

Parameterization of a process-based tree-growth model: Comparison of optimization, MCMC and Particle Filtering algorithms

C. Gauchere^{a,*}, F. Campillo^b, L. Misson^c, J. Guiot^a, J.-J. Boreux^d

^a CEREGE, UMR 6635 CNRS, Université P. Cézanne, BP 80, F-13545 Aix-en-Provence Cedex 4, France

^b IRISA, INRIA, Campus de Baulieu, F-35042 Rennes Cedex, France

^c ESPM Department, 151 Hilgard Hall, University of California, Berkeley, CA 94720-3110, USA

^d ULg, Dép. Sciences et Gestion de l'Environnement, Av. de Longwy, 185, B-6700 Arlon, Belgium

ARTICLE INFO

Article history:

Received 15 January 2007

Received in revised form 3 March 2008

Accepted 6 March 2008

Available online 22 May 2008

Keywords:

Bayesian calibration

Dendrochronology

Sensitivity analysis

Tree-growth model

ABSTRACT

Finely tuned process-based tree-growth models are of considerable help in understanding the variations of biomass increments measured in the dendrochronological series. Using site and species parameters, as well as daily climate variables, the MAIDEN model computes the water balance at ecosystem level and the daily increment of carbon storage in the stem through photosynthesis processes to reproduce the structure of the tree-ring series. In this paper, we use three techniques to calibrate this model with *Pinus halepensis* data sampled in the Mediterranean part of France: a standard optimization (PEST), Monte Carlo Markov Chains (MCMC) and Particle Filtering (PF). Contrary to PEST, which tries to find an optimum fit (giving the lowest error between observations and simulations), the principle of MCMC and PF is to walk, from *a priori* distributions, in the parameter space according to particular statistical rules to compute each parameter distribution. The PEST and MCMC calibrations of our dendrochronological series lead to rather similar adjustments between simulations and observations. PF and MCMC calibrations give different parameter distributions, showing how complementary are these methods, with a better fit for MCMC. Yet, independent validations over 11 independent meteorological years show a higher efficiency of the recent PF method over the others.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

When conceiving a model, environmental modelers have to identify the major sociological or ecological processes, to write the mathematical equation of the model or to find an algorithm able to provide solution(s), to estimate its overall sensitivity and to calibrate (to parameterize) it, i.e. to estimate the unknown parameters chosen. Assigning values to the parameters is a difficult task. When enough data are available, modelers use optimization techniques which lead to best fit values. Some recent intensive computing techniques inspired from Bayesian statistics greatly improved this task (Gelman et al., 1995), as it can be noted in hydrological disciplines (Pelletier et al., 2006; Dorner et al., 2007). However, the different existing techniques must be compared in a common framework and illustrated in other environmental disciplines such as ecology (Zhou et al., 2004; Bugg et al., 2006).

In dendrochronology, statistical methods – often linear – have been developed for several decades to answer various ecological

questions at large time and spatial scales (Fritts and Swetnam, 1989; Bugg et al., 2006). However, tree-growth processes are not linear and, to better understand the tree-growth, more mechanistic models are necessary. Process-based dendrochronological models are examples of complex multi-parameter models (Fritts and Swetnam, 1989; Fritts et al., 1991). Recently, a model (MAIDEN) was proposed in which parameters can be tuned using both ecophysiological measurements and dendrochronological time series (Misson, 2004; Misson et al., 2004). The authors have determined the parameters needed by the model on the basis of literature and calibration by trial and error. We propose in this paper to test different numerical algorithms and inference approaches, with the objective of comparing their observation/simulation fits (the calibration phase) and their errors on independent observations (validation phase). Bayesian approach to these models is rather novel and we expect new environmental insights, further discussed, by this mean.

There exist numerous optimization routines in the literature, helping to adjust model parameters by minimizing a quantity related to the simulation/observation residuals (Yapo et al., 1998; Dorner et al., 2007). Often based on a gradient of the minimization function, they all give optimal parameter values sometimes with a confidence interval, but without any idea of the parameter distribution. The PEST program for example, one of the most

* Corresponding author. INRA, EFPA, UMR AMAP, TA40/PS2, 34398 Montpellier Cedex 5, France. Tel.: +33 (0) 4 67 61 56 08; fax: +33 (0) 4 67 61 56 68.

E-mail address: gauchere@cirad.fr (C. Gauchere).

commonly used algorithms set for optimization (Doherty, 2003), adjusts model parameters until the fit between model outputs and observations is optimized in the weighted least squares sense. For example, PEST has been recently used with success in order to optimize parameter estimation in hydrological studies (Gallagher and Doherty, 2007; Tischler et al., 2007). PEST is a nonlinear estimator using the Gauss–Marquardt–Levenberg method, needing fewer runs than most of the other estimation methods.

Bayesian techniques such as Monte Carlo Markov Chains (MCMC) offer the opportunity to estimate the (*a posteriori*) parameter distribution, the calibration being done, starting from their predefined (*a priori*) statistical distributions (Gelman et al., 1995; Andrieu et al., 2001). Such approaches could nevertheless lead to unrealistic results, in case of inappropriate *a priori* information or non-relevant parameter distribution extractions. To bypass this difficulty, other kinds of procedures such as Particle Filtering (PF) have been recently imagined on the basis of genetic algorithms (Doucet et al., 2001; Pelletier et al., 2006). Instead of MCMC approaches, the procedure still consists of going through the parameter space, but with a high number of initial choices called the particles and defined by their parameter values (their coordinates). Where MCMC algorithms work with one particle, PF ones make evolving many of them simultaneously, selecting sequentially the best in terms of observation/simulation fits.

The aim of this article is to compare three calibration approaches, namely PEST, the MCMC and PF algorithms, in the specific case of the MAIDEN model (Misson, 2004), where 11 new parameters need to be assigned. Detailed sensitivity and elasticity analyses are performed to compare calibrated parameter values and distributions and to analyze the performances of each technique and their ecological implications.

2. Data and model

2.1. Data

We use dendrochronological data from 21 *Pinus halepensis* (Aleppo pine) stands located on calcareous soil in Provence, southeast France (Nicault, 1999). Based on dendrochronological cores, time series of annual earlywood width (E_w), latewood width (L_w), earlywood density (E_d) and latewood density (L_d) were used to calculate an index ($W(t)$) of annual wood increment that takes into account not only stem radial growth but also wood density (see Rathgeber et al., 2000 for a complete description). Based on time series of $W(t)$ built for each stands, arithmetic means were computed for each year (over 34 years) to generate a series covering the entire region. This regional dendrochronological series is the main observational data that we will use to calibrate MAIDEN. Botanical, ecological and topographical details were also recorded at each stands (Nicault, 1999). Physiological data come from a pine–oak experimental mixed forest. Tree transpiration has been measured using the sap flow technique and is used here at daily time step during year 2004 on the Aleppo pine species (Granier et al., 1996).

The Aix-en-Provence meteorological station (MétéoFrance) provides maximum and minimum air temperatures and precipitation data at daily time steps for the 1961–1994 period. The distance between this station and the sites varies between 5 and 65 km. For the independent validation stage, we used the daily temperatures and precipitations that have been recorded at the neighboring station Istres during the 1948–1960 period. Years 1956 and 1985 are characterized by extremely low winter temperature for the region (freeze), which induced partial damaging of the cambium, which had an effect on the growth during 3 years. We have then removed the years 1956–1958 and 1985–1987 from all of the observation/simulation comparisons made in this study because MAIDEN does not integrate such processes (see Misson et al., 2004).

2.2. Model

The MAIDEN model is extensively described in Misson (2004) and analyses of Aleppo pine dendrochronological series can be found in Misson et al. (2004). The model needs site and species parameters such as altitude, latitude, maximum absolute LAI, Specific Leaf Area (SLA), initial bole biomass, soil thickness and soil textural classes. In addition, the model has 11 unknown parameters that can be tuned to adjust the outputs of the model to ecophysiological and dendrochronological data. The main output of the model gives the annual increment of bole carbon reservoir (hereafter named M_t) that will be used further to calibrate the model. Climatic driving variables are daily minimum and maximum temperature and precipitation. MAIDEN uses the MT-CLIM algorithm (Running et al., 1987) to estimate daily solar radiation and vapor pressure deficit from daily temperature and precipitation. The model calculates processes such as photosynthesis, stomatal conductance and carbon allocation. The water balance is computed at the ecosystem level, including canopy water interception, transpiration, soil evaporation, soil water transfer, drainage and runoff. MAIDEN partitions daily net primary production (NPP) between carbon reservoirs (leaf, bole, root and storage) according to phenological phase-dependent ratios. These phases are (1) winter: no activity, (2) spring: leaf and root expansion, (3) summer: bolewood production, (4) early fall: carbohydrate-reserve accumulation, (5) late fall: leaf and root senescence. An original modeling procedure of carbon storage and mobilization was developed to mimic the autocorrelation structure of tree-ring series. The annual increment of bole carbon reservoir at stand level is the modeled variable that will be compared with the dendrochronological tree-ring series.

We use in this study a new version of the MAIDEN model from which some of the ecophysiological processes simulated have been improved. In particular, we modified the carbon allocation dynamics, allowing the store reservoir to give some carbon to the bole reservoir during the whole summer and fall months. Concerning the physiological processes, we first updated the temperature dependencies of some photosynthetic parameters to equations proposed by Bernacchi et al. (2003), secondly we removed the p parameter used in the earlier version of the MAIDEN model and calculated the net photosynthesis as the minimum between the rate of carboxylation when ribulose biphosphate is saturated (W_c) and the rate of carboxylation when ribulose biphosphate regeneration is limited by electron transport (W_j), as initially proposed by Farquhar et al. (1980). This new version uses 11 “calibration” parameters, whose only seven are in common with the previous version (Table 1). Some uncertainties remain in the exact value to be used for some parameters which were not calibrated in the first version of the model. In order to improve the control on ecophysiological processes during the calibration stage, we parameterized these parameters: the maximum Leaf Area Index of the site (LAI_{max}), the V_{cmax} photosynthetic parameter, the photoperiod at which the phenological phase 5 begins ($Photo_{per}$) and the parameter R_{day} linking leaf dark respiration rate to V_{cmax} .

3. Inference methods

In this problem the time series is short (34 years), and the observations are very noisy. Moreover, this experiment is not renewable. It is thus clear that we are in a Bayesian context. The Bayesian approach meets a strong success, especially in environmental sciences, after Markov Chain Monte Carlo algorithms were applied (Robert and Casella, 1999; Bernier et al., 2000). We will describe two methods of numerical Bayesian inference, namely the Markov Chain Monte Carlo (MCMC) and the particle filtering (PF). As a comparison, we also used a non-Bayesian method, the optimization algorithm PEST.

Table 1

List of the 11 MAIDEN parameter values after the three MAIDEN calibrations (PEST, MCMC and PF); the parameters are defined as follows

G_{dd}	Growing degree-days in °C
a_1	Leuning factor, linking the stomatal conductance and the net primary production
D_0	Coefficient linking the stomatal conductance and the humidity deficit at the leaf surface, in Pa
θ_c	Soil water stress parameter, in mm
c_2	Temperature function coefficient controlling the fraction of daily net primary production allocated to the store reservoir during fall
$f_{store \rightarrow bole}$	Fraction of stored carbon to bole during summer
$f_{npp \rightarrow bole}$	Fraction of net primary production to bole during summer
LAI_{max}	Maximum Leaf Area Index of the site
VC_{max}	Maximum Vc CO ₂ limitation of the photosynthesis of the species
$Photo_{per}$	Threshold of duration day-light for growing start
R_{day}	Daily sun radiation coefficient
t_{max}	Parameter modulating the maximum daily temperatures
t_{min}	Parameter modulating the minimum daily temperatures
p	Parameter modulating the daily precipitation

The fit (r^2) is computed on the calibration period (1963–1994) and the RMSE on the Istres independent years (1950–1962). The lower and upper bounds provide the *a priori* intervals. The bounds after ecophysiological calibration are given for the six parameters concerned. For each calibration we give the mode of the distribution and the 95% confidence intervals.

Here we suppose that the observation data are of the following form:

$$y_t = M_t(\theta, u_t) + \sigma_t v_t \text{ for } t = 1 \dots T \quad (1)$$

where v_t is a white Gaussian noise $N(0,1)$, u_t is the daily climatic variable, $M_t(\theta, u_t)$ is the nonlinear annual bole increment, $\theta \in \mathbb{R}^d$ is the unknown parameter vector, the observation noise variance σ^2 is also unknown.

In the Bayesian context we want to get a numerical approximation of the following conditional law:

$$\pi_T(\theta, \sigma^2) \stackrel{\text{def}}{=} [\theta, \sigma^2 | y_1 \dots y_T] \quad (2)$$

where $[\cdot]$ defines a probability. A first possibility is given by the MCMC methods: we build an ergodic Markov Chain whose limit distribution is π_T (Tierney, 1994; Gilks et al., 1995; Robert and Casella, 1999). Another possibility is to suppose that (θ, σ^2) are unobserved state variables, then π_T appears to be the solution of a nonlinear filter problem whose solution can be approximated by means of Particle Filtering (Gordon et al., 1993; Doucet et al., 2001).

3.1. Monte Carlo Markov Chains

Here the target law is given by Eq. (2). In Eq. (1) we assume that $v_t, \sigma^2, \theta^1, \dots, \theta^d$ are mutually independent. Hence, the parameters prior law satisfies

$$[\sigma^2, \theta^1, \dots, \theta^d] = [\sigma^2] \times [\theta^1] \times \dots \times [\theta^d]$$

and we choose uniform priors:

$$q_0(\sigma^2) \stackrel{\text{def}}{=} [\sigma^2] \propto 1_{[\sigma_{\min}^2, \sigma_{\max}^2]}[\sigma^2], \quad q_p(\theta^p) \stackrel{\text{def}}{=} [\theta^p] \propto 1_{[\theta_{\min}^p, \theta_{\max}^p]}(\theta^p)$$

where $0 < \sigma_{\min}^2 < \sigma_{\max}^2 < +\infty$ and $-\infty < \theta_{\min}^p < \theta_{\max}^p < +\infty$.

Bayes formula gives

$$\pi_T(\theta, \sigma^2) \propto [y_1 \dots y_T | \theta, \sigma^2] \times [\theta, \sigma^2]$$

According to Eq. (1), the likelihood function is

$$[y_1 \dots y_T | \theta, \sigma^2] = \frac{1}{(2\pi\sigma^2)^{T/2}} \times \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^T [y_t - M_t(\theta, u_t)]^2\right)$$

so that

$$\pi_T(\theta, \sigma^2) \propto f_T(\theta, \sigma^2) \stackrel{\text{def}}{=} \frac{1}{(\sigma^2)^{T/2}} \times \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^T [y_t - M_t(\theta, u_t)]^2\right) \quad (3)$$

if $\theta^p \in [\theta_{\min}^p, \theta_{\max}^p]$, $p = 1 \dots d$, and $\sigma^2 \in [\sigma_{\min}^2, \sigma_{\max}^2]$; $f_t(\theta, \sigma^2) = 0$ otherwise.

Here we propose a simple Metropolis algorithm with symmetric proposal distribution. We sample a new candidate according to

$$\tilde{\theta}^p = \theta^p + \gamma^p w^p, \quad p = 1 \dots d, \quad \tilde{\sigma}^2 = \sigma^2 + \gamma^0 w^0 \quad (4)$$

where w^0, w^1, \dots, w^d are independent $N(0,1)$ variables (we restrict these random walks to the prescribed intervals). The new candidate is then accepted according to the following probability:

$$r \stackrel{\text{def}}{=} \min\left(\frac{f_T(\tilde{\theta}, \tilde{\sigma}^2)}{f_T(\theta, \sigma^2)}, 1\right)$$

Here we adapt the variance $(\gamma^p)^2$ of the proposal Gaussian distributions Eq. (4) in order to balance the trade-off between a high acceptance rate and a low mixing. To do so, we checked the averaged fit between simulations and observations every 100 iterations and divided (respectively multiplied) by 2 the variance if this fit decreased (respectively increased). This takes into account the previous values of the Metropolis chain. Note that in this case the resulting process is no more Markovian, see Haario et al. (2001) for details.

3.2. Particle Filtering

The original model Eq. (1) can be rewritten as a nonlinear state-space model:

$$\theta_t^p = \theta_{t-1}^p + \gamma^p w_t^p, \quad \theta_0^p \sim q_p, \quad p = 1 \dots d \quad (5.1)$$

$$\sigma_t^2 = \sigma_{t-1}^2 + \gamma^0 w_t^0, \quad \sigma_0^2 \sim q_0 \quad (5.2)$$

$$y_t = M_t(\theta_t, u_t) + \sigma_t v_t \quad (5.3)$$

Then, $x_t \stackrel{\text{def}}{=} (\theta_t, \sigma_t^2)$ is now a non-observed state process. We want to determine the conditional law:

$$\pi_t = \pi_t(\theta_t, \sigma_t^2) \stackrel{\text{def}}{=} [\theta_t, \sigma_t^2 | y_1 \dots y_t] \quad (6)$$

When $\gamma^p = 0$, the model Eq. (5) is equivalent to the original one Eq. (1) and Eq. (6) is consistent with the definition (2). It is well known that standard Particle Filter implementation cannot handle such a situation because of degeneracy of particles. Taking $\gamma^p > 0$ induces an artificial dynamics of the parameters and increases the variance errors. This point will be discussed later.

The conditional law Eq. (6) is a solution of an infinite-dimensional nonlinear filter. A tractable approximation of this nonlinear filter is achieved by Particle Filtering. This approximation is of the form of an empirical probability measure:

$$\pi_t(\theta_t, \sigma_t^2) \cong \pi_t^N(\theta_t, \sigma_t^2) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i=1}^N \delta_{(\theta_t^i, \sigma_t^{i2})}(\theta_t, \sigma_t^2) \quad (7)$$

where $\delta_{(\theta_t^{[i]}, \sigma_t^{2,[i]})}(\theta_t, \sigma_t^2)$ denotes the Dirac measure at point $(\theta_t^{[i]}, \sigma_t^{2,[i]})$. The N particle positions $(\theta_t^{[i]}, \sigma_t^{2,[i]})_{i=1 \dots N}$ are then given by a Particle Filter method, for which the bootstrap algorithm is presented here.

3.2.1. Initialization

The initial set of particles is sampled from the priors:

$$\theta_t^{p,[i]} \sim q_p, \sigma_t^{2,[i]} \sim q_0 \text{ for } p = 1 \dots d, i = 1 \dots N$$

Then we repeat iteratively the two classical steps.

3.2.2. Mutation

The particles mimic the state Eqs. (5.1) and (5.2):

$$\theta_{t+1}^{p,[i]} \sim N(\theta_t^{p,[i]}, (\gamma^p)^2), \sigma_{t+1}^{2,[i]} \sim N(\sigma_t^{2,[i]}, (\gamma^0)^2) \quad (8)$$

for $p = 1 \dots d, i = 1 \dots N$

3.2.3. Selection

Using the new observation y_t and the model Eq. (5.3), we compute the likelihood of each particle:

$$\omega_t^{[i]} \propto f_t(\theta_t^{p,[i]}, \sigma_t^{2,[i]})$$

where f_t is given by Eq. (3) and $\sum_{i=1}^N \omega_t^{[i]} = 1$. We resample N new particles according to

$$(\theta_{t+1}^{p,[i]}, \sigma_{t+1}^{2,[i]}) \sim \sum_{j=1}^N \omega_t^{[j]} \delta_{(\theta_t^{p,[j]}, \sigma_t^{2,[j]})} \quad (9)$$

The multinomial sampling Eq. (9) could be cumbersome for N large. The number of particles is stable all along the process and often ranges from $\sim 10^2$ to 10^6 (we used 10^3 particles here). In practice, we use special procedures like the residual resampling one (Liu and Chen, 1998).

3.3. Nonlinear least squares (PEST)

One can try to infer this problem through a non-Bayesian approach like the nonlinear least squares one to get a referenced approach. The goal here is to find θ which minimizes the following criterion:

$$\sum_{t=1}^T (y_t - M_t(\theta, u_t))^2$$

We use the PEST package (Model Independent Parameter Estimation), a public domain model independent parameter estimator based on the Gauss–Marquardt–Levenberg method (Doherty, 2004). Note that PEST also provides estimations of the parameter σ^2 as a coefficient of proportionality of the covariance error matrix, and of the confidence intervals for the θ^p s.

3.4. Calibration scheme

We had at our disposal only one series of ecophysiological (transpiration) data in our studied region, coming from the Lamanon Aleppo pine stands. Hence, in order to improve the model calibration, we basically proceeded in two main stages. The first stage involves a calibration of the model on the basis of the six parameters controlling the ecophysiological processes and the transpiration data. This first stage used the MCMC method to deduce the parameter distributions and boundaries (the 5 and 95 percentiles of each parameter) calibrated for the studied region. These boundaries serve to define the *a priori* laws used in the

second stage. The second stage consists of a complete calibration based on the 11 model parameters (including the six ones calibrated on the transpiration data) and the dendrochronological time series. This calibration stage is the one used to compare the three previously described calibration methods. Modal fits and parameter distributions are based on five draws being merged. More precisely, they are computed on the last 80% of the 5000 iterations composing each draw (i.e. a 1000 iteration burn' in). For each of them, adapted convergence tests have been implemented. While the PEST algorithm tested the non-uniqueness of the parameter values obtained, the MCMC algorithm has been coupled with several convergence tests proposed by the BUGS project (Spiegelhalter et al., 1993). We used convergence tests dedicated to isolate draw for the PF algorithm. For the MCMC algorithm, we used the Gelman–Rubin ratio, which compares the inter- and inner-variances of the various chains (draws) that have been simulated in parallel (Gelman and Rubin, 1992). This ratio tends to unity when there is convergence.

In order to improve the method comparison, we performed detailed elasticity tests of the model to test possible misspecifications of model's components, as well as validations on independent meteorological data to detect possible over-parameterization. First, a sensitivity analysis has been performed after parameterization, consisting of varying the values of the calibration parameters one after the other ($\pm 5\%$ variations, with all other parameters being kept constant at their calibrated values) and observing the influence on the relative change of model predictions (Hamby, 1994). To be comparable between each other, these sensitivity values have to be normalized following the elasticity formula:

$$e_p = \frac{dM_t}{M_t} \cdot \frac{p}{dp_{5\%}}$$

where the elasticity e_p of the parameter p depends on the main model output, the bole increment M_t . We test also the sensitivity of the model to meteorological conditions (daily minimum temperature, daily maximum temperature and precipitation). For precipitations, we used the daily values of an “average” year (1975) to capture the stochastic behavior of this variable. This averaged meteorological year also helped us to choose the burn' in period necessary before the effective simulation runs, because of the temporal autocorrelation imposed by the existence of carbon and water pools in MAIDEN. The validation stage of the three calibrations over independent meteorological data uses the Root Mean Squared Errors (RMSE) to quantify the differences between observed and simulated bole increments between 1948 and 1961 (after removing the years 1956–1958 having extremely low winter temperatures).

4. Results

4.1. Physiological calibration

A MCMC calibration of the MAIDEN model has been performed on the daily transpiration measurements with the six parameters involved, starting from random θ -parameter values. Only the parameter boundaries are necessary for the rest of the study (Table 1, third column). The aim of this first stage was not to compare the calibration methods, but rather to decompose the global calibration scheme into ecophysiological and carbon allocation processes. We choose the MCMC method for this stage, as it was the most straightforward one to get the ecophysiological parameter distributions (we further use the boundaries rather than the modes) and to understand the model responses. We launch five runs in parallel to test the convergence. The determination coefficient r^2 between transpiration observations and MAIDEN simulations for the pine

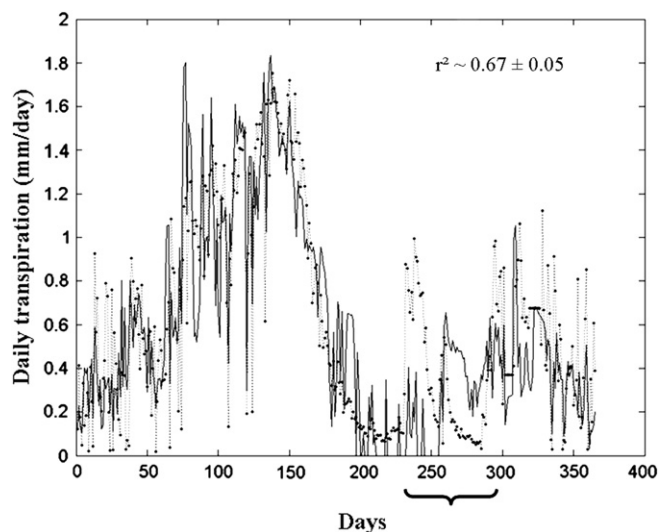


Fig. 1. The Lamanon stand transpiration curve with its final RMSE mode. Observations (dotted line) are measured at Lamanon site and simulations (plain line) are done with the MCMC algorithm, both for year 2004.

species at the Lamanon site after a MCMC calibration is equal to $r^2 = 0.67 \pm 0.05$, corresponding to a $\text{RMSE} = 0.28 \pm 0.03$ (Web Fig. 1a). The Gelman–Rubin ratio of the five draws tends to 1 and proves that the convergence has been reached (Web Fig. 1b).

The transpiration simulation based on this calibration and the corresponding observations shows the same pattern (Fig. 1), successively involving an increase in February (day 50), a decrease at the end of May (day 150), some discrepancies in September and October (days 240–290), and similar variations in November and December (days 300–365). The agreement is around 4%. Considering the water cycle, the discrepancy could be due to a high LAI_{max} value and/or low soil and vapor pressure deficit (D_0) stresses on stomatal conductance. However, this discrepancy does not seem to be due to the parameter values as both observations' and simulations' curves present the same daily variations, but rather due to a high soil depth inducing a low soil water stress. Considering the carbon cycle, the gross primary productivity (GPP) associated to this run seems relatively low ($\sim 970 \text{ gC/m}^2$) with a maximum carbon assimilation for the canopy at 6.5 gC/m^2 in May (day 140). After that, stomatal conductance decreases, thus reversing the assimilation trend. The assimilation maximum is not synchronous with the VC_{max} maximum in July (day 200). Finally, the six parameter probability distributions have normal shapes (not shown) and define the most probable limits (Table 1) used for subsequent calibrations of this study.

4.2. Dendrochronological calibrations

The MCMC calibration of MAIDEN with the PACA dendrochronological series has been done starting from random parameter values within the initial boundaries chosen. It shows a modal fit of $r^2 = 0.37 \pm 0.04$ (Web Fig. 2a). The fit evolution is progressive on most of the draws with a reasonable ($\sim 17\%$) acceptance/rejection rate of the Metropolis–Hastings criteria for 11 parameters. The MCMC run did globally reach a stationary state as confirmed by the convergence tests (Web Fig. 2b), leading to the 11 parameter distributions (Web Fig. 3b, Table 1 columns 6–7). Most of the parameter distributions have normal shapes, while the uniform or multi-modal distributions (parameters D_0 , LAI_{max} , VC_{max} and R_{day}) are mainly due to the running time constraints not allowing to reach a stationary state (Web Fig. 3a for D_0). Differences between the calibrated bole increment simulation and the observations are

quite regular, showing the ability of the model to simulate rather different growing situations and hence, carbon storage in the tree compartments (Fig. 2a).

The PEST optimization of the model with the PACA dendrochronological series gives a fit equal to $r^2 = 0.46$. This optimum fit is in a good agreement with the MCMC final fit distribution, as the maximum MCMC fit value reached is equal to $r^2 \sim 0.52$ (while $r^2_{\text{MCMC}} = 0.37$ is the modal fit). PEST needed 16 iterations to reach this value. Most of the lower parameter boundaries of the 95% confidence intervals are negative and too large to be realistic and useful.

Finally, for the PF algorithm, we start from random values for θ in the same intervals used for MCMC algorithm. As a sequential algorithm, PF tests each new observation year on the basis of mean-squares (r^2). The last step of this calibration run gives a final fit equal to $r^2 = 0.25 \pm 0.02$ with 1000 particles, a fit lower than for the other calibration techniques (Web Fig. 4). Years 1987 and 1994 for example are constraining for this calibration, in the sense that they bring new data information that forces the model to adapt, while the fit is improved for years around 1974 and for 1992. The last

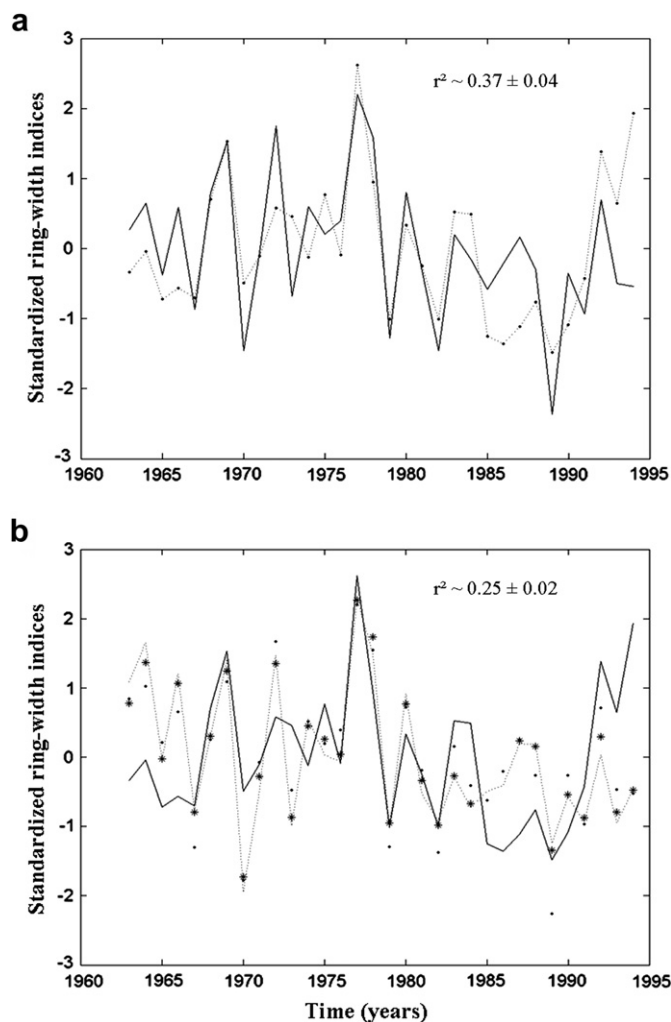


Fig. 2. The standardized bole increment evolutions simulated by MAIDEN over 1963–1994 years for the MCMC (a) and PF (b) calibration methods and their observation/simulation fits. The MCMC dendrochronological time series shows the measures (dotted line) and the simulation (plain line). The PF dendrochronological time series is superimposed with observations (dots), the simulations using the initial parameters (plain line), the simulated values during the run at each new step (stars) and the simulations using the final values at the last step (dashed line). The absence of stars for the freezing years 1984–1985 indicates that they were not used during this calibration.

calibration stage provides the *a posteriori* distributions (Fig. 3b) and their associated confidence intervals (Table 1, columns 8 and 9). Most of the parameter distributions are normally shaped, indicating a real update of our knowledge on these parameters whereas for the others ($f_{\text{store} \rightarrow \text{bole}}$, $f_{\text{npp} \rightarrow \text{bole}}$ and R_{day}) it seems that one gained nothing in terms of information. Fig. 2b compares bole

increment observations and simulations before (plain line), during (stars) and after (dashed line) the calibration. This highlights in particular the progressive improvement of the parameter set choices.

As a validation step, the bole increment simulation on independent Istres meteorological data after the MCMC calibration

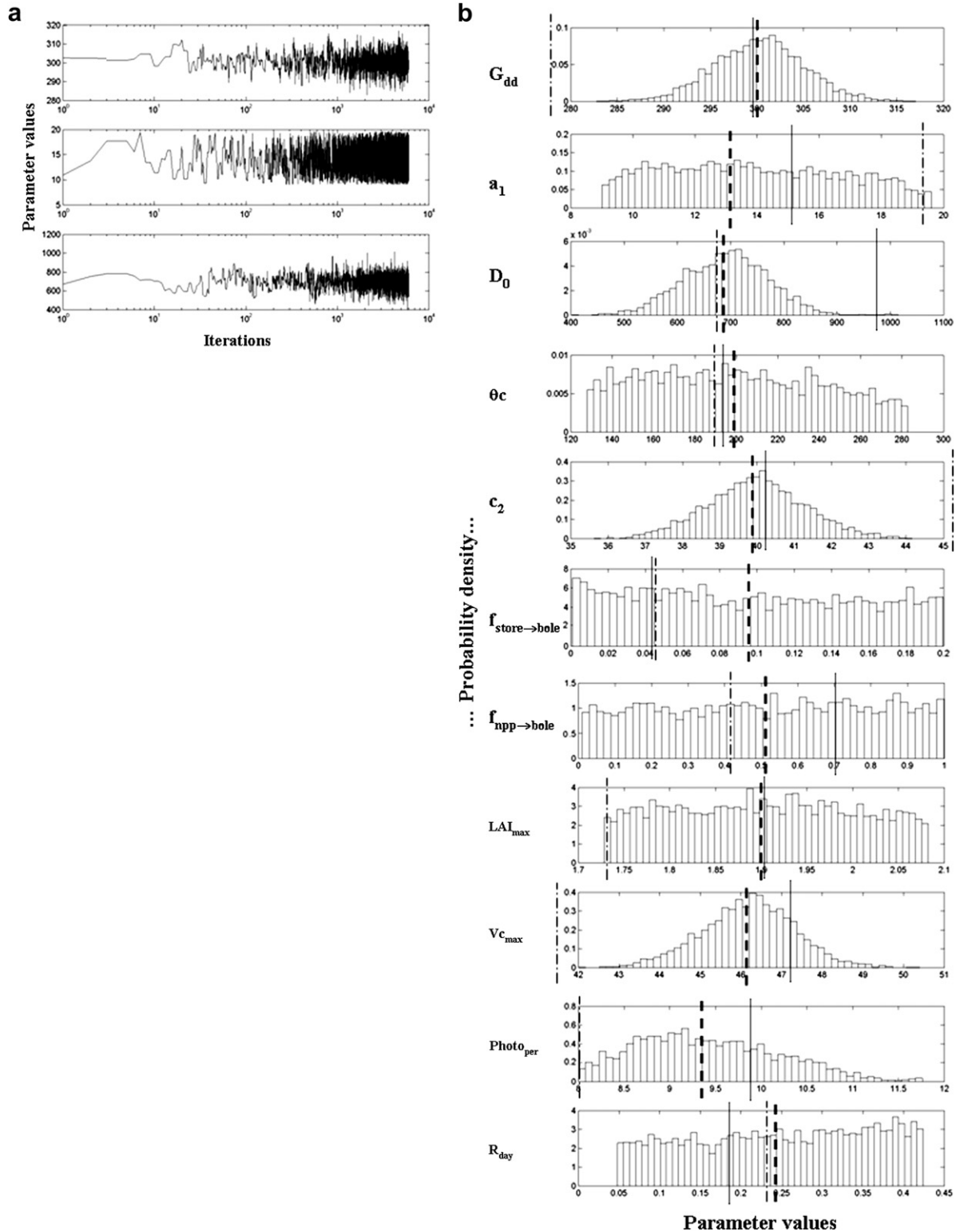


Fig. 3. The 11 MAIDEN parameter probability distributions (b) and the value evolution of the first three during the PF calibration (a). The whole run values are shown (6000 particles). The heavy dashed lines indicate the optimum parameter value, the dash-dotted lines indicate the 5 and 95 percentiles, while the light dashed lines show the MCMC parameter modes for comparison.

leads to a RMSE ~ 0.90 , a value lower than 1 standard deviation. It is acceptable but not negligible. The bole increment curve built from the Istres meteorological data corresponding to the PEST calibration appears to be very similar to the MCMC (Web Fig. 5). It leads to a RMSE ~ 0.91 . The final PF calibration RMSE over the independent Istres meteorological data is equal to ~ 0.79 , lower than the other technique ones. Hence, this better bole increment simulation of the PF algorithm on the basis of the earlier period (before 1960) seems to compensate its rather low calibration score.

4.3. Calibration comparison

Significant differences appear between the MCMC, PF and PEST calibrations, concerning either the parameter modes and therefore conditioning the bole increment simulations, or concerning the parameter distributions and limits (Table 1).

Concerning the calibration modes of the model parameters, PEST is often in disagreement (20% difference in average and up to $\sim 35\%$ for a_1) with the other method values: in particular parameters G_{dd} , a_1 , c_2 , LAI_{max} , V_{cmax} , $Photo_{per}$, R_{day} . Most of the parameter modes are in agreement between the MCMC and PF methods, except D_0 , for which the MCMC mode is outside the PF limits. As a consequence, the dendrochronological simulations using the MCMC and PEST calibration modes seem rather similar, while the parameters and intermediate ecophysiological simulated variables are different. Three main differences are identified. MCMC run shows a lower total water stress (atmosphere + soil), mainly due to a lower soil water stress and thus higher transpiration and GPP (although with distinct V_{cmax}). A second difference is observed for the carbon allocation, the PEST run having a much lower $Photo_{per}$ ($=8$) inducing a reduced but still effective bole growth between the end of the summer and the beginning of the winter. Moreover, the stored carbon (in the storage reservoir) and its use for the bole growth are higher for the PEST run than for the MCMC run. The third difference is observed on the G_{dd} values being lower for the PEST run, thus starting the growing season in average 4 days in advance. The PF run, while being rather similar to the others, does not simulate some inter-annual variations (the PF coefficient variation is equal to ~ 8.36 instead of ~ 7.87 for PEST and ~ 10.75 for MCMC). The use of the carbon reserves (the storage reservoir $f_{store \rightarrow bole}$) is higher, the soil water stress θ_c is higher and the Leuning coefficient a_1 lower for the PF run. This explains why the soil water stress starts quite early in the year and why the transpiration is too low, leaving too much water unused in the soil. Most of the bole increment variability comes thus from the reserve use, leading to a slightly flatter curve than for the MCMC run.

Concerning the probability distributions of the model parameters, most of the PF distributions appear more regular and normally shaped than the MCMC ones. Parameters $f_{store \rightarrow bole}$, $f_{npp \rightarrow bole}$, R_{day} are the only one showing non-informative (uniform) distributions after PF calibration, and truncated normally shaped distributions after MCMC calibration. The associated MCMC and PF parameter confidence intervals are often in agreement, while the PEST uncertainty intervals appear to be unrealistically large. Only G_{dd} and V_{cmax} have confidence intervals physically acceptable. PF confidence intervals are usually narrower than the MCMC ones, indicating a better ability to converge into a relevant (high observation/simulation fits) region of the parameter space. Three of the 11 MAIDEN parameters shown rather different values than what was expected given the calibration done using trial-and-error method with a previous version of the model (Misson et al., 2004). The G_{dd} is found around 300 ± 50 degree-days, which is about 100 degree-days higher than the previous value calibrated for *P. halepensis*. This suggests a bud-burst triggered about 1 week later during the year. The θ_c is found around 200 ± 50 mm, which is about 100 mm higher than the previous value. This suggests, other

things being equal, a soil water stress arriving earlier in the season than previously expected. The LAI_{max} is found around 1.9 ± 0.2 , which is lower than the constant value of 2.7 used in the earlier version of the model. This suggests less productive pine sites than previously expected. It is possible that convergence towards lower LAI_{max} has influenced a higher θ_c in order to trigger consistent soil water stress on stomatal conductance in both the earlier version and the present version of the model.

4.4. Elasticity comparison

Elasticity and validation tests are necessary to evaluate the calibration methods. The $Photo_{per}$ imposing the start of the fall period, the c_2 parameter controlling the fraction of daily NPP allocated to the store reservoir, the t_{max} parameter modulating the maximum daily temperatures, and the V_{cmax} photosynthetic parameter are the parameters having the highest impact, roughly of the same order of magnitude (Web Table 1). Only t_{max} has a recurrent high elasticity across all the calibration methods and seems to be the most influent parameter overall. The previous four-parameter elasticities roughly stay at the same relative order with higher parameter variations up to $\pm 50\%$ for the MCMC calibration for example. Most of the other model parameters have a weak and comparable influence. The p parameter modulating the precipitations appears to have significant influence, mainly for the MCMC calibration.

4.5. Independent verification

When applied to independent meteorological data (Istres: 1950–1962), the three calibration parameter sets (PEST, MCMC, PF) lead to simulated bole increment variations being rather similar (Web Fig. 5). Residuals between the standardized final simulations and the observations over the 11 years of validation show the difficulties of the MAIDEN model to simulate some of the years: namely the freezing years 1956/1958 and following, the rainy year 1963 and years 1993–1994 (Fig. 4). When looking at the residuals between the three final bole increment series and the corresponding observations, it appears that MCMC and PEST present

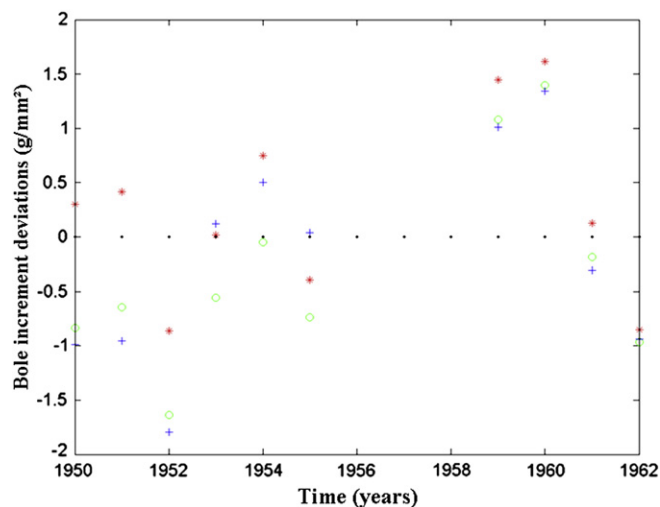


Fig. 4. The residuals of the regional calibrated bole increment curves. The comparison of the standardized bole increment residuals are deduced from the three calibration methods and the corresponding observations, on the basis of the Istres meteorological data for independent years (1950–1962) after removing the 1956–1958 freezing years. The calibration methods are figured by different colors: PEST: green, PF: red, MCMC: blue and observations in black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

rather similar behavior, while distinct from that of the PF algorithm. The final observation/simulation RMSE confirms this (Table 1). The main contribution to these residuals comes from the first 3 years.

5. Discussion

In this study we used three different approaches to calibrate (parameterize) a complex nonlinear ecophysiological model: PEST optimization, MCMC and PF algorithms. We showed that they have respective advantages and disadvantages that could be exploited depending on the nature of the environmental model and the computing power available. We discuss here the results according to ecophysiological and statistical points of view.

5.1. Ecophysiology

The physiological part of the MAIDEN model has been calibrated first using 1 year of observed transpiration measurements. The observation/simulation fit is seasonally correct, while some differences exist for shorter-term variations. The low residuals (of about 4% over the year) could be due to the maybe high soil thickness in the model, leading to an inefficient soil water stress on transpiration. In addition, the seasonality of the $V_{c_{max}}$ parameter is related to variations in air temperature only, which could be inappropriate in the very dry conditions characterizing the experimental site where the data were collected. Other factors should be taken into account, however, this is still badly understood and suggest new observations and experiments (Misson et al., 2006). It must be pointed out that our simulations correspond to an “averaged” carbon allocation for *P. halepensis* on a typical regional site and then that the site conditions cannot be realistic. This weakness is confirmed by the sensitivity analyses done above which showed that the t_{max} parameter was always the most sensitive one. The water stress parameters are more dependent on the local conditions.

Then, the dendrochronological part of the MAIDEN model has been calibrated. MCMC and PEST simulations are rather similar and slightly different than the PF ones. Simulations are realistic from 1963 to 1984, but the periods 1985–1987 and 1991–1994 are poorly simulated (Web Fig. 5). For the first period, it could be explained by the cold winters (low t_{min}) that damaged the cells, a process that MAIDEN is not able to take into account. There is no clear explanation for the second period. The PF simulation is slightly less correlated with the observations: it shows roughly the same inter-annual variations than for the two other calibrations, but in a lower extent. It could be due to a too high carbon allocation in the storage compartment. The PF run is also the one having the highest soil water stress θ_c and the lowest Leuning coefficient a_l , leading to a weaker and unrealistic transpiration. Nevertheless, these runs correspond to “averaged” carbon allocations for *P. halepensis* on a typical regional site (Abbas, 1986).

5.2. Bayesian vs. nonlinear least square methods

The MCMC and PEST algorithms globally show the same ability to calibrate a complex tree-growth model, although built on different inference bases. Both gave relatively similar results, with slightly higher observation/simulation fits than the PF procedure (Table 1). PEST gives unrealistic confidence intervals; the MCMC and PF algorithms are much better in this respect. The counterpart is a much longer computer time. PEST also suffers from a high dependence on the noise in the observations. The fact that it only gives a best fit estimate instead of probability distributions makes that we are not able to appreciate the errors. For example, this fact is probably responsible for the low $Photo_{per}$ value (=8) as estimated by PEST. This value, leading to a bole growth during fall, is close to the lower boundary and should not have a high probability. This

nonlinear least square approach is not comparable to the Bayesian approach and each of them would be selected only if we need a quick estimate of the parameters. PEST has the best fit ($r^2 \sim 0.45$). It is logical, as an optimization procedure leads to the best fit, reached several times by MCMC. While PEST retains the highest, MCMC and PF retain the value the most frequently reached, which is more robust and less dependent on the observation noise. The independent validation confirms that a too high value of r^2 leads to a more important risk of over-parameterization. PF has the lowest r^2 and the best prediction RMSE (~ 0.8).

5.3. MCMC vs. PF

Although both MCMC and PF are Monte Carlo methods, they are very different. MCMC is founded on the ergodic theorem, whereas PF is founded on the law of the large numbers. The MCMC methods are iterative, whereas PF are not. The problem of convergence of these methods is a key issue (Cowles and Bradley, 1996; Brooks and Gareth, 1998). MCMC are batch procedures, whereas PF are sequential. Indeed, if a new measurement y_{T+1} is made available, MCMC methods cannot treat it alone, the procedures must be restarted from the beginning. PF methods, also called sequential Monte Carlo, treat the observations available progressively. This is one of the major issues in nonlinear filtering. In our problem this point is not essential. One can use a more sophisticated mutation procedure than Eq. (8), for example an acceptance/rejection scheme. There exist PF methods which are more adapted to the present situation, like the kernel ones (Liu and West, 2001; Doucet and Tadic, 2003; Campillo and Rossi, 2006). They make it possible to avoid adding the artificial diffusion terms in Eqs. (5.1) and (5.2). Those approaches will be developed in a future work.

Practically in the present study, the MCMC algorithm coupled with the convergence tests (Gelman–Rubin) remains robust. Yet, this Bayesian parameterization procedure suffers from a real tuning difficulty, at the opposite to the PF algorithm. The *a priori* distributions (and boundary variations to zoom into the parameter space) and the algorithm convergence require a good knowledge of this algorithm and even a long experience of their use in various cases. The main disadvantage of the PF algorithm is that the particle number is important (from $\sim 10^2$ to 10^6) for the algorithm to be efficient. At counterpart, PF algorithm has several advantages: i) based on a genetic algorithm concept easy to understand, it is easy to develop and to implement; ii) no burn-in period and no parameter distribution filtering are necessary as the *a posteriori* distributions are directly given at the last running step; iii) It can be parallelized in case of high number of particles; iv) and almost no tuning algorithm is necessary. As a sequential approach, PF algorithm has the ability to indicate the very constraining years, those which poorly fit simulations and observations. As a last advantage, PF algorithm based on the law of the large numbers obtains smoother and more regular (normal) parameter distributions being thus easier to analyze and interpret. Yet, we emphasize that all of these calibration methods could benefit to socio-ecological scientists.

Acknowledgements

This research has been done in the frame of project REFORME funded by the French Ministère de l'Ecologie et du Développement Durable (program GICC2). We thank E. Rivot for his advices and R. Huc for the Lamanon data.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.envsoft.2008.03.003.

References

- Abbas, H., 1986. La productivité des forêts de pin d'Alep dans le Sud-Est méditerranéen français. Analyse écodendrométriques. CIHEAM-Options Méditerranéennes, 127–156.
- Andrieu, C., Djuric, P.M., Doucet, A., 2001. Model selection by MCMC computation. *Signal Processing* 81, 19–37.
- Bernacchi, C.J., Pimentel, C., Long, S.P., 2003. In vivo temperature response functions of parameters required to model RuBP-limited photosynthesis. *Plant, Cell and Environment* 26, 1419–1430.
- Bernier, J., Parent, E., Boreux, J.-J., 2000. Statistique pour l'environnement: Traitement bayésien de l'incertitude. Tech. & Doc. Lavoisier.
- Brooks, S.P., Gareth, O.R., 1998. Convergence assessment techniques for Markov chain Monte Carlo. *Statistics and Computing* 8, 319–335.
- Bugg, A.L., Landsberg, J., Laughlin, G., Tickle, P., Frakes, I., 2006. Application of three modelling approaches to simulating tree growth in central NSW, Australia. *Environmental Modelling and Software* 21, 310–319.
- Campillo, F., Rossi, V., 2006. Convolution Particle Filters for Parameter Estimation in General State-space Models. Research Report No. 5939, INRIA.
- Cowles, M.K., Bradley, P.C., 1996. Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association* 91, 883–904.
- Doherty, J., 2003. Ground water model calibration using pilot points and regularization. *Ground Water* 41, 170–177.
- Doherty, J., 2004. PEST: Model-independent Parameter Estimation. User Manual. Watermark Numerical Computing.
- Dorner, S., Shi, J., Swayne, D., 2007. Multi-objective modelling and decision support using a Bayesian network approximation to a non-point source pollution model. *Environmental Modelling and Software* 22, 211–222.
- Doucet, A., Tadic, V.B., 2003. Parameter estimation in general state-space models using particle methods. *Annals of the Institute of Statistical Mathematics* 55, 409–422.
- Doucet, A., de Freitas, N., Gordon, N.J., 2001. Sequential Monte Carlo Methods in Practice. Statistics for Engineering and Information Science. Springer-Verlag, New York, USA.
- Farquhar, G.D., Von Caemmerer, S., Berry, J.A., 1980. A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta* 149, 78–90.
- Fritts, H.C., Swetnam, T.W., 1989. Dendroecology – a tool for evaluating variations in past and present forest environments. *Advances in Ecological Research* 19, 111–188.
- Fritts, H.C., Vaganov, E.A., Sviderskaya, I.V., Shashkin, A.V., 1991. Climatic variation and tree-ring structure in conifers: empirical and mechanistic models of tree-ring width, number of cells, cell size, cell-wall thickness and wood density. *Climate Research* 1, 97–116.
- Gallagher, M., Doherty, J., 2007. Parameter estimation and uncertainty analysis for a watershed model. *Environmental Modelling and Software* 22, 1000–1020.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7, 457–511.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis. Chapman & Hall, London, 526 pp.
- Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1995. Markov Chain Monte Carlo in Practice. Chapman & Hall, London.
- Gordon, N.J., Salmond, D.J., Smith, A.F.M., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings Part F* 140 (2), 107–113.
- Granier, A., Huc, R., Barigah, S.T., 1996. Transpiration of natural rainforest and its dependence on climatic factors. *Agricultural and Forest Meteorology* 78, 19–29.
- Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. *Bernoulli* 7 (2), 223–242.
- Hamby, D.M., 1994. A review of techniques for parameter sensitivity analysis of environmental-models. *Environmental Monitoring and Assessment* 32, 135–154.
- Liu, J., West, M., 2001. Combined parameter and state estimation in simulation-based filtering. In: Doucet, A., de Freitas, N., Gordon, N.J. (Eds.), *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, pp. 197–223.
- Liu, J.S., Chen, R., 1998. Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association* 93, 1032–1044.
- Misson, L., 2004. MAIDEN: a model for analyzing ecosystem processes in dendroecology. *Canadian Journal of Forest Research. (Revue Canadienne De Recherche Forestiere)* 34, 874–887.
- Misson, L., Rathgeber, C., Guiot, J., 2004. Dendroecological analysis of climatic effects on *Quercus petraea* and *Pinus halepensis* radial growth using the process-based MAIDEN model. *Canadian Journal of Forest Research. (Revue Canadienne De Recherche Forestiere)* 34, 888–898.
- Misson, L., Tu, K., Boniello, R., Goldstein, A.H., 2006. Seasonality of photosynthetic parameters in a multi-specific and vertically complex forest ecosystem in the Sierra Nevada of California. *Tree Physiology* 26, 729–741.
- Nicault, A., 1999. Analyse de l'influence du climat sur les variations inter- et intra-annuelles de la croissance radiale du pin d'Alep (*Pinus halepensis* Mill.) en provenance calcaire. Aix-Marseille III, Marseille, FR.
- Pelletier, G.J., Chapra, S.C., Tao, H., 2006. QUAL2Kw – a framework for modeling water quality in streams and rivers using a genetic algorithm for calibration. *Environmental Modelling and Software* 21, 419–425.
- Rathgeber, C., Nicault, A., Guiot, J., Keller, T., Guibal, F., Roche, P., 2000. Simulated responses of *Pinus halepensis* forest productivity to climatic change and CO₂ increase using a statistical model. *Global and Planetary Change* 26, 405–421.
- Robert, C.P., Casella, G., 1999. Monte Carlo Statistical Methods. Springer-Verlag, Berlin.
- Running, S.W., Nemani, R.R., Hungerford, R.D., 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Canadian Journal of Forest Research. (Revue Canadienne De Recherche Forestiere)* 17, 472–483.
- Spiegelhalter, D.J., Thomas, A., Gilks, W.R., 1993. BUGS, Bayesian Inference Using Gibbs Sampling. MRC Biostatistics Unit, Cambridge, UK.
- Tierney, L., 1994. Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics* 22, 1701–1728.
- Tischler, M., Garcia, M., Peters-Lidard, C., Moran, M.S., Miller, S., Thoma, D., Kumar, S., Geiger, J., 2007. A GIS framework for surface-layer soil moisture estimation combining satellite radar measurements and land surface modeling with soil physical property estimation. *Environmental Modelling and Software* 22, 891–898.
- Yapo, P.O., Gupta, H.V., Sorooshian, S., 1998. Multi-objective global optimization for hydrologic models. *Journal of Hydrology* 204, 83–97.
- Zhou, X.L., Peng, C.H., Dang, Q.L., 2004. Assessing the generality and accuracy of the TRIPLEX model using in situ data of boreal forests in central Canada. *Environmental Modelling and Software* 19, 35–46.