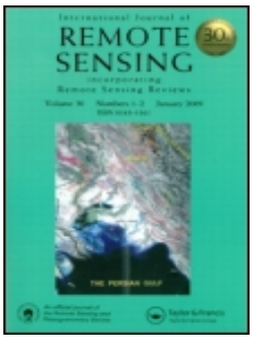

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: www.tandfonline.com/journals/tres20

Integrating remote sensing datasets into ecological modelling: a Bayesian approach

G. Patenaude, R. Milne, M. Van Oijen, C. S. Rowland & R. A. Hill

To cite this article: G. Patenaude, R. Milne, M. Van Oijen, C. S. Rowland & R. A. Hill (2008) Integrating remote sensing datasets into ecological modelling: a Bayesian approach, International Journal of Remote Sensing, 29:5, 1295-1315, DOI: 10.1080/01431160701736414

To link to this article: https://doi.org/10.1080/01431160701736414

International Journal of Remote Sensing
Vol. 29, No. 5, 10 March 2008, 1295–1315

Taylor & Francis
Taylor & Francis Group

Integrating remote sensing datasets into ecological modelling: a Bayesian approach

G. PATENAUDE*†, R. MILNE‡, M. VAN OIJEN‡, C. S. ROWLAND§ and R. A. HILL§

†School of Geosciences, University of Edinburgh, Drummond Street, Edinburgh EH3 9XP, UK

‡CEH Edinburgh, Bush Estate, Penicuik, Midlothian EH26 0QB, UK

§CEH Monks Wood, Abbots Ripton, Huntingdon, Cambridgeshire PE28 2LS, UK

Process-based models have been used to simulate 3-dimensional complexities of forest ecosystems and their temporal changes, but their extensive data requirement and complex parameterisation have often limited their use for practical management applications. Increasingly, information retrieved using remote sensing techniques can help in model parameterisation and data collection by providing spatially and temporally resolved forest information. In this paper, we illustrate the potential of Bayesian calibration for integrating such data sources to simulate forest production. As an example, we use the 3-PG model combined with hyperspectral, LiDAR, SAR and field-based data to simulate the growth of UK Corsican pine stands. Hyperspectral, LiDAR and SAR data are used to estimate LAI dynamics, tree height and above ground biomass, respectively, while the Bayesian calibration provides estimates of uncertainties to model parameters and outputs. The Bayesian calibration contrasts with goodness-of-fit approaches, which do not provide uncertainties to parameters and model outputs. Parameters and the data used in the calibration process are presented in the form of probability distributions, reflecting our degree of certainty about them. After the calibration, the distributions are updated. To approximate posterior distributions (of outputs and parameters), a Markov Chain Monte Carlo sampling approach is used (25 000 steps). A sensitivity analysis is also conducted between parameters and outputs. Overall, the results illustrate the potential of a Bayesian framework for truly integrative work, both in the consideration of field-based and remotely sensed datasets available and in estimating parameter and model output uncertainties.

1. Introduction

Process-based models are widely used in the fields of forest physiology and forest ecology as they enable deeper insights into the drivers of forest production and growth and offer higher flexibility than conventional production tables (Landsberg and Waring 1997). This flexibility enables the quantification and prediction of forest 2-dimensional and 3-dimensional structural variables due to deterministic, mechanistic and/or stochastic algorithms simulating the processes affecting growth.

*Corresponding author. Email: genevieve.patenaude@ed.ac.uk

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160701736414

However, their practical value has often been limited due to (a) their extensive data requirement and (b) their complexity and the difficulty in quantifying parameters and model output uncertainty (e.g. Gertner *et al.* 1999).

Remote sensing technology is increasingly exploited for forest inventorying and monitoring (e.g. Baulies and Pons 1995, Hyyppä *et al.* 2000) as it can provide insights into the spatial and temporal variability of forests, information which is seldom available from ground surveys alone. Spatially resolved ground based data can be time consuming, expensive but also logistically difficult to acquire where access to forested land is limited. Comparatively, highly resolved remote sensing data can be obtained at relatively low costs. Additionally, novel approaches now supply estimates of forest structural variables of accuracy equivalent if not superior to traditional measurement techniques (e.g. Magnussen and Boudewyn 1998, Hyyppä *et al.* 2001). Remote sensing may therefore help meet forest ecophysiologists and modellers' data requirements.

In this context, we present Bayesian Calibration (BC) as a means to integrate remotely acquired datasets into ecological models. BC provides a probability distribution over parameter values that are possible: it is not restricted to identifying only a single 'optimal' parameter set. BC is a systematic procedure for adjusting the parameter distribution when new data arrive. This approach therefore offers a number of advantages in comparison with goodness-of-fit and optimization approaches:

- BC facilitates the integration of data of varying degrees of uncertainty (Van Oijen *et al.* 2005) including that retrieved remotely.
- BC enables the quantification of uncertainty associated with parameters and model outputs, an important requirement for practical applications of models (Green *et al.* 2000, Van Oijen *et al.* 2005).
- Parameters and data used in the calibration process are presented in the form of probability distributions, reflecting our degree of certainty about them (Jansen 1999).
- BC enables the updating of distributions as further information is gained (Ghazoul and McAllister 2003).

In this paper, we demonstrate the usefulness of the approach by calibrating the 3-PG model (Physiological Processes Predicting Growth, Landsberg and Waring 1997, Sands and Landsberg 2002) for UK Corsican pine stands (*Pinus nigra* car. *maritima* (AIT.) Melv.). 3-PG is built on a combination of process-based calculations, several key simplifying assumptions and few empirical relationships. The model predicts gross and net primary production as well as biomass allocation to different pools. Over the years, it has been increasingly and successfully applied to new species worldwide (Landsberg and Waring 1997, Law *et al.* 2000, Waring 2000, Coops and Waring 2001, Coops *et al.* 2001, Sands and Landsberg 2002, Almeida *et al.* 2004, Stape *et al.* 2004). However, the parameterisation of the model for new species remains a challenge. As stated by Sands (2004, p.3): 'In only a few cases have parameters characterizing a species been rigorously determined, and even then this has been largely by a process of trial and error.' In this context, the aim of this paper is to illustrate the potential of BC as a means to (i) calibrate models for novel species (ii) integrate multi-source datasets and (iii) quantify model parameters and outputs along with uncertainty.

2. Structure of the 3-PG model

The 3-PG model has monthly or annual time steps and entails five state variables—foliage, stem biomass, root biomass, stocking density and available soil water, in conjunction with five submodels—biomass production, biomass allocation, soil water availability and evapotranspiration, mortality, and inventory variables. The required climatic data are monthly average values of solar radiation ($MJ\,m^{-2}\,d^{-1}$), atmospheric water pressure deficit (mbar), mean air temperature (°C), rainfall ($mm\,month^{-1}$) and frost days. Other input variables include site latitude, an estimate of soil fertility, maximum available soil water (mm per depth of rooting zone, in metres) and a general description of soil texture. Soil fertility is a dimensionless parameter (0–1). Near 0 values characterize poor soils (e.g. sandy outwash), and near 1, fertile soils (e.g. loamy soils). 3-PG outputs considered in the calibration were leaf area index (LAI, projected), above ground biomass (ABG biomass, $t\,ha^{-1}$), stem biomass ($t\,ha^{-1}$), foliage biomass ($t\,ha^{-1}$), root biomass ($t\,ha^{-1}$) and stem height (m).

2.1 *Biomass production*

The biomass submodel converts solar radiation into dry matter. The interception of radiation is defined by Beer's law and canopy LAI. The amount of photosynthetically active radiation intercepted by a stand (ϕ_{pa}, $mol\,MJ^{-1}$) is then converted into carbohydrates by means of a canopy quantum efficiency coefficient (α_{cx}, $mol\,mol^{-1}$) and a conversion factor converting carbohydrates into dry matter (parameter values and conversion factors used in this study are provided in §4). Further constraints on assimilation are then applied by dimensionless environmental factors varying between 0 and 1 (1 indicates optimal conditions). These factors, also referred to as modifiers, are multiplicative and are derived as functions of a combination of the following influences: vapour pressure deficit (D) or soil moisture, whichever is the most limiting, mean air temperature (T), frost, and soil nutrition on photosynthetic assimilation. Equations are provided in Sands (2004). Gross primary productivity (P_g, $t\,ha^{-1}\,d^{-1}$) is then converted to net primary productivity (P_n, $t\,ha^{-1}\,d^{-1}$) using a P_n/P_g ratio (Y) (Waring *et al.* 1998).

2.2 *Biomass allocation and mortality*

P_n is then allocated to the different plant components (roots, foliage and stems including branches) at each time step. Allocation to roots is proportional to the harshness of the environment. It is influenced by site fertility, stand age and either water pressure deficit or soil moisture (whichever of the latter two is most limiting), but does not fall below or exceed set values of minimum and maximum allocation to roots. The remaining P_n is shared between stems and foliage through a foliage-to-stem allocation ratio, given by an allometric relationship with mean diameter at breast height (Sands and Landsberg 2002, Sands 2004). DBH is itself obtained from an allometric relationship with stem biomass. Whereas P_n partitioning parameters must generally be estimated from fitting methods, those pertaining to the allometric relationship between stem biomass and diameter can be derived from forest mensuration (Sands and Landsberg 2002). Mortality is applied through the self thinning 3/2 law, which defines the mean plant biomass as proportional to a $-3/2$ power of their densities. This sets an upper limit to the mean single-tree stem mass at a given stocking level.

2.3 *Soil water balance*

Available soil water Θ (mm month^{-1}) is governed by rainfall interception by the canopy (i_R), rainfall (R_P, mm month^{-1}) and evapotranspiration (E_T, mm month^{-1}). If the maximum available water at saturation is exceeded, the excess of water is lost as runoff.

$$\Theta = (1 - i_R)R_P - E_T \tag{1}$$

Rainfall interception increases with canopy LAI and is taken as a fraction of rainfall. E_T is calculated using the Penman Monteith equation controlled by the canopy conductance, solar radiation and D. Canopy conductance (g_C, m s^{-1}) increases with LAI but is bounded by the LAI value at which conductance is at a maximum (g_{Cx}, m s^{-1}). The relationship between g_C and LAI is further controlled by age and either vapour pressure deficit or soil moisture, whichever is most limiting stomatal aperture. Further details on 3-PG can be found in Landsberg and Waring (1997) and Sands and Landsberg (2002).

3. Materials and methods

3.1 *Study site and available datasets*

The calibration of the 3-PG model was conducted for Corsican pine stands of yield class 14 using existing data from a 20 000 ha forest plantation, East Anglia, UK (Thetford forest, 52°30′ N, 0°30′ E). Stands in the UK are typically managed under an intermediate spacing, intermediate thinning and 80-year rotation regime (Edwards and Christie 1981). This management regime was also assumed for Thetford stands.

3.2 *Field based datasets*

The following datasets were used in the calibration: (i) the UK Forestry Commission GIS database, a spatially exhaustive catalogue comprising of stand level information on species, yield class, planting year, planting density and stemwood volume, (ii) the Maestro-1 1989 campaign and the 2000 Synthetic Aperture Radar and Hyperspectral Airbone Campain (SHAC) campaign datasets (Baker 1992, Baker *et al.* 1994, Skinner and Luckman 2000), which consist of ground data collected on stand level information (each sampled stand was allocated a Forestry Commission code maintaining consistency with the GIS database) and (iii) datasets collected in Thetford over the years (e.g. Ovington 1957, Corbett 1973, Roberts 1976, Beadle *et al.* 1982, 1985a,b,c, Stewart 1988, Mencuccini and Grace 1996). A specific description on the use of these datasets in the calibration (with their uncertainties) is provided in §4.2

The model was initialized for a stand aged 15 years using chronosequenced biomass data obtained from the Maestro dataset (Baker 1992, Baker *et al.* 1994). Initializing the model at this age removes the need for extra parameterisation required by early growth processes while still enabling the calibration of key parameters. Root, stem and foliage biomass were 7.1 t ha^{-1}, 22 t ha^{-1} and 9.8 t ha^{-1}, respectively. Initial stocking of 3955 trees per hectare was obtained from the production tables (Edwards and Christie 1981).

The required climatic data were derived from the Climate Research Unit datasets and the Cambridge botanical garden meteorological station (New *et al.* 2000, http://

Table 1. Climate database for Thetford, UK.

Monthly climate data	Mean Tmax[a] (°C)	Mean Tmin [a] (°C)	Rain [b] (mm)	Solar rad [b] ($MJ\,m^{-2}d^{-1}$)	Frost days [a]
January	6.47	1.08	55.0	2.52	10.7
February	7.29	0.97	42.4	4.53	11.2
March	10.18	2.13	51.9	8.26	7.9
April	13.13	3.95	48.0	13.10	3.3
May	16.86	6.70	55.0	16.58	0.8
June	20.08	9.73	55.0	18.43	0.0
July	22.31	11.83	54.0	16.64	0.0
August	22.15	11.58	58.0	14.42	0.0
September	19.17	9.64	61.1	10.00	0.0
October	14.99	6.66	61.1	5.80	1.5
November	10.07	3.51	69.0	2.86	5.7
December	12.61	2.04	61.1	1.96	9.1

[a]70 years average derived from the British Atmospheric Data Centre.
[b]Rain and solar radiation derived from University of East Anglia climate research unit (New *et al.* 2000).

badc.nerc.ac.uk/home/index.html). The area is characterized by a relatively flat topography and insignificant climatic variations within the site were assumed (Ovington 1957). Long term average climatic conditions are summarized in table 1.

Other input variables include site latitude, an estimate of soil fertility and texture, as well as available soil water (mm per depth of rooting zone, in metres). The soils of the plantation are of poor quality, predominantly sandy with deep alkaline chalky bedrock and drain freely throughout the forest (Corbett 1973, Mencuccini and Grace 1996). Minimum available water was estimated based on field measurements taken during the drought year of 1976. During the drought, measurements have shown that at least 170 mm soil water was available (Roberts *et al.* 1982). The maximum available water was assumed as 250 mm based on: (a) the assumption that storage capacity for sandy soils is approximately 150 mm per metre of soil with a permanent wilting point of 50 mm and (b) field measurements taken in Thetford, showing that 95% of roots are located in the first metre of soil (Roberts 1976). Given the documented deep bedrock (Corbett 1973), we assumed a 2 m soil layer.

All runs were made with 3-PGpjs, a Visual Basic implementation of 3-PG in Excel available at http://www.ffp.csiro.au/fap/3pg/download_details.htm.

3.3 *Remote sensing datasets and processing*

SAR, Hyperspectral, and LiDAR datasets acquired in 2000 were included in the calibration. These datasets were used instead of alternative empirical, approximate yield based tables. They have the advantage of being more site-specific and enable us to quantify the variability in the biophysical variables of interest.

3.3.1 SAR. A multi-frequency, Synthetic Aperture Radar instrument (E-SAR) was flown on 31 May 2000 in wide swath mode, with data collected at L-HH, L-HV, L-VV, X-VV, plus repeat-pass L-band fully polarimetric data. The mean stand backscatter coefficient, σ^0 (dB), and the mean stand interferometric coherence were calculated for the L-HH, L-VV and L-HV polarizations from the geocoded E-SAR data. Although InSAR data were available, only the interferometric coherence and backscatter were used for the work described here. A neural network was trained to

estimate stand top height in Corsican pine stands from the E-SAR backscatter and coherence data. Top height is defined as the average height of the hundred trees of largest diameter per hectare. The data were divided in half, with half used as a training dataset to train the neural network and the other half used as a testing dataset, to assess the ability of the proposed relationships against unseen data. The inputs to the network were the three mean stand values for coherence (L-HH, L-HV, L-VV), plus the three mean stand values for backscatter (L-HH, L-HV, L-VV). The neural network was a 1-hidden layer network trained with a Levenberg-Marquardt based learning algorithm. Two network structures were investigated, with 2 and 11 nodes in the hidden layer, respectively. To ensure that the best network was selected, 50 trained networks were generated, with the best network selected based on minimum RMSE against the test dataset. The lowest error was produced by a network with two nodes in the hidden layer resulting in a R^2 of 0.90 and a RMSE of 2.51 m when tested against the test dataset (Rowland *et al.* 2003).

3.3.2 Hyperspectral. Hyperspectral data were acquired using the SHAC HyMAP imaging spectrometer in June 2000 (126 contiguous bands, 436–2486 nm at 15 nm spectral resolution, 4 m spatial resolution, see http://www.hyvista.com). The HyMap data were supplied by the National Remote Sensing Centre (NRSC), UK (now Infoterra Ltd, UK) as a 126-waveband raster image with DN values converted to radiance. The imagery were atmospherically corrected utilizing the HyCorr algorithm. The overlapping scenes were also georectified, mosaicked and normalized to minimize the effect of sensor look angle. Signal-to-noise ratio analysis was conducted to remove noisy atmospheric water absorption bands from the original dataset.

3.3.3 LiDAR. E-SAR and hyperspectral datasets were complemented in June 2000 with first and last return data acquired by means of a small footprint Airborne Laser Terrain Mapper (Optech ALTM 1210). The ALTM emits laser pulses at a wavelength of 1047 nm (NIR) where vegetation is highly reflective. The data were collected at footprint size of 0.05 m^2. A $\pm 10^\circ$ scanning orientation perpendicular to the flight path was selected, which generated irregular ground measurements ranging between 2.80 m^2 to 6.50 m^2. The precision of the instrument was estimated at 0.60 m in the x and y positions and 0.15 m in the z position (www.optech.on.ca).

A digital canopy height model (DCHM) was obtained by subtraction of a digital terrain model (DTM) from a digital surface model (DSM). The DSM and the DTM were derived from the first and last significant LiDAR returns, respectively (methodology described in Gaveau and Hill 2003, Rowland *et al.* 2003 and Patenaude *et al.* 2004). Both the first and last returns were converted from a point to a gridded format. The DTM was then produced by applying a minimum value filter to identify local height minima in the gridded LiDAR last return product. Top height per stand was extracted from the DCHM based on the maximum canopy height per stand (R^2=0.94, RMSE 1.68 m, bias 0.48 m). The use of percentiles was also tested (90th, 95th, 97.5th and 99th). However, whilst they may be appropriate for mean stand height, they were found to underestimate canopy top height for the Thetford stands (Rowland *et al.* 2003).

4. Bayesian calibration

In Bayesian statistics, probability is interpreted as the degree of certainty for some quantity, conditional to available data and knowledge. As model parameter values

are not precisely known, this uncertainty can be represented as a probability distribution over the parameters. Thus, if we define θ as a parameter vector for 3-PG, then $P(\theta)$ represents its probability distribution and $P(f(\theta))$ the uncertainty in model outputs ($f(\theta)$) generated by the uncertainty in the parameters. In this context, Bayesian calibration is a method enabling $P(\theta)$ to be updated as new data come in.

Given a dataset D, we can derive $P(\theta|D)$ from $P(\theta)$ by applying Bayes Theorem:

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

In Bayesian terminology, $P(\theta|D)$ is the updated or posterior parameter distribution; $P(\theta)$ is the original distribution, referred to as the prior; $P(D|\theta)$ is the conditional probability of the data for a given parameterisation, called the likelihood; and $P(D)$ is a normalization constant that may be referred to as the evidence.

4.1 *The prior*

The prior distribution is built from marginal distributions, which reflect our current knowledge of parameters. The distribution that best describes the available information about parameters must be used. Following the example by Van Oijen *et al.* (2005), we used uniform distributions, bounded by biophysically or biologically reasonable maximum and minimum values for each parameter given the limited information available on parameter values for our study site. Table 2 presents the parameter values selected for calibration.

Boundaries to the prior were obtained from direct observation on Corsican pine stands in Thetford (CP-T), from literature on Corsican or other pine species (P-L), from surrogate species or 3-PG set default values (D) or finally as best guess estimates or fitting approaches (F) (table 2). Key parameters difficult to measure in the field and for which little information was available were included in the calibration. Constant values were prescribed for the remaining parameters (table 2).

4.2 *The likelihood*

A total of 28 data points was used in the calibration exercise: LiDAR derived heights (4); E-SAR and field based estimates of total above-ground biomass (4 and

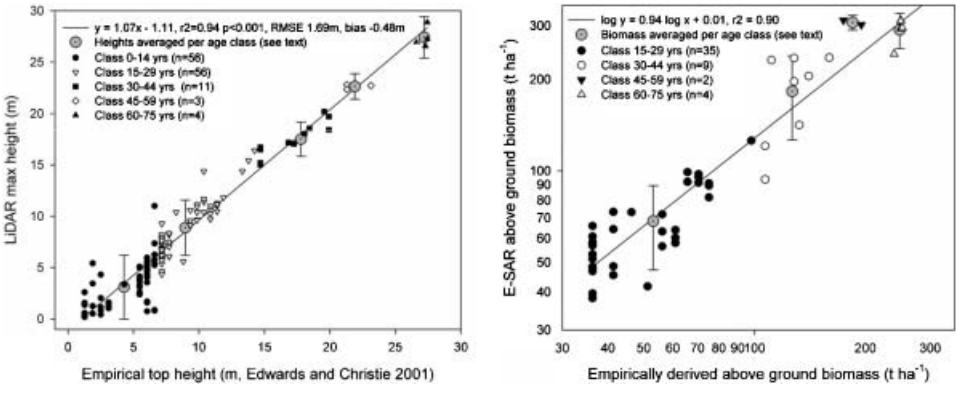

Figure 1. LiDAR heights and E-SAR biomass in relation to empirical data. Note: (i) averaged height and biomass data points were aggregated and averaged per age class (ii) logarithmic scale on E-SAR biomass graph (due to heterocedasticity of variances).

Table 2. Parameters of the 3-PG model, fixed values and ranges of values for the prior. ([a]Sensitivity class as defined by Sands (2004). L=Low, M=Medium, H=High. [b]CP-T, data on Corsican pine derived from studies conducted in Thetford; CP-L, data on Corsican pine derived from the literature; P-L, Data for pine species; D, default values from surrogate species or value specific to 3-PG.)

3-PG symbol: Description (units)	S class[a]	Prescribed parameter values	Calibrated parameters: ranges of values in the prior		Data class[b]	Source/comment
			θ_{min}	θ_{max}		
a_S: Constant in stem mass versus diameter relationship	M	0.02	–	–	CP-T	Baker 1992, Baker *et al.* 1994, Edwards and Christie 1981
c_θ: Moisture ratio deficit giving f_θ=0.5	H	0.7	–	–	D	Default for sandy soils
f_{N0}: Value of f_N when FR=0	M	0.6	–	–	D	
FR: Fertility rating	?	–	0.2	0.6	CP-T	Soil fertility (Corbett 1973, Roberts *et al.* 1982)
g_B: Canopy boundary layer conductance ($m\,s^{-1}$)	L	0.2	–	–	D	
g_{Cx}: Maximum canopy conductance ($m\,s^{-1}$)	H	–	0.015	0.03	P-L	Kelliher *et al.* 1995
i_{Rx}: Maximum fraction of rainfall intercepted by canopy	M	0.15	–	–	D	
k: Extinction coefficient for PAR absorption by canopy	M	–	0.4	0.7	P-L	Stenberg *et al.* 1994, Mencuccini and Grace 1996
k_F: Number of days production lost for each frost day	L	1	–	–	D	
L_{Cx}: Canopy LAI for maximum canopy conductance ($m^2\,m^{-2}$)	L	3.33	–	–	P-L	Kelliher *et al.* 1995, Mencuccini and Grace, 1996
L_{ix}: LAI for maximum rainfall interception ($m^2\,m^{-2}$)	L	0	–	–	D	
m_0: Value of m when FR = 0	?	0	–	–	D	
m_F: Fraction of mean foliage biomass per dying tree	L	0	–	–	D	
m_R: Fraction of mean root biomass per dying tree	L	0.2	–	–	P-L	Empirical data (Edwards and Christie 1981, Ovington 1957)

Table 2. Continued

3-PG symbol: Description (units)	S class[a]	Prescribed parameter values	Calibrated parameters: ranges of values in the prior		Data class[b]	Source/comment
			θ_{min}	θ_{max}		
m_S: Fractions of mean stem biomass per dying tree	L	0.2	–	–	CP-L	Empirical data (Edwards and Christie 1981)
n_{age}: Power of relative age in f_{age}	L	4	–	–	D	
n_N : Power in self thinning law	L	1.5	–	–	P-L	Theoretical scaling laws and observation
n_{fN}: Power of (1-FR) in f_N	L	1	–	–	D	
n_S: Power in stem mass versus diameter relationship	H	2.88	–	–	CP-T	Baker 1992, Baker *et al.* (1994), Edwards and Christie 1981
n_θ: Power of moisture ratio deficit in f_θ	L	9	–	–	D	Default for sandy soils
p_2: Ratio of foliage:stem partitioning at $B = 2$ (cm)	H	–	0.5	1	P-L	Gower *et al.* 1994
p_{20}: Ratio of foliage:stem partitioning at $B = 20$ (cm)	H	–	0.1	0.5	P-L	Gower *et al.* 1994
p_{BB0}: Branch and bark fraction at stand age 0	L	0.5	–	–	P-L	Default 3-PG values for *P. radiata*
p_{BB1}: Branch and bark fraction for mature aged stands	L	0.1	–	–	P-L	Default 3-PG values for *P. radiata*
r_{age}: Relative age to give $f_{age} = 0.5$	L	0.95	–	–	D	
t_{BB} : Age at which $p_{BB} = \frac{1}{2}(p_{BB0}+p_{BB1})$	L	5	–	–	P-L	Default 3-PG values for *P. radiata*
t_c: Age at full canopy cover (years)	M	0	–	–	P-L	
T_{max}: Maximum temperature for growth (°C)	L	35	–	–	P-L	
T_{min}: Minimum temperature for growth (°C)	L	0	–	–	D	
T_{opt}: Optimum temperature for growth (°C)	M	–	18	22	P-L	Waring and Running 1998
t_x: Maximum stand age used to compute relative age (years)	L	–	60	100	D	10% of age at maximum height (Waring, personal communication)

Table 2. Continued

3-PG symbol: Description (units)	S class[a]	Prescribed parameter values	Calibrated parameters: ranges of values in the prior		Data class[b]	Source/comment
			θ_{min}	θ_{max}		
$t_{\gamma F}$: Age at which litterfall rate has median value (months)	L	36	–	–	D	
t_{σ} : Age at which specific leaf area = ½($\sigma_0 + \sigma_1$) (years)	L	2.5	–	–	D	
W_{Sx1000}: Maximum stem mass per tree at 1000 trees ha^{-1}		–	160	400	CP-L	Live stem numbers time-series (Edwards and Christie 1981)
Y: Ratio NPP/GPP	H	0.47	–	–	P-L	Waring and Running 1998
α_{Cx}: Maximum canopy quantum efficiency (mol mol^{-1})	H	–	0.045	0.065	P-L	Range for temperate species in 3-PG (e.g. Stenberg *et al*. 1994, Law *et al.* 2000, Waring 2000, Waring *et al.* 2002, Wang *et al*. 2004)
γ_{F0}: Litterfall rate at t=0 (month^{-1})	L	0.001	–	–	D	
γ_{F1}: Litterfall rate for mature stands (month^{-1})	H	–	0.025	0.035	P-T	Beadle *et al*. 1982, Cousens 1988
γ_R: Average monthly root turnover rate (month^{-1})	L	–	0.006	0.015	P-L	Gill and Jackson 2000
η_{Rn}: Minimum fraction of NPP to roots	M	–	0.20	0.50	P-T	Ovington 1957, Levy *et al*. 2004
η_{Rx}: Maximum fraction of NPP to roots	M	–	0.50	0.80	P-T	Ovington 1957, Levy *et al*. 2004
σ_0 : Specific leaf area at stand age 0 (m^2 kg^{-1})	L	5			P-L	Van Hees and Bartelink 1993
σ_1: Specific leaf area for mature aged stands (m^2 kg^{-1})	H	–	4	8	P-L	Van Hees and Bartelink 1993
ρ_1: Basic density	H	0.43	–	–	CP-L	Hamilton 1975

1, respectively); field based estimates of stem, foliage and root biomass (3, 3 and 5, respectively); and Hyperspectral and ground based LAI estimates (7 and 1, respectively).

4.2.1 LiDAR heights. LiDAR heights were taken as surrogates of top heights (§3.3.3). These were aggregated and averaged per 15 years age classes (figure 1). Uncertainty was estimated as standard deviations to height averages per class. An additional ± 0.5 m error was added to small samples ($n<9$).

4.2.2 E-SAR and field based above-ground biomass. Above-ground biomass data were derived from E-SAR top height estimates. Conversion of top height to above ground biomass involved two stages of calculations (Rowland *et al.* 2003, summarized here): (i) conversion of top height to stemwood volume using an empirical relationship derived from Edwards and Christie (1981) (ii) conversion of stemwood volume to biomass using a biomass expansion factor and a generic basic density coefficient (1.5 for temperate pine species, Milne 1992, IPCC 2003; 0.43 t m^{-3}, Hamilton 1975, respectively). E-SAR biomass estimates were plotted against yield table estimates (log transformed, figure 1). Untransformed standard deviations of biomass (aggregated and averaged per 15 years age classes) were used as error estimates. An additional ± 10 t ha^{-1} uncertainty was added to small samples ($n<9$).

4.2.3 Stem, foliage and root biomass. Stem and foliage biomass data points were derived from Baker (1992) and Baker *et al.* (1994). For root biomass, a root to shoot ratio was derived from destructive measurements made in six mature Scot pine stands (Ovington 1957). The ratio below to above-ground across ages (0.3, Std 0.05) was assumed representative to that of Corsican pine. This value is also consistent with that given by the IPCC (2003) for temperate coniferous forests. Five root biomass points were derived. A $\pm 10\%$ relative error was assumed.

4.2.4 LAI. Given the absence of ground based or alternative sources, LAI data points were derived from hyperspectral data. LAI in pine plantations generally exhibit a growth pattern expressed as (e.g. Mencuccini and Grace 1996):

$$LAI = ae^{-0.5\left(\frac{\ln(x/x_0)}{b}\right)^2} \tag{3}$$

Where a represents the maximum LAI reached by a stand, x_0 is the age at which this maximum is reached, and b is a parameter controlling the tailing off of the LAI curve. Equation (3) was solved in a three way procedure:

(i) Corsican pine stands in the GIS database were co-registered to the image allowing the chronosequencing of leaf area index (LAI) throughout the rotation.

(ii) Based on the results by Lee *et al.* (2004) and Pu and Gong (2004) where close proportionality was found between LAI and the primary axis of a principal component analysis (PCA) for the different wavelengths, PCA was tested to estimate LAI growth patterns in Thetford CP stands. In our study, PC1 contained 91% of the variance of all wavebands; the strongest loadings were on NIR bands; and eigenmatrix loadings were of opposite sign between the IR and visible bands. This confirmed that in our case study, PC1 truly represents vegetation. Averaged values per stand were plotted against stand age using the GIS attribute database. The x_0 and b parameters, which pertain

to the shape of the curve only, not the magnitude of LAI, were solved by minimizing the distance between chronological PCA points and equation (3) (figure 2).

(iii) Conversion of PCA values to LAI was completed using the available projected LAI datum (Ovington 1957). Large relative uncertainties (30%) were assumed.

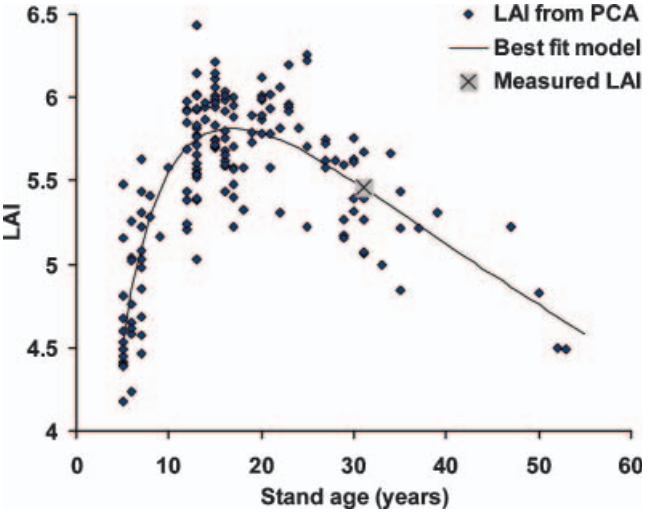

Figure 2. Leaf Area Index (LAI) growth dynamic derived from HyMap data

4.2.4 Estimating the likelihood. To calculate the likelihood, i.e. the probability of the data given a model parameterisation $P(D|\theta)$, information about measurement error must be available. Assuming that the errors associated with our data are independent and Gaussian, $P(D|\theta)$ then follows from the comparison of each data point D_i with the corresponding model output $f_i(\theta)$ as (Van Oijen *et al.* 2005):

$$P(D|\theta) = \prod_{i}^{n} \varphi(D_i - \mathrm{f_i}(\theta); 0, SD_i) \tag{4}$$

where, φ symbolizes a Gaussian function with 0 and SD_i as mean and standard deviation of errors, and n=28, the number of points in the data sample.

4.3 *The posterior: a Monte Carlo estimation of the posterior distribution*

The application of Bayes Theorem to process-based models has traditionally been hampered by two problems: (i) the models cannot be solved analytically, so a sampling method to explore the parameter space is required (we define parameter space as the space of all possible parameter vectors) (ii) the models need to be run at every sampled point in the parameter space (to calculate the likelihood), a highly time consuming and computer intensive process. In recent years, Markov Chain Monte Carlo (MCMC) methods have been found useful to resolve this type of problem (Van Oijen *et al.* 2005). Here, we used the MCMC Metropolis Hastings Random Walk. We used a chain length of 25 000, so the MCMC produced a representative sample of 25 000 parameter vectors. The proof that MCMC always

converges properly to the correct posterior probability density function is provided in Robert and Casella (1999). Because the posterior distribution cannot be described analytically, the results are presented in the form of marginal distributions using descriptive statistics. As suggested by Van Oijen *et al.* (2005), in addition to means ($\overline{\theta(i)}$) and standard deviations, we present the maximum *a posteriori* (MAP) estimate of θ, considered as the single 'best' parameter value estimated from the MCMC sample.

$$\theta_{MAP} = \arg \max_{\theta} p(\theta|D) \tag{7}$$

Although this should not be interpreted as an optimized parameter vector, this nevertheless provides information as to what vector has the highest probability density given the available data.

4.4 *Sensitivity analysis*

The sensitivity of a given model output with respect to a parameter (and vice versa) has also been estimated from partial correlations calculated between the 25 000 parameter (14) and output (28) vectors. This resulted in a 14×28 partial correlation matrix.

5. Results and discussion

A 25 000 vectors sample was generated from the posterior distribution using the MCMC sampling approach. Figure 3 shows an example of MCMC trace plot and the resulting marginal posterior distribution for the fertility rating (*FR*) parameter.

Summary statistics for the marginal distributions of parameters are presented in table 3, which include the mean ($\overline{\theta(i)}$) and standard deviation and the vector of highest *a posteriori* probability density (θ_{MAP}). Figure 4 shows the mean model outputs from the 25 000 estimates, the 3-PG outputs from θ_{MAP} (best fit) and the datasets used in the likelihood.

The results from the sensitivity analysis (partial correlations between parameters and outputs) are presented in figure 5. These are illustrated in the form of a colour fingerprint between the 14 parameters included in the calibration (table 2, 4th column) and the 28 model outputs for which we have calibration data (figure 4). High negative correlations are shown in dark blue and high positive correlations in dark red. Light regions indicate weak or no correlation (white). While parameters

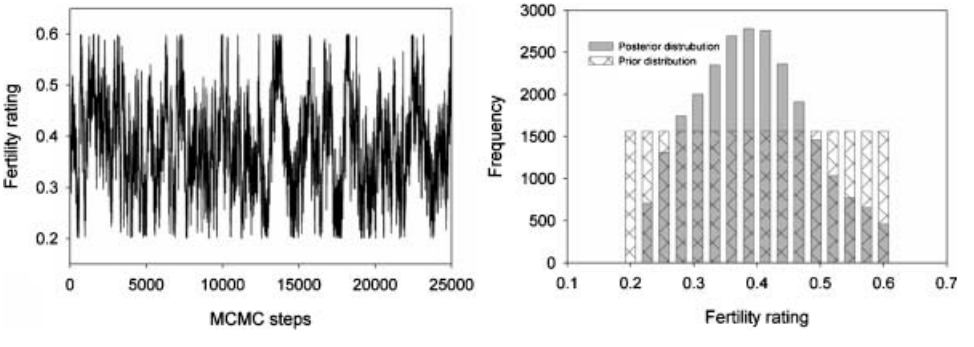

Figure 3. (*a*) Markov Chain Monte Carlo (MCMC) trace plot. (*b*) Histogram: sample resulting from MCMC.

Table 3. Statistics to the parameters posterior distribution. The statistics are the means $\overline{\theta(i)}$ and standard deviations to the marginal distributions of parameters and the vector of highest *a posteriori* probability (θ_{MAP}).

$\theta(i)$	$\overline{\theta(i)}$	SD	θ_{MAPx}
FR	0.380	0.090	0.391
g_{Cx}	0.023	0.004	0.023
K	0.539	0.087	0.439
p_2	0.694	0.133	0.502
p_{20}	0.441	0.045	0.497
T_{opt}	20.893	0.911	20.42
t_x	90.656	6.978	95.86
w_{Sx1000}	182.826	17.504	165.0
α_{Cx}	0.047	0.002	0.046
γ_{FI}	0.028	0.003	0.026
γ_R	0.013	0.002	0.013
η_{Rn}	0.237	0.028	0.221
η_{Rx}	0.580	0.067	0.557
σ_I	5.711	1.029	4.539

are listed on the vertical axis, the model outputs considered are given on the horizontal axis. Note that the outputs sequentially follow the rotation of a stand. As an example, root biomass (indicated as Rts_M) is consecutively given for ages 20, 30, 50, 60 and 70.

In table 3, one can observe the close similarity between the θ_{MAP} vector and the posterior mean $\overline{\theta(i)}$, which suggests that the posterior is unimodal and symmetric. Corresponding model outputs are shown in figure 4. Outputs from θ_{MAP} (best fit) and $\overline{\theta(i)}$ (posterior mean) also lie closely to data error bounds. Note the smaller error bounds to the posterior in comparison with the data. While above ground, stem and foliage biomass model dynamics closely match those observed on the ground, allocation to roots appears to level towards 25 years into the rotation, and decrease thereafter. Additional pilot MCMC analyses (results not shown here) were conducted to explore the influence of the model structure on the model output dynamics. The data distributions used in the calibration were set close to uniform, to allow an exploration across the range of permittable values. Likewise, large but realistic ranges of parameter values (with uniform distributions) were given. By doing so, the underlying model structure became apparent, not masked by the influence of priors. Preliminary results suggested that the root dynamics observed here are an artefact of the model structure, rather than a result of parameterisation. Further investigations are required however to investigate the foundation to this model behaviour.

The results from the sensitivity analysis presented in figure 5 also provide strong insights into the multivariate interactions imbedded in the model. Correlations between parameters and outputs may be consistent (nearly constantly) decreasing or increasing throughout the rotation. The optimum temperature for growth, T_{opt} , is consistently negatively correlated with all the model outputs considered here (dark blue grid cells throughout). This implies that as T_{opt} increases, productivity decreases. Moreover, table 3 shows that θ_{MAP} and $\overline{\theta(i)}$ values for T_{opt} are $\approx$20°C. Arguably, these results suggest that stands in Thetford grow outside their 'optimum climatic zone'. Indeed, this species is endemic to Mediterranean regions, where hot

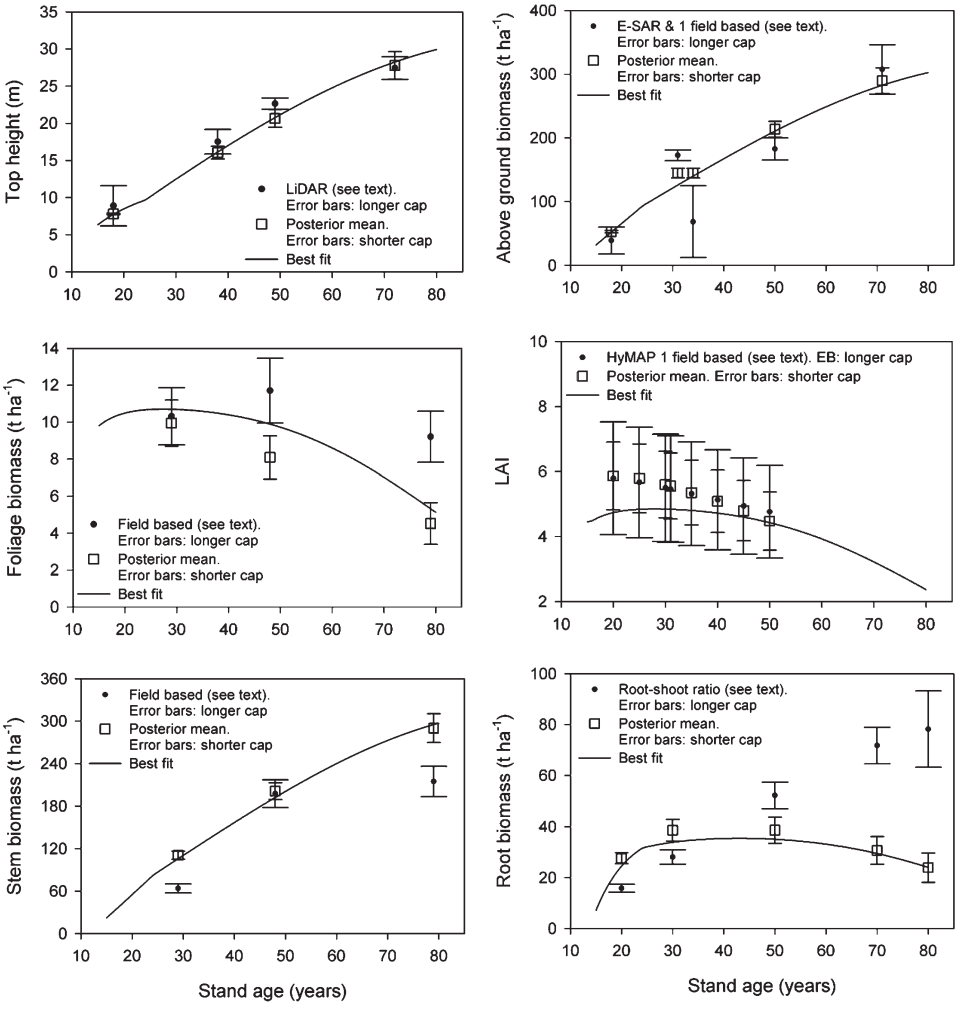

Figure 4. Results from the Bayesian calibration ($\overline{\theta(i)}$and best fit θ_{MAP}). Error bars to the posterior mean are standard deviations. The data used in the likelihood are also shown (uncertainties stated in text). Legends indicate origin of data.

summer days (>25°C) are four times more frequent than in Lowland Britain (Brown 1960, Kerr 2000). Consistently positive correlations are observed between the maximum canopy quantum efficiency (α_{Cx}) and all biomass outputs; an expected result, as net primary production (and biomass production) are proportional to the product of the maximum canopy quantum efficiency (α_{Cx}), the P_n/P_g ratio (Y), light interception and environmental constraints. Other consistent correlations are also noted between specific leaf area (σ_1), litterfall rate (γ_{F1}), the ratio of foliage to stem partitioning at maturity (p_{20}) and LAI; between the fertility rating (FR) and above ground components (biomass and LAI); and between root turnover (γ_R) and root biomass.

Conversely, decreasing or increasing correlations can either indicate specific periods where the influence of a parameter is most significant, or equally, how data collected at specific moments during the rotation may be useful in calibrating and

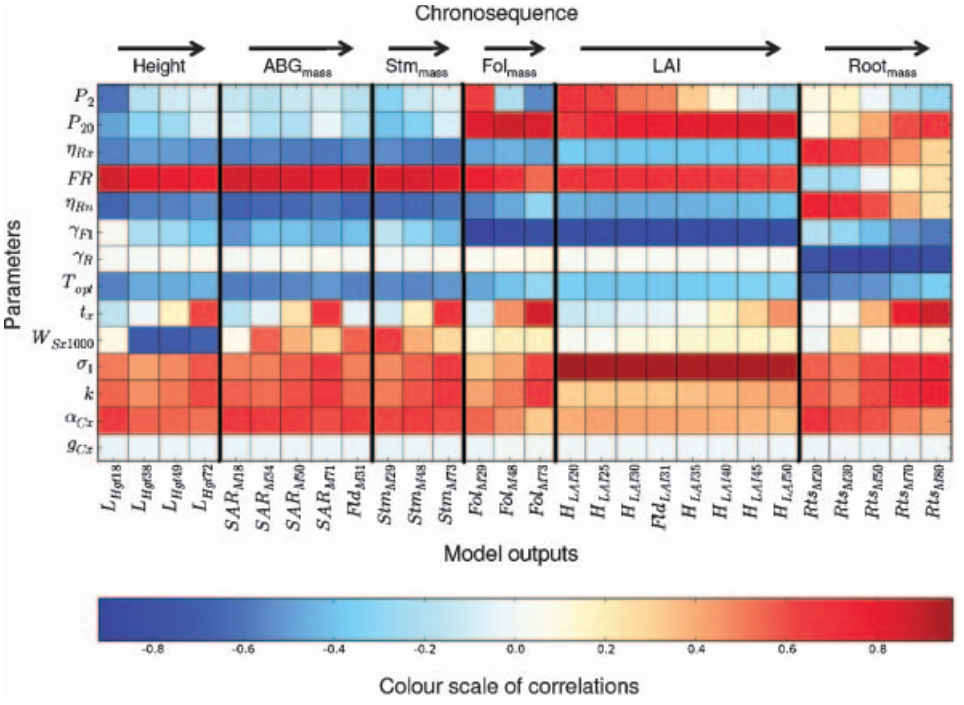

Figure 5. Partial correlation fingerprint between 3-PG model parameters (14) and model outputs (28) used in the calibration. Dark blue: high negative correlations. Dark red: high positive correlations. White: no correlation. Parameter symbols are provided in table 2.

reducing the uncertainty for a given parameter. Such results allow the rapid identification of parameters most critical for determining uncertainty. In this study, these parameters—selected as those of highest average correlation (above 0.5) in absolute terms over all model outputs (see figure 5, horizontal correlations)—were the maximum canopy quantum efficiency (α_{Cx}), the specific leaf area for mature stands (σ_1), litterfall rate (γ_{F1}), the fertility rating (FR), and the extinction coefficient for PAR absorption by canopy (k). By identifying these parameters, one can specifically target remote sensing data collection for uncertainty reduction. Accurate and precise height and stem biomass data, for instance (notably early in the rotation), would reduce uncertainties in the maximum canopy quantum efficiency (α_{Cx}). Similarly, these datasets, when collected notably for mature stands, would help in reducing uncertainties to the extinction coefficient for PAR absorption by canopy (k). In figure 5, one can also observe that foliage biomass and above all LAI estimates are critical for reducing the uncertainty to the specific leaf area for mature aged stands (σ_1), litterfall rate (γ_{F1}) and the fertility rating (FR). These critical results illustrate the potential of BC, not only to provide uncertainty estimation but importantly, as a means to plan future data acquisition by testing the influence of different data collection scenarios (with related uncertainties) on resulting model outputs.

To evaluate where most uncertainty reduction was achieved, post calibration standard deviations were compared to pre calibration standard deviations for a given output type (LAI, root biomass etc., figure 5). Uncertainties were reduced by

approximately 40% for height, 65% for total above-ground biomass, 21% for stem biomass, 24% for foliage biomass, 38% for LAI and 28% for root biomass (figure 4).

Together, these results serve three purposes. First, through this modelling exercise and with the use of the LiDAR, E-SAR and HyMap remote sensing datasets, we present the parameterisation of 3-PG for Corsican pine stands in Thetford. While occupying more than 30 000 ha of the UK territory (Forestry Commission 2001), relatively limited information is available on Corsican pines, as compared to more economically viable species such as Sitka spruce and Scots pine. This provides a first attempt to quantify parameters for this species, which may have been difficult to obtain otherwise. Secondly, we illustrated the ability of BC as a framework to integrate remote sensing datasets into ecological modelling, thereby helping to further our understanding of forest ecosystem functioning. This is a critical result, as remote sensing often provides the only source of data available at the spatial and temporal scales required. Thirdly, with BC, data of poor precision yet of relevant information content may still be used in the calibration. One should note however that while the posterior will at first be strongly influenced by the data, this influence decreases as (i) new data come in and (ii) uncertainty in the prior decreases. Fourthly, we have shown that given ever increasing computing power and speed, uncertainty quantification and model parameterisation can be achieved with relative ease using BC. Despite the fact that the probability density of a scalar model output or parameter is nearly nil, process based models used in forestry are too often parameterised by adjusting selected parameters to discrete values for the model output to fit the data time series, without any indication of parameter and output uncertainties. The parameterisation of 3-PG for novel species has unfortunately been no exception (e.g. White *et al.* 2000, Sands and Landsberg 2002, Sands 2004, Stape *et al.* 2004 and Almeida *et al.* 2004).

6. Conclusion

In summary:

- Provided that uncertainty estimates to remote sensing datasets are available, Bayesian calibration is an efficient means of integrating remote sensing data into forest ecosystem modelling.
- From our sensitivity analysis, critical parameters determining uncertainty were found to be the maximum canopy quantum efficiency (α_{Cx}), the specific leaf area for mature aged stands (σ_1), litterfall rate (γ_{F1}), the fertility rating (FR) and the extinction coefficient for PAR absorption by canopy (k). Given the potential of remote sensing to retrieve the latter, such investigations should be encouraged.
- Our sensitivity analysis showed that accurate and precise height and stem biomass data would help reduce uncertainties to the maximum canopy quantum efficiency (α_{Cx}) and the extinction coefficient for PAR absorption by canopy (k), while LAI estimates are critical for reducing the uncertainty to the specific leaf area for mature aged stands (σ_1), litterfall rate (γ_{F1}) and the fertility rating (FR). This contributes to (i) prioritizing data collection and (ii) identifying best suited remote sensing (LiDAR, E-SAR, Hyperspectral or others) and ground data collection to reduce uncertainty in parameters and model outputs.

- Application of Bayesian calibration also showed that uncertainties were reduced strongly on all model outputs, but particularly for height and total above ground biomass.
- Preliminary analysis of model outputs across the range of permittable parameter and data values suggests that the 3PG model may be structurally unable to capture the observed root dynamics for Corsican pine in Thetford. Further MCMC analyses are required to investigate the influence of the model structure on root dynamics.

While both optimization and Bayesian approaches address the need to test whether a model can predict available data or not, in optimization, parameter values are adjusted such that the model yields outputs closest to the data. This reduces the usefulness of uncertain datasets, ancillary or remotely sensed, which provide information on variables not currently or commonly compiled. Conversely, Bayesian calibration advocates the quantification of uncertainties to parameters, thereby yielding uncertainties in model outputs, over the derivation of an optimized set of parameters based on a goodness-of-fit approach. By doing so, Bayesian calibration provides a means to conduct truly integrative work, across disciplines involved in research on forest ecosystem functioning. It systematically provides the framework for quantifying model output and parameter uncertainty, while considering all the existing information, including that enclosed in the model itself.

Acknowledgments

The authors would like to thank Maurizio Mencuccini (University of Edinburgh), David Cameron (CEH Edinburgh), Richard Waring (Oregon State University), and Peter Savill (University of Oxford) for sharing data and advice. Thanks are due to the British Atmospheric Data Centre and the Forestry Commission for providing access to databases, to the Environment Agency for making available the LiDAR data, and to all those involved in the Thetford fieldwork campaigns in 1989 and 2000. The E-SAR data were acquired during the SHAC 2000 campaign by the Natural Environmental Research Council and the British National Space Centre.

References

ALMEIDA, A.C., LANDSBERG, J.J. and SANDS, P.J., 2004, Parameterisation of 3-PG model for fast-growing Eucalyptus grandis plantations. *Forest Ecology and Management*, **193**, pp. 179–195.

BAKER, J.R., 1992, The UK element of the Maestro-1 SAR campaign. *International Journal of Remote Sensing*, **13**, pp. 1593–1608.

BAKER, J.R., MITCHELL, P.L., CORDEY, R.A., GROOM, G.B., SETTLE, J.J. and STILEMAN, M.R., 1994, Relationships between physical characteristics and polarimetric radar backscatter for Corsican pine stands in Thetford Forest, UK. *International Journal of Remote Sensing*, **15**, pp. 2827–2849.

BAULIES, X. and PONS, X., 1995, Approach to forestry inventory and mapping by means of multispectral airborne data. *International Journal of Remote Sensing*, **16**, pp. 61–80.

BEADLE, C.L., TALBOT, H. and JARVIS, P.G., 1982, Canopy structure and leaf-area index in a mature Scots pine forest. *Forestry*, **55**, pp. 105–123.

BEADLE, C.L., JARVIS, P.G., TALBOT, H. and NEILSON, R.E., 1985a, Stomatal conductance and photosynthesis in a mature Scots pine forest .2. Dependence on environmental variables of single shoots. *Journal of Applied Ecology*, **22**, pp. 573–586.

Beadle, C.L., Neilson, R.E., Talbot, H. and Jarvis, P.G., 1985b, Stomatal conductance and photosynthesis in a mature Scots pine forest. 1. Diurnal, seasonal and spatial variation in shoots. *Journal of Applied Ecology*, **22**, pp. 557–571.

Beadle, C.L., Talbot, H., Neilson, R.E. and Jarvis, P.G., 1985c, Stomatal conductance and photosynthesis in a mature Scots pine forest. 3. Variation in canopy conductance and canopy photosynthesis. *Journal of Applied Ecology*, **22**, pp. 587–595.

Brown, J.M.B., 1960, The Corsican pine in its native island, Part 2. *Empire Forestry Review*, **39**, pp. 422–436.

Coops, N.C. and Waring, R.H., 2001, The use of multiscale remote sensing imagery to derive regional estimates of forest growth capacity using 3-PGS. *Remote Sensing of Environment*, **75**, pp. 324–334.

Coops, N.C., Waring, R.H. and Landsberg, J.J., 2001, Estimation of potential forest productivity across the Oregon transect using satellite data and monthly weather records. *International Journal of Remote Sensing*, **22**, pp. 3797–3812.

Corbett, W.M., 1973, Breckland forest soils: special survey no.7, Rothamsted Experimental Station, Harpenden, Herts, UK.

Cousens, J.E., 1988, Report of a 12-year study of litter fall and productivity in a stand of mature Scots pine. *Forestry*, **61**, pp. 255–266.

Edwards, P.N. and Christie, J.M., 1981, Yield models for forest management, Forestry Commission booklet no. 48, HMSO, London, 274 p.

Forestry Commission, 2001, National inventory of woodlands and trees. Report No. http://www.forestry.gov.uk/forestry/HCOU–54PG9U, Forest Research, Forestry Commission, Edinburgh, UK.

Gaveau, D.L.A. and Hill, R.A., 2003, Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing*, **29**, pp. 650–657.

Gertner, G.Z., Fang, S.F. and Skovsgaard, J.P., 1999, A Bayesian approach for estimating the parameters of a forest process model based on long-term growth data. *Ecological Modelling*, **119**, pp. 249–265.

Ghazoul, J. and McAllister, M., 2003, Communicating complexity and uncertainty in decision making contexts: Bayesian approaches to forest research. *International Forestry Review*, **5**, pp. 9–19.

Gill, R.A. and Jackson, R.B., 2000, Global patterns of root turnover for terrestrial ecosystems. *New Phytologist*, **147**, pp. 13–31.

Gower, S.T., Gholz, H.L., Kaneyuki, N. and Baldwin, V.C., 1994, Production and carbon allocation patterns of pine forests. *Ecological Bulletins*, **43**, pp. 115–135.

Green, E.J., MacFarlane, D.W. and Valentine, H.T., 2000, Bayesian synthesis for quantifying uncertainty in predictions from process models. *Tree Physiology*, **20**, pp. 415–419.

Hamilton, G.J., 1975, Forest mensuration handbook. Forestry Commission booklet no. 39, HMSO, London, 274 p.

Hyyppä, J., Hyyppä, H., Inkinen, M., Schardt, M. and Ziegler, M., 2000, Forest inventory based on laser scanning and aerial photography. *Laser Radar Technology and Applications*, V, pp. 106–118.

Hyyppä, J., Kelle, O., Lehikoinen, M. and Inkinen, M.A., 2001, A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, pp. 969–975.

IPPC, 2003, *IPPC Good Practice Guide for Land Use, Land Use Change and Forestry*. Available online at: http://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf-contents.htm (accessed 31 December 2007) (Japan: Institute for Global Environment Strategies).

Jansen, M.J.W., 1999, Data use and Bayesian statistics for model calibration. In A. Stein and F.W.T. Penning de Vries (Eds). *Data and Models in Action*, pp. 69–80 (Dordrecht: Kluwer).

Kelliher, F.M., Leuning, R. and Schulze, E.D., 1993, Evaporation and canopy characteristics of coniferous forests and grasslands. *Oecologia*, **95**, pp. 153–163.

Kelliher, F.M., Leuning, R., Raupach, M.R. and Schulze, E.D., 1995, Maximum conductances for evaporation from global vegetation types. *Agricultural and Forest Meteorology*, **73**, pp. 1–16.

Kerr, G., 2000, Natural regeneration of Corsican pine (Pinus nigra subsp laricio) in Great Britain. *Forestry*, **73**, pp. 479–488.

Landsberg, J.J. and Waring, R.H., 1997, A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. *Forest Ecology and Management*, **95**, pp. 209–228.

Law, B.E., Waring, R.H., Anthoni, P.M. and Aber, J.D., 2000, Measurements of gross and net ecosystem productivity and water vapour exchange of a *Pinus ponderosa* ecosystem, and an evaluation of two generalized models. *Global Change Biology*, **6**, pp. 155–168.

Lee, K.S., Cohen, W.B., Kennedy, R.E., Maiersperger, T.K. and Gower, S.T., 2004, Hyperspectral versus multispectral data for estimating leaf area index in four different biomes. *Remote Sensing of Environment*, **91**, pp. 508–520.

Levy, P.E., Hale, S.E. and Nicoll, B.C., 2004, Biomass expansion factors and root: shoot ratios for coniferous tree species in Great Britain. *Forestry*, **77**, pp. 421–430.

Magnussen, S. and Boudewyn, P., 1998, Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research*, **28**, pp. 1016–1031.

Mencuccini, M. and Grace, J., 1996, Hydraulic conductance, light interception and needle nutrient concentration in Scots pine stands and their relations with net primary productivity. *Tree Physiology*, **16**, pp. 459–468.

Milne, R., 1992, The carbon content of vegetation and its geographical distribution in Great Britain. In *Carbon Sequestration by Vegetation in the UK*. Interim report to the Department of the Environment, Contract No. PECD 7/12/79, Midlothian, 107 pp.

New, M., Hulme, M. and Jones, P., 2000, Representing twentieth-century space-time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *Journal of Climate*, **13**, pp. 2217–2238.

Ovington, J.D., 1957, Dry-matter production by *Pinus sylvestris* L. *Annals of Botany*, **XXI**, pp. 288–314.

Patenaude, G., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.B.J. and Dawson, T.P., 2004, Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sensing of Environment*, **93**, pp. 368–380.

Pu, R.L. and Gong, P., 2004, Wavelet transform applied to EO-1 hyperspectral data for forest LAI and crown closure mapping. *Remote Sensing of Environment*, **91**, pp. 212–224.

Roberts, J., Pitman, R.M. and Wallace, J.S., 1982, A comparison of evaperation from stands of scots pine and corsican pine in Thetford Chase, East Anglia. *Journal of Applied Ecology*, **19**, pp. 859–872.

Roberts, J.M., 1976, A study of root distribution and growth in a *Pinus sylvestris* L. (Scots pine) plantation in East Anglia. *Plant and Soil*, **44**, pp. 607–621.

Robert, C.P. and Casella, G., 1999, *Monte Carlo Statistical Methods* (New York: Springer-Verlag).

Rowland, C.S., Patenaude, G., Balzter, H., Dawson, T.P., Hill, R.A., Luckman, A., Milne, R. and Skinner, L., 2003, Forest carbon estimation: a comparison of techniques. *Proceedings of the RSPSoc Annual Conference*, 10–12 September 2003, Nottingham, CD-ROM.

Running, S.W. and Coughlan, J.C., 1988, A general model of forest ecosystem processes for regional applications I. Hydrological balance canopy gas exchange and primary production processes. *Ecological Modelling*, **42**, pp. 125–154.

SANDS, P.J., 2004, Adaptation of 3-PG to novel species: guidelines for data collection and parameter assignment. Technical Report. No. 141, CSIRO, CRC Sustainable Production Forestry, Hobart.

SANDS, P.J. and LANDSBERG, J.J., 2002, Parameterisation of 3-PG for plantation grown Eucalyptus globulus. *Forest Ecology and Management*, **163**, pp. 273–292.

STAPE, J.L., RYAN, M.G. and BINKLEY, D., 2004, Testing the utility of the 3-PG model for growth of Eucalyptus grandis X urophylla with natural and manipulated supplies of water and nutrients. *Forest Ecology and Management*, **193**, pp. 219–234.

STENBERG, P., KUULUVAINEN, T., KELLOMAKI, S., GRACE, J.C., JOKELA, E.C. and GHOLZ, H.L., 1994, Crown structure, light interception and productivity of pine trees and stands. *Ecological Bulletins*, **43**, pp. 20–34.

STEWART, J.B., 1988, Modeling surface conductance of pine forest. *Agricultural and Forest Meteorology*, **43**, pp. 19–35.

VANCLAY, J.K., 1989, A growth model for North Queensland rainforests. *Forest Ecology and Management*, **27**, pp. 245–271.

VAN HEES, A.F.M. and BARTELINK, H.H., 1993, Needle area relationships of Scots pine in the Netherlands. *Forest Ecology and Management*, **58**, pp. 19–31.

VAN OIJEN, M., ROUGIER, J. and SMITH, R., 2005, Bayesian calibration of process-based models: bridging the gap between models and data. *Tree Physiology*, **25**, pp. 915–927.

WANG, K.Y., KELLOMAKI, S., ZHA, T.S. and PELTOLA, H., 2004, Component carbon fluxes and their contribution to ecosystem carbon exchange in a pine forest: an assessment based on eddy covariance measurements and an integrated model. *Tree Physiology*, **24**, pp. 19–34.

WARING, R.H., 2000, A process model analysis of environmental limitations on the growth of Sitka spruce plantations in Great Britain. *Forestry*, **73**, pp. 65–79.

WARING, R.H., LANDSBERG, J.J. and WILLIAMS, M., 1998, Net primary production of forests: a constant fraction of gross primary production? *Tree Physiology*, **18**, pp. 129–134.

WARING, R.H. and RUNNING, S.W., 1998, *Forest Ecosystems: Analysis at multiple scales*, 2nd edn (London: Academic Press).

WHITE, J.D., COOPS, N.C. and SCOTT, N.A., 2000, Estimates of New Zeland forest and scrub biomass from the 3–PG model. *Ecological Modelling*, **131**, pp. 175–190.