Geosci. Model Dev. Discuss., 7, 6893–6948, 2014
www.geosci-model-dev-discuss.net/7/6893/2014/
doi:10.5194/gmdd-7-6893-2014

Geoscientific Model Development Discussions
Open Access

Global sensitivity analysis, probabilistic calibration, and predictive assessment for the Data Assimilation Linked Ecosystem Carbon model

C. Safta[1], D. Ricciuto[2], K. Sargsyan[1], B. Debusschere[1], H. N. Najm[1], M. Williams[3], and P. Thornton[2]

[1]Sandia National Labs, Livermore, CA 94551, USA
[2]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[3]School of GeoSciences and National Centre for Earth Observation, University of Edinburgh, EH9 EJN, UK

Received: 1 September 2014 – Accepted: 25 September 2014 – Published: 15 October 2014

Correspondence to: C. Safta (csafta@sandia.gov)

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

Abstract

In this paper we propose a probabilistic framework for an uncertainty quantification study of a carbon cycle model. A Global Sensitivity Analysis (GSA) study indicates the parameters and parameter couplings that are important at different times of the year for Quantities of Interest obtained with the Data Assimilation Linked Ecosystem Carbon (DALEC) model. We then employ a Bayesian approach to calibrate the parameters of DALEC using net ecosystem exchange observations at the Harvard Forest site. The calibration exercise is guided by GSA and by Fisher information matrix results that quantify the amount of information carried by the experimental data about specific model parameters. The calibration results are employed in the second part of the paper to assess the predictive skill of the model via posterior predictive checks. These checks show a better performance for the non-steady state model during the growing season compared to the one employing steady state assumptions. Overall, this study leads to a 40 % improvement in the predictive skill of DALEC and highlights the importance of considering correlations in the model parameters as informed by the data.

1 Introduction

Climate studies strongly depend on the modeling of the Carbon cycle. Carbon cycle models, in turn, strongly depend on the capability of current land models to simulate the terrestrial ecosystem and to capture C exchanges between land and atmosphere. There have been a significant number of studies looking to leverage the increasing amount of experimental observations and calibrate parameters in several terrestrial ecosystem models. These studies have faced a number of challenges related to handling data and measurement errors from multiple sources, formalizing model error, dealing with parameter observability and data sparsity, to name a few. In this paper we propose a probabilistic framework to estimate parameters for a process-based

GMDD
7, 6893–6948, 2014

ecosystem model. Representative studies, both probabilistic and non-probabilistic, are reviewed below.

Over the past two decades several studies employed data assimilation techniques to calibrate Carbon cycle models. Here we discuss the works that motivated the current study. Kaminski et al. (2002) used an adjoint approach to infer model parameters for a Simple Diagnostic Biosphere Model. The variational data assimilation problem was formulated based on Bayes formula with both the likelihood and the prior presumed Gaussian. This results in a quadratic cost function that employs an L_2 regularization of the model parameters. This formulation led to optimal values for model parameters. The width of the approximate Gaussian distributions around these optimal values was sensitive to the covariance matrices assumed in the cost function. More recently, Kaminski et al. (2012) employed a similar framework to calibrate the process parameters of a terrestrial biosphere model against two observational data streams. The model employing optimized parameters shows clear improvements when checked against independent observations compared to non-optimized parameters. A similar approach was applied by Rayner et al. (2005) to study the space–time distribution of terrestrial carbon fluxes generated by a terrestrial carbon cycle data assimilation system. Tjiputra et al. (2007) employed an adjoint approach to estimate optimal values for 10 ecosystem control variables in an ocean general circulation model coupled with a carbon cycle model. The optimization problem is based on a quadratic misfit between the simulated surface chlorophyll and observations. Kuppel et al. (2012) used measurements of net CO_2 fluxes (NEE) and latent heat fluxes (LE) to constrain the parameters of a biogeochemical vegetation model. The optimization employed an L-BFGS algorithm for a quadratic cost function similar to the study by Kaminski et al. (2002). They found that the simulation results are improved by using data from multiple sites, compared to single-site parameter optimization.

Some of the above studies start from a Bayesian framework when setting the cost function for a least-square fitting procedure. These studies are based on a Gaussian assumption for the discrepancy between model outputs and observations, and they

GMDD
7, 6893–6948, 2014

also employ Gaussian priors to help regularize the problem. However, the resulting probability densities for model parameters are approximated as multivariate Gaussian distributions near the Maximum a Posteriori (MAP) estimate of the parameter values. This assumption is valid only in the vicinity of MAP values, unless the model is linear in all parameters. In this paper we propose to employ a Bayesian framework to estimate parameters in the Data Assimilation Linked Ecosystem Carbon (DALEC) model (Williams et al., 2005), without relying on Gaussian assumptions for posterior distributions. Several studies in the past decade, some of which mentioned below, employed sampling techniques to explore non-Gaussian posterior distributions for parameters in ecosystem models.

Knorr and Kattge (2005) employed a Bayesian framework to calibrate the parameters of a Terrestrial Ecosystem Model (TEM). A Metropolis–Hastings Markov Chain Monte Carlo (MCMC) approach was used to sample the posterior distribution of model parameters given a Gaussian likelihood based on eddy covariance measurements of carbon and water fluxes. It was found that about 5 parameters were constrained by the available data and that uniform prior ranges had a strong impact on the posterior distributions. Braswell et al. (2005) performed a synthetic analysis of Net Ecosystem Exchange (NEE) of CO_2 at Harvard Forest using a simplified photosynthesis and evapotranspiration model. In a Bayesian framework, they employed independent Gaussian daily discrepancies between model predictions and observations. The posterior distributions for modeled parameters, sampled with MCMC, were compared for several synthetic data sets to determine how much information the NEE observations carry about each parameter. A Bayesian framework was also employed by Xu et al. (2006) to study the posterior distributions of C transfer coefficients and pool sizes in a TEM, based on several data sets from the Duke Forest Free-Air CO_2 site. Tang and Zhuang (2009) employed both Global Sensitivity Analysis (GSA) and a Bayesian framework to improve parameterization of a Terrestrial Ecosystem Model. This study employed Latin Hypercube Sampling from the prior distributions of model parameters, and sample importance resampling, to construct posterior distributions for model parameters, and to

GMDD
7, 6893–6948, 2014

identify key parameters for the ecosystem model and their effect on seasonal C dynamics. Ricciuto et al. (2008) employed an MCMC approach to sample the posterior densities of key parameters for combined global-scale terrestrial and ocean carbon cycle models. The study found that temporal correlation has a significant impact on the calibrated parameters and subsequently on model predictions.

Several studies compared probabilistic and non-probabilistic parameter estimation methods for terrestrial biogeochemical models. Several participants to the OptIC project (Trudinger et al., 2007) presented results employing optimization, variational, and probabilistic methods. The main conclusion of the study was that modeling choices, i.e. the type of cost function for optimization methods, or the choice of densities for probabilistic methods, had a greater impact on the results than the choice of solution methods. Similarly, the REFLEX project (Fox et al., 2009) selected the DALEC v1 model (Williams et al., 2005) to assess the performance of several parameter estimation algorithms, using both synthetic and observed NEE and LAI data. This study found that it is difficult to analyze the performance of parameter estimation methods in the presence of noisy and sparse data, and that all methodologies should employ uncertainty models that are consistent with observations. More recently, Ziehn et al. (2012) compared variational and probabilistic techniques to estimate parameters for BETHY, a process-based model of the terrestrial biosphere. It was found that the Gaussian approximation is reasonable for most parameters. This study also indicates that probabilistic approaches can be prohibitively expensive for complex ecosystem models.

From this review, we noted a set of critical outstanding research questions in the field of constraining C cycle models. First, few, if any, C cycle models have had a complete parameter sensitivity analysis, particularly with respect to temporal dynamics. Such analyses are vital for organising effective parameter calibration. Second, few, if any, calibration studies have investigated steady state/non-steady state assumptions. It is also important for the ecological community to understand how information content depends on model assumption, e.g. steady state. Currently, there are no agreed

GMDD
7, 6893–6948, 2014

approaches in this community for quantifying information content of data on parameters, or for estimating the predictive skill of ecosystem models.

In this paper we propose a framework for the estimation of uncertainties in ecosystem land model parameters followed by a forward Uncertainty Quantification (UQ) study to examine the predictive capabilities of the model given the calibrated set of parameters. Figure 1 shows a schematic of this framework, consisting of two intrinsically connected workflows, for *Parameter Estimation* and *Forward UQ*. In this schematic, the same ecosystem Carbon model is used for both the "Measurement Model" $g()$ and the "Computational Model" $m()$. The Carbon model is based on a modified version of the DALEC v1 model (Williams et al., 2005; Fox et al., 2009). This version of DALEC has been modified to facilitate comparisons with the Community Land Model (Thornton et al., 2007), and with the Local Terrestrial Ecosystem Carbon Model (Ricciuto et al., 2011).[1] The joint probability density for input parameters is estimated in a Bayesian framework. Bayesian methods provide a flexible framework for handling heterogeneous information, and allow for sequential updates of posterior distributions as the prior information is revised.

To facilitate the estimation of a high-dimensional posterior density for model parameters, we undertake parameter sensitivity tests using a variety of methods. First, parameters are ranked using Sobol indices (Sobol, 1993; Campolongo et al., 2000). Posterior densities are estimated first for the most important parameters, while less important parameters are fixed at their nominal values. This constraint is subsequently relaxed to arrive at a joint posterior distribution over the entire parameter space. Second, since the GSA does not consider the error model when ranking parameters, we complement the GSA results with an analysis of the Fisher Information Matrix (FIM) (Lehmann and Casella, 2003). The FIM results quantify the amount of information the experimental observations carry about the set of DALEC parameters for a particular setting for the discrepancy between model predictions and data. This study also allows

[1]The source code for the modified DALEC version is available upon request from Daniel Ricciuto (ricciutodm@ornl.gov)

GMDD
7, 6893–6948, 2014

an investigation of the information content of data based on steady state vs. non-steady state assumptions.

Finally, we undertake a Bayesian posterior predictive check (Lynch and Western, 2004) to assess the adequacy of the calibrated Carbon model to predict the experimental observations. The predictive skill of this model is further assessed via Continuous Rank Predictive Score (Gneiting and Raftery, 2007) computations.

This paper is organized as follows. Section 2 provides a description of the processes comprising DALEC and of their associated parameters. Section 3 presents the GSA results, including first order effects, in Sect. 3.1, and joint effects, in Sect. 3.2. FIM results and posterior distributions for model parameters are explored in Sect. 4 and the predictive capabilities are estimated in Sect. 5. We end with conclusions in Sect. 6. The methods employed in this paper are part of UQTk v3.0.[2]

2 Description of carbon cycle model

The schematic in Fig. 2 shows a 1 day time step consisting of a sequence of process-based submodels shown with green boxes. These submodels are connected via fluxes and interact with five major Carbon (C) pools. The fluxes calculated on any given day impact C pools and processes in subsequent days. The blue arrows in this figure indicate C pools or model variables that are input parameters to specific sub-models, while green arrows indicate the C pools or model variables affected by a particular sub process.

This version of DALEC used in this study is modified from DALEC v1 used in Fox et al. (2009). Both versions of the model consist of three vegetation C pools, for leaf, stem, and root, and two soil C pools, for soil organic matter and litter. The photosynthesis is driven by the Aggregate Canopy Model (ACM) (Williams et al., 2005), which itself

[2]http://www.sandia.gov/UQToolkit. UQTk v3.0 is currently undergoing formal review. In the meantime, the source code is available upon request from Bert Debusschere (bjdebus@sandia.gov)

GMDD
7, 6893–6948, 2014

is calibrated to the Soil-Plant-Atmosphere (SPA) model (Williams et al., 1996). The following modifications were made: An update was made to employ a temperature-based deciduous phenology used in Ricciuto et al. (2011), driven by the six parameters shown in Fig. 2. Spring phenology is driven by a linear relationship to growing degree days, while senescence is driven by mean air temperature. To reduce model complexity, the plant labile pool was removed and stem carbon is used to support springtime leaf flush given the spring phenology and the maximum leaf area index parameter. Given the importance of maintenance respiration in other sensitivity analyses (Sargsyan et al., 2014), this process was added along with parameters controlling the base rate and temperature sensitivity.

In this version of DALEC, ACM shares one parameter, the specific leaf area (*lma*), with the deciduous phenology and employs two additional parameters, leaf C : N ratio (*leafcn*) and Nitrogen use efficiency (nue). The autotrophic respiration model computes the growth and maintenance respiration components and is controlled by three parameters: the growth respiration fraction (*rg_frac*), and the base rate at 25 °C (*br_mr*) and temperature sensitivity for maintenance respiration (*q10_mr*), respectively. The allocation sub-model partitions C to several vegetation C pools. Leaf allocation is first determined by the phenology submodel, and the remaining available C is allocated to the root and stem pools depending on the fractional stem allocation parameter (*astem*). The "Litterfall" submodel redistributes the C content from vegetation pools to soil pools and is based on the turnover times for stem (*tstem*), root (*troot*), and leaves (*tleaf*). The sequence of sub-models concludes with the "Decomposition" which models the heterotrophic respiration component and the decomposition of litter into soil organic matter (SOM). This sub-model is driven by temperature sensitivity for heterotrophic respiration (*q10_hr*), the base turnover times for litter and SOM at 25 °C (*br_lit*, *br_som*), respectively, and by the decomposition rate (*dr*) from litter to SOM.

Model parameters and their nominal values are provided in Table 1. These parameters are grouped according to the sub-model that employs them. Except for leaf mass per unit area (*lma*) which impacts both the deciduous leaf phenology and ACM, all

GMDD
7, 6893–6948, 2014

other parameters are employed in single submodels. The numerical ranges for these parameters are also provided in the table. These ranges, corresponding to the Harvard Forest site (Urbanski et al., 2007), are set to capture a broad range of reasonable values used in past studies (Fox et al., 2009; White et al., 2000) and will serve as a base for the GSA study presented in the next section. In addition to the model parameters, several processes are driven by the observed air temperature, solar radiation, vapor pressure deficit, and CO_2 concentration at the flux tower site.

3 Global sensitivity analysis

GSA formally connects uncertainties in model output to the underlying uncertainties present in the model inputs. We employ variance-decomposition methods where the variance of the model output is decomposed into fractions associated with input factors and their interactions. The primary quantity of interest (QoI) for GSA is NEE, for which we have experimental observations available. We explore GSA for several other QoI's to understand the role each parameter or set of parameters play on other DALEC outputs in addition to NEE. Specifically we consider the Gross Primary Production (GPP), the Total Vegetation Carbon (TVC), and the Total Soil Carbon (TSC).

The effects of input parameters $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_{N_\theta}\}$ and their interactions on a model output $y = m(\boldsymbol{\theta})$, are quantified through Sobol indices (Sobol, 1993; Campolongo et al., 2000). The first order Sobol indices are given by

$$S_i = \frac{\mathrm{Var}_{\theta_i}[E_{\theta_{\sim i}}(m(\boldsymbol{\theta})|\theta_i)]}{\mathrm{Var}_{\theta}[m(\theta)]}, i = 1, \ldots, N_\theta \quad (1)$$

where $\theta_{\sim i} = \{\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_{N_\theta}\}$, $E_{\theta_{\sim i}}[\cdot]$ is the expectation with respect to $\theta_{\sim i}$, and $\mathrm{Var}_{\theta_i}[\cdot]$ is the variance with respect to θ_i.

GMDD
7, 6893–6948, 2014

Similarly, the joint sensitivity indices S_{ij} are

$$S_{ij} = \frac{\mathrm{Var}_{\theta_i,\theta_j}[E_{\theta_{\sim(i,j)}}(m(\theta)|\theta_i,\theta_j)]}{\mathrm{Var}_{\theta}[m(\theta)]} - S_i - S_j, \quad i,j = 1,\ldots,N_\theta. \tag{2}$$

While interactions between three or more parameters can be defined in a similar fashion, for most physical models these higher-order interactions are typically negligible.

The sensitivity index S_i can be interpreted as the fraction of the variance in the QoI that can be attributed to the ith input parameter only, while S_{ij} is the variance fraction that is due to the joint contribution of the ith and jth input parameters. The Sobol indices (1) and (2) can be written in integral forms, but these integrals will not be analytically tractable when the input parameter space is high-dimensional. In order to evaluate these indices numerically we employ a Monte-Carlo approach enhanced by techniques described by Saltelli (2002) and modified by Kucherenko et al. (2012) to account for parameter dependencies. This method employs sampling the input parameters from their prior distributions and an efficient re-use of model evaluations to reduce the computational cost of estimation of the above conditional variances.

We employ maximum entropy (MaxEnt) arguments to choose prior distributions for the model parameters, since we are only given prior information on parameter bounds (Table 1). The MaxEnt principle states that the maximum entropy distribution is the least informative distribution (Jaynes, 1968). Among distributions with finite support, the uniform distribution has the largest entropy, hence, given available prior information, we choose uniform prior distributions for all model parameters, with bounds provided in Table 1. The prior distributions for these parameters are independent, except for the spring phenology parameters *gdd_min* and *gdd_max*, which are bound by the inequality constraint *gdd_min* < *gdd_max*.

For each of the QoIs mentioned above, we compute monthly averages corresponding to the entire simulation, i.e. the January average is computed using the January daily QoI values for all available years. Global averages for all QoIs are also analyzed for comparison purposes. The simulations are driven by daily minimum and maximum

GMDD
7, 6893–6948, 2014

temperatures, global radiation, and CO_2 concentration for years 1992–2006, at the Harvard Forest site (Urbanski et al., 2007).

3.1 First-order effects

Figures 3–6 show matrices of first-order Sobol indices for the four QoIs mentioned above. The colormap changes from red for large Sobol index values to blue for Sobol indices $\approx 1\%$. The grayscale corresponds to Sobol index values from 1% down to 0.1%, while blank cells indicate values smaller the 0.1%. Each row in these matrices shows the Sobol indices corresponding to a particular average QoI. The sum of these values on each row indicates the sum of variance contributions due to individual parameters to the total variance of that particular average QoI. For example, in Fig. 3, the first order Sobol indices for the September average NEE sum up to 0.73. The remaining 0.27 fraction of the total variance for this month is due to pairwise interactions between parameters or higher order interactions.

Different parameters have larger impacts at certain times of the year. For NEE, in Fig. 3, phenology parameters *tsmin* and *leaffall*, which control the senescence of leaves in the Fall, have a significant impact on NEE during this period only. Specifically, *tsmin*, which is the critical temperature at which *leaffall* begins, mainly affects NEE in October. Similar behavior is seen for parameters that control GPP. Parameter *gdd_min*, which is the number of growing degree days at which leaf budbreak occurs, has the most impact in March and April. The strong dependence of these fluxes on phenology parameters highlights the importance of an accurate phenology model, as has been shown in other modeling studies, e.g. (Richardson et al., 2012). On the other hand, the Nitrogen use efficiency *nue*, which controls the amount of GPP per unit leaf Nitrogen, is important throughout most of the growing season (June–September). This is broadly consistent with other sensitivity studies that have shown strong sensitivity to leaf nitrogen, e.g. Sargsyan et al. (2014).

TVC and TSC are carbon pools and tend to vary on a much larger timescale than GPP or NEE, which are fluxes. Therefore, the Sobol indices do not exhibit significant

GMDD
7, 6893–6948, 2014

seasonal variability. TVC is most strongly controlled by the base rate of maintenance respiration *br_mr*, which represents a Carbon cost plants must continuously spend during their lifetime. TSC is most strongly controlled by the base rate of decomposition for soil organic matter *br_som*, which effectively determines the pool residence time. Given the same inputs, a pool with a longer residence time will contain more Carbon.

3.2 Joint effects

Figures 7–8 show relevant joint sensitivity indices corresponding to NEE and GPP which exhibit seasonal variability for the first-order Sobol indices. In these figures, each node shows relevant parameters while the label on each link corresponds to the joint Sobol index S_{ij}, in % units. In this figure the joint Sobol index values are rounded to the nearest integer for clarity. The results for these months, selected based on the relevant active processes affecting these two model outputs, show that parameter interactions are also important. For example, during Spring, the interaction between *gdd_min* and other ACM and AR model parameters account for around 10–20 % of the total variance for both NEE and GPP. Conversely, in the Fall, *tsmin* and *br_mr* have important interactions with several other phenology, AR, and Decomposition model parameters. These interactions account for about 15 % of the total variance in both NEE and GPP and play an important role in determining the evolution of the Carbon cycle during the senescence period.

4 Parameter calibration

We employ a Bayesian framework to compute posterior probabilities for the model parameters discussed in the previous sections. This framework is well-suited for dealing with uncertainties from different sources, including parametric and model uncertainty and experimental errors (Sivia, 1996). In the Bayesian approach, the probability density

GMDD
7, 6893–6948, 2014

for the model parameters is characterized as

$$p(\boldsymbol{\theta}|\mathcal{D}) = \frac{L_{\mathcal{D}}(\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathcal{D})} \tag{3}$$

Here $p(\boldsymbol{\theta})$ and $p(\boldsymbol{\theta}|\mathcal{D})$ are the prior and posterior probability densities, respectively, for model parameters $\boldsymbol{\theta}$. These densities represent our knowledge of $\boldsymbol{\theta}$ before and after learning from the data $\mathcal{D}$. The likelihood function $L_{\mathcal{D}}(\boldsymbol{\theta}) = p(\mathcal{D}|\boldsymbol{\theta})$ is the likelihood of the data $\mathcal{D}$ for a particular instance of model parameters $\boldsymbol{\theta}$. The denominator in Eq. (3), $p(\mathcal{D})$, is the "evidence", computed by integrating the numerator over the support of $\boldsymbol{\theta}$. It plays a role of a normalizing constant in the parameter estimation context, and is not computed here.

The data available for the calibration of model parameters consists of the Harvard Forest's daily Net Ecosystem Exchange daily values processed for the North American Carbon Program Site Synthesis study (Barr et al., 2013). Flux data were measured by the site PI's (Urbanski et al., 2007). Hill et al. (2012) estimated that daily NEE estimates follow a normal distribution. Consistent with Hill et al. (2012), systematic biases are not included in the present study. We further assume that daily measurement noise/errors, ϵ_{d}, are independent. Daily measurement errors or standard deviations are provided by the North American Carbon Program (NACP) interim synthesis (Barr et al., 2009). For this study we neglect the model error, ϵ_{g} in Fig. 1. Given these assumptions, the likelihood $L_{\mathcal{D}}(\boldsymbol{\theta})$ is written as

$$L_{\mathcal{D}}(\boldsymbol{\theta}) = \prod_{k=1}^{N_d} \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(y_k - \mathcal{D}_k)^2}{2\sigma_k^2}\right) \tag{4}$$

where $\mathcal{D}_k$ and σ_k are the observed NEE value and its standard deviation for day k, while $y_k = m_k(\boldsymbol{\theta})$ is the corresponding NEE value predicted by the DALEC model. The daily observations cover a period of 15 years starting with year 1992. A snapshot of these observations, including the magnitude of the observation error, is provided in

GMDD
7, 6893–6948, 2014

Fig. 9. The standard deviations for the daily NEE values were estimated using a bootstrapping technique using half-hourly NEE data (Papale et al., 2006; Barr et al., 2009).

For this study we consider two approaches for running the forward model and generating the output needed for the computation of the likelihood in Eq. (4). The first approach employs a steady state assumption, with DALEC run in a spinup mode until a steady state is reached. This takes typically 30–50 cycles of the 1992–2006 meteorology (450–750 total years). In this context, each cycle corresponds to running the model for 15 years with the meteorology inputs of 1992–2006. At the start of the first cycle, the Carbon pools are empty. For subsequent cycles, the C pools are initialized with the final state from the previous cycle. The daily model-predicted NEE values used for parameter estimation are those of the first cycle after the system reaches a steady state. This approach follows the protocol for NACP interim synthesis simulations, but fails to capture the large negative NEE observed at Harvard Forest. In the second approach, the initial values of the C pools in January 1991 are added to the set of model parameters to be estimated. This approach employs unsteady assumptions and, for any given set of parameter values, DALEC is run one cycle only, for 1992–2006. The resulting model output values are employed to compute the likelihood. The model evaluations are cheaper compared to the first approach, however the dimensionality of the parameter space is increased by 5, 3 vegetation C pools and 2 soil Carbon pools, from 18 to 23 parameters. Henceforth, we will refer to these two approaches as D18 and D23.

4.1 Fisher information matrix

We first proceed to estimate the amount of information datasets consisting of NEE observations are expected to carry about the DALEC model parameters. This is quantified via Fisher Information (Fisher, 1973), which is defined as the amount of information the observable NEE carries, as a random vector, about the unknown parameter vector θ. Let Z be a random vector of NEE observations. For this work, the probability density for Z is the multivariate normal likelihood defined by Eq. (4). In the current context, the

GMDD
7, 6893–6948, 2014

specific dataset of NEE observations, $\mathcal{D}$, is replaced by a random vector of NEE observations Z. The Fisher Information Matrix (FIM) is defined as (Lehmann and Casella, 2003)

$$\mathcal{I}(\boldsymbol{\theta})_{i,j} = -E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log L_Z(\boldsymbol{\theta})|\boldsymbol{\theta}\right] = -\int_{\Omega(Z)}\frac{\partial^2\log L_Z(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}L_Z(\boldsymbol{\theta})\mathrm{d}Z \tag{5}$$

where $\Omega(Z)$ represents the space of all possible values of Z. Since $L_Z(\boldsymbol{\theta})$ is a multivariate normal, $Z \sim N(\boldsymbol{m}(\boldsymbol{\theta}), \boldsymbol{\Sigma})$, with a constant covariance matrix $\boldsymbol{\Sigma}$, the Fisher Information Matrix (FIM) entries in Eq. (5) can be shown to be

$$\mathcal{I}(\boldsymbol{\theta})_{i,j} = \frac{\partial\boldsymbol{m}^T}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{m}}{\partial\theta_j}. \tag{6}$$

Here, $\boldsymbol{m}(\boldsymbol{\theta})$ is the vector of daily NEE values output by DALEC, $\boldsymbol{m}(\boldsymbol{\theta}) = \left(m_1(\boldsymbol{\theta}), \ldots, m_{N_d}(\boldsymbol{\theta})\right)^T$, with $m_k = \mathrm{NEE}_k$ for day k. The covariance matrix $\boldsymbol{\Sigma}$ is diagonal, constructed with daily variance values, σ_k^2, on the diagonal, see also Eq. (4). In order to compare the FIM entries corresponding to different parameters, the parameter values are normalized by their corresponding prior range. The normalized FIM entries are then computed as $\mathcal{I}^*(\boldsymbol{\theta})_{i,j} = \mathcal{I}(\boldsymbol{\theta})_{i,j}\Delta_{\theta_i}\Delta_{\theta_j}$. Here Δ_{θ_i} is the range corresponding to θ_i, computed based on values given in Table 1.

The FIM values are expectations over the data, computed for specific values for model parameters $\boldsymbol{\theta}$. Thus, an uncertain $\boldsymbol{\theta}$ leads to uncertain FIM entries. We use Monte Carlo sampling to generate random samples from the MaxEnt-derived priors on $\boldsymbol{\theta}$. This yields an ensemble of FIM values, from which we can construct histograms for each FIM component. Since the model output dependence on $\boldsymbol{\theta}$ is not given

GMDD
7, 6893–6948, 2014

analytically, we compute the partial derivatives in Eq. (6) by numerical differentiation[3]

$$\frac{\partial m_k}{\partial \theta_i} \approx \frac{\delta m_k}{\delta \theta_i} = \frac{m_k(\ldots, \theta_{i-1}, \theta_i + \delta\theta_i, \theta_{i+1}, \ldots) - m_k(\boldsymbol{\theta})}{\delta \theta_i}. \quad (7)$$

The perturbation $\delta\theta_i$ is set to $\sqrt{\epsilon_m}\theta_i$, where ϵ_m is the upper bound of relative error due to rounding in floating-point arithmetic, and is typically of the order of 2.2×10^{-16} for double precision (64 bit) computations.

Normalized histograms for select diagonal entries of the FIM are shown in Figs. 10 and 11, using the model setup for D18. In these figures, FIM_{θ_i} stands for $\mathcal{J}(\boldsymbol{\theta})^*_{i,i}$, defined above. Convergence tests, results not shown, indicate that about 10^4 Monte Carlo samples are sufficient to generate converged histograms. The results were grouped, by visual inspection, according to the magnitude of $\log(\mathrm{FIM}_{\theta_i})$. Figure 10 shows model parameters with larger FIM diagonal entries while Fig. 11 shows parameters corresponding to generally smaller values. This indicates that NEE observations are informative about the set of parameters shown in Fig. 10 (see the figure caption for the list of parameter names), and that, consequently, the probability distributions for these parameters are likely to be significantly updated through the Bayesian parameter estimation discussed in this section. It is interesting to note that these parameters were found to be important, based on the GSA results in the previous section, for NEE. Specifically, *gdd_min*, *q10_mr*, *br_mr*, and *rg_frac* exhibit relevant first order effects, shown in Fig. 3, while *gdd_max* and *leaffall* are relevant mostly through interactions with other parameters, in Fig. 7.

Conversely, the FIM values shown in Fig. 11 are much smaller, hence the calibration exercise is not expected to update the probability distributions for these parameters significantly. Among these parameters, *br_som* was found to be the most important for the TSC. However, based on the FIM results, the NEE observations do not carry

[3]Alternatively, a set of ordinary differential equations for $\partial m_k/\partial\theta_i$ can be derived and solved numerically.

GMDD
7, 6893–6948, 2014

information about this parameter and its prior density will likely not be updated by using the NEE data only. This likely results from the long residence time of the SOM pool – the 15 year NEE record is not long enough to constrain it sufficiently. These conjectures will be verified using the calibration results presented in the next section.

Based on the FIM results presented in this section and on the GSA results in Sect. 3, we separate the DALEC parameters into three groups. In the first group we include *gdd_min*, *gdd_max*, *q10_mr*, *br_mr*, *rg_frac*, and *leaffall*. These parameters were highligted both by the GSA and FIM results as being important for NEE. In the second group we include *tsmin*, *q10_hr*, *br_lit*, and *lma*. These parameters were selected since either the GSA or the FIM results suggested they are relevant to NEE. Finally, we place the remaining parameters in the third group. In the next section the posterior distributions for model parameters are constructed sequentially starting with the most important group of parameters, then gradually adding parameters, one group at a time.

4.2 Posterior distributions via MCMC

A Markov Chain Monte Carlo (MCMC) algorithm is used to sample from the posterior probability density $p(\boldsymbol{\theta}|\mathcal{D})$. MCMC is a class of techniques that allows sampling from a probability density by constructing a Markov Chain that has the target density as its stationary distribution (Gamerman, 1997; Gilks et al., 1996). In particular, we employ an adaptive Metropolis algorithm (Haario et al., 2001), which uses the covariance of the previously visited chain states to find better proposal distributions, allowing it to explore the posterior distribution in an efficient manner. Haario et al. (2001) shows that, for Gaussian distributions, the adaptive sampling algorithm is similar in performance to the Metropolis algorithm. For non-Gaussian posterior densities, the adaptive procedure is superior to non-adaptive procedures, however the adaptive procedure is challenged by the dimensionality of the parameter space.

To facilitate the convergence of the adaptive MCMC algorithm we proceed gradually, starting with the first group of parameters mentioned in the previous section. The schematic in Fig. 12 shows one iteration in the sequence of MCMC simulations. For

GMDD
7, 6893–6948, 2014

the first iteration,

$$\theta^{(1)} = \{gdd\_min, gdd\_max, q10\_mr, br\_mr, rg\_frac, leaffall\}$$

with initial values $\theta^{(1)}_{ini}$ set to the nominal conditions provided in Table 1. The rest of parameters are held constant at their nominal values. The initial covariance matrix, $\mathbf{C}^{(1)}_{ini}$, allows the MCMC algorithm to explore a number of possible states before adapting the sample covariance based on the sample history. For this study we found that a diagonal covariance matrix with entries set to a fraction of about 1/16 of the variances for the corresponding prior distributions provided a good start for the MCMC algorithm.

The MCMC states obtained during the first iteration are used to compute the covariance matrix corresponding to the first set of parameters $\mathbf{C}^{(1)}$ which is then used to construct the initial covariance matrix for the second iteration, $\mathbf{C}^{(2)}_{ini}$. This process is shown schematically in Fig. 12. The initial parameter values for the 2nd iteration consist of the Maximum A Posteriori (MAP) for $\theta^{(1)}$ augmented with the nominal values for

$$\theta^{(2)\backslash(1)} = \{tsmin, q10\_hr, br\_lit, lma\}$$

The iterative process is completed after the third iteration, with $\theta^{(3)\backslash(2)}$ containing the rest of DALEC parameters.

We employ the Raftery–Lewis diagnostic (Raftery and Lewis, 1992) to determine when the MCMC samples converge to stationary posterior distributions. For D18, approximately 4×10^6 samples are necessary to predict the 5, 50, and 95% quantiles of all parameters to within ±1% accuracy with 95% probability. For D23, the Raftery–Lewis diagnostic test shows that 6×10^6 are necessary for converged posterior distributions. Given 5×10^6 MCMC samples, the Effective Sample Size (Kass et al., 1998) (ESS) for D18 varies between 10000 and 15000 samples depending on each parameter, while for D23, ESS is between 8000 and 12000. This shows that there is significant autocorrelation between chain samples, which is somewhat typical for MCMC samplers in high-dimensional spaces. To ensure converged posterior distributions, and since the

GMDD
7, 6893–6948, 2014

computational model is cheap, results presented below are based on 10^7 MCMC samples for both D18 and D23.

We first proceed to analyse the model calibration results for D18, when DALEC is run to a quasi-steady state for each parameter sample. In order to measure the degree of dependence in the joint posterior distribution for the 18 model parameters we examine the "distance correlation" values (Székely et al., 2007) estimated based on the MCMC samples. The distance correlation is a measure of *dependence* between two random variables, being zero when they are independent. Given random variables X and Y with finite first moments, the distance correlation $\mathcal{R}(X,Y) \in [0,1]$ is defined as

$$\mathcal{R}(X,Y) = \frac{\vartheta^2(X,Y)}{\sqrt{\vartheta^2(X)\vartheta^2(Y)}} \tag{8}$$

where $\vartheta^2(X,Y)$ is the "distance covariance" between X and Y and $\vartheta^2(X)$ is the "distance variance", $\vartheta^2(X) = \vartheta^2(X,X)$. The distance covariance $\vartheta^2(X,Y)$ is defined as

$$\begin{aligned}\vartheta^2(X,Y) =& E(\|X-X'\|\|Y-Y'\|) + E(\|X-X'\|)E(\|Y-Y'\|) \\ & - 2E(\|X-X'\|\|Y-Y''\|)\end{aligned} \tag{9}$$

where (X',Y'), (X'',Y'') are independent and identically distributed random variables, drawn from the same joint density as (X,Y). Székely et al. (2007) provide numerical algorithms to compute $\mathcal{R}(X,Y)$ given samples of random variables X and Y. The results are shown in Table 2. In this table, parameters are grouped in blocks according to the sub-model they participate in. The entries in the diagonal blocks show dependencies between parameters in the same sub-model while the entries in off-diagonal blocks indicate dependencies between parameters from different sub-models.

The most important statistical dependencies are between *nue* and *lma* that control the gross photosynthesis (ACM) and between *rg_frac* and *nue* that control net photosynthesis. Relevant dependencies are also observed between *q10_mr*, a parameter of the autotrophic respiration process, and *q10_hr* which participates in the heterotrophic

GMDD
7, 6893–6948, 2014

respiration process. In order to further understand the dependencies between model parameters we compute 1-D and 2-D joint marginal densities, via Kernel Density Estimates (KDE) (Scott, 1992; Silverman, 1986), for the model parameters that exhibit distance correlation factors greater than 0.3. These results are shown in Fig. 13. The statistical dependencies identified above through $\mathcal{R}$ are also evident in 2-D joint marginal densities for the same parameters.

Figure 14 shows 1-D marginal densities for the rest of the parameters. These parameters show little dependence on other parameters and so the 1-D marginal distribution is sufficient to characterize their density. Some parameters are well constrained towards the center of the prior range, for instance *br_mr* and *br_lit*; these parameters control the basal autotrophic and litter respiration rates, which occur on timescales for which the NEE data have high information content. For *tleaf* and *br_som*, the NEE observations are not informative, see Sect. 4.1, and their posterior densities remain nearly uniform, the same as their prior densities. For *tleaf*, the lack of information is due to the fact that the effects of leaf turnover on net fluxes are much more strongly controlled by their timing, as determined by the phenology parameters, than by the background turnover rate. For *br_som*, its turnover rate is slow enough such that the NEE data contain little useful information. The posterior densities for other parameters, e.g. *laimax*, are tilted toward one end of their prior range. This might indicate that the calibration process attempts to compensate for structural discrepancies between observations and model predictions by pushing some parameters toward either the minimum or the maximum value of their prior range.

While the posterior distribution for *br_mr* is well updated in the calibration, the $\mathcal{R}$ values between this parameter and all other parameters are smaller than the threshold used to select parameters that show mutual dependencies. The posterior distribution for *tsmin* is piecewise uniform. This is due to the fact that minimum daily temperatures, in degrees Celsius, are provided with one decimal digit accuracy and this parameter is a threshold for leaf drop, i.e. its participation in the computational model is through

GMDD
7, 6893–6948, 2014

an "if" statement. Hence all samples between successive one-digit accurate thresholds are equally probable during the MCMC sampling process.

Next we examine the departure of each parameter's density from its uniform prior as a result of the Bayesian update via Eq. (3). We quantify these changes via the Kullback–Leibler (KL) divergence between prior and marginal posterior densities,

$$D_{KL}(p\|q) = \int_{-\infty}^{\infty} p(x) \ln\left(\frac{p(x)}{q(x)}\right) \mathrm{d}x, \tag{10}$$

where p is the posterior densty and q is the prior density. KL divergence results for certain paramaters are presented in Fig. 15. In this figure, parameters are sorted in ascending order based on their D_{KL} values. Also shown in the figure is the inverse of the scaled standard deviation based on the MCMC sample values for each model parameter. This is obtained from the standard deviation of the MCMC samples for each parameter, σ_i, by scaling with the standard deviation based on the corresponding prior density, $\sigma_i^* = \sigma_{p,i}/\sigma_{q,i}$. The calibration exercise had negligible effect on the probability density for the first two parameters, *tleaf* and *br_som*. $D_{KL}(p\|q)$ values for these two parameters were less than 10^{-2} and were rounded to zero. The corresponding σ^* values are close to 1, indicating little change from their prior uniform densities. This result confirms that NEE data contain little information on the turnover rate of SOM, or on the rapidity of leaf drop (rather than the timing of leaf drop, see below). At the other end of the spectrum, *br_mr*, *gdd_min*, and *q10_mr*, exhibit the largest $D_{KL}(p\|q)$, indicating a larger departure of their posterior densities from their prior density. These parameters are well constrained by the NEE data, reflecting the useful information in the flux data on timing of phenological events (*gdd_min*) and the dynamics of autotrophic respiration *(br_mr*, *q10_mr*). The large $D_{KL}(p\|q)$ values for these parameters are accompanied by small σ^* values, or large values for $1/\sigma^*$ as shown in Fig. 15, indicators that the marginal posterior density is significantly narrower than their prior

GMDD
7, 6893–6948, 2014

density. Back-of-the-envelope regression tests empirically suggest a power law dependence between $D_{KL}(p\|q)$ and σ^*.

To further assess the importance of statistical dependencies between model parameters, we revisit the GSA exercise to determine the relative importance of model parameters, based on the posterior densities of model parameters instead of the prior ones. Figure 16 shows select first order Sobol indices given posterior distributions based on D18. In this figure, several parameters are grouped together if the distance correlation values, shown in Table 2, imply that their mutual dependence is significant. These groups, named G1, G2, and G3, respectively, consist of the following parameters:

- G1: lma, nue, rg_frac
- G2: q10_mr, q10_hr
- G3: gdd_min, gdd_max

The results shown with red bars in Fig. 16 are based on joint posterior distributions for G1, G2 and G3, and marginal distributions for the rest of parameters. For the results shown with blue bars, the joint distributions for G1 through G3 are products of marginal distributions of each parameter in the group, hence neglecting any statistical dependence. Visual inspection of the relative importance of parameters or groups of parameters to the total variance of the average monthly NEE values shows that neglecting joint dependecies between parameters can significantly alter the results. This is true both for parameters that show significant dependence, e.g. see group G1, and for parameters that show little dependence on other parameters, e.g. *br_mr*.

Next, we analyze the calibration results for D23. For this model setup, the initial values for the C pools at the beginning of year 1991 are part of the set of model parameters and each DALEC simulation consists of only one cycle, for the time span 1992–2006. Table 3 shows $\mathcal{R}$ values for D23 results. Only the parameters common between D23 and D18 are listed in this table. The $\mathcal{R}$ values observed in this table for D23 are similar to the ones observed above for D18 with the exception of pair (*q10_mr*,

GMDD
7, 6893–6948, 2014

q10_hr). While, for D18, these two parameters exhibit significant dependence, for D23 they are nearly independent. Figure 17 shows 1-D and 2-D joint marginal densities for parameters with distance correlation factors greater than 0.3 based on D23 results. In general, these marginal densities are similar to the ones based on D18.

Finally, Fig. 18 shows marginal densities for two Carbon pools that were updated in the calibration exercise D23. *vpool2* corrresponds to the stem C while *spool2* corresponds to the soil organic matter. While *dr* shows little dependence on other model parameters, it has a dominant role in the conversion of the litter C pool into soil organic matter, and the distance correlation between this parameter and *spool2* is about 0.5.

5 Predictive assessment

In this section we explore the predictive skill given the posterior distributions for the model parameters for D18 and D23. First, we employ the Bayesian posterior predictive distribution (Lynch and Western, 2004) to assess the adequacy of the calibrated DALEC model, and the Gaussian data noise model, for prediction of the NEE observations. Specifically, the posterior distribution for the *predicted* NEE data, $p(y|\mathcal{D})$, is computed by marginalization of the likelihood over the posterior distribution of model parameters and hyperparameters, here θ:

$$p\left(y|\mathcal{D}\right) = \int_{\theta} p(y|\boldsymbol{\theta})p(\boldsymbol{\theta}|\mathcal{D})\mathrm{d}\theta. \tag{11}$$

For the present work, $y|\boldsymbol{\theta} \sim N(\boldsymbol{m}(\boldsymbol{\theta}), \boldsymbol{\Sigma})$, where $\boldsymbol{y} = \{y_k|k = 1 \ldots N_d\}$ is a N_d-dimensional vector with NEE predictions over a range of N_d days, and Σ is a diagonal covariance matrix with variances of daily NEE observations on the diagonal. The 1-D-marginal posterior predictive distributions for daily NEE values for a two-year snapshot around 1995 are shown in Fig. 19. These distributions were computed by sampling, using the MCMC samples behind the posterior densities presented in the previous

GMDD
7, 6893–6948, 2014

section. For computational efficiency, $p(y|\boldsymbol{\theta})$ was computed with KDE for each θ sample. This density was then interpolated on a uniform grid centered around the daily observed NEE value. It is worth to note that the variance of the posterior predictive distributions can also be estimated analytically as the sum of the measurement error variance and the pushed-forward variance, i.e. the variance of the output quantity of interest with respect to posterior variability.

The top frame in Fig. 19 corresponds to D18 and the bottom frame to D23. Generally, the predicted data spread covers well the observed NEE values except for a time range around May-June-July when the observations, with red line, are frequently outside the 5–95 % band, shown in blue. This discrepancy occurs mostly for years 1993–1996. For other years, the predicted data covers the May-June-July observational data well.

In order to quantitatively compare the predictive capability of the calibrated models for D18 and D23, we adopt a probabilistic score based on the predictive cumulative distribution function (CDF). The Continuous Rank Predictive Score (CRPS) (Gneiting and Raftery, 2007) measures the difference between the CDF of the provided data and that of the forecast/predicted data, i.e. data generated based on the posterior predictive distribution. Thus,

$$\mathrm{CRPS}(\mathcal{F},\mathcal{D}) = \frac{1}{N_d}\sum_{k=1}^{N_d}\int_{-\infty}^{\infty}\left(\mathcal{F}_k(y_k|\mathcal{D}) - \mathcal{H}_{\mathcal{D}_k}(y_k)\right)^2 \mathrm{d}y_k \tag{12}$$

Here, $\mathcal{F}_k(y_k|\mathcal{D})$ is the 1-D marginal posterior predictive CDF for day/component k computed using 1-D marginal posterior predictive distributions

$$\mathcal{F}_k(y_k|\mathcal{D}) = \int_{-\infty}^{y_k} p_k\left(y'_k|\mathcal{D}\right)\mathrm{d}y'_k \tag{13}$$

where

$$p_k\left(y_k|\mathcal{D}\right) = \int p\left(y_1, y_2 \ldots, y_{N_d}|\mathcal{D}\right)\mathrm{d}y_1 \cdots \mathrm{d}y_{k-1}\mathrm{d}y_{k+1}\cdots\mathrm{d}y_{N_d}. \tag{14}$$

GMDD
7, 6893–6948, 2014

is the 1-D marginal posterior predictive distribution corresponding to day k. The CDF of the provided data is approximated as a Heaviside function centered at the observation value $\mathcal{D}_k$ (Hersbach, 2000), $\mathcal{H}_{\mathcal{D}_k}(y_k) = \mathbb{1}_{y_k \geq \mathcal{D}_k}$.

Table 4 displays CRPS values based on posterior distributions obtained by averaging over several time ranges. The first row shows the values corresponding to a 60 day time frame, from mid-June to mid-August, while the second row corresponds to the remainder of the year. The last row shows the aggregated values, considering the entire year. The averages in Eq. (12) are taken over all years considered in this study. The lower values for D23 compared to D18 indicate a better predictive skill for the setup when DALEC is run for one cycle and the C pools are treated as parameters. The results reveal that the largest improvement, about 6 %, occurs for the June–August time frame, while for the rest of the year the improvement is about 2 % only.

In order to measure the effect of calibration on the predictive capability of DALEC we employ the Continuous Rank Predictive Skill Score (CRPSS) (Wilks, 2011)

$$\mathrm{CRPSS} = \frac{\mathrm{CRPS}_{\mathrm{psp}} - \mathrm{CRPS}_{\mathrm{prp}}}{\mathrm{CRPS}_{\mathrm{prf}} - \mathrm{CRPS}_{\mathrm{prp}}} \tag{15}$$

where $\mathrm{CRPS}_{\mathrm{psp}}$ is the CRPS computed above based on the posterior predictive distribution, $\mathrm{CRPS}_{\mathrm{prp}}$ is based on the prior predictive distribution, and $\mathrm{CRPS}_{\mathrm{prf}}$ is the CRPS based on "perfect" predictions. For the current study, the "perfect" predictions have a multivariate normal distribution centered on the observations and diagonal covariance matrix Σ defined above. The prior predictive distribution is defined analogous to the posterior predictive distribution in Eq. (11), with $p(\theta|\mathcal{D})$ being replaced by $p(\theta)$, the prior density for model parameters θ.

A CRPSS value of 0 implies no improvement of the predictive skill for the calibrated model parameters compared to the predictions based on the prior information, while a value of 1 can be achieved when the posterior distribution reduces to a point and the model prediction is the same as the corresponding experimental data. For the current study, $\mathrm{CRPS}_{\mathrm{prp}} = 2.38$ and $\mathrm{CRPS}_{\mathrm{prf}} = 0.53$ for D18. This leads to CRPSS = 0.4,

GMDD
7, 6893–6948, 2014

indicating a 40 % improvement in the predictive skill of DALEC as a result of calibration. We only show here the CRPSS values for D18 since for D23, the C pools employed improper priors for which the $CRPS_{prp}$ is not well defined.

6 Conclusions

We presented uncertainty quantification results for a process-based ecosystem Carbon model. We assembled several probabilistic methodologies in a framework that tackles the connected problems of parameter estimation and forward propagation of input uncertainties. Depending on the simulation setup, the model employs either steady state or non-steady assumptions, respectively, and it is driven by meteorological data corresponding to years 1992–2006 at the Harvard Forest site. Daily Net Ecosystem Exchange (NEE) observations were available to calibrate the model parameters and test the performance of the model.

We first discussed global sensitivity analysis (GSA) results for the complete set of input parameters. Based on their contribution to the variance, we find that different parameters have larger impacts for NEE at certain times of the year when the processes they control become important. One example is the *tsmin* parameter, which is the critical temperature at which leaf fall begins, and mainly affects NEE in October. We also found that parameter interactions can also be relevant to the variability of NEE or Gross Primary Production (GPP). Unlike NEE and GPP which are fluxes, the Carbon pools, either vegetation or soil, tend to vary more slowly and their month-to-month variability depends on a small subset of parameters.

We also employed Fisher Information Matrix (FIM) computations to estimate the relative information the NEE data contains on the model parameters. To our knowledge this type of study is employed for the first time in the context of a Carbon model. We ranked model parameters according to the relative magnitude of the diagonal entries in the FIM and generally found that most “informed” parameters are also ranked as important based on the GSA results.

GMDD
7, 6893–6948, 2014

We then proceeded to calibrate the model parameters in a Bayesian framework. In this context we examined both steady and unsteady assumptions for the Carbon model simulations. The daily discrepancies between measured and predicted NEE values were modeled as independent and identically distributed Gaussians with prescribed daily variance according to the recorded instrument error. All model parameters were assumed to have uninformative priors with bounds set according to expert opinion.

The posterior distribution of model parameters was sampled sequentially by first considering the most relevant parameters and then progressively adding less important parameters, according to GSA and FIM results. The posterior samples, obtained with a Markov Chain Monte Carlo algorithm, exhibit significant dependencies between some of the model parameters. Further, a GSA analysis based on marginal posterior distributions shows the importance of considering parameter dependencies when establishing the importance of each parameter or set of parameters for given quantities of interest.

The predictive capabilities of the model, employing the parameters' posterior distribution, were assessed qualitatively through posterior predictive checks and quantitatively through Continuous Rank Predictive Score (CRPS) computations. Based on the CRPS values, the unsteady model setup, for which C pools are set as simulation parameters, performed slightly better, in particular during the growing seasons, compared to model setup assuming steady state conditions.

The analysis presented in this paper considered a single data series at one site only. However, the Bayesian framework employed in the parameter calibrations is well-suited to deal with both heterogenous data and models. We are currently exploring avenues to extend this work to multi-site studies together with employing multiple data streams to better constrain the model parameters.

The framework presented here encompasses robust statistical methodologies that can be employed in the development and analysis of more detailed models like the Community Land Model (CLM). Since some of these methodologies are sampling-based, their application is restricted to computationally inexpensive models. To this end

GMDD
7, 6893–6948, 2014

we are currently working on developing efficient surrogate models that can be used in place of expensive models like CLM. With a surrogate model approach in place, one can proceed to study individual CLM sub-models as well as the CLM model as a whole and potentially improve its predictive capabilities.

Acknowledgements. C. Safta, D. Ricciuto, K. Sargsyan, H. N. Najm, B. Debusschere, and P. Thornton were supported by the US Department of Energy, Office of Science, under the project "Climate Science for a Sustainable Energy Future", funded by the Biological and Environmental Research (BER) program. M. Williams was supported by NERC National Centre for Earth Observation. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the US Department of Energy's National Nuclear Security Administration under contract DE-AC04-94-AL85000. Oak Ridge National Laboratory is managed by UT-BATTELLE for DOE under contract DE-AC05-00OR22725.

References

Barr, A., Hollinger, D., and Richardson, A.: CO_2 Flux Measurement Uncertainty Estimates for NACP, AGU Fall Meeting, 2009. 6905, 6906

Barr, A., Ricciuto, D., Schaefer, K., Richardson, A., Agarwal, D., Thornton, P., Davis, K., Jackson, B., Cook, R., Hollinger, D., van Ingen, C., Amiro, B., Arain, A. A. M., Baldocchi, D., Black, T., Bolstad, P., Curtis, P., Desai, A., Dragoni, D., Flanagan, L., Gu, L., Katul, G., Law, B., Lafleur, P., Margolis, H., Matamala, R., Meyers, T., McCaughey, H., Monson, R., Munger, J., Oechel, W., Oren, R., Roulet, N., Torn, M., and Verma, S.: NACP Site: Tower Meteorology, Flux Observations with Uncertainty, and Ancillary Data, available at: http://daac.ornl.gov (last access: 10 October 2014), Oak Ridge National Laboratory Distributed Active Archive Center, doi:10.3334/ORNLDAAC/1178, 2013. 6905

Braswell, B., Sacks, W., Linder, E., and Schimel, D.: Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations, Global Change Biology, 11, 335–355, doi:10.1111/j.1365-2486.2005.00897.x, 2005. 6896

GMDD

7, 6893–6948, 2014

Campolongo, F., Saltelli, A., Sørensen, T., and Tarantola, S.: Hitchhiker's guide to sensitivity analysis, in: Sensitivity Analysis, edited by: Saltelli, A., Chan, K., and Scott, E., Wiley, Chicester, 2000. 6898, 6901

Fisher, R.: Statistical Methods and Scientific Inference, Macmillan Pub Co, 1973. 6906

Fox, A., Williams, M., Richardson, A. D., Cameron, D., Gove, J. H., Quaife, T., Ricciuto, D., Reichstein, M., Tomelleri, E., Trudinger, C. M., and Wijk, M. T. V.: The REFLEX project: comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data, Agr. Forest. Meteorol., 149, 1597–1615, doi:10.1016/j.agrformet.2009.05.002, 2009. 6897, 6898, 6899, 6901

Gamerman, D.: Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Chapman & Hall, London, 1997. 6909

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J.: Markov Chain Monte Carlo in Practice, Chapman & Hall, London, 1996. 6909

Gneiting, T. and Raftery, A.: Strictly proper scoring rules, prediction, and estimation, J. Am. Stat. Assoc., 102, 359–378, 2007. 6899, 6916

Haario, H., Saksman, E., and Tamminen, J.: An adaptive Metropolis algorithm, Bernoulli, 7, 223–242, 2001. 6909

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems., Weather Forecast., 15, 559–570, 2000. 6917

Hill, T., Ryan, E., and Williams, M.: The use of CO_2 flux time series for parameter and carbon stock estimation in carbon cycle research, Glob. Change Biol., 18, 179–193, doi:10.1111/j.1365-2486.2011.02511.x, 2012. 6905

Jaynes, E. T.: Prior probabilities, IEEE Transactions on Systems Science and Cybernetics, 4, 227–241, doi:10.1109/TSSC.1968.300117, 1968. 6902

Kaminski, T., Knorr, W., Rayner, P. J., and Heimann, M.: Assimilating atmospheric data into a terrestrial biosphere model: a case study of the seasonal cycle, Global. Biogeochem. Cy., 16, 14-1–14-16, doi:10.1029/2001GB001463, 2002. 6895

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Consistent assimilation of MERIS FAPAR and atmospheric CO_2 into a terrestrial vegetation model and interactive mission benefit analysis, Biogeosciences, 9, 3173–3184, doi:10.5194/bg-9-3173-2012, 2012. 6895

GMDD
7, 6893–6948, 2014

Kass, R., Carlin, B., Gelman, A., and Neal, R.: Markov Chain Monte Carlo in practice: a roundtable discussion, Am. Stat., 52, 93–100, doi:10.1080/00031305.1998.10480547, 1998. 6910

Knorr, W. and Kattge, J.: Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling, Glob. Change Biol., 11, 1333–1351, doi:10.1111/j.1365-2486.2005.00977.x, 2005. 6896

Kucherenko, S., Tarantola, S., and Annoni, P.: Estimation of global sensitivity indices for models with dependent variables, Comput. Phys. Commun., 183, 937–946, doi:10.1016/j.cpc.2011.12.020, 2012. 6902

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, Biogeosciences, 9, 3757–3776, doi:10.5194/bg-9-3757-2012, 2012. 6895

Lehmann, E. and Casella, G.: Theory of Point Estimation, Springer Texts in Statistics, Springer, 2003. 6898, 6907

Lynch, S. and Western, B.: Bayesian posterior predictive checks for complex models, Sociol. Method. Res., 32, 301–335, doi:10.1177/0049124103257303, 2004. 6899, 6915

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, Biogeosciences, 3, 571–583, doi:10.5194/bg-3-571-2006, 2006. 6906

Raftery, A. and Lewis, S.: How many iterations in the Gibbs sampler?, in: Bayesian Statistics 4, Oxford University Press, 763–773, 1992. 6910

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), Global Biogeochem. Cy., 19, GB2026, doi:10.1029/2004GB002254, 2005. 6895

Ricciuto, D., Davis, K., and Keller, K.: A Bayesian calibration of a simple carbon cycle model: the role of observations in estimating and reducing uncertainty, Global Biogeochem. Cy., 22, doi:10.1029/2006GB002908, 2008. 6897

Ricciuto, D. M., King, A. W., Dragoni, D., and Post, W. M.: Parameter and prediction uncertainty in an optimized terrestrial carbon cycle model: effects of constraining variables and data record length, J. Geophys. Res.-Biogeo., 116, G01033, doi:10.1029/2010JG001400, 2011. 6898, 6900

GMDD
7, 6893–6948, 2014

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Migliavacca, M., Monson, R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., and Xue, Y.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, Glob. Change Biol., 18, 566–584, doi:10.1111/j.1365-2486.2011.02562.x, 2012. 6903

Saltelli, A.: Making best use of model evaluations to compute sensitivity indices, Comput. Phys. Commun., 145, 280–297, doi:10.1016/S0010-4655(02)00280-1, 2002. 6902

Sargsyan, K., Safta, C., Najm, H. N., Debusschere, B., Ricciuto, D., and Thornton, P.: Dimensionality reduction for complex models via Bayesian compressive sensing, International Journal of Uncertainty Quantification, 4, 63–93, doi:10.1615/Int.J.UncertaintyQuantification.2013006821, 2014. 6903

Scott, D.: Multivariate Density Estimation. Theory, Practice and Visualization, Wiley, New York, 1992. 6912

Silverman, B.: Density Estimation for Statistics and Data Analysis, Chapman and Hall, London, 1986. 6912

Sivia, D.: Data Analysis: A Bayesian Tutorial, Oxford Science, 1996. 6904

Sobol, I. M.: Sensitivity estimates for nonlinear mathematical models, Math. Modeling Comput. Exper., 1, 407–414, 1993. 6898, 6901

Székely, G., Rizzo, M., and Bakirov, N.: Measuring and testing dependence by correlation of distances, Ann. Stat., 35, 2769–2794, doi:10.1214/009053607000000505, 2007. 6911

Tang, J. and Zhuang, Q.: A global sensitivity analysis and Bayesian inference framework for improving the parameter estimation and prediction of a process-based terrestrial ecosystem model, J. Geophys. Res.-Atmos., 114, D15303, doi:10.1029/2009JD011724, 2009. 6896

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A., and Mahowald, N. M.: Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability, Global Biogeochem. Cy., 21, GB4018, doi:10.1029/2006GB002868, 2007. 6898

Tjiputra, J., Polzin, D., and Winguth, A.: Assimilation of seasonal chlorophyll and nutrient data into an adjoint three-dimensional ocean carbon cycle model: sensitivity analysis and ecosystem parameter optimization, Global Biogeochem. Cy., 21, GB1001, doi:10.1029/2006GB002745, 2007. 6895

GMDD
7, 6893–6948, 2014

Trudinger, C. M., Raupach, M. R., Rayner, P. J., Kattge, J., Liu, Q., Pak, B., Reichstein, M., Renzullo, L., Richardson, A. D., Roxburgh, S. H., Styles, J., Wang, Y. P., Briggs, P., Barrett, D., and Nikolova, S.: OptIC project: an intercomparison of optimization techniques for parameter estimation in terrestrial biogeochemical models, J. Geophys. Res.-Biogeo., 112, G02027, doi:10.1029/2006JG000367, 2007. 6897

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., and Munger, J. W.: Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest, J. Geophys. Res.-Biogeo., 112, 1–25, doi:10.1029/2006JG000293, 2007. 6901, 6903, 6905

White, M. A., Thornton, P. E., Running, S. W., and Nemani, R. R.: Parameterization and sensitivity analysis of the BIOME–BGC terrestrial ecosystem model: net primary production controls, Earth Interac., 4, 1–85, doi:10.1175/1087-3562(2000)004<0003:PASAOT>2.0.CO;2, 2000. 6901

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, 2011. 6917

Williams, M., Rastetter, E., Fernandes, D., Goulden, M., Wofsy, S., Shaver, G., Melillo, J., Munger, J., Fan, S.-M., and Nadelhoffer, K.: Modelling the soil-plant-atmosphere continuum in a Quercus–Acer stand at Harvard Forest: the regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties, Plant Cell Environ., 19, 911–927, doi:10.1111/j.1365-3040.1996.tb00456.x, 1996. 6900

Williams, M., Schwarz, P., Law, B., Irvine, J., and Kurpius, M.: An improved analysis of forest carbon dynamics using data assimilation, Glob. Change Biol., 11, 89–105, doi:10.1111/j.1365-2486.2004.00891.x, 2005. 6896, 6897, 6898, 6899

Xu, T., White, L., Hui, D., and Luo, Y.: Probabilistic inversion of a terrestrial ecosystem model: analysis of uncertainty in parameter estimation and model prediction, Global Biogeochem. Cy., 20, GB2007, doi:10.1029/2005GB002468, 2006. 6896

Ziehn, T., Scholze, M., and Knorr, W.: On the capability of Monte Carlo and adjoint inversion techniques to derive posterior parameter uncertainties in terrestrial ecosystem models, Global Biogeochem. Cy., 26, GB3025, doi:10.1029/2011GB004185, 2012. 6897

GMDD
7, 6893–6948, 2014

Table 1. Description of model parameters.

	Param.	Nom. val.	Range	Description	Units
Decid. Phen.	gdd_min	100	10···250	threshold for leafout	[°C day]
	gdd_max	200	50···500	threshold for max. LAI	[°C day]
	tsmin	5	0···10	Temperature for leaffall	[°C]
	laimax	4	2···7	Seasonal max. LAI	[m^2 leaf/m^2]
	leaffall	0.1	0.03···0.95	rate of leaffall	[day^{-1}]
	lma	80	20···150	specific leaf area	[gC/m^2 leaf]
ACM	leafcn	25	*fixed*	leaf C : N ratio	[gC/gN]
	nue	7	1···20	Nitrogen use efficiency	[]
A.R.	q10_mr	2	1···4	Maintenance resp. T-sensitivity	[]
	br_mr	10^{-4}	10^{-5}···10^{-2}	Base rate for maintenance resp.	[gC m^{-2} day^{-1}/gC biomass]
	rg_frac	0.2	0.05···0.5	growth respiration fraction	[]
A.	astem	0.7	0.1···0.95	Allocation to plant stem pool	[]
Litter.	tstem	$\frac{1}{50\times365}$	$\frac{1}{250\times365}$···$\frac{1}{10\times365}$	stem turnover time	[day^{-1}]
	troot	$\frac{1}{5\times365}$	$\frac{1}{25\times365}$···$\frac{1}{365}$	root turnover time	[day^{-1}]
	tleaf	10^{-2}	10^{-3}···10^{-1}	leaf turnover time	[day^{-1}]
Decomp.	q10_hr	2	1···4	Heterotrophic resp. T-sensitivity	[]
	br_lit	$\frac{1}{2\times365}$	$\frac{1}{5\times365}$···$\frac{10}{5\times365}$	base turnover for litter	[gC m^{-2} day^{-1}/gC litter]
	br_som	$\frac{1}{30\times365}$	$\frac{1}{100\times365}$···$\frac{1}{10\times365}$	base turnover for SOM	[gC m^{-2} day^{-1}/gC SOM]
	dr	10^{-3}	10^{-4}···10^{-2}	decomposition rate	[day^{-1}]

GMDD
7, 6893–6948, 2014

Table 2. Distance correlation factors for D18. The diagonal blocks are marked according to the process the parameters contribute to, see also Fig. 2 and Table 1. The entries in the diagonal block show dependencies between parameters from the same process, while the entries in the off-diagonal block show dependencies between parameters from different processes.

	gdd_min	gdd_max	tsmin	laimax	leaffall	lma	nue	q10_mr	br_mr	rg_frac	astem	tstem	troot	tleaf	q10_hr	br_lit	br_som	dr
gdd_min	1																	
gdd_max	0.4	1																
tsmin	0	0	1															
laimax	0	0	0	1														
leaffall	0.1	0.1	0	0.1	1													
lma	0.4	0.1	0	0	0.1	1												
nue	0.1	0.2	0	0	0.3	0.7	1											
q10_mr	0.1	0.2	0	0	0	0.1	0.2	1										
br_mr	0.2	0.3	0	0	0.2	0.3	0.1	0.3	1									
rg_frac	0.2	0.3	0	0	0.4	0.2	0.7	0.3	0.3	1								
astem	0	0.1	0	0	0.1	0	0	0.2	0.1	0	1							
tstem	0	0	0	0	0.1	0	0	0.1	0.1	0	0	1						
troot	0	0	0	0	0	0	0	0.1	0	0	0	0	1					
tleaf	0	0	0	0	0	0	0	0	0	0	0	0	0	1				
q10_hr	0.1	0.1	0	0	0.1	0.2	0	0.6	0.3	0.3	0.3	0.2	0	0	1			
br_lit	0	0.2	0.1	0	0.1	0.1	0	0	0.1	0.1	0.2	0	0.1	0	0.1	1		
br_som	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
dr	0	0.1	0	0	0	0	0	0	0.1	0	0.1	0	0.1	0	0	0.2	0	1

GMDD
7, 6893–6948, 2014

Table 3. Distance correlation factors for D23. The diagonal blocks are marked according to the process the parameters contribute to, see also Fig. 2 and Table 1. The entries in the diagonal block show dependencies between parameters from the same process, while the entries in the off-diagonal block show dependencies between parameters from different processes.

	gdd_min	gdd_max	tsmin	laimax	leaffall	lma	nue	q10_mr	br_mr	rg_frac	astem	tstem	troot	tleaf	q10_hr	br_lit	br_som	dr
gdd_min	1																	
gdd_max	0.4	1																
tsmin	0	0	1															
laimax	0	0.1	0	1														
leaffall	0	0	0	0.1	1													
lma	0.3	0.1	0	0	0.2	1												
nue	0.1	0.3	0	0	0.3	0.8	1											
q10_mr	0.1	0.1	0	0	0.1	0.2	0.2	1										
br_mr	0.1	0.3	0	0	0.3	0.3	0.1	0.2	1									
rg_frac	0.2	0.3	0	0	0.3	0.4	0.8	0.1	0.1	1								
astem	0	0	0	0	0	0	0	0.2	0.1	0	1							
tstem	0	0	0	0	0	0.1	0.1	0.2	0.1	0	0	1						
troot	0.1	0	0	0	0	0	0	0.5	0	0	0.1	0	1					
tleaf	0	0	0	0	0	0	0	0.1	0	0	0	0	0	1				
q10_hr	0.1	0	0	0	0.1	0.1	0.1	0.1	0	0.1	0.1	0.1	0.3	0	1			
br_lit	0	0.1	0	0	0	0	0	0.1	0	0	0.1	0.1	0.1	0	0	1		
br_som	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0	0	1	
dr	0	0	0	0	0	0	0	0.1	0.1	0	0	0.1	0.3	0	0.1	0.1	0.3	1

GMDD
7, 6893–6948, 2014

Table 4. CRPS values for D18 and D23.

Period	CRPS-D18	CRPS-D23	% change
Jun–Aug	1.26	1.18	6
Rest of year	1.27	1.25	2
Overall	1.27	1.24	2.5

GMDD
7, 6893–6948, 2014

Table 5. Nomenclature.

ACM	Aggregate Canopy Model
CRPS	Continuous Rank Predictive Score
CRPSS	Continuous Rank Predictive Skill Score
DALEC	Data Assimilation Linked Ecosystem Carbon
FIM	Fisher Information Matrix
GPP	Gross Primary Production
GSA	Global Sensitivity Analysis
MCMC	Markov Chain Monte Carlo
NEE	Net Ecosystem Exchange
QoI	Quantity of Interest
TSC	Total Soil Carbon
TVC	Total Vegetation Carbon
$D_{KL}(p\|\|q)$	Kullback–Leibler divergence between probability densities q and p
$L_{\mathcal{D}} = p(\mathcal{D}\|\boldsymbol{\theta})$	Likelihood of the data $\mathcal{D}$ for a particular instance of model parameters $\boldsymbol{\theta}$
$p(\boldsymbol{\theta})$, $p(\boldsymbol{\theta}\|\mathcal{D})$	prior and posterior probability densities, respectively, for model parameters $\boldsymbol{\theta}$
$p(y\|\mathcal{D})$	posterior distribution for the predicted NEE data y
$p_k(y_k\|\mathcal{D})$	marginal posterior distribution for the predicted NEE component y_k
$\mathcal{R}(X,Y)$	Distance correlation between random variables X and Y
S_i	First-order Sobol index for parameter i
S_{ij}	Joint Sobol index for parameters i and j
$\boldsymbol{\theta}$	Vector of parameters for DALEC

GMDD
7, 6893–6948, 2014

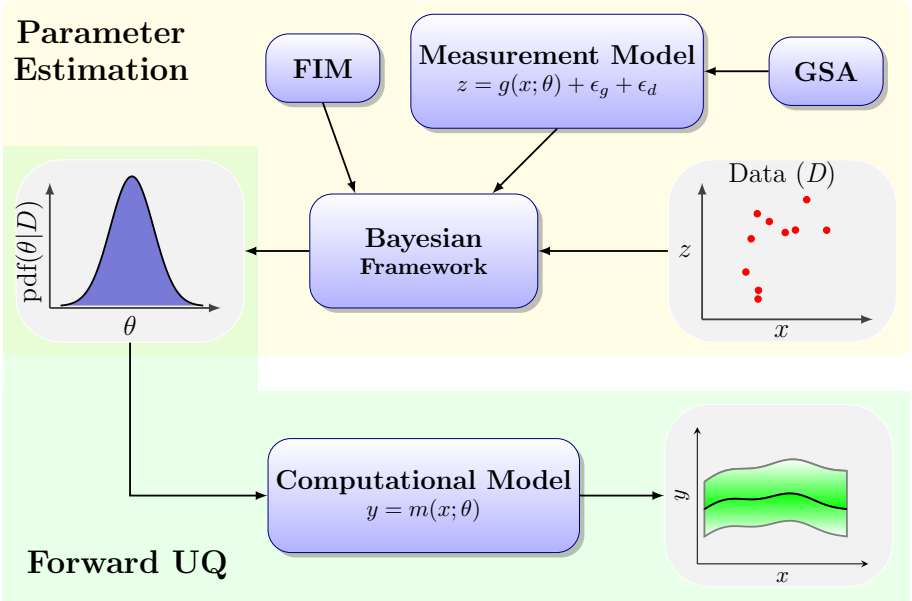

Figure 1. Schematic of parameter estimation, on yellow background, and forward UQ workflows, on green background. For this work DALEC is used as both "measurement model", g, and as "computational model", m. In the Bayesian framework the parameters estimation depends both on the model error ϵ_g and on the measurement error ϵ_d.

GMDD
7, 6893–6948, 2014

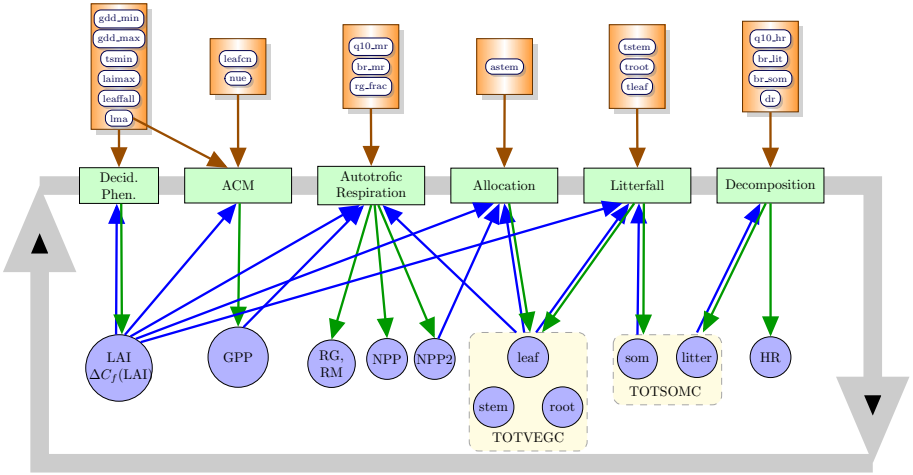

Figure 2. Schematic of processes, shown with green boxes, in DALEC with associated parameters, listed in orange boxes. The blue arrows indicate how internal parameters and QoIs, shown with blue circles, impact DALEC processes, while while the green arrows show the impact of processes on the QoI and other internal parameters.

GMDD
7, 6893–6948, 2014

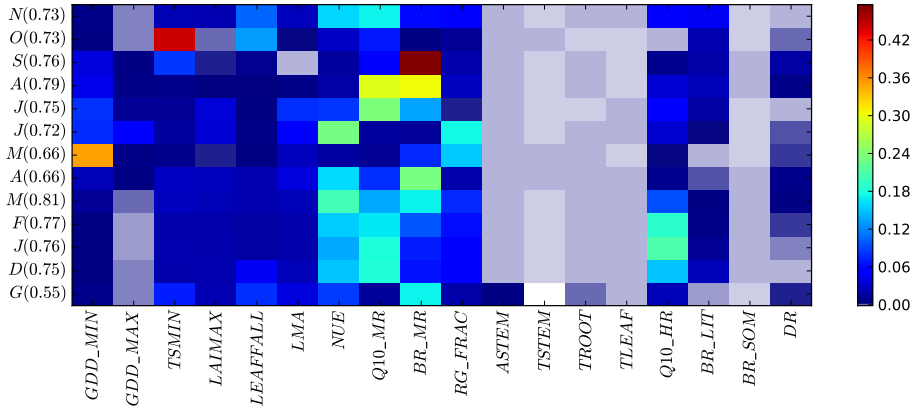

Figure 3. Matrices with first-order Sobol indices for monthly averages of NEE. Also shown are the main Sobol indices for the global average (G). The largest value $S_{\text{br\_mr}} = 0.49$ occurs for September average NEE. The sum of first-order Sobol indices for each month is shown in parentheses.

GMDD
7, 6893–6948, 2014

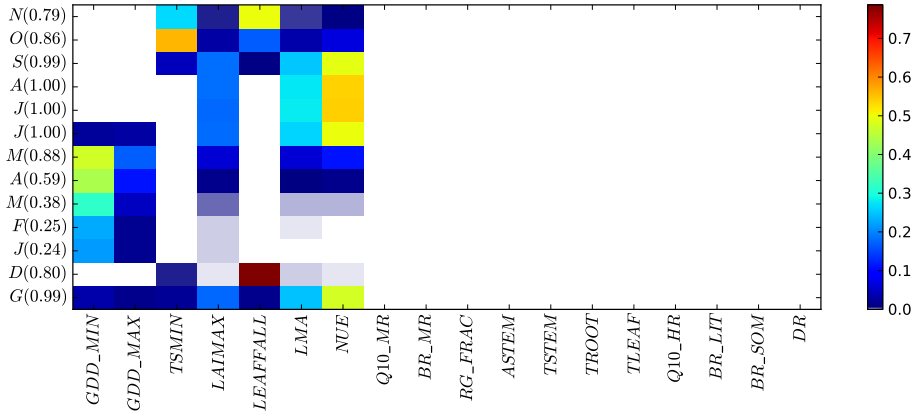

Figure 4. Matrices with main Sobol indices for monthly averages of GPP. Also shown are the main Sobol indices for the global average (G). The largest value $S_{\text{leaffall}} = 0.77$ occurs for December average GPP. The sum of first-order Sobol indices for each month is shown in parentheses.

GMDD
7, 6893–6948, 2014

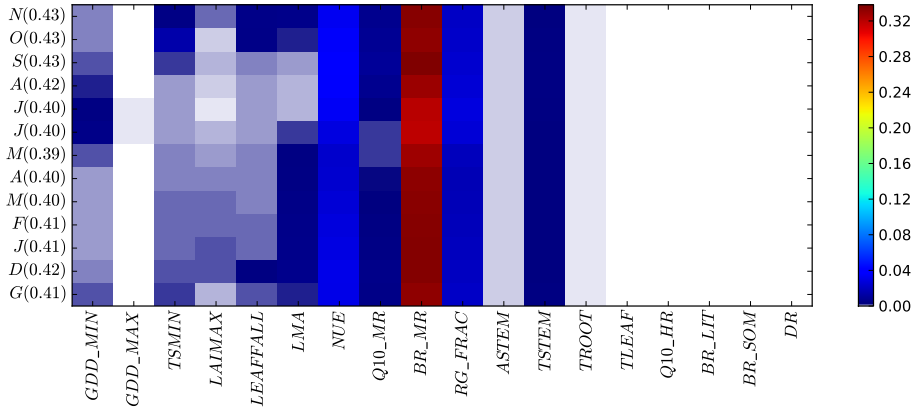

Figure 5. Matrices with main Sobol indices for monthly averages of TVC. Also shown are the main Sobol indices for the global average (G). The largest value $S_{\text{br\_mr}} = 0.36$ occurs for several TVC monthly averages. The sum of first-order Sobol indices for each month is shown in parentheses.

GMDD
7, 6893–6948, 2014

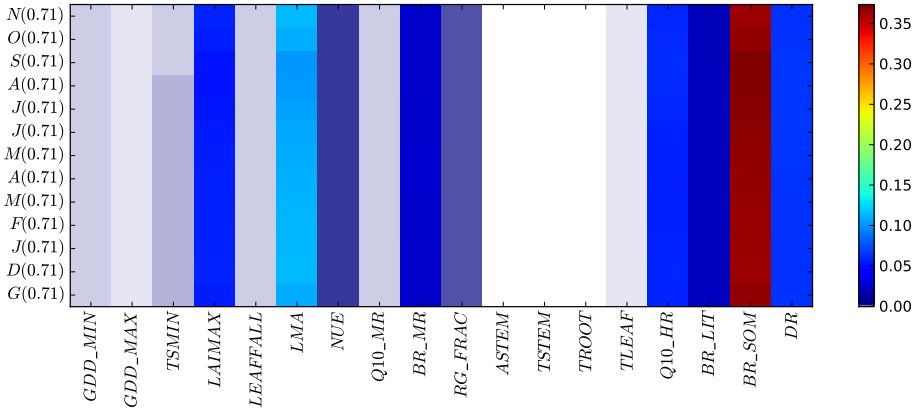

Figure 6. Matrices with main Sobol indices for monthly averages of TSC. Also shown are the main Sobol indices for the global average (G). The largest value $S_{\text{br\_som}} = 0.38$ occurs for several TSC monthly averages. The sum of first-order Sobol indices for each month is shown in parentheses.

GMDD
7, 6893–6948, 2014

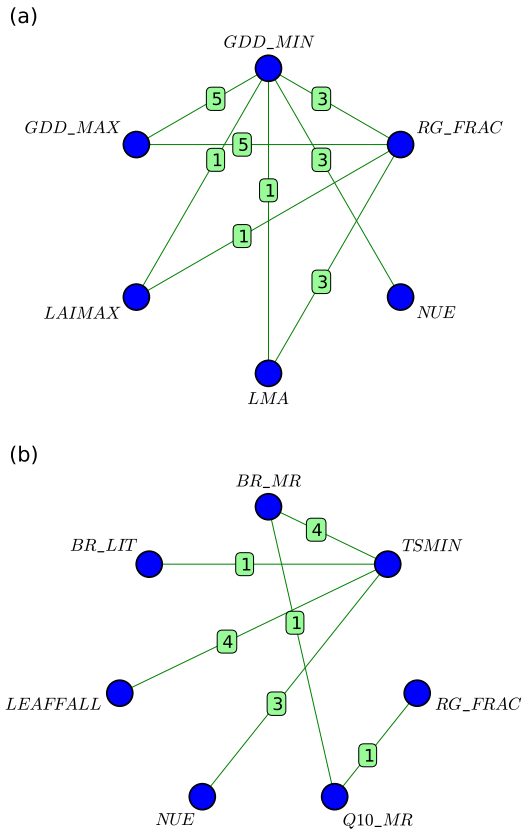

Figure 7. Relevant joint Sobol indices corresponding to monthly NEE averages for **(a)** May and **(b)** October. The labels on each line shows the magnitude, in %, of Sobol indices for the corresponding parameter pairs.

GMDD
7, 6893–6948, 2014

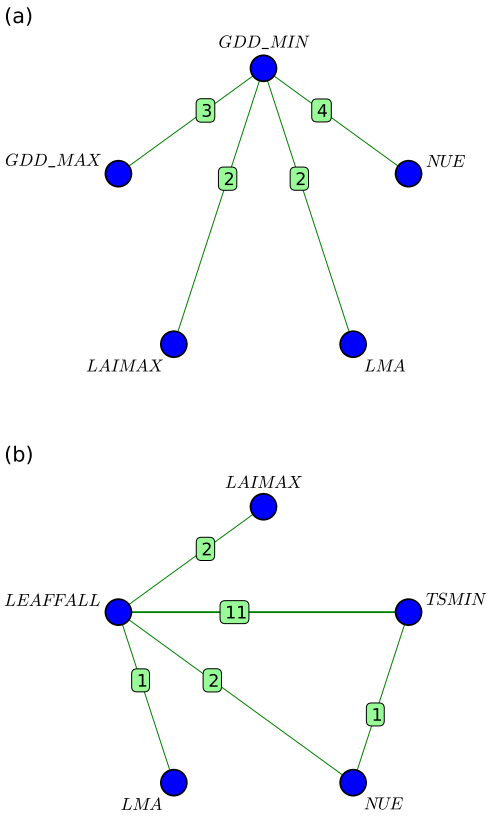

Figure 8. Relevant joint Sobol indices corresponding to monthly GPP averages for **(a)** May and **(b)** November. The labels on each line shows the magnitude, in %, of Sobol indices for the corresponding parameter pairs.

GMDD
7, 6893–6948, 2014

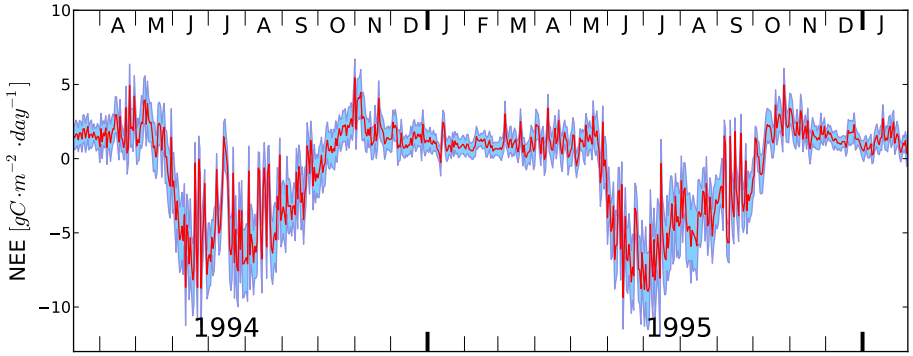

Figure 9. Snapshot of NEE observations (with red line) for the Harvard Forest site. The light blue region, bordered by thick blue lines corresponds to $\pm 2\sigma$ around the daily NEE values.

GMDD
7, 6893–6948, 2014

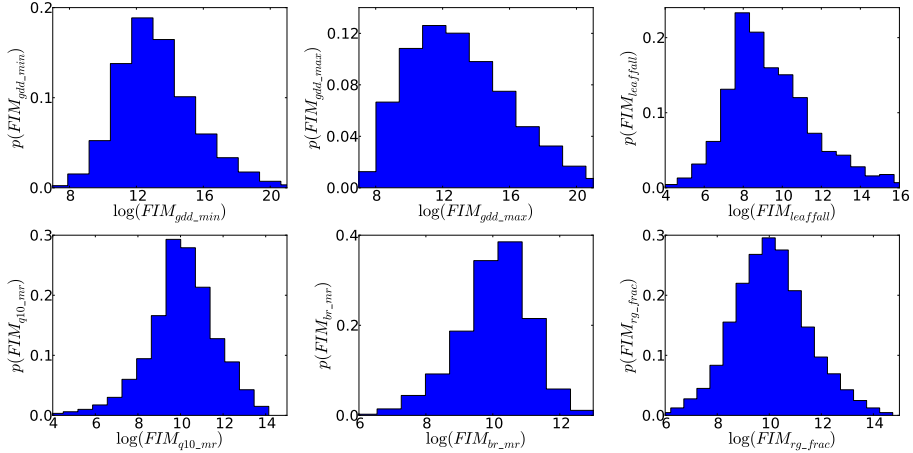

Figure 10. Normalized histograms for diagonal entries of the Fisher Information Matrix corresponding to parameters *gdd_min*, *gdd_max*, *leaffall*, *q10_mr*, *br_mr*, and *rg_frac*. Results are based on NEE and an ensemble of parameter values drawn from the corresponding prior distributions.

GMDD
7, 6893–6948, 2014

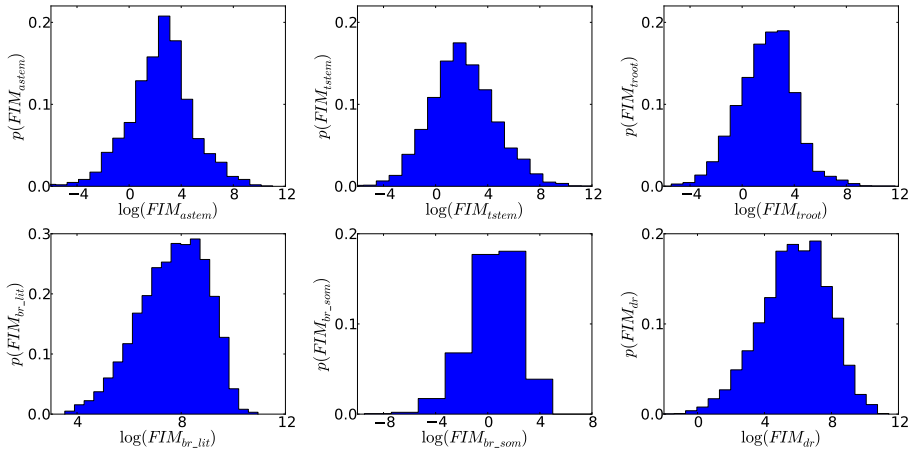

Figure 11. Normalized histograms for diagonal entries of the Fisher Information Matrix corresponding to parameters *astem*, *tstem*, *troot*, *br_lit*, *br_som*, and *dr*. Results are based on NEE and an ensemble of parameter values drawn from the corresponding prior distributions.

GMDD
7, 6893–6948, 2014

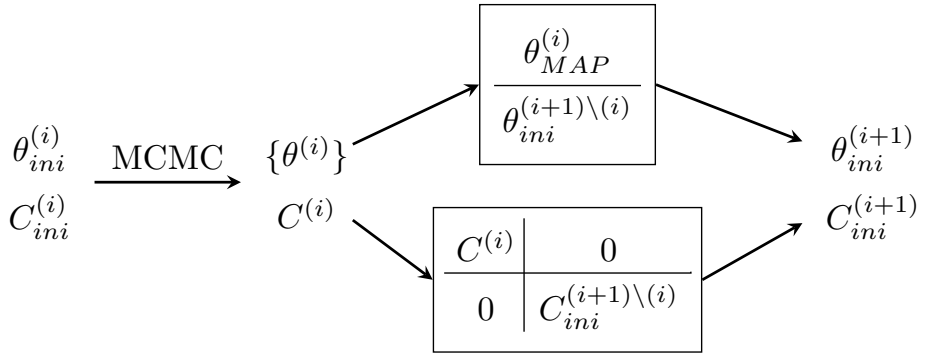

Figure 12. Schematic of the iterative process for parameter calibration. The MCMC sampling of the joint density for the set of parameters $\theta^{(i)}$ starts at $\theta_{\text{ini}}^{(i)}$ using an initial proposal covariance $\mathbf{C}_{\text{ini}}^{(i)}$. For the following iteration, $(i+1)$, the initial condition is constructed using the MAP estimate for $\theta^{(i)}$, augmented with initial conditions, in this case the nominal values, for the rest of parameters, $\theta^{(i+1)\setminus(i)}$. The initial proposal covariance $\mathbf{C}_{\text{ini}}^{(i+1)}$ is constructed based on the sample covariance matrix for $\theta^{(i)}$, augmented with an initial proposal covariance for $\theta^{(i+1)\setminus(i)}$, $\mathbf{C}_{\text{ini}}^{(i+1)\setminus(i)}$.

GMDD
7, 6893–6948, 2014

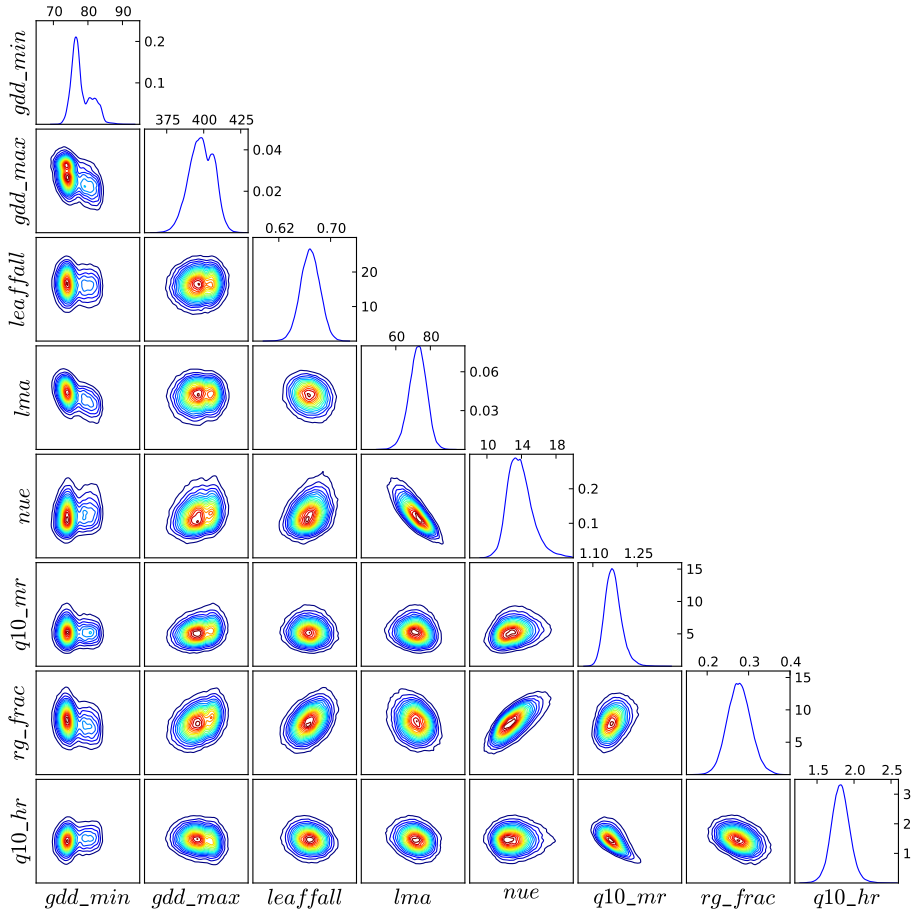

Figure 13. D18-problem: 1-D marginal and 2-D joint marginal PDFs for parameters showing distance correlation factors above 0.3, see also Table 2.

GMDD
7, 6893–6948, 2014

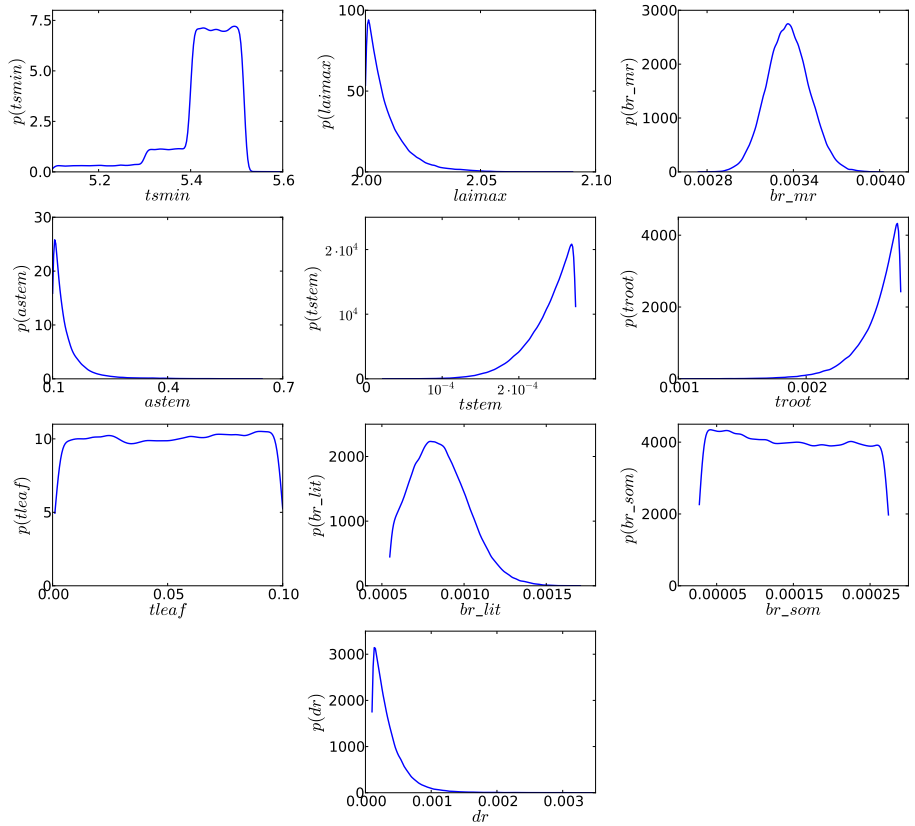

Figure 14. D18-problem: 1-D marginal PDFs for parameters showing distance correlation factors less than 0.3 with other parameters, see also Table 2.

GMDD
7, 6893–6948, 2014

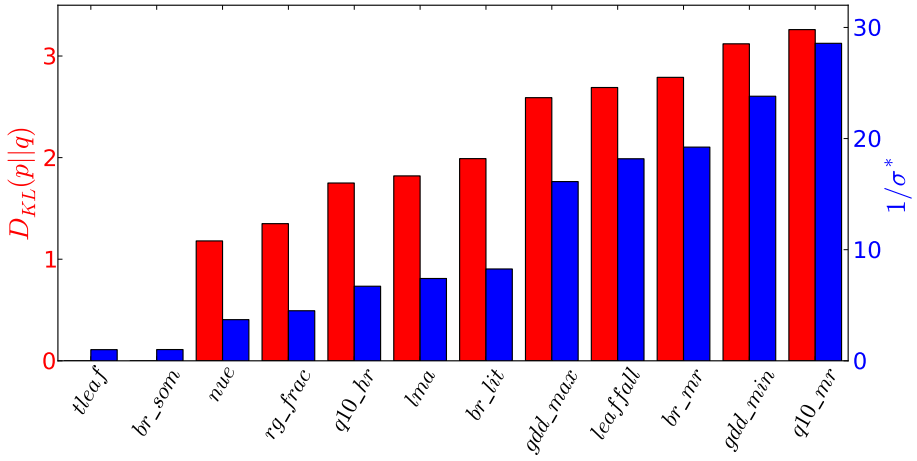

Figure 15. Kulback–Leibler divergence, $D_{KL}(p||q)$, between prior q and posterior p densities and scaled inverse standard deviation, $1/\sigma^* = \sigma_q/\sigma_p$ for select parameters.

GMDD
7, 6893–6948, 2014

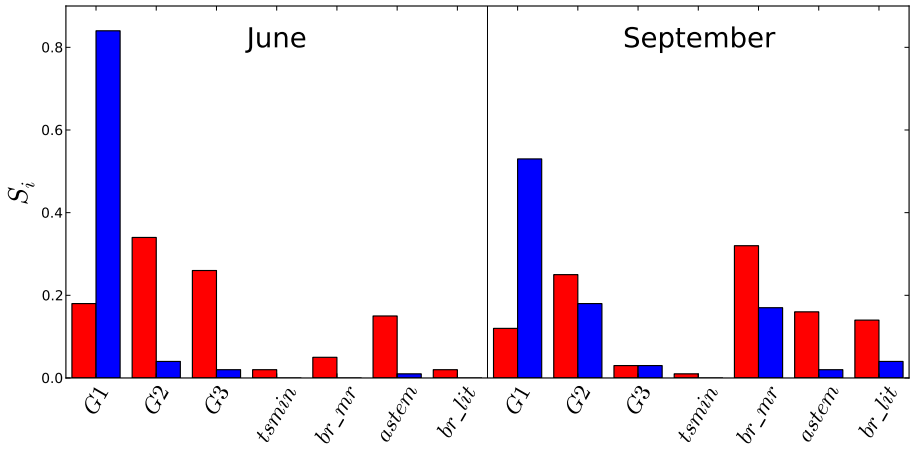

Figure 16. First order Sobol indices for June and September average NEE values. The red bars show results based on joint 3-D (for group G1) and 2-D (for groups G2 and G3) posterior distributions and 1-D marginal posterior distributions for the rest; blue bars show results based on 1-D marginal posterior distributions for all parameters; G1 = {*lma*, *nue*, *rg_frac}*, G2 = {*q10_mr*, *q10_hr}*, G3 = {*gdd_min*, *gdd_max}*.

GMDD
7, 6893–6948, 2014

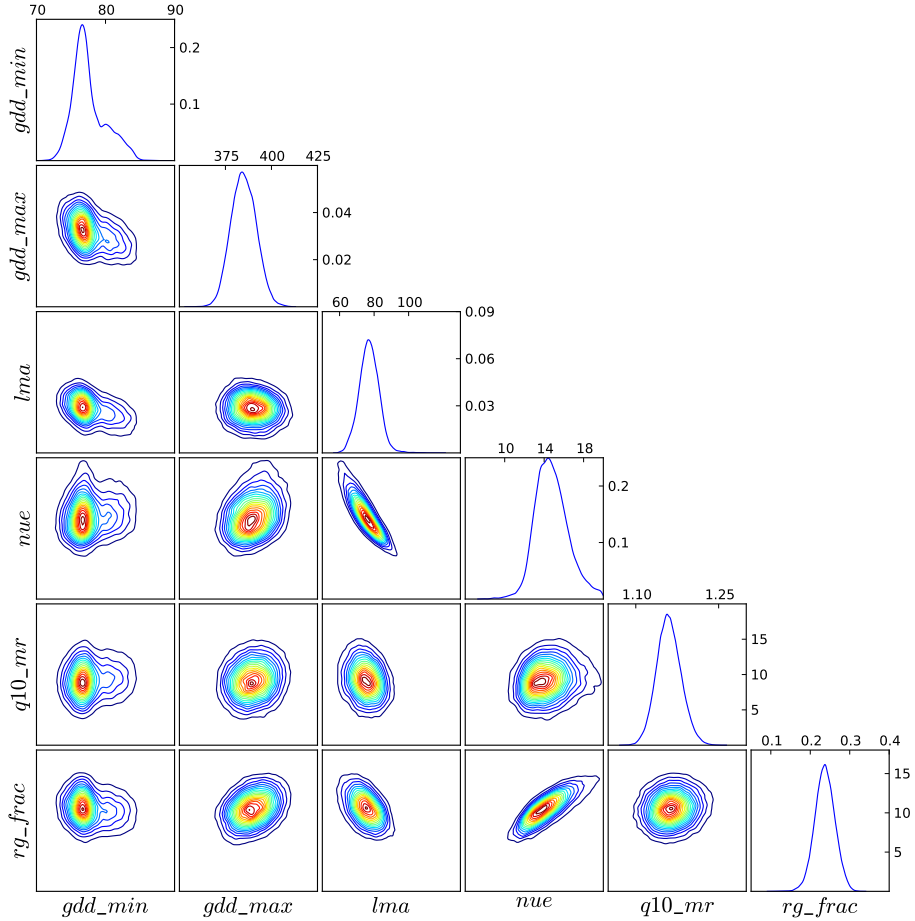

Figure 17. D23-problem: 1-D marginal and 2-D joint marginal PDFs for parameters showing distance correlation factors above 0.3, see also Table 3.

GMDD
7, 6893–6948, 2014

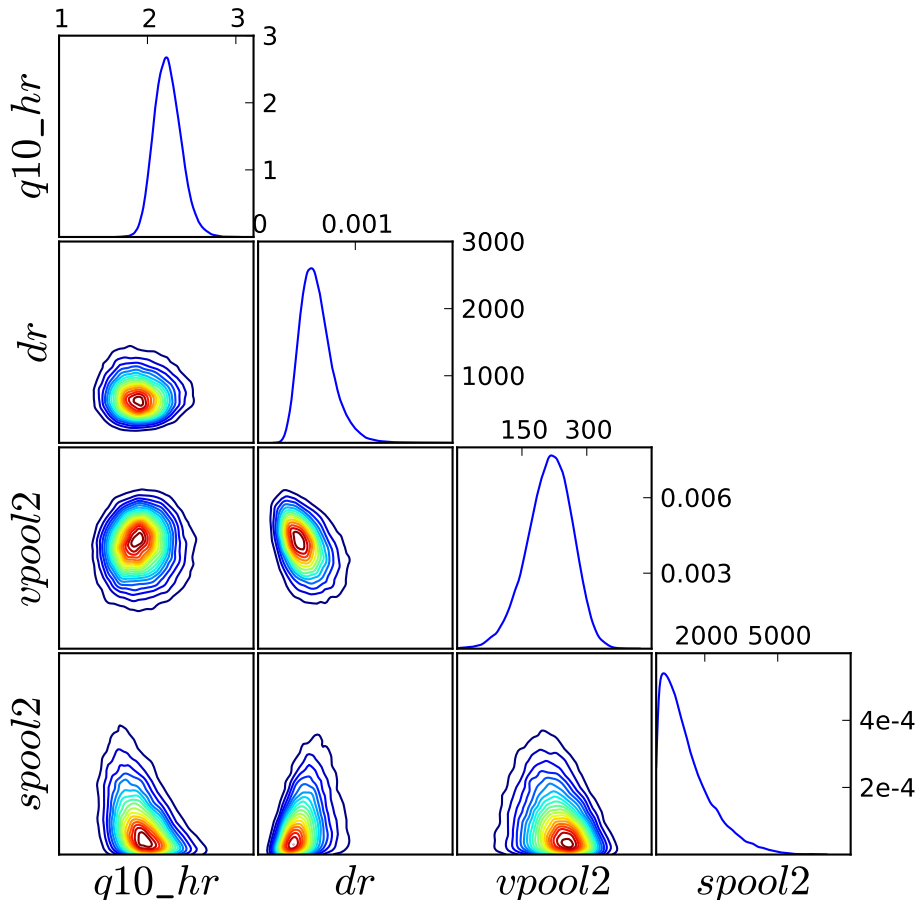

Figure 18. D23-problem: 1-D marginal and 2-D joint marginal PDFs for parameters correlated with the Carbon pools.

GMDD
7, 6893–6948, 2014

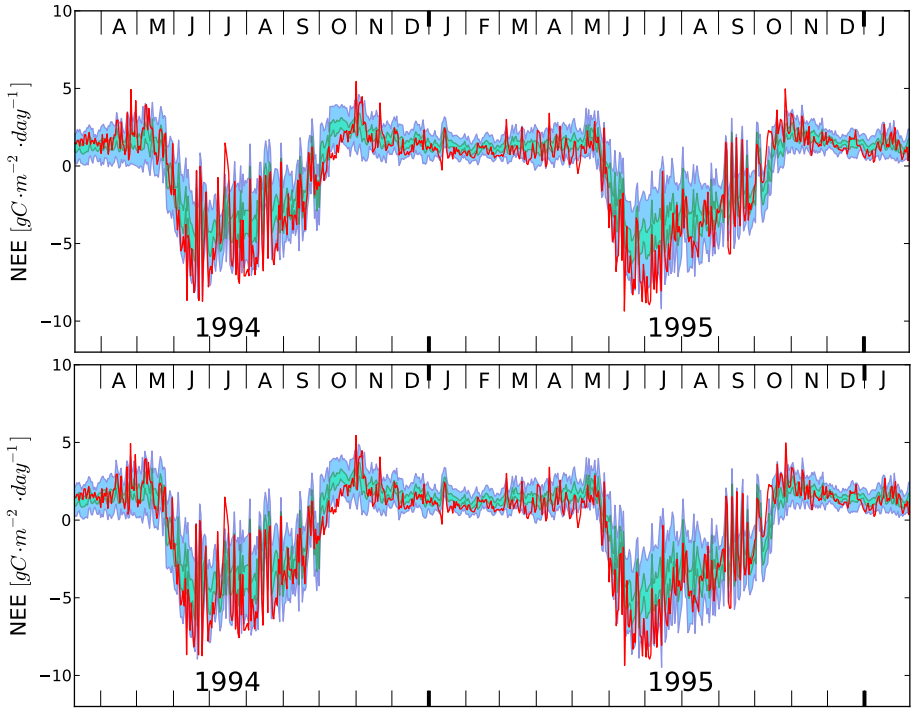

Figure 19. Posterior predictive distributions using the calibration results for D18 (top frame) and D23 (bottom frame) presented in Sect. 4. The blue regions correspond to the daily 5–95 % quantile range and the green regions to 25–75 % quantile range. The red line shows the daily NEE observations.

GMDD
7, 6893–6948, 2014