Geosci. Model Dev. Discuss., 7, 6997–7031, 2014
www.geosci-model-dev-discuss.net/7/6997/2014/
doi:10.5194/gmdd-7-6997-2014

Geoscientific Model Development Discussions
Open Access

A user-friendly forest model with a multiplicative mathematical structure: a Bayesian approach to calibration

M. Bagnara[1,2], M. Van Oijen[3], D. Cameron[3], D. Gianelle[4], F. Magnani[2], and M. Sottocornola[1,5]

[1]Fondazione Edmund Mach, Research and Innovation Centre, Sustainable Agro-ecosystems and Bioresources Department, via E. Mach, 1, S. Michele all'Adige (TN), Italy
[2]Department of Agricultural Sciences, University of Bologna, Viale Fanin 46, Bologna, Italy
[3]Centre for Ecology and Hydrology, CEH-Edinburgh, Bush Estate, Penicuik EH26 0QB, UK
[4]FOXLAB, Research and Innovation Centre, via E. Mach, 1, S. Michele all'Adige (TN), Italy
[5]Department of Chemical and Life Sciences, Waterford Institute of Technology, Cork road, Waterford, Ireland

Received: 12 September 2014 – Accepted: 25 September 2014 – Published: 22 October 2014

Correspondence to: M. Bagnara (maurizio.bagnara@fmach.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Forest models are being increasingly used to study ecosystem functioning, through the reproduction of carbon fluxes and productivity in very different forests all over the world. Over the last two decades, the need for simple and "easy to use" models for practical applications, characterized by few parameters and equations, has become clear, and some have been developed for this purpose. These models aim to represent the main drivers underlying forest ecosystem processes while being applicable to the widest possible range of forest ecosystems. Recently, it has also become clear that model performance should not be assessed only in terms of accuracy of estimations and predictions, but also in terms of estimates of model uncertainties. Therefore, the Bayesian approach has increasingly been applied to calibrate forest models, with the aim of estimating the uncertainty of their results, and of comparing their performances.

Some forest models, considered to be user-friendly, rely on a multiplicative or quasi-multiplicative mathematical structure, which is known to cause problems during the calibration process, mainly due to high correlations between parameters. In a Bayesian framework using a Markov Chain Monte Carlo sampling this is likely to impair the reaching of a proper convergence of the chains and the sampling from the correct posterior distribution.

Here we show two methods to reach proper convergence when using a forest model with a multiplicative structure, applying different algorithms with different number of iterations during the Markov Chain Monte Carlo or a two-steps calibration. The results showed that recently proposed algorithms for adaptive calibration do not confer a clear advantage over the Metropolis–Hastings Random Walk algorithm for the forest model used here. Moreover, the calibration remains time consuming and mathematically difficult, so advantages of using a fast and user-friendly model can be lost due to the calibration process that is needed to obtain reliable results.

GMDD

7, 6997–7031, 2014

1 Introduction

Gross Primary Production (GPP) is a key component of the carbon balance. Therefore, it is the central output of many forest ecosystem models (Mäkelä et al., 2000; Tjiputra et al., 2013; De Weirdt et al., 2012), and is being increasingly targeted by remote sensing applications as a proxy to predict global carbon fluxes and plant light-use efficiency at large spatial scales (Still et al., 2004; Wisskirchen et al., 2013; Zhang and Kondragunta, 2006). It can also be estimated by the Eddy-covariance technique: this micrometeorological method computes the net CO_2 turbulent flux between a given ecosystem and the atmosphere from the covariance between the fluctuations of vertical wind velocity and CO_2 concentrations, averaged at a half-hour time scale. This can be used to estimate both ecosystem respiration (Re) and GPP. Currently, a global network of more than 500 EC stations exist worldwide to continuously monitor the CO_2 and energy exchange between ecosystems and the atmosphere, whose homogeneity is ensured by similar standardized procedures (Baldocchi, 2008). Despite extensive efforts and several techniques tested, GPP quantification remains challenging in most ecosystems given its dependence on several meteorological, environmental and internal drivers at several time scales.

Most of the models of forest growth and biogeochemical cycles are detailed, multivariable models that need much environmental information and careful parameterisation before they can be run (Landsberg and Waring, 1997). To make them suitable for a wider range of purposes and accessible to a wider range of users, a process of simplification started in the 90's (White and Running, 1994; Landsberg and Waring, 1997) with the aim of developing models that could be of use in applied forest management.

One step in this direction was represented by the creation of hybrid models (e.g. FORCYTE-11, Kimmins, 1986), combining the predictive power of process-based models with the short-term believability of mensuration-based models (Kimmins et al., 1999; Medlyn et al., 2003). Unlike full process-based models, hybrid models are based on the principle that only the processes that are expected to change would be included

GMDD
7, 6997–7031, 2014

in the modelling effort, reducing the number of processes taken into account and resulting in a simplification of the overall model structure (Kimmins et al., 2008).

A widely used group of simple models for GPP is based on the concept of Light Use Efficiency (LUE), defined as the ratio of GPP to Absorbed Photosynthetically Active Radiation (APAR). They mainly rely on simplified physiological processes and empirical parameters, require little information to be run, the computations are usually fast, and their mathematical structure is often quasi- or totally multiplicative. These models assume that vegetation has a potential LUE, which can be described as the ability of plants to use light for photosynthesis in absence of limiting factors, decreased by modifying factors that account for suboptimal conditions for photosynthesis (Landsberg and Waring, 1997; McMurtrie et al., 1994).

Some examples of these models are 3PG (Landsberg and Waring, 1997), C-Fix (Veroustraete et al., 1994), the model developed by Horn and Schulz (2011b), and Prelued (Mäkelä et al., 2008a).

Despite relying on a multiplicative mathematical structure and on several empirical parameters, of which little is known in the literature, Prelued has been successfully applied to several ecosystems all across the world (Bagnara et al., 2014; Mäkelä et al., 2008a; Peltoniemi et al., 2012). Compared to the majority of the LUE-based models that work at monthly or annual time scale, relying on a linear relationship between GPP and APAR and on a parabolic effect of temperature, the Prelued model calculates GPP at a daily time scale, basing the calculations on a nonlinear relationship between APAR and GPP (Medlyn et al., 2003; Turner et al., 2003), a saturating effect of average daily temperature (which simulates the ecosystem "acclimation" to temperature, Mäkelä et al., 2004), and daily meteorological and environmental variables. The importance of these environmental variables has been recently highlighted by McCallum et al. (2013), applying the Prelued model at four Eddy-Covariance sites in Russia: this clearly demonstrates the improved fit of the model when considering both temperature acclimation and nonlinearity in response to APAR, especially in temperature-controlled ecosystems.

GMDD
7, 6997–7031, 2014

The Bayesian approach to model calibration has become more and more popular in the last few years to obtain insights on both model predictions and uncertainties. The main characteristic of a Bayesian calibration is that it quantifies model inputs and outputs in the form of probability distributions, and applies the rules of probability theory to update the distributions when new data are obtained (van Oijen et al., 2005; Sivia, 1996). This approach has been widely used in the past in different fields, and recently it has been applied to different kind of ecosystem models, focusing on both crop (Zhu et al., 2014) and forest (van Oijen et al., 2005, 2011, 2013; Svensson et al., 2008; Chevallier et al., 2006). Even so, the application of the bayesian method to LUE-based models is not as common as its application to process-based models, with a very few studies heading in this direction (Bagnara et al., 2014; Still et al., 2004; Xenakis et al., 2008).

Implementations of Bayesian calibration rely mainly on Markov Chain Monte Carlo (MCMC) algorithms for sampling the parameter space to obtain posterior probability distributions for the model parameters. The efficiency of the MCMC technique is highly dependent on the model structure. The high correlations between parameters that are induced by a multiplicative model structure generally make the convergence of the MCMC more difficult, impairing the reliability of the results of the calibration. Different methods have been implemented to avoid or reduce such problems: the use of very long chains (Gilks et al., 1996), model re-parameterization to avoid strong correlations (Buzzi-Ferraris and Manenti, 2010; Gilks et al., 1996), and the use of more efficient algorithms (ter Braak, 2006; Gilks et al., 1996). In this study, three algorithms characterized by increasing complexity and efficiency were applied: the Metropolis–Hastings Random Walk (MHRW), the Adaptive Metropolis (AM), and the Differential Evolution Markov Chain (DEMC).

The Metropolis–Hastings random walk (MHRW) (Robert and Casella, 1999) algorithm produces a walk through the parameter space such that the collection of visited points forms the desired sample from the posterior distribution, discarding some initial values (van Oijen et al., 2005). At each iteration of the algorithm, a new candidate

GMDD
7, 6997–7031, 2014

parameter vector is proposed stochastically, i.e. the jump from the current point to the proposed next one follows a probability distribution. The most commonly used proposal distribution is the multivariate Gaussian. Whether the proposal is accepted, depends on the prior probabilities and likelihoods of the current and proposed parameter vectors. In the MHRW, the proposal distribution itself does not change, so average proposed jump directions and distances remain the same throughout the random walk. This is different in the next two MCMC algorithms. The adaptive Metropolis algorithm (AM) is a modification of the MHRW. The key attribute of the AM algorithm is the continuous adaptation of its proposal distribution. The adaptation consists of gradual convergence of the covariance matrix of the proposal distribution to the covariance matrix of the parameters visited so far in the chain (Haario et al., 2001; Smith and Marshall, 2008). The differential evolution Markov chain algorithm (DEMC) is formed by combining the differential evolution algorithm of Storn and Price (1997), designed for global optimization in real parameter spaces, with MCMC sampling, utilizing standard Metropolis principles. The result is a population MCMC algorithm, where multiple chains are run in parallel and allowed to learn from each other. This combination intends to overcome the difficulties common to MCMC methods of choosing an appropriate scale and orientation (respectively the size of each jump in the MCMC sampling and its direction in the parameter space) for the proposal distribution, while also addressing issues of computational efficiency related to the time to reach convergence (ter Braak, 2006; Smith and Marshall, 2008).

This work aims at testing different procedures that could be successfully applied to the variety of forest models with similar structure to reach proper convergence during the MCMC sampling. We applied a Bayesian calibration with different algorithms and number of iterations, as well as reparameterization and multiple steps calibration, to the Prelued model, employed as a case study, using one year of daily GPP data from an Eddy-Covariance (EC) tower in the Italian Alps.

GMDD
7, 6997–7031, 2014

2 Materials and methods

2.1 Model formulation

The model used as a case study is a modified version of a LUE-type model of daily photosynthetic production of the canopy (Mäkelä et al., 2008a):

$$\mathrm{GPP}_j = \beta \mathrm{APAR}_j \prod_i F_{ij} \quad (1)$$

where GPP_j is canopy Gross Primary Production ($\mathrm{g\,C\,m^{-2}}$) during day j, β is potential Light Use Efficiency ($\mathrm{g\,C\,mol^{-1}}$), APAR_j is Absorbed Photosynthetically Active Radiation ($\mathrm{mol\,m^{-2}}$) during day j, and $F_{ij} \in [0,1]$ are modifying factors accounting for suboptimal conditions on day j. The actual LUE of the canopy on day j is the product of β and the current values of the modifiers.

To account for the nonlinearity in the response to APAR, a light modifier FL was defined so as to yield the rectangular hyperbola when multiplied with the linear response included in the LUE model:

$$\mathrm{FL}_j = 1/(\gamma \mathrm{APAR}_j + 1) \quad (2)$$

where γ ($\mathrm{m^2\,mol^{-1}}$) is an empirical parameter. The effect of temperature on daily GPP was modelled using the concept of state of acclimation, S_j (°C) (Mäkelä et al., 2004), a piecewise linear function of X_j (°C) calculated from the mean daily ambient temperature, T_j (°C), using a first-order dynamic delay model:

$$X_j = X_{j-1} + (1/\tau)(T_j - X_{j-1}), \quad X_1 = T_1 \quad (3)$$

$$S_j = \max\{X_j - X_0, 0\} \quad (4)$$

where τ (days) is the time constant of the delay process and X_0 (°C) is a threshold value of the delayed temperature. The modifying function FS is defined as

$$\mathrm{FS}_j = \min\{S_j/S_{\max}, 1\} \quad (5)$$

GMDD
7, 6997–7031, 2014

where the empirical parameter S_{max} (°C) determines the value of S_j at which the temperature modifier attains its saturating level.

Following Landsberg and Waring (1997) the Vapour Pressure Deficit (VPD) modifier was defined as

$$FD_j = e^{\kappa VPD_j} \quad (6)$$

where VPD_j (kPa) is VPD in day j and κ (kPa^{-1}) is an empirical parameter assuming typically negative values.

2.2 Data

The data for the Italian Eddy Covariance site of Lavarone for the year 2004 has been downloaded from the European Fluxes Database Cluster (www.europe-fluxdata.eu). Lavarone is a ca. 130 years old alpine coniferous forest, dominated by Silver fir (*Abies alba* Mill.) and Norway spruce (*Picea abies* (L.) Karst.), with minor presence of European beech (*Fagus sylvatica* L.) and located at 1350 m a.s.l. in the Trento province, western Italian Alps. The site characteristics of Lavarone are described in detail in Rodeghiero and Cescatti (2005).

Daily air temperature, relative humidity (RH) and PAR were used as input data. Daily VPD was calculated from RH and air temperature following Allen et al. (1998), while daily GPP was used to calibrate the model. Daily APAR was calculated following Mäkelä et al. (2008a), using Normalized Difference Vegetation Index (NDVI) data as a proxy for fAPAR: for that purpose, NDVI data with 0.25 km spatial grid and 16 days time-step were downloaded from the MODIS repository (http://daac.ornl.gov/cgi-bin/MODIS/GLBVIZ_1_Glb/modis_subset_order_global_col5.pl).

Missing data for either a weather variable or GPP resulted in a missing outcome of the model for that day j. Therefore, 292 data points were actually used to calibrate the model.

The Bayesian calibration requires an estimate of the uncertainties around the data used during the calibration (van Oijen et al., 2005). Uncertainties around GPP_j were

GMDD
7, 6997–7031, 2014

calculated as follows:

$$\mathrm{GPP}_j = \mathrm{GPP}_j \pm y_j \quad (7)$$

where y_j is the maximum of $1\,\mathrm{g\,C\,m^{-2}}$ and a random value in the interval $[1, 0.3 \times \mathrm{GPP}_j]$. The lower bound of $1\,\mathrm{g\,C\,m^{-2}}$ is necessary to ensure that low values of GPP_j would not get an overwhelming weight during the calibration procedure.

2.3 Bayesian calibration

The prior parameter distributions were set based on the information made available by (Mäkelä et al., 2008a) and Peltoniemi et al. (2012). Since the parameter distributions were partly unknown, and since many parameters are empirical and without physiological meaning, we set the prior distributions as uniform distributions (i.e. any value has the same probability to occur) and wide enough to cover a very wide range of possible values (Table 1).

To investigate in detail the model behavior during a Bayesian calibration and to tackle the issues related to slow convergence, we tested four different procedures:

1. *Single-step calibration:* for each of the three algorithms applied (MHRW, AM, DEMC) different simulations with an increasing number of iterations were performed to test the efficiency of each algorithm in reaching convergence. Three simulations were run, with 10^4, 10^5 and 10^6 iterations in total for each algorithm. An initial burn-in phase was set to 30 % of the total number of iterations for all the algorithms. For the DEMC algorithm, 100 chains were considered, making the number of iterations per chain respectively 10^2, 10^3 and 10^4.

2. *Model comparison:* we ran a second LUE-based model with slightly different structure, on the same data and with the same calibration settings. The model chosen for this purpose was the model developed by Horn and Schulz (2011a) as described in Horn and Schulz (2011b). An initial burn-in phase was set to 30 % of

GMDD
7, 6997–7031, 2014

the total number of iterations for all the algorithms. For the DEMC algorithm, 100 chains were considered, making the number of iterations per chain respectively 10^2, 10^3 and 10^4. The main difference in the mathematical structure between the two models is that while in Prelued GPP is calculated as a product of potential LUE, APAR, and modifiers (Eq. 1), in Horn and Schulz (2011b) GPP is calculated as:

$$\mathrm{GPP}_j = \mathrm{LUE}\,\mathrm{APAR}_j[p\mathrm{FT}_j + (1-p)\mathrm{FW}_j] \tag{8}$$

with GPP_j ($\mathrm{g\,C\,m^{-2}}$) denoting the gross flux of carbon uptake in day j, LUE ($\mathrm{g\,C\,MJ^{-1}}$) being the maximum attained Light Use Efficiency, APAR_j ($\mathrm{MJ\,m^{-2}}$) the Absorbed Photosynthetically Active Radiation in day j, and p a weighting factor for the modifiers FT and FW.

FT is a sigmoidal peak function defined as:

$$\mathrm{FT} = 4e^{-(T_\mathrm{s}-T_\mathrm{opt})/kT}/(1+e^{-(T_\mathrm{s}-T_\mathrm{opt})/kT})^2 \tag{9}$$

where T_s is the soil temperature (°C), T_opt (°C) is the temperature at which the light use efficiency is maximum, and kT ($^\circ\mathrm{C}^{-1}$) is the rate of change from the lower level of FT to its maximum.

FW is defined as following sigmoidal function:

$$\mathrm{FW} = 1/(1+e^{kW(\mathrm{W}-Wi)}) \tag{10}$$

where W is a moisture surrogate (in our case the Soil Water Content ($\mathrm{m^3\,m^{-3}}$)), kW is the constant rate of change between lower and upper level (set to −13.1 following Horn and Schulz, 2011b) and Wi is the inflection point with units depending on the choice of W.

Following Jarvis et al. (2004), a lag function was applied to T_s:

$$\mathrm{ZF}_j = (1-\alpha)T_{\mathrm{s}j} + \alpha\mathrm{ZF}_{j-1} \tag{11}$$

GMDD
7, 6997–7031, 2014

where α (–) is the lag parameter. Equation (11) is only applied to T_s, considered the dominant driver of the vegetation stands; this main driver is expected to trigger the start and end of dormant periods after which the vegetation has to regenerate and redevelop green tissue (Horn and Schulz, 2011a).

FT and FW are scaled between 0 and 1 and describe the dependence of the Light Use Efficiency on the soil temperature and a moisture surrogate. Overall, the structure of this latter model is less multiplicative than Prelued, which should make its calibration easier. The prior distributions for this model have been derived from Table 2 in Horn and Schulz (2011a), using the minimum and maximum value for each parameter as boundaries and keeping the distributions uniform.

3. *Reparameterization:* we reformulated four parameters of Prelued out of six, changing their meaning and the model formulation accordingly:

$$\beta' = \beta/\gamma \quad (12)$$

$$\gamma' = 1/\gamma \quad (13)$$

$$S_{max}' = 1/S_{max} \quad (14)$$

$$X_0' = X_0/S_{max} \quad (15)$$

Given the purpose of this approach was reaching faster the convergence only two calibrations instead of three (10^4 and 10^5 iterations) were performed.

4. *Two-steps calibration:* in this procedure, the posterior correlations between parameters found in the first step were used to reduce the number of parameters involved in the second step. If two parameters were strongly correlated, the coefficients of the linear regression between them were used to estimate one parameter as a function of the second, reducing the number of calibrated parameters. These coefficients were calculated for each number of iterations, and used in a second calibration with the same length: the coefficients calculated after the 10^4 iterations calibration were used for a second 10^4 iterations calibration with reduced number

GMDD
7, 6997–7031, 2014

of parameters, and the same approach was used for the 10^5 iterations calibration. Since convergence was reached after the first step for the longest calibration, only two second steps were run (10^4 and 10^5 iterations).

2.4 Measure of convergence

The reaching of the convergence region was visually assessed, along with the different behavior of the Markov Chain between different numbers of iterations and their similar behavior between algorithms. In order to obtain a quantitative measure of convergence of the chains to the posterior distribution, the last 50 % of the longest chain for each algorithm were split in half and the means and variances of the two halves were compared. For the DEMC algorithm, only the chain with maximum log-likelihood was chosen for this purpose.

3 Results

3.1 Bayesian calibration

3.1.1 Single-step calibration

For all the three algorithms of increasing complexity used in this study (MHRW, AM, DEMC), the Markov Chain Monte Carlo did not reach convergence at 10^4 iterations, approached convergence at 10^5 iterations, and reached good convergence at 10^6 iterations (Fig. 1). The same pattern emerged from the analysis of the posterior distributions: for many parameters, the posterior distributions were bimodal, shifted, or as broad as the priors at 10^4 iterations, while becoming leptokurtic at 10^6 iterations for all the parameters. With the latter number of iterations, the posterior distribution thus narrowed the parameter space, converging in the same region (Fig. 2). The parameter sets with best log-likelihood (Table 2) and the posterior correlation coefficients between parameters (Table 3) were very similar between algorithms with only few exceptions.

GMDD
7, 6997–7031, 2014

This confirmed the convergence on the same joint posterior distribution and not only on the marginal distributions for each parameter. The strongest correlation was found for the threshold value of the delayed temperature X_0 and the empirical parameter S_{max} (correlation coefficient varies from −0.923 to −0.928 depending on the algorithm), both involved in the response to temperature. Strong correlation existed also between the Potential LUE, β, and the empirical parameter γ (correlation coefficient varies from 0.89 to 0.91 depending on the algorithm), which were both involved in the response to APAR. Concerning the log-likelihood values of the best parameter set, the MHRW algorithm showed the best result compared to the AM and the DEMC (Table 2).

3.1.2 Model comparison

The application of the less multiplicative LUE-based model developed by Horn and Schulz (2011b) to the same dataset did not show better results compared to Prelued, in terms of reaching proper convergence, even at a high number of iterations. For all the three algorithms of increasing complexity used in this study (MHRW, AM, DEMC), the Markov Chain Monte Carlo did not reach convergence at 10^4 and 10^5 iterations, and reached convergence at 10^6 iterations for some parameters only (Fig. 3). The analysis of the posterior distributions showed the same trends as in Prelued: for many parameters, the posterior distributions were bimodal, shifted, or as broad as the priors at 10^4 iterations, while narrowing the parameter space at 10^6 iterations and converging in the same region (Fig. 4).

3.1.3 Reparameterization

The alternative formulation proposed to overcome the convergence problems with the calibration of Prelued did not result in faster convergence. For all the three algorithms of increasing complexity used in this study (MHRW, AM, DEMC), the Markov Chain Monte Carlo did not reach convergence at 10^4 and 10^5 iterations (Fig. 5). The analysis of the posterior distributions showed a situation far from convergence in every case:

GMDD
7, 6997–7031, 2014

for many parameters, the posterior distributions were bimodal, shifted, or as broad as the priors at both 10^4 and 10^5 iterations, sometimes exploring different regions of the parameter space (Fig. 6).

3.1.4 Two-steps calibration

The posterior correlations found in the first step of calibration described in paragraph 3.1 (parameters β and γ, and parameters τ and S_{max}) were used to reduce the number of parameters estimated in the second step. In particular, γ was estimated as a linear function of β, and S_{max} as a linear function of τ. For all the three algorithms of increasing complexity used in this study (MHRW, AM, DEMC), the Markov Chain Monte Carlo did not reach convergence at 10^4 iterations, but reached good convergence at 10^5 iterations for all the parameters (Fig. 7). The analysis of the posterior distributions showed well defined distributions at both 10^4 and 10^5 iterations, and in the latter case they converged in the same region of the parameter space (Fig. 8).

3.2 Quantitative measure of convergence

For the single-step calibration with 10^6 iterations the means of the first- and second-halves of the MCMC were within 1.5 % of each other except for parameter τ (time constant of the delay process in response to temperature) using the DEMC algorithm (2.9 %). The variances of the first- and second-halves were within 15 % except for parameter X_0 (i.e., the threshold value of the delayed temperature) using the DEMC algorithm (23.2 %) (Table 4). This proves that the longest chains for each parameter and algorithm are converging on the same region of parameter space (Fig. 9).

4 Discussion

In disagreement with the expectation, given their increasing complexity and efficiency, all three MCMC-methods tested in this study were similarly effective. Although this

GMDD
7, 6997–7031, 2014

similarity in behavior between algorithms was a surprising result, the main output of this study was that a very high number of iterations was required for each of the three calibration algorithms to stabilize in the convergence region. This is especially true considering the simplicity of the Prelued model. This 6-parameters empirical model required 10^6 iterations to reach convergence, whereas a 39-parameter mechanistic forest model was calibrated with chains of length 10^5 (van Oijen et al., 2005), and 10^5 iterations were enough to allow proper convergence for 4 process-based models with higher complexity than Prelued (van Oijen et al., 2011). Despite the high number of iterations required, all three algorithms reached convergence since they all explored the same parameter space (Fig. 3). This demonstrates that the three algorithms were not attracted in different regions of the parameter space, which excludes the risk of undiagnosed slow convergence (Gilks et al., 1996).

The LUE model by Horn and Schulz (2011b) showed the same convergence problems as Prelued when calibrated with a Bayesian approach, despite a less multiplicative structure (Fig. 3). Therefore, the comparison of these two models suggested that the multiplicative structure of Prelued was probably the main factor responsible for the difficulties in the calibration, but is unlikely to be the only one. It should also be considered that even if LUE-type models are largely empirical, they usually rely on parameters with physiological meaning. Their analysis thus gives insights about the ecosystem characteristics and behavior, and allows for comparison between different models. For example, the well-known and widely applied 3PG model (Landsberg and Waring, 1997) has the same mathematical properties of Prelued, even if not so extreme. Therefore, beside the strong multiplicative mathematical structure, the problems in calibrating Prelued were likely due to the indefinite nature of the empirical parameters, neither ecological nor physiological. This renders the prior distributions difficult to specify and forces the MCMC to investigate a broad parameter space, delaying the identification of the convergence region.

The reparameterization procedure applied to Prelued in order to reach faster convergence proved to be ineffective (Fig. 5). This result should not be surprising given

GMDD
7, 6997–7031, 2014

the simple mathematical formulation of the model, which does not allow the users to considerably change the parameter meaning and the model structure. Even if our approach allowed to reformulate four parameters out of six, this change in the parameters formulation did not lead to a substantial change in the overall model formulation, and this is likely to be the main reason of the ineffectiveness of this kind of procedure in this particular case.

It is not uncommon for data-based modeling exercises to show issues related to equifinality: frequently, the optimal parameter set is not uniquely defined. Instead, there may be many sets of parameters that all fit the data more or less equally well (Franks and Beven, 1997; Hollinger and Richardson, 2005; Schulz et al., 2001). This usually results in a delayed convergence, and can be due to high posterior correlation between parameters. These correlations could also be due to model overparameterization, which is known to lead to slow convergence (Rannala, 2002). An alternative solution to the issue of slow convergence was a two-steps calibration, using the posterior correlations between parameters resulting from the first step to reduce the number of parameters calibrated in the second step (thus reducing the dimensionality of the parameter space): this procedure allowed to reduce the number of parameters estimated, tackling both the issue of overparameterization and of equifinality, and reaching convergence with a smaller number of iterations (Fig. 7).

Since it was shown to be the same, the efficiency of the three considered algorithms should not drive their choice. The MHRW provided the parameter vector with best log-likelihood, but this did not result in better model performances over all. We suggest the DEMC algorithm as the best choice in this case study, due to the automatic computation of both the scale and orientation of the MCMC sampling. These are both user-defined in the MHRW algorithm, while only orientation is internally computed in the AM leaving scale as a user-defined setting. Since the optimal combination of scale and orientation is dependent on the prior distributions and on the data, the user might need several attempts to find it, making the calibration process even more time-consuming. It is also important to note that, once the optimal settings have been

GMDD
7, 6997–7031, 2014

decided, the computational effort was the same for all the algorithms, even if other studies suggest that the DEMC algorithm is slower and requires more computational power than the others used in this case study (ter Braak, 2006).

5 Conclusions

In this study we compared the performance of three different Markov Chain Monte Carlo-based algorithms within a Bayesian framework to calibrate a Light Use Efficiency model (Prelued). The application of the three different algorithms of increasing complexity (Metropolis–Hastings Random Walk, Adaptive Metropolis, Differential Evolution Markov Chain) with different number of iterations showed that all three MCMC-methods were similarly effective in reaching convergence. For all of them a very high number of iterations (10^6) is required for the Markov Chain to stabilize in the convergence region. This is due to the combination of at least two different factors: a strongly multiplicative mathematical structure, coupled with empirical parameters with neither ecological nor physiological meaning and parameters about which little is known in the literature.

An alternative solution to a very high number of iterations was a two-steps calibration, using the posterior correlations between parameters resulting from the first step to reduce the number of parameters calibrated in the second step. This approach reduced the computational effort necessary to reach proper convergence and was less time consuming than the previous one.

We suggest the DEMC algorithm as the best choice in this case study, even if its efficiency has proven to be similar to the other algorithms used, due to the automatic computation of both the scale and orientation of the MCMC sampling.

Acknowledgements. Maurizio Bagnara's Ph.D. fellowship "AM07 – Forest Modelling" was funded by the FIRST FEM International Ph.D. School Trentino. We thank Mauro Cavagna and Roberto Zampedri for maintaining the instrumentation, Francesco Minunno, University of Lisbon; for providing the R code for DEMC algorithm, Jeroen Pullens for the comments on an

GMDD
7, 6997–7031, 2014

earlier versions of the draft. Matteo Sottocornola acknowledges funding by the Marie-Curie FP7 – PCOFUND-GA-2008-226070, “Progetto Trentino”, CfPAT project.

References

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration guidelines for computing crop water requirements, Irrig. Drain. Pap. 65, 300 pp., United Nations Food Agric. Organ., Rome, Italy, 1998.

Bagnara, M., Sottocornola, M., Cescatti, A., Minerbi, S., Montagnani, L., Gianelle, D., and Magnani, F.: Bayesian optimization of a light use efficiency model for the estimation of daily gross primary productivity in a range of Italian forest ecosystems, Ecol. Model., doi:10.1016/j.ecolmodel.2014.09.021, in press, 2014.

Baldocchi, D. D.: “Breathing” of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems, Aust. J. Bot., 56, 1–26, doi:10.1071/BT07151, 2008.

Buzzi-Ferraris, G. and Manenti, F.: Better reformulation of kinetic models, Comput. Chem. Eng., 34, 1904–1906, 2010.

Chevallier, F., Viovy, N., Reichstein, M., and Ciais, P.: On the assignment of prior errors in Bayesian inversions of CO_2 surface fluxes, Geophys. Res. Lett., 33, 1–5, doi:10.1029/2006GL026496, 2006.

De Weirdt, M., Verbeeck, H., Maignan, F., Peylin, P., Poulter, B., Bonal, D., Ciais, P., and Steppe, K.: Seasonal leaf dynamics for tropical evergreen forests in a process-based global ecosystem model, Geosci. Model Dev., 5, 1091–1108, doi:10.5194/gmd-5-1091-2012, 2012.

Franks, S. W. and Beven, K. J.: Bayesian estimation of uncertainty in land surface–atmosphere flux predictions, J. Geophys. Res., 102, 23991–23999, 1997.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J.: Markov Chain Monte Carlo in Practice, CRC press, London, 1996.

Haario, H., Saksman, E., and Tamminen, J.: An adaptive Metropolis algorithm, Bernoulli, 7, 223–242, 2001.

Hollinger, D. Y. and Richardson, A. D.: Uncertainty in eddy covariance measurements and its application to physiological models, Tree Physiol., 25, 873–885, doi:10.1093/treephys/25.7.873, 2005.

GMDD

7, 6997–7031, 2014

Horn, J. E. and Schulz, K.: Identification of a general light use efficiency model for gross primary production, Biogeosciences, 8, 999–1021, doi:10.5194/bg-8-999-2011, 2011a.
Horn, J. E. and Schulz, K.: Spatial extrapolation of light use efficiency model parameters to predict gross primary production, J. Adv. Model. Earth Syst., 3, M12001, doi:10.1029/2011MS000070, 2011b.
Jarvis, A. J., Stauch, V. J., Schulz, K., and Young, P. C.: The seasonal temperature dependency of photosynthesis and respiration in two deciduous forests, Glob. Change Biol., 10, 939–950, doi:10.1111/j.1529-8817.2003.00743.x, 2004.
Kimmins, J. P.: Predicting the consequences of intensive forest harvesting on long-term productivity: the need for a hybrid model such as FORCYTE-11, Rapp. Institutionen Ekol. och Miljovard, Sveriges Lantbruksuniversitet, 26, 31–84, 1986.
Kimmins, J. P., Mailly, D., and Seely, B.: Modelling forest ecosystem net primary production: the hybrid simulation approach used in forecast, Ecol. Model., 122, 195–224, doi:10.1016/S0304-3800(99)00138-6, 1999.
Kimmins, J. P., Blanco, J. A., Seely, B., Welham, C., and Scoullar, K.: Complexity in modelling forest ecosystems: how much is enough?, Forest Ecol. Manag., 256, 1646–1658, doi:10.1016/j.foreco.2008.03.011, 2008.
Landsberg, J. and Waring, R. H.: A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning, Forest Ecol. Manag., 95, 209–228, doi:10.1016/S0378-1127(97)00026-1, 1997.
Mäkelä, A., Landsberg, J., Ek, A. R., Burk, T. E., Ter-Mikaelian, M., Agren, G. I., Oliver, C. D., and Puttonen, P.: Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation, Tree Physiol., 20, 289–298, 2000.
Mäkelä, A., Hari, P., Berninger, F., Hänninen, H., and Nikinmaa, E.: Acclimation of photosynthetic capacity in Scots pine to the annual cycle of temperature, Tree Physiol., 24, 369–376, doi:10.1093/treephys/24.4.369, 2004.
Mäkelä, A., Pulkkinen, M., Kolari, P., Lagergren, F., Berbigier, P., Lindroth, A., Loustau, D., Nikinmaa, E., Vesala, T., and Hari, P.: Developing an empirical model of stand GPP with the LUE approach: analysis of eddy covariance data at five contrasting conifer sites in Europe, Glob. Change Biol., 14, 92–108, doi:10.1111/j.1365-2486.2007.01463.x, 2008.
McCallum, I., Franklin, O., Moltchanova, E., Merbold, L., Schmullius, C., Shvidenko, A., Schepaschenko, D., and Fritz, S.: Improved light and temperature responses for light-use-efficiency-based GPP models, Biogeosciences, 10, 6577–6590, doi:10.5194/bg-10-6577-2013, 2013.

GMDD
7, 6997–7031, 2014

McMurtrie, R. E., Gholz, H. L., Linder, S., and Gower, S. T.: Climatic factors controlling the productivity of pine stands?: a model-based analysis, Ecol. Bull., 43, 173–188, 1994.
Medlyn, B., Barrett, D., Landsberg, J., Sands, P., and Clement, R.: Conversion of canopy intercepted radiation to photosynthate: a review of modelling approaches for regional scales, Funct. Plant Biol., 30, 153–169, doi:10.1071/FP02088_CO, 2003.
Peltoniemi, M., Pulkkinen, M., Kolari, P., Duursma, R. A., Montagnani, L., Wharton, S., Lagergren, F., Takagi, K., Verbeeck, H., and Christensen, T.: Does canopy mean nitrogen concentration explain variation in canopy light use efficiency across 14 contrasting forest sites?, Tree Physiol., 32, 200–218, doi:10.1093/treephys/tpr140, 2012.
Rannala, B.: Identifiability of parameters in MCMC Bayesian inference of phylogeny, Syst. Biol., 51, 754–760, doi:10.1080/10635150290102429, 2002.
Robert, C. P. and Casella, G. : Monte Carlo Statistical Methods, Springer, New York, 1999.
Rodeghiero, M. and Cescatti, A.: Main determinants of forest soil respiration along an elevation/temperature gradient in the Italian Alps, Glob. Change Biol., 11, 1024–1041, doi:10.1111/j.1365-2486.2005.00963.x, 2005.
Schulz, K., Jarvis, A. J., Beven, K. J., and Soegaard, H.: The predictive uncertainty of land surface fluxes in response to increasing ambient carbon dioxide, J. Climate, 14, 2551–2562, doi:10.1175/1520-0442(2001)014<2551:TPUOLS>2.0.CO;2, 2001.
Sivia, D. S.: Data Analysis: a Bayesian Tutorial, Oxford University Press, Oxford, 1996.
Smith, T. J. and Marshall, L. A.: Bayesian methods in hydrologic modeling?: a study of recent advancements in Markov chain Monte Carlo techniques, 44, 1–9, doi:10.1029/2007WR006705, 2008.
Still, C. J., Randerson, J. T., and Fung, I. Y.: Large-scale plant light-use efficiency inferred from the seasonal cycle of atmospheric CO_2, Glob. Change Biol., 10, 1240–1252, doi:10.1111/j.1365-2486.2004.00802.x, 2004.
Storn, R. and Price, K.: Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces, J. Global Optim., 11, 341–359, doi:10.1023/A:1008202821328, 1997.
Svensson, M., Jansson, P., Gustafsson, D., Kleja, D., Langvall, O., and Lindroth, A.: Bayesian calibration of a model describing carbon, water and heat fluxes for a Swedish boreal forest stand, Ecol. Model., 213, 331–344, doi:10.1016/j.ecolmodel.2008.01.001, 2008.

GMDD
7, 6997–7031, 2014

ter Braak, C. J. F.: A Markov chain Monte Carlo version of the genetic algorithm differential evolution: easy Bayesian computing for real parameter spaces., Stat. Comput., 16, 239–249, doi:10.1007/s11222-006-8769-1, 2006.

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., Seland, Ø., and Heinze, C.: Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM), Geosci. Model Dev., 6, 301–325, doi:10.5194/gmd-6-301-2013, 2013.

Turner, D. P., Urbanski, S., Bremer, D., Wofsy, S. C., Meyers, T., Gower, S. T., and Gregory, M.: A cross-biome comparison of daily light use efficiency for gross primary production, Glob. Change Biol., 9, 383–395, doi:10.1046/j.1365-2486.2003.00573.x, 2003.

Van Oijen, M., Rougier, J., and Smith, R.: Bayesian calibration of process-based forest models: bridging the gap between models and data, Tree Physiol., 25, 915–27, doi:10.1093/treephys/25.7.915, 2005.

Van Oijen, M., Cameron, D. R., Butterbach-Bahl, K., Farahbakhshazad, N., Jansson, P.-E., Kiese, R., Rahn, K.-H., Werner, C., and Yeluripati, J. B.: A Bayesian framework for model calibration, comparison and analysis: application to four models for the biogeochemistry of a Norway spruce forest, Agr. Forest Meteorol., 151, 1609–1621, doi:10.1016/j.agrformet.2011.06.017, 2011.

Veroustraete, F., Patyn, J., and Myneni, R. B.: Forcing of a simple ecosystem model with fAPAR and climatic data to estimate regional scale photosynthetic assimilation, edited by: Veroustraete F. and Ceulemans, R., VGT, modelling and climate change effects, 151–177, The Hague, the Netherlands, Academic Publishing, 1994.

White, J. D. and Running, S. W.: Testing scale dependent assumptions in regional ecosystem simulations, J. Veg. Sci., 5, 687–702, doi:10.2307/3235883, 1994.

Wißkirchen, K., Tum, M., Günther, K. P., Niklaus, M., Eisfelder, C., and Knorr, W.: Quantifying the carbon uptake by vegetation for Europe on a 1 km^2 resolution using a remote sensing driven vegetation model, Geosci. Model Dev., 6, 1623–1640, doi:10.5194/gmd-6-1623-2013, 2013.

Xenakis, G., Ray, D., and Mencuccini, M.: Sensitivity and uncertainty analysis from a coupled 3-PG and soil organic matter decomposition model, Ecol. Model., 219, 1–16, doi:10.1016/j.ecolmodel.2008.07.020, 2008.

Zhang, X. and Kondragunta, S.: Estimating forest biomass in the USA using generalized allometric models and MODIS land products, Geophys. Res. Lett., 33, 1–5, 2006.

GMDD
7, 6997–7031, 2014

Zhu, G. F., Li, X., Su, Y. H., Zhang, K., Bai, Y., Ma, J. Z., Li, C. B., Hu, X. L., and He, J. H.: Simultaneously assimilating multivariate data sets into the two-source evapotranspiration model by Bayesian approach: application to spring maize in an arid region of northwestern China, Geosci. Model Dev., 7, 1467–1482, doi:10.5194/gmd-7-1467-2014, 2014.

GMDD
7, 6997–7031, 2014

Table 1. Prior probability distributions for each parameter in the Prelued model.

Parameter	Unit	Prior min.	Prior max.
β	$g\,C\,mol^{-1}$	0.0	1.5
γ	$m^2\,mol^{-1}$	0.0	0.1
κ	kPa^{-1}	−10.0	0.0
X_0	°C	−100.0	0.0
τ	°C	0.0	100.0
S_{max}	°C	0.0	100.0

GMDD
7, 6997–7031, 2014

Table 2. Best parameter sets and log-likelihood values for the three MCMC algorithms applied (single-step calibration of Prelued), compared with the optimized parameter values found by Mäkelä et al. (2008) in similar ecosystems.

Site	Year	Dominant species	Algorithm	Best parameter vector/ Optimized parameter value						Log-likelihood	Reference
				β	γ	κ	X_0	τ	S_{max}		
Lavarone	2004	*A. alba* + *P. abies*	MHRW	0.55	0.02	−0.92	−7.01	9.51	13.28	−117.78	–
			AM	0.56	0.02	−0.93	−6.89	9.19	12.91	−124.41	
			DEMC	0.56	0.02	−0.93	−6.60	9.52	12.21	−134.14	
Norunda	1999	*P. abies* + *P. sylvestris*	–	0.49	0.002	−0.39	−10.0	5.0	29.0	–	Mäkelä et al. (2008)
Tharandt	2003	*P. abies*		0.66	0.016	−0.70	−5.0	2.0	19.50		
Bray	2001	*P. pinaster*		0.49	0.021	−0.06	−1.0	2.0	19.0		

GMDD
7, 6997–7031, 2014

Table 3. Posterior coefficients of correlation between parameters (single-step calibration of Prelued). Coefficients that differ more than one order of magnitude or have different sign between algorithms are highlighted in bold text.

Algorithm	Parameter	β	γ	κ	X_0	τ	S_{max}
MHRW	$\boldsymbol{\beta}$	1	0.91	0.135	−0.15	−0.262	0.369
AM		1	0.89	**0.039**	−0.095	−0.269	0.294
DEMC		1	0.896	0.156	−0.106	−0.257	0.291
MHRW	$\boldsymbol{\gamma}$	0.91	1	0.471	−0.13	−0.226	0.325
AM		0.89	1	0.417	−0.106	−0.218	0.272
DEMC		0.896	1	0.512	−0.116	−0.263	0.27
MHRW	$\boldsymbol{\kappa}$	0.135	0.471	1	**0.006**	0.012	0.072
AM		**0.039**	0.417	1	−0.021	0.067	0.062
DEMC		0.156	0.512	1	−0.017	**−0.07**	0.081
MHRW	$\boldsymbol{X_0}$	−0.15	−0.13	**0.006**	1	0.434	−0.923
AM		−0.095	−0.106	−0.021	1	0.483	−0.928
DEMC		−0.106	−0.116	−0.017	1	0.418	−0.926
MHRW	$\boldsymbol{\tau}$	−0.262	−0.226	0.012	0.434	1	−0.512
AM		−0.269	−0.218	0.067	0.483	1	−0.578
DEMC		−0.257	−0.263	**−0.07**	0.418	1	−0.529
MHRW	$\boldsymbol{S_{max}}$	0.369	0.325	0.072	−0.923	−0.512	1
AM		0.294	0.272	0.062	−0.928	−0.578	1
DEMC		0.291	0.27	0.081	−0.926	−0.529	1

GMDD
7, 6997–7031, 2014

Table 4. Means and variances of the first and second half of last 50 % of the longest chain for each algorithm (single-step calibration of Prelued).

Algorithm	Parameter	Mean of first half	Mean of second half	Variance of first half	Variance of second half
MHRW	β	0.55036	0.54845	0.0011	0.00116
	γ	0.01799	0.01779	0.00001	0.00001
	κ	0.924	0.92421	0.00816	0.00816
	X_0	6.96082	7.0221	0.54956	0.59826
	τ	9.60252	9.60196	1.96344	2.06421
	S_{max}	13.15563	13.24499	1.74995	1.92255
AM	β	0.56413	0.56479	0.00105	0.00101
	γ	0.01903	0.01898	0.00001	0.00001
	κ	0.92784	0.93448	0.00752	0.00758
	X_0	6.89452	6.90664	0.53949	0.53714
	τ	9.18685	9.07216	1.97048	1.82252
	S_{max}	12.93946	12.9531	1.55456	1.57706
DEMC	β	0.55799	0.55883	0.00083	0.00076
	γ	0.01849	0.0187	0.00001	0.00001
	κ	0.92787	0.92596	0.00675	0.00689
	X_0	6.7482	6.75782	0.45696	0.35082
	τ	9.54574	9.26533	1.35935	1.26707
	S_{max}	12.46423	12.52407	1.24835	1.08612

GMDD
7, 6997–7031, 2014

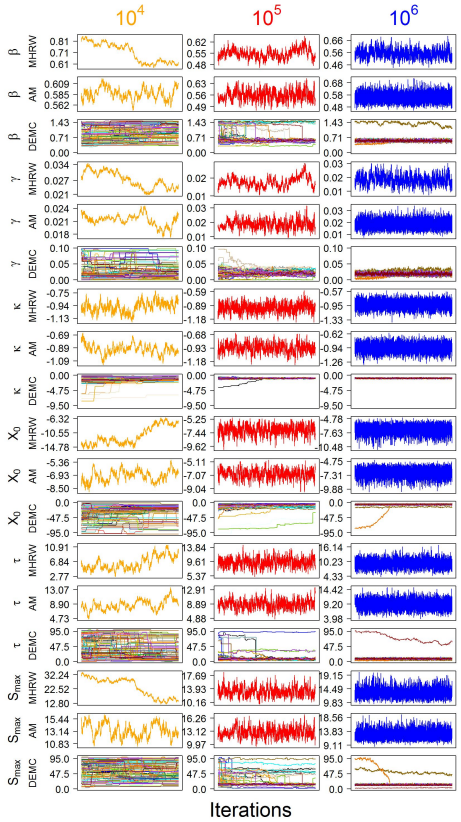

Figure 1. Traceplots of the post burn-in MCMC sampling for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the single-step calibration of the Prelued model.

GMDD
7, 6997–7031, 2014

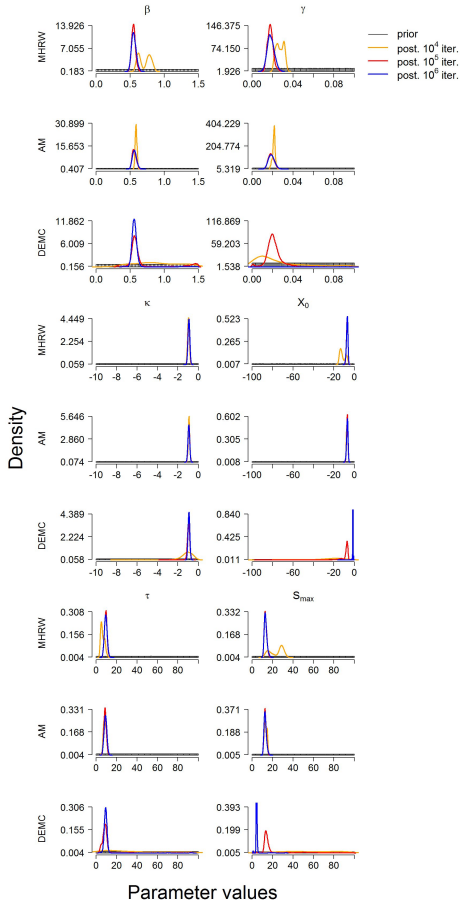

Figure 2. Posterior probability distributions of parameters for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the single-step calibration of the Prelued model.

GMDD
7, 6997–7031, 2014

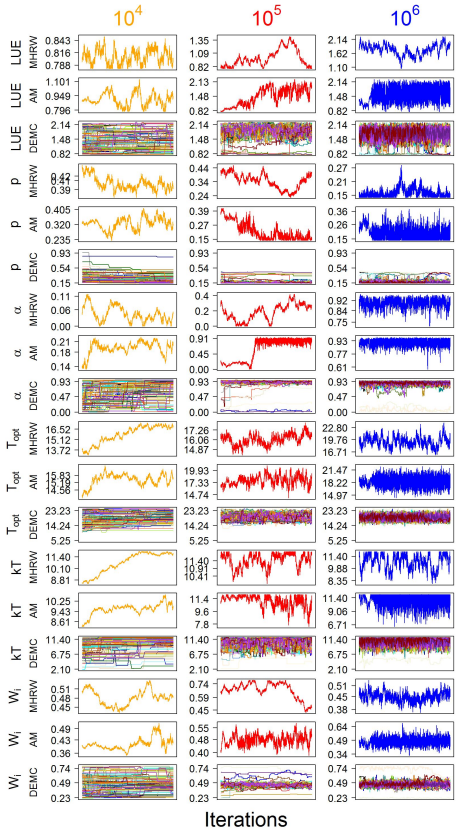

Figure 3. Traceplots of the post burn-in MCMC sampling for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the model by Horn and Schulz (2011b).

GMDD
7, 6997–7031, 2014

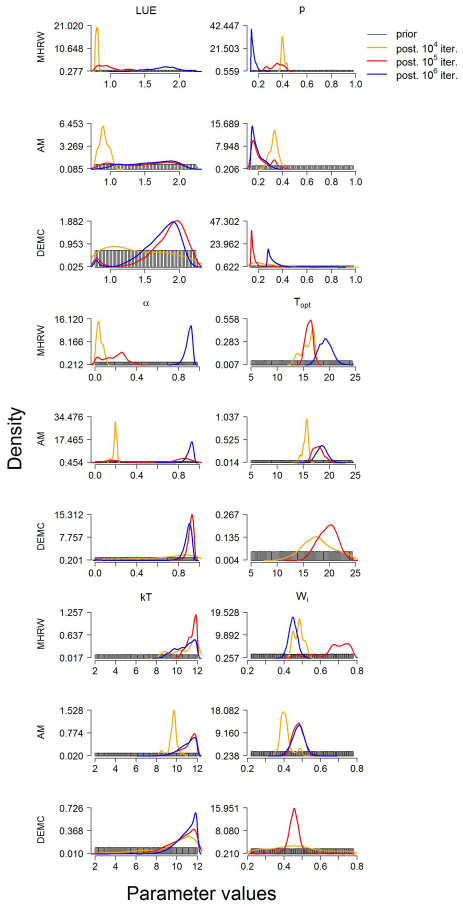

Figure 4. Posterior probability distributions of parameters for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the model by Horn and Schulz (2011b).

GMDD
7, 6997–7031, 2014

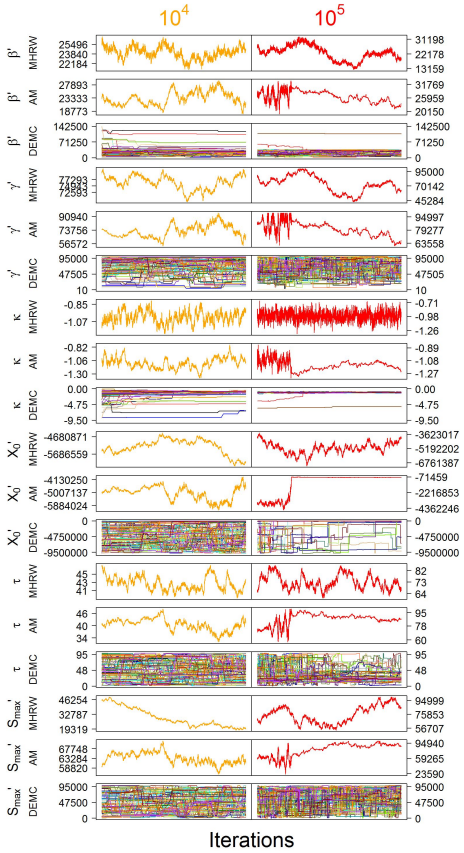

Figure 5. Traceplots of the post burn-in MCMC sampling for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the reparameterized Prelued model.

GMDD
7, 6997–7031, 2014

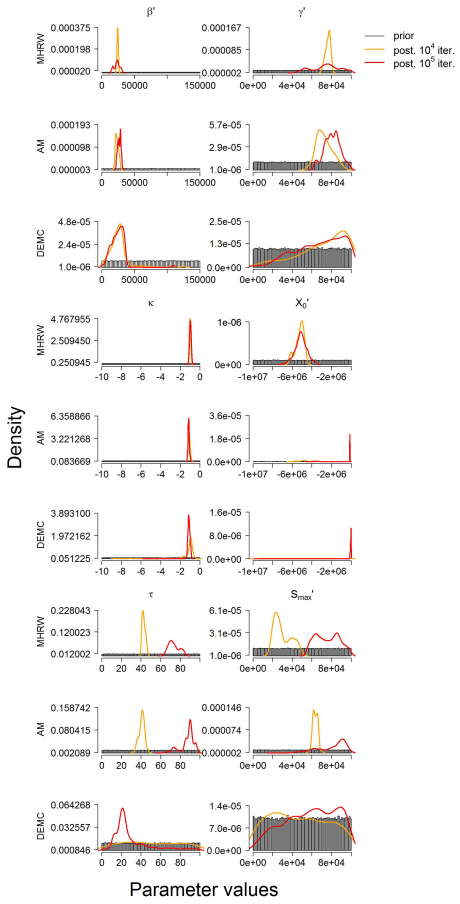

Figure 6. Posterior probability distributions of parameters for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the reparameterized Prelued model.

GMDD
7, 6997–7031, 2014

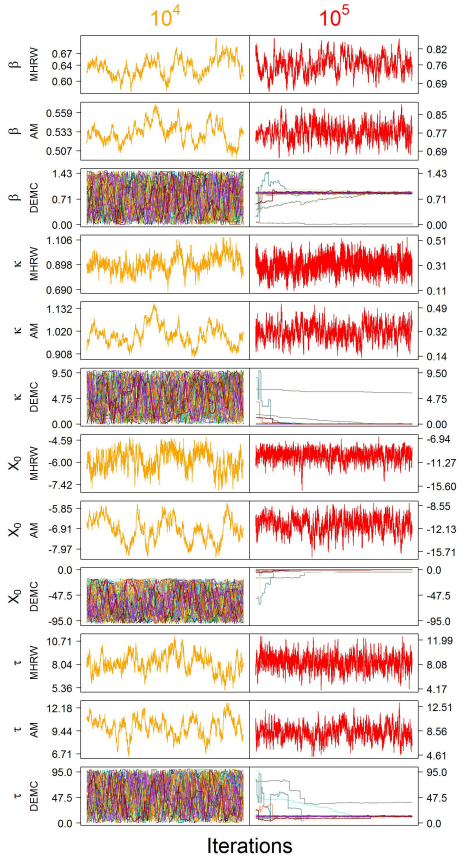

Figure 7. Traceplots of the post burn-in MCMC sampling for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the two-steps calibration of the Prelued model.

GMDD
7, 6997–7031, 2014

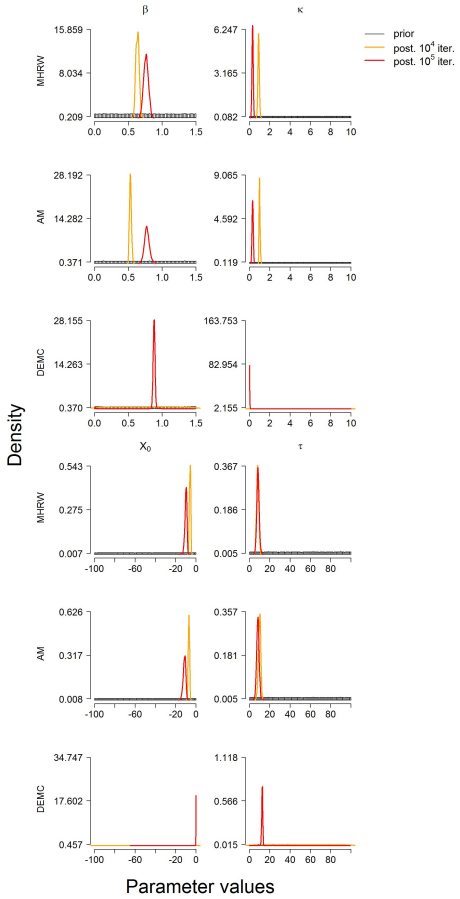

Figure 8. Posterior probability distributions of parameters for all the applied algorithms (MHRW, AM, DEMC) with different number of iterations, for the two-steps calibration of the Prelued model.

GMDD
7, 6997–7031, 2014

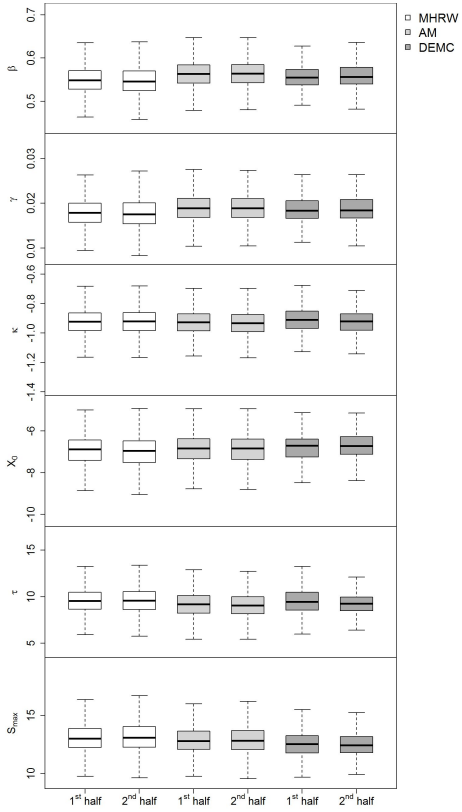

Figure 9. Boxplot of the first and second half of the last 50 % of the longest chain for each algorithm for the single-step calibration of the Prelued model.

GMDD
7, 6997–7031, 2014