuedo-R^2 as well as the BMC. The model using existing site YC as a measure of productivity (Model 2) performs better than the alternative model, especially for Scots pine due to lower RMSE and R^2 values closer to 1. Conversely, the BMC results indicate the model structure that uses the current state to estimate productivity (Model 1) has a higher probability for both species. It should be noted that the differences in performance are small between the two models according to all of the benchmarking statistics used here. Nonetheless, as the full calibration uses Bayesian methods the decision was made to perform calibrations for both species using the Model 1 structure, which showed the highest probability over all variables and in total according to the BMC.

Model calibration

For both species the full calibration reached convergence after the three million iterations with Gelman–Rubin statistics <1.1 and convergence observed in a graphical assessment of the Markov-Chain. Three million iterations took 50 min for Scots pine and

Table 2 Model percentage probability and fit statistics of validation data to model runs using most probable parameter vectors for each state variable for both Sitka spruce and Scots pine. Model 1 refers to the model that uses current state to estimate site productivity. Model 2 refers to the model that uses estimated yield class (YC) to define site productivity

Model	Top height, H (m)			Stem density, N (stems ha^{-1})			Tubular volume, W (m^3)			Total model
	Probability	RMSE	R^2	Probability	RMSE	R^2	Probability	RMSE	R^2	Probability
Sitka spruce										
1	0.73	1.56	0.89	0.53	427	0.88	0.51	128	0.88	0.64
2	0.27	1.09	0.95	0.47	361	0.89	0.49	106	0.93	0.36
Scots pine										
1	0.69	0.64	0.98	0.51	173	0.98	0.50	65.0	0.95	0.58
2	0.31	0.59	0.98	0.50	151	0.98	0.50	83.9	0.90	0.42

Table 3 Uniform prior bounds and posterior parameter estimates for Sitka spruce and Scots pine. Most probable parameters are those that had the highest product of likelihood and prior vector in the Markov-Chain. There is no distribution for b_1 as it is a local parameter calculated as a function of b_2, b_3 and stand height at measured time

	Priors		Posteriors	SD	Most probable	95% credible interval	
	Min	Max	Mean			Min	Max
Sitka spruce							
b_2	0.00	1.00	0.0115	3.42×10^{-4}	0.0115	0.0112	0.0124
b_3	0.00	1.00	0.863	0.00126	0.863	0.860	0.865
b_4	0.00	1.00	3.64×10^{-5}	1.57×10^{-5}	3.64×10^{-5}	1.71×10^{-5}	8.07×10^{-5}
b_5	0.00	10.0	0.704	0.0339	0.704	0.638	0.769
b_6	0.00	10.0	1.76	0.0455	1.76	1.66	1.84
b_7	0.00	50.0	9.71	3.99	9.71	3.38	18.6
b_8	0.00	100	0.0584	0.0566	0.0583	2.09×10^{-3}	0.21
b_9	0.00	1.00	0.375	0.0738	0.375	0.24	0.536
Scots pine							
b_2	0.00	1.00	0.129	5.06×10^{-4}	0.0124	0.0124	0.0143
b_3	0.00	1.00	0.818	2.47×10^{-3}	0.820	0.812	0.822
b_4	0.00	1.00	8.70×10^{-3}	8.24×10^{-3}	4.03×10^{-3}	1.62×10^{-3}	0.0297
b_5	0.00	10.0	0.587	0.0698	0.620	0.446	0.711
b_6	0.00	10.0	1.21	0.0812	1.25	1.04	1.35
b_7	0.00	50.0	18.2	10.9	14.8	3.39	44.0
b_8	0.00	100	0.206	0.154	0.0157	0.0104	0.581
b_9	0.00	1.00	0.184	0.122	0.269	9.99×10^{-3}	0.458

60 min for Sitka spruce (all computations performed on 1.6 GHz Intel Core i5-2467M CPU). The most probable parameter vector is found in the chain as the Markov-Chain as the highest product of likelihood and prior vectors. The most probable parameter estimates are and their credible intervals for both species are given in Table 3.

The result of using the most probable parameter vector (see Table 3) to predict the observations in the full dataset can be seen in Figure 1. The state variable fit statistics using the full dataset for each species are given in Table 4. The model closely fits observations for both species in all state variables indicated by the R^2 values. However, slopes significantly different from 1 indicate that in Scots pine predictions of top height and number of stems tend to be overestimated, while tubular volume has a tendency to be underestimated for both species.

Example forecasts over a range of productivity classes using the estimated parameters with their uncertainty are given in Figures 2 and 3. The forecasts are made using a draw of 5000 parameter vectors from the Markov-Chain. The plots illustrate forecasts from the last measurement of a randomly selected stand from the dataset until the stand reaches 120 years. It is possible to see the skew in probable predictions in the location of the most probable predictions (dashed lines) differing from the centre of the shaded probability intervals. The predictions are mostly consistent with the existing yield tables of Edwards and Christie (1981) (dotted lines) which are commonly used for growth predictions for these species in UK forests. The yield tables do not agree particularly well with stem number and tubular volume measurements in the case of Scots pine (Figure 3).

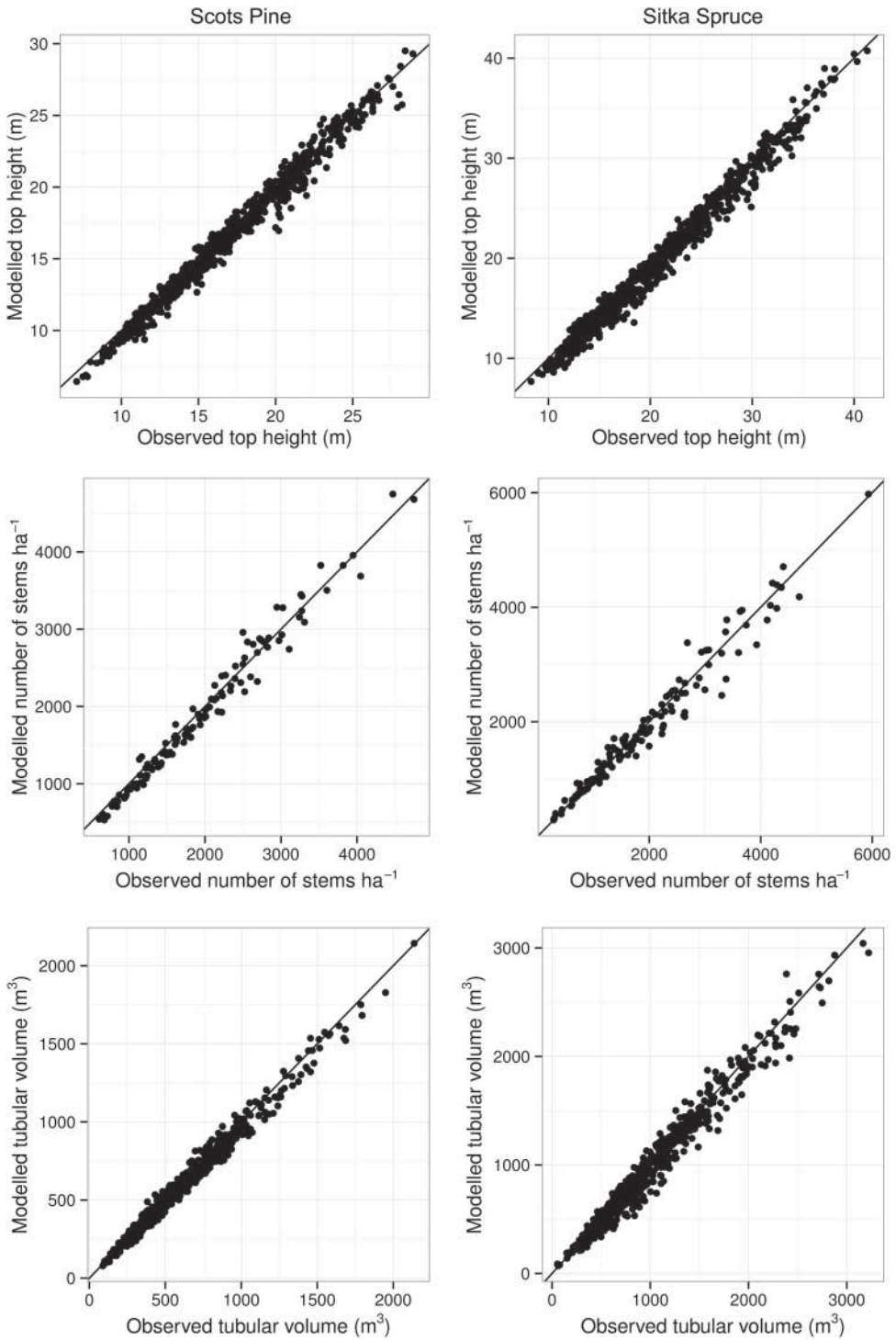

Figure 1 Model fit for both Scots pine and Sitka spruce SLeDG models using the most probable parameter vector given by the DE-MC calibration.

Discussion

The results presented here demonstrate that Bayesian calibration provides a way to calibrate dynamic forest models based on complete posterior distributions of model parameters and any model outputs of interest. Similarly to Li *et al.* (2011), we would not suggest that Bayesian methods replace frequentist methods, which have been shown to be adequate for parameterizing SLeDG models (e.g. García, 2013; Lonsdale *et al.*, 2014). However, there are several advantages in the Bayesian approach when considering

model uncertainty. The posterior parameter distributions provide a much better understanding of parameter uncertainty resulting in a better illustration of prediction uncertainty as can be seen in Figures 2 and 3. Performing model comparisons by using the posterior probability distributions is useful when considering the model structural uncertainty (Kass and Raftery, 1995).

Table 4 State variable fit statistics for Scots pine and Sitka spruce models using the most probable parameter vector against the full datasets. Parameter vectors were estimated using a full dataset calibration

	Top height (m)	Stems ha^{-1}	Tubular volume (m^3)
Sitka spruce			
RMSE	0.916	206	93.8
R^2	0.981	0.964	0.970
Slope	0.997	0.964	0.967*
Scots pine			
RMSE	0.596	137	42.1
R^2	0.983	0.978	0.982
Slope	1.02*	1.04*	0.955*

*Slopes significantly different to 1 ($P < 0.01$).

Here, the two different model structures could be considered demonstrations of 'stand site index' versus 'site site index' as described by García (2006). Model 1, which assumes the productivity is defined by the current state of a stand growing on a site is an example of stand site index. Conversely, Model 2, which has a fixed productivity throughout a rotation predefined by a site's estimated YC, thus a property of the site itself (not the trees growing on the site) is an example of site site index. The BMC results suggest better performance of Model 1 with the potential to reduce uncertainty; however, the RMSE and R^2 values from the validation show close to equal performance of the models. This suggests that the two additional parameters in Model 2 (α_i) are what cause the reduced model probability, equivalent to the penalization imposed by an increased number of parameters in AIC-based approaches. Values of YC for the vast majority of forests in the UK are available in the Forestry Commission's sub-compartment database (SCDB) and thus Model 2 provides a method to spatialize predictions across the UK. However, while measures such as local yield class (Edwards and Christie, 1981) may go some way to adjusting site quality measured over a rotation, only limited measurements of site productivity may be available due to cost of providing repeated stand measurements. The increased prevalence of remote sensing techniques such as Light detection and ranging (LiDAR) may reduce this cost, making Model 2 a favourable approach. Indeed, additional data

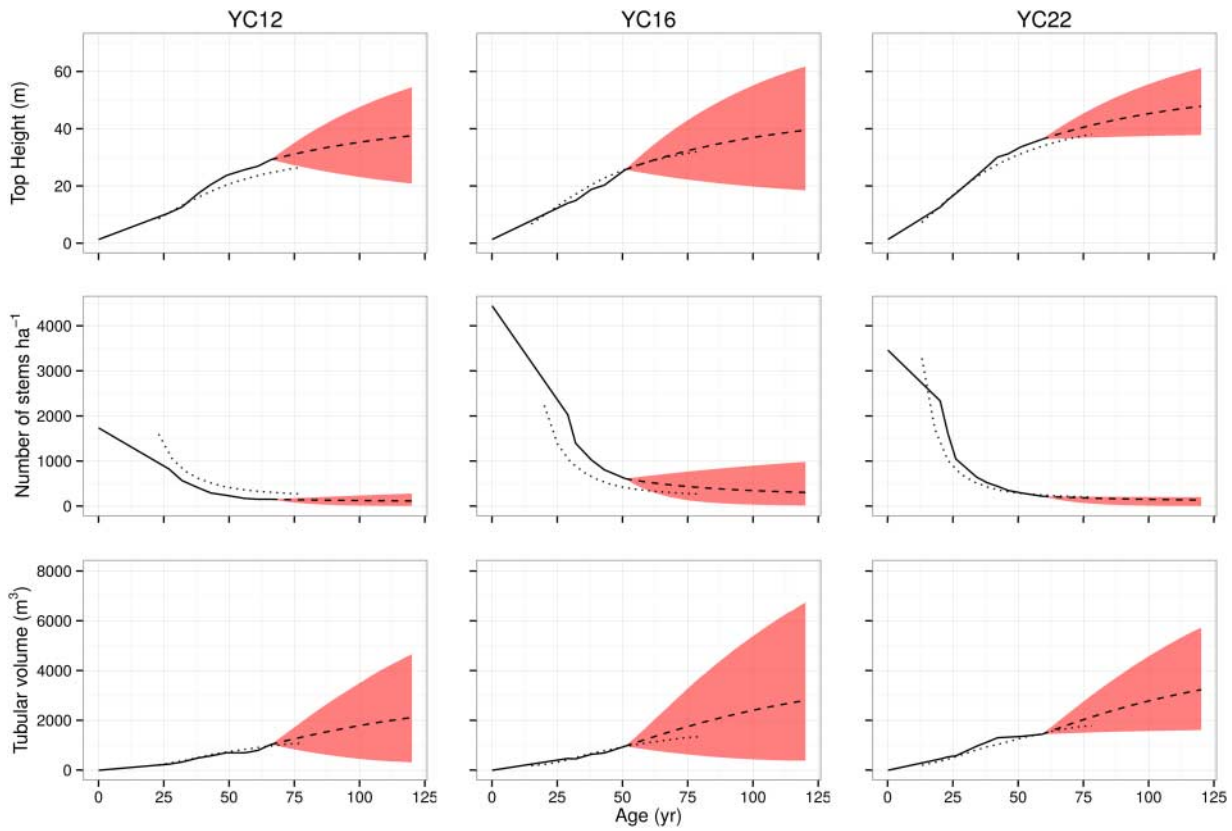

Figure 2 Example forecasts of state variables (continuous line) for three randomly selected Sitka spruce stands (Yield Classes 12, 16 and 22) using 5000 parameter draws from the Markov-Chain. Dashed line indicates most probale parameter vector prediction, and shaded area indicates 95% probable interval. Dotted line indicates forecast based on the site index and the yield tables of Edwards and Christie (1981).

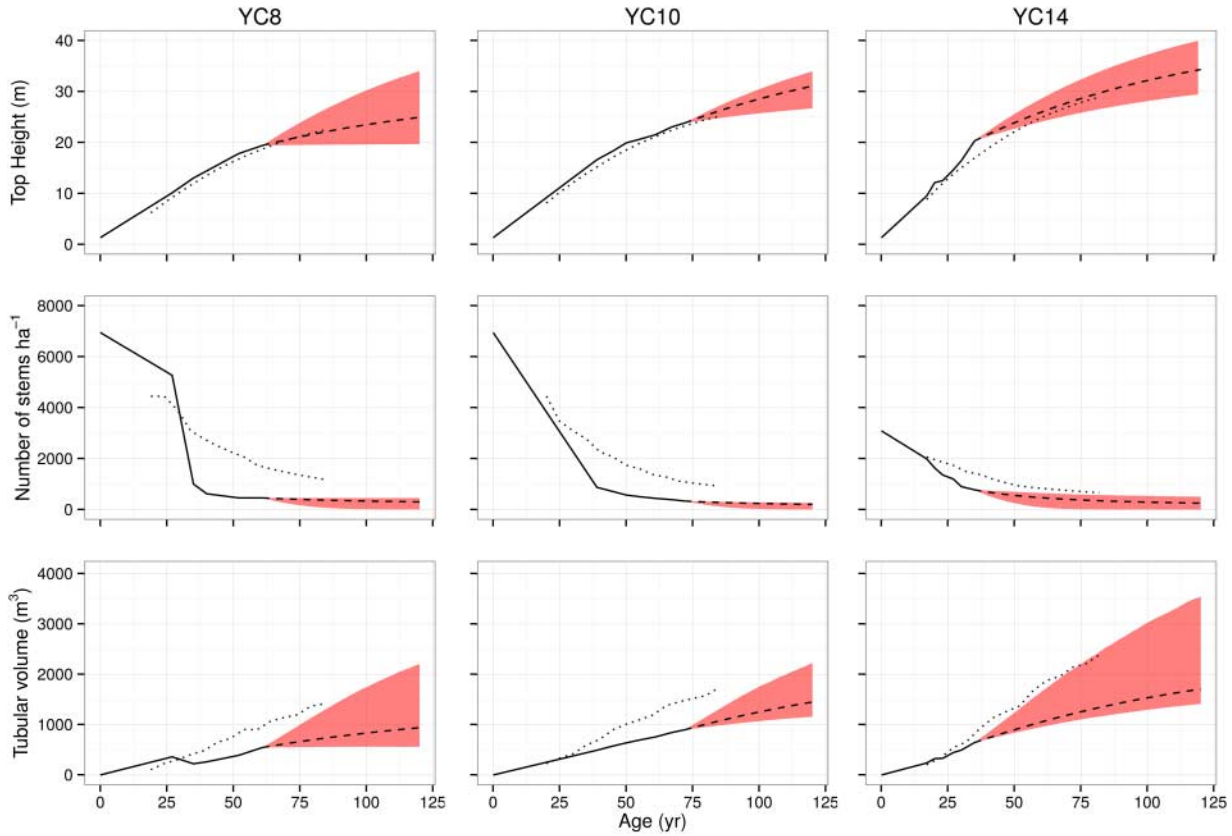

Figure 3 Example forecasts of state variables (continuous line) for three randomly selected Scots pine stands (Yield Classes 8, 10 and 14) using 5000 parameter draws from the Markov chain. Dashed line indicates most probable parameter vector prediction, and shaded area indicates 95% probable interval. Dotted line indicates forecast based on the site index and the yield tables of Edwards and Christie (1981).

provided by LiDAR measurements could be added to the existing dataset to further refine the growth model in a form of data assimilation similar to Patenaude *et al.* (2008), potentially further reducing parameter uncertainty. Overall, consideration of scale and data availability will also be of importance when selecting a model for making predictions. Large-scale prediction using Model 1 requires further data collection and processing for productivity estimation whereas Model 2 may be applied at the UK scale using the currently available data, although with increased uncertainty.

The final model calibrations for both species showed excellent fit between the observed and most probable parameter vector-based predictions for each of the state variables. To some extent this is to be expected as fits are shown for data with which the calibration was completed. However, one of the main arguments for the use of Bayesian calibrations is that both model and dataset may be fully utilized. The under- and over-prediction seen in certain state variables are small (Table 4), but may be improved by changing the structure of the sub-models (Equations (1a)–(1c)). For example García and Ruiz (2003) used a simpler mortality model than presented here, and alternative forms of the von Bertalanffy equation have been previously compared for site index (Li *et al.*, 2011). This would be straightforward to compare using a BMC, as previously suggested.

The broad prior distributions for the parameters allowed both species to converge on posterior parameter distributions within three million iterations. Accordingly, the same prior distributions could be employed to obtain posterior distributions and thus parameter estimates for other species in other regions given comparable PSP data. The prior distribution provided an adequate balance between exploring parameter space and time to compute the Markov chain. Alternative functions for prior distributions could be used and compared with a Bayesian sensitivity analysis (Insua *et al.*, 2000) but this was beyond the scope of this study. A simpler Markov-Chain method could have been used, but the number of iterations required to reach convergence would have been greater. Thus the DE-MC can be considered a computationally efficient approach to Bayesian calibration. Alternative Bayesian sampling methods could equally be used, such as Gibbs sampling (Gelman and Hill, 2007). The number of open source packages to perform such sampling are rapidly increasing (e.g JAGS Plummer, 2013, STAN Stan Development Team, 2014) allowing for rapid development of forest models in a Bayesian framework.

Conclusions

This study has shown how Bayesian calibration can provide a flexible all-in-one framework for developing forest growth models, and calibrating their parameters. Both model structure and parameters can be simultaneously investigated and probabilities assigned to each. Additionally, uncertainty estimates for both parameters and forecasts provided by posterior distributions enhance both model understanding and enhance accounting efforts. The successful calibration of the SLeDG model demonstrated here could easily be adapted for other species, and it is entirely possible to use this method to calibrate other forms of growth model. We would recommend this method for any future efforts in developing forest growth models.

Acknowledgements

We thank the Forest Research Forest Mensuration, Modelling and Forecasting Group for the provision of ForestYield predictions and provision of PSP data.

Conflict of interest statement

None declared.

Funding

This work was supported by the National Environmental Research Council; and the Forestry Commission. Funding to pay the Open Access publication charges for this article was provided by the Natural Environment Research Council [grant number NE/G00725x/1].

References

Broad, L.R. and Lynch, T. 2006 Growth models for Sitka spruce in Ireland. *Irish For.* **63**, 53–79.

Bullock, B.P. and Boone, E.L. 2007 Deriving tree diameter distributions using Bayesian model averaging. *For. Ecol. Manag.* **242**, 127–132.

Clark, J., Wolosin, M. and Dietze, M. 2007 Tree growth inference and prediction from diameter censuses and ring widths. *Ecol. Appl.* **17**, 1942–1953.

Edwards, P.N. and Christie, J. 1981 *Yield Models for Forest Management*. HMSO.

Forestry Commission 2011 Standing timber volume for coniferous trees in Britain: National Forest Inventory Report. *Tech. rep.*, Forestry Commission.

García, O. 1979 Modelling stand development with stochastic differential equations. In *Mensuration for Management Planning of Exotic Forest plantations. New Zealand Forest Service, FRI Symposium*. D. Elliott (ed.), no. 20.

Garcia, O. 1983 A stochastic differential equation model for the height growth of forest stands. *Biometrics* **39**, 1059–1072.

García, O. 2006 Site Index: concept and methods. In *The 2nd International Conference on Forest Measurements and Quantitative Methods and Management*. C.J. Cieszewski and M. Strub (eds). Hot Springs.

García, O. 2009 A simple and effective forest stand mortality model. *Math. Comput. For. Nat. Resource Sci.* **1**, 1.

García, O. 2010a A parsimonious dynamic stand model for interior spruce in British Columbia. *For. Sci.* **57**, 265–280.

García, O. 2010b Scube, a growth model for interior spruce in the SBS zone of British Columbia Working document.

García, O. 2013 Building a dynamic growth model for trembling aspen in western Canada without age data. *Can. J. For. Res.* **43**, 256–265.

García, O. and Ruiz, F. 2003 A growth model for eucalypt in Galicia, Spain. *For. Ecol. Manag.* **173**, 49–62.

García, O., Burkhart, H.E. and Amateis, R.L. 2011 A biologically-consistent stand growth model for loblolly pine in the Piedmont physiographic region, USA. *For. Ecol. Manag.* **262**, 2035–2041.

Gelman, A. and Hill, J. 2007 *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gelman, A. and Rubin, D. 1992 Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7**, 457–472.

Green, E. and Strawderman, W. 1996 Predictive posterior distributions from a Bayesian version of a slash pine yield model. *For. Sci.* **42**, 456–464.

Hu, Z. and García, O. 2010 A height-growth and site-index model for interior spruce in the Sub-Boreal Spruce biogeoclimatic zone of British Columbia. *Can. J. For. Res.* **40**, 1175–1183.

Insua, D., Ruggeri, F. and Martin, J. 2000 Bayesian sensitivity analysis. In *Sensitivity Analysis*. A. Satelli, K. Chan and E. Scott (eds), chap. 10. Wiley.

Kass, R. and Raftery, A. 1995 Bayes factors. *J. Am. Statist. Assoc.* **90**, 773–795.

Kennedy, M.C. and O'Hagan, A. 2001 Bayesian calibration of computer models. *J. R. Statist. Soc. B (Statist. Methodol.)* **63**, 425–464.

Laloy, E. and Vrugt, J.A. 2012 High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing. *Water Resources Res.* **48**, W01526.

Li, R., Stewart, B. and Weiskittel, A. 2011 A Bayesian approach for modelling non-linear longitudinal/hierarchical data with random effects in forestry. *Forestry* **85**, 17–25.

Lonsdale, J., Xenakis, G., Mencuccini, M. and Perks, M. 2014 A comparison of methods for quantifying standing carbon in UK Scots pine forests. In *Reducing Uncertainty in UK Forest Carbon Estimates: A Data Fusion Approach*. Ph.D. thesis. J. Lonsdale (ed.), chap. 2. University of Edinburgh, Edinburgh.

Metcalf, C.J.E., McMahon, S.M. and Clark, J.S. 2009 Overcoming data sparseness and parametric constraints in modeling of tree mortality: a new nonparametric Bayesian model. *Can. J. For. Res.* **39**, 1677–1687.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. 1953 Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087.

Minunno, F., van Oijen, M., Cameron, D. and Pereira, J. 2013 Selecting parameters for Bayesian calibration of a process-based model: a methodology based on canonical correlation analysis. *SIAM/ASA J. Uncertainty Quanti.* **1**, 370–385.

Patenaude, G., Milne, R., Van Oijen, M., Rowland, C.S. and Hill, R.A. 2008 Integrating remote sensing datasets into ecological modelling: a Bayesian approach. *Int. J. Remote Sensing* **29**, 1295–1315.

Plummer, M. 2013 *rjags: Bayesian graphical models using MCMC*. Comprehensive R Archive Network (CRAN), http://cran.r-project.org/web/packages/rjags/rjags.pdf (accessed on 20 February, 2015).

Rennolls, K. 1995 Forest height growth modelling. *For. Ecol. Manag.* **71**, 217–225.

Stan Development Team 2014 *Stan Modeling Language User's Guide and Reference Manual*. Stan Development Team http://mc-stan.org/manual.html (accessed on 20 February, 2015).

ter Braak, C.J.F. 2006 A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. *Statist. Comput.* **16**, 239–249.

ter Braak, C.J.F. and Vrugt, J.A. 2008 Differential evolution Markov Chain with snooker updater and fewer chains. *Statist. Comput.* **18**, 435–446.

Tuomi, M., Vanhala, P., Karhu, K., Fritze, H. and Liski, J. 2008 Heterotrophic soil respiration–Comparison of different models describing its temperature dependence. *Ecol. Model.* **211**, 182–190.

van Oijen, M., Rougier, J. and Smith, R. 2005 Bayesian calibration of process-based forest models: bridging the gap between models and data. *Tree Physiol.* **25**, 915–927.

van Oijen, M., Reyer, C., Bohn, F., Cameron, D., Deckmyn, G., Flechsig, M. *et al*. 2013 Bayesian calibration, comparison and averaging of six forest models, using data from Scots pine stands across Europe. *For. Ecol. Manag.* **289**, 255–268.

Vanclay, J. and Skovsgaard, J. 1997 Evaluating forest growth models. *Ecol. Model.* **98**, 1–12.