

Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling

WOLFGANG KNORR*† and JENS KATTGE†

*QUEST Department of Earth Sciences, Wills Memorial Building, University of Bristol, Queen's Road BS8 1RJ, UK,

†Max-Planck Institute for Biogeochemistry Hans-Knöll-Str. 10 07745 Jena, Germany

Abstract

Effective measures to counter the rising levels of carbon dioxide in the Earth's atmosphere require that we better understand the functioning of the global carbon cycle. Uncertainties about, in particular, the terrestrial carbon cycle's response to climate change remain high. We use a well-known stochastic inversion technique originally developed in nuclear physics, the Metropolis algorithm, to determine the full probability density functions (PDFs) of parameters of a terrestrial ecosystem model. By thus assimilating half-hourly eddy covariance measurements of CO₂ and water fluxes, we can substantially reduce the uncertainty of approximately five model parameters, depending on prior uncertainties. Further analysis of the posterior PDF shows that almost all parameters are nearly Gaussian distributed, and reveals some distinct groups of parameters that are constrained together. We show that after assimilating only 7 days of measurements, uncertainties for net carbon uptake over 2 years for the forest site can be substantially reduced, with the median estimate in excellent agreement with measurements.

Keywords: carbon cycle, climate change, ecosystem models, eddy covariance, Monte Carlo, parameter estimation, photosynthesis, probability density function, respiration

Received 28 June 2004; revised version received 24 January 2005; and accepted 14 February 2005

Introduction

Only about half of the increasing emissions of CO₂ from human activities currently remain in the atmosphere (Prentice *et al.*, 2001). The remainder is taken up by both the oceans and the terrestrial biosphere, to roughly equal amounts (Joos *et al.*, 2003). This current carbon sink in the terrestrial biosphere is, by some models at least, predicted to turn into a source (Cox *et al.*, 2000; Cramer *et al.*, 2001; Friedlingstein *et al.*, 2003). Better quantification of the exchange fluxes of CO₂ between the terrestrial biosphere and the atmosphere and better understanding of the underlying processes are therefore of foremost importance for the design of efficient climate protection strategies. Terrestrial ecosystem models (TEMs) have been used extensively to study the processes leading to either carbon loss or gain by

the land biota (McGuire *et al.*, 2001; Prentice *et al.*, 2001). However, results still vary significantly because of differences between models (Cramer *et al.*, 1999). While only very few studies using TEMs have considered uncertainties in fluxes as a result of parameter uncertainties, Knorr & Heimann (2001a, b) have shown that uncertainties of TEM process parameters lead at least to the same spread of simulated atmosphere–vegetation carbon fluxes as intermodel differences.

More recently, Kaminski *et al.* (2002) have shown that TEMs can be combined with atmospheric transport inversion techniques. By using an additional process model and a Bayesian approach to parameter inversion, such inversions are both better constrained than transport inversions and allow inferences about the underlying processes. An example of a more complex Carbon Cycle Data Assimilation System (CCDAS) is given by Rayner *et al.* (2005). CCDAS requires one to specify prior means and error covariance matrices of model parameters, as an approximation of the prior probability density function (PDF) of parameters. To

Correspondence: Wolfgang Knorr, tel. +44-117-3315133, fax: +44 117 925 3385; e-mail: wolfgang.knorr@bristol.ac.uk

generate and analyze such a PDF is one purpose of the present study.

Few attempts exist at quantifying uncertainty ranges based directly on experimental data (Knorr, 2000; White *et al.*, 2000; Knorr & Heimann, 2001a). It is, therefore, of general interest to utilize the still growing amount of eddy covariance measurements of CO₂ and water fluxes (FLUXNET, Global Carbon Project, 2003) for ecosystem model parameter estimation. Wang *et al.* (2001) used a non-Bayesian parameter optimization and showed that for their model, up to five parameters could be estimated on the basis of eddy covariance measurements of CO₂, water, heat and ground heat fluxes. Prior knowledge of parameter values was used to initialize the parameters that were optimized, to set the parameters that remained unaffected by the optimization, and to determine reasonable limits for the space of parameter solutions allowed. The result is a set of model parameters that are either based fully on prior estimates, or fully on the inversion against measurements.

Here, Bayesian methods offer a more consistent approach by combining prior knowledge with the additional information gained from the inversion. This does not only allow the simultaneous determination of all parameters, it also allows considering prior knowledge consistently for all parameters. Weakly constrained parameters are, therefore, given an appropriate uncertainty range instead of being excluded *a priori* from the optimization. The method can be applied to global scale inversions (Rayner *et al.*, 2005), or to sites using flux measurements as a model constraint.

With this contribution, we will demonstrate a general method for Bayesian parameter estimation of complex, process-based TEM, where parameter uncertainty ranges are derived from systematic sampling of the complete PDF. By comparing prior and posterior uncertainty ranges of parameters, it will be determined which parameters can be constrained by eddy covariance measurements of CO₂ and water fluxes for a given set of prior parameter uncertainties and for a given measurement error, using a particular TEM. The analysis of covariances is then used to determine which parameter values cannot be determined independently by the data. Finally, simulations with the constrained model – using both the complete PDF or its first two moments – are carried out for much longer time series than those used during the parameter estimation, to test the validity of the parameterization across time. Here, we also assess whether an approximation to the full PDF as used by CCDAS (means and error covariances) sufficiently represents uncertainties in net CO₂ fluxes. The

method is thus presented as a prototype for an initial step of CCDAS that allows the exploitation of widely available site-specific flux data through constraining model parameters.

Methods

Monte Carlo inversion

The given task of this study is to determine the probability distribution of a vector of model parameters \mathbf{p} , given a set of measurements \mathbf{f} , in this case fluxes. Whether a given vector \mathbf{p} agrees with \mathbf{f} is determined by running the model \mathbf{M} , such that

$$\mathbf{f}_M(\mathbf{p}) = \mathbf{M}(\mathbf{p}, \mathbf{c}, \mathbf{s}). \quad (1)$$

\mathbf{f}_M is the vector of model-simulated measurements, and \mathbf{c} and \mathbf{s} vectors of environmental boundary conditions and model state variables, respectively. \mathbf{f} , \mathbf{f}_M , \mathbf{c} and \mathbf{s} contain values across both time and types of data (CO₂, water and heat fluxes; temperature, solar radiation, humidity; soil moisture and leaf area index (LAI)), while \mathbf{p} is assumed invariable in time. For a process-based TEM, \mathbf{M} is usually nonlinear and too complex to be expressed as a set of standard mathematical functions. According to Mosegaard (1998), this amounts to a *general inverse problem* that can most efficiently be solved by direct sampling of the PDF in parameter space using Monte Carlo techniques. Developed for applications in nuclear physics (Metropolis *et al.*, 1953), and later geophysics (Mosegaard & Tarantola, 1995; Mosegaard, 1998; Mosegaard & Rygaard-Hjalsted, 1999), it is now widely used in other fields of environmental modeling. It consists of a stochastic technique that generates a random set of points $\mathbf{p}^1, \dots, \mathbf{p}^N$ in parameter space with a distribution that approximates any given function $f(\mathbf{p})$ for large values of N . For a Bayesian inversion, this function is chosen as the posterior PDF of model parameters, given by

$$f(\mathbf{p}) = \nu L(\mathbf{p})\rho(\mathbf{p}) \quad (2)$$

with a normalization constant, ν (Mosegaard & Sambridge, 2002). $L(\mathbf{p})$ is the likelihood function, which expresses the misfit between model-derived values and measurements in relation to measurement error, and $\rho(\mathbf{p})$ is the prior probability distribution of parameters. Errors representing missing or incorrect processes were neglected in this study. The likelihood function is expressed as the negative exponential of the misfit against measurements, $J_f(\mathbf{p})$, such that

$$L(\mathbf{p}) = \exp\{-J_f(\mathbf{p})\} \quad (3a)$$

with

$$J_f(\mathbf{p}) = \frac{1}{2}(\mathbf{M}(\mathbf{p}) - \mathbf{f})^T \mathbf{C}_f^{-1}(\mathbf{M}(\mathbf{p}) - \mathbf{f}). \quad (3b)$$

C_f is the error covariance matrix of the measurements, and the superscript T denotes the transposed vector. Similarly, the prior probability, $\rho(\mathbf{p})$, can be written as

$$\rho(\mathbf{p}) = \exp\{-J_p(\mathbf{p})\} \quad (4a)$$

and

$$J_p(\mathbf{p}) = \frac{1}{2}(\mathbf{p} - \mathbf{p}_0)^T \mathbf{C}_p^{-1}(\mathbf{p} - \mathbf{p}_0) \quad (4b)$$

with \mathbf{p}_0 , the vector of prior parameter values, and \mathbf{C}_p , the error covariance matrix of the priors. The formulation assumes Gaussian uncertainties for both prior parameters, \mathbf{p} , and the assimilated data, \mathbf{f} .

Standard inversion techniques aim at inferring the global minimum of the function, $J(\mathbf{p}) = \exp\{-f(\mathbf{p})\}$. In the case of Monte Carlo inversion, the generated series of sample points, $\mathbf{p}^1, \dots, \mathbf{p}^N$, simply has a distribution with its highest density in the vicinity of the maximum of $f(\mathbf{p})$. If the objective is to gain understanding of the probability distribution of parameters rather than to find the exact optimum, this technique has obvious advantages. The sampled distribution can subsequently be used to compute the expected value of any desired variable or expression, x , under the predefined PDF, $f(\mathbf{p})$:

$$\langle x \rangle = \int x(\mathbf{p})f(\mathbf{p}) \, d\mathbf{p} \cong \frac{1}{N} \sum_{i=1}^N x(\mathbf{p}^i). \quad (5)$$

To assess to what degree the distribution of model parameters deviates from a Gaussian one, it is also possible to compute the projection of the multidimensional PDF onto the dimension of a single parameter (importance sampling) from

$$\begin{aligned} f_i(p) &= \int f(\mathbf{p}) \, dp_1 \dots dp_{i-1} dp_{i+1} \dots dp_n \\ &\cong \frac{1}{\varepsilon N} \sum_{j=1}^N I_{p-\varepsilon/2, p+\varepsilon/2}(\mathbf{p}^j). \end{aligned} \quad (6)$$

$I_{a,b}(x)$ denotes the interval function, which is 1 if $a \leq x < b$, else 0, and ε an appropriately chosen resolution parameter.

The complete method of Monte Carlo inversion is described in detail by Mosegaard & Tarantola (1995) and reviewed by Mosegaard & Sambridge (2002). We always perform one iteration starting from the prior set of parameters (i.e. $\mathbf{p}^1 = \mathbf{p}_0$). For some cases (see Results), we add an ensemble of Monte Carlo integrations with varying starting points in the way suggested by Gelman *et al.* (1995). To generate subsequent values $\mathbf{p}^2, \mathbf{p}^3, \dots$ in each series, a new point is tried by varying all vector elements by some step, $\Delta\mathbf{p}$, chosen with a Gaussian distributed random number generator with mean zero and standard deviation set for each parameter separately to the prior uncertainty times an appropriately chosen step-length factor. The new point,

$\mathbf{p}^i + \Delta\mathbf{p}$ at step i of the iteration, is accepted or rejected according to a two-step version of the Metropolis algorithm: The first step is always accepted, if $\rho(\mathbf{p}^i + \Delta\mathbf{p})/\rho(\mathbf{p}) \geq 1$, and it is accepted with a probability of $\rho(\mathbf{p}^i + \Delta\mathbf{p})/\rho(\mathbf{p})$ if $\rho(\mathbf{p}^i + \Delta\mathbf{p})/\rho(\mathbf{p}) < 1$. Acceptance with a probability of < 1 – the latter case – is carried out by generating a uniformly distributed random number between 0 and 1. Only if this number is less than the chosen probability, the first step is accepted. The second step is assessed in the same way as the first, except that the prior probability $\rho(\mathbf{p})$ is replaced by the likelihood function $L(\mathbf{p})$. Only if both steps are accepted, the next point in the series is $\mathbf{p}^{i+1} = \mathbf{p}^i + \Delta\mathbf{p}$, else $\mathbf{p}^{i+1} = \mathbf{p}^i$ (i.e. the previous point is chosen again). We adjust the step length for each parameter to values which lead to an average acceptance rate of the new points around 0.3 (Gelman *et al.*, 1995). Only the second step requires model execution.

Simulations

As a demonstration of the Monte Carlo method, we chose two different photosynthesis models and two setups with a reduced and a more extensive part of the Biosphere Energy-Transfer Hydrology (BETHY) model. The reduced version of BETHY is used together with the C4 photosynthesis model and excludes the heterotrophic respiration part. Compared with the C3 version with heterotrophic respiration, this reduces the number of free parameters from 23 to 14. The C4 version uses eddy covariance measurements, by Kim & Verma (1991), from the first ISLSCP field experiment (FIFE) grassland experimental site in Kansas, and the C3 version data from the Loobos pine forest site in the Netherlands (Dolman *et al.*, 2002).

Input and flux data

The FIFE site in north-eastern Kansas, USA (39°03'N, 96°32'W) was dominated by the C4 tallgrass species *Andropogon gerardii*, *Sorghastrum nutans* and *Panicum virgatum*. The implementation of BETHY for this site is also described by Knorr (1997). In this case, we assimilated daytime data of net canopy assimilation (gross primary productivity (GPP) minus total-canopy leaf respiration) derived from eddy covariance measurements of net ecosystem exchange (NEE) by subtracting soil and plant, excluding leaf, respiration rates derived from night-time CO₂ fluxes, for 4 days representative of the 1987 growing season: June 5, July 2, July 30, and August 20. July 30 was the only date with severe drought. We also assimilated daytime canopy conductance values for the same dates and times that were obtained through inversion of the

Penman–Monteith equation against daytime latent energy flux measurements. Photosynthetically active radiation (PAR), air temperature, vapor pressure deficit (VPD) and relative plant available soil moisture (w/w_m , Eqns (A12) and (A17)) were used as input data. All data were taken from Kim & Verma (1991). Global radiation was computed from Julian day, longitude and latitude, while wind speed and free-air CO₂ concentration were left constant at 3 ms⁻¹ and 355 ppm, respectively. We used a relative uncertainty of 20% for both net canopy assimilation and canopy conductance, with a threshold of 3.0 μmol m⁻² s⁻¹ and 1.5 mm s⁻¹, respectively.

The vegetation at the Loobos site (the Netherlands, 52°10'N, 5°74'E) was dominated by *Pinus sylvestris* with an understory of the grass *Deschampsia flexuosa* (Dolman *et al.*, 2002). Global radiation, PAR, air temperature, ambient CO₂ concentration, wind speed, VPD and total soil water content, w_{tot} , were used as input data. Soil water content at wilting point (2.5% vol.) and at field capacity (12.4% vol.) were estimated from soil texture information. We assimilated half hourly values of NEE and latent energy flux (LE) from 7 days, 5 in 1997 (January 14, March 3, July 9, September 24, October 25), and 2 in 1998 (May 15, August 9). Those days were chosen to represent typical conditions of the various seasons, have as complete data coverage as possible, and at least some cloud-free conditions during daytime. We also required that there was no rainfall during the day and the day before so that soil and canopy evaporation could be neglected. The uncertainty of NEE was taken to be 20% of NEE during day and 50% of NEE during night, accounting for low wind speed and little turbulence during night-times. The minimum uncertainty threshold was set to 3.0 μmol m⁻² s⁻¹. Uncertainties of LE were considered to be 20% of LE, with a threshold of 22.0 W m⁻². Uncertainties of input data were not considered for either site.

Prior model parameter values and uncertainties

All model parameters and their prior values are listed in Table 1. Their choice is based on the model description of BETHY (Knorr, 2000), with a few exceptions: the value for $r_{j_m v_m}$ was derived from data by Wullschlegel (1993), Medlyn *et al.* (2002) and Leuning (2002); k^{25} and E_k follow Knorr (1997); E_{R_d} was set to the value cited by Kim & Verma (1991); $f_{R,leaf}$ was modified for one plant respiration rate instead of separate maintenance and growth respiration; R_{het}^0 was set to a value for which the heterotrophic respiration model (at *a priori* parameter values) driven with data from the Loobos site reproduces the range of measured soil respiration rates given in Raich *et al.* (2002) and

Reichstein *et al.* (2003); Q_{10} follows Raich *et al.* (2002); w_{pwp} was derived from soil texture information and soil water potential relations from Schachtschabel *et al.* (1992); and a_v was set to the upper bound of values given by Knorr (2000).

We distinguish between model parameters (Table 1) and parameters used by the Metropolis algorithm written $\mathbf{p} = \{p_1, \dots, p_N\}$. For the model parameters, we assume *a priori* that (a) they are always greater than zero and (b) that they have a log-normal distribution. The parameters used by the optimization will be called log-normal parameters and have a prior value of 1 and a prior uncertainty of 0.125, 0.25, and 0.5. The log-normalized parameters have a prior distribution that is Gaussian (Eqn (4b)). To transform from log-normalized parameter, \mathbf{p} , to model parameter $\underline{\mathbf{p}}$ (see previous section), we use the following equation:

$$\underline{p}_i = \exp(p_i - 1) \underline{p}_{i,0} \quad (7)$$

$\underline{p}_{i,0}$ is the prior value in model parameter space as listed in Table 1. Note that throughout the remainder of this analysis, we will show and discuss exclusively the log-normalized parameters (with the exception of Table 3), because their prior values are identical (1.0) and their posterior values easier to intercompare.

An exceptional model parameter is f_{C_i} , for which we require $0 \leq f_C \leq 1$. Instead of a log-normal distribution in model parameter space, we choose a probability distribution that is defined by a normal distribution but is cut off at 0 and 1. $f_{C_i,0}$ is the prior estimate of f_{C_i} , and $f_{C_i} = p_k f_{C_i,0}$ replaces Eqn (7), where k is the parameter index for f_{C_i} .

The vector of prior log-normalized parameters is thus $\mathbf{p}_0 = \{1, \dots, 1\}$, and \mathbf{C}_p , the error covariance matrix of the priors:

$$C_{p_{i,j}} = \begin{cases} x^2 & \text{if } i = j \\ 0 & \text{else,} \end{cases}$$

where x is 0.125, 0.25, or 0.5, as above. Covariances for priors are assumed to be zero. For the prior probability distribution, $\rho(\mathbf{p})$ (Eqns (4)), we have the additional condition

$$\rho(\mathbf{p}) = \begin{cases} 0 & \text{if } p_k \leq 0 \text{ or } p_k \geq 1/f_{C_i,0}, \\ \exp\{-J_p(\mathbf{p})\} & \text{else.} \end{cases}$$

Results

We will, first, show results to demonstrate convergence of the algorithm. Next, optimized parameter values will be described by their means, standard errors, and covariances, all in the space of log-normalized parameters (cf. Eqn (7)). Comparison with prior means and errors indicates about how many parameters we have learned something through the assimilation of the eddy

covariance data. We also assess for which parameters the posterior PDF differs from the Gaussian distribution assumed for the prior PDF. For the Loobos site, we eventually compute the cumulative NEE with and without optimized parameters over a period of 2 years to test the validity of parameterizations across time and to assess to what degree the inversion has lead to a constraint on the modeled longer-term ecosystem carbon balance.

Convergence of the algorithm

To insure convergence, we performed rather long integrations with 500 000 iterations (and more in one case). For the two cases with 0.25 prior uncertainty, we produced a series of six independent simulations starting from different points in parameter space: the

prior parameter vector, $\mathbf{p}_0 = \{1, \dots, 1\}$ in the space of log-normalized parameters, and points shifted away from the estimated posterior optimum, \mathbf{p}' , by one to several times the posterior standard deviations, $\sigma' = \{\sigma'_1, \dots, \sigma'_n\}$ estimated from preliminary simulations. For FIFE, the starting points were \mathbf{p}_0 , $\mathbf{p}' + \sigma'$, $\mathbf{p}' + 2\sigma'$, $\mathbf{p}' + 3\sigma'$, $\mathbf{p}' - \sigma'$, $\mathbf{p}' - 2\sigma'$, for Loobos \mathbf{p}_0 , $\mathbf{p}' + 2\sigma'$, $\mathbf{p}' - 2\sigma'$, $\mathbf{p}' + 4\sigma'$, $\mathbf{p}' + 4\{\sigma'_1, -\sigma'_2, +\sigma'_3, \dots\}$ and $\mathbf{p}' - 4\{\sigma'_1, -\sigma'_2, +\sigma'_3, \dots\}$. Sampling was done at every 10th iteration to avoid correlations between subsequent samplings (i.e. 500 000 iterations yielded 50 000 samplings of parameters of interest sampled according to Eqn (5)). To determine at which iteration the sequences have converged to a common maximum, as opposed to sampling around local maxima, we applied Gelman's criterion of convergence (Gelman *et al.*, 1995) for all parameters. This test of convergence,

Table 1 Parameters of BETHY that were used as priors in the inversion against eddy covariance measurements

Symbol	Description	Value	Unit	Eqn	C3	C4
<i>Photosynthesis</i>						
α_q	Quantum efficiency of photon capture (C3)	0.28	$\text{mol}(\text{e}^-)\text{mol}^{-1}$	A1c	X	
V_m^{25}	Maximum carboxylation rate at 25 °C (C3)	29	$\mu\text{mol m}^{-2}\text{s}^{-1}$	A2	X	
V_m^{25}	Maximum carboxylation rate at 25 °C (C4)	8	$\mu\text{mol m}^{-2}\text{s}^{-1}$	A2		X
E_{V_m}	Activation energy of V_m	58 520	J mol^{-1}	A2	X	X
r_{m,V_m}	Ratio of J_m to V_m at 25 °C	1.79	–	A3	X	
Γ_*^{25}	CO ₂ compensation point without dark respiration at 25 °C	42.5	$\mu\text{mol mol}^{-1}$	A4	X	
K_C^{25}	Michaelis–Menten constant for carboxylation at 25 °C	460	$\mu\text{mol mol}^{-1}$	A5	X	
E_{K_C}	Activation energy of K_C	59 356	J mol^{-1}	A5	X	
K_O^{25}	Michaelis–Menten constant for carboxylation at 25 °C	0.33	mol mol^{-1}	A6	X	
E_{K_O}	Activation energy of K_O	35 948	J mol^{-1}	A6	X	
α_i	Quantum efficiency of photon capture (C4)	0.04	$\text{mol}(\text{CO}_2)\text{mol}^{-1}$	A7		X
k^{25}	CO ₂ specificity at 25 °C	0.14	$\text{mol m}^{-2}\text{s}^{-1}$	A8		X
E_k	Activation energy of k	50 967	J mol^{-1}	A8		X
<i>Carbon balance</i>						
f_{R_d}	Ratio of leaf dark respiration at 25 °C and V_m^{25} (C3)	0.011	–	A10	X	
f_{R_d}	Ratio of leaf dark respiration at 25 °C and V_m^{25} (C4)	0.042	–	A10		X
E_{R_d}	Activation energy of leaf dark respiration	45 000	J mol^{-1}	A10	X	X
$f_{R,\text{leaf}}$	Ratio of canopy to total plant respiration	0.5	–	A11	X	
R_{het}^0	Heterotrophic respiration at 0 °C and field capacity	2.07	$\mu\text{mol m}^{-2}\text{s}^{-1}$	A12	X	
κ	Soil moisture factor of heterotrophic respiration	1	–	A12	X	
Q_{10}	Temperature dependency of heterotrophic respiration	1.72	–	A12	X	
<i>Stomatal control</i>						
w_{pwp}	Soil water content at permanent wilting point	2.5	vol%	–	X	
f_{C_i}	Non water limited ratio of $C_{i,0}$ and C_a (C)	0.87	–	A14	X	
f_{C_i}	Non water limited ratio of $C_{i,0}$ and C_a (C4)	0.67	–	A14		X
c_w	Maximum water supply rate of root system	1	mm h^{-1}	A17	X	X
<i>Energy and radiation balance</i>						
ω	Single scattering albedo of leaves	0.12	–	–	X	X
a_v	Albedo of close vegetation surface cover	0.2	–	A18	X	X
a_s	Fraction of solar radiation absorbed by soil under close canopy	0.05	–	A18	X	X
ϵ_s	Sky emissivity factor	0.64	–	A19	X	X
$g_{a,v}$	Vegetation factor of atmospheric conductance	0.04	–	A20	X	

BETHY, Biosphere Energy-Transfer Hydrology.

designed for practical purposes, yields a reduction factor defined as the square root of the ratio of the mean variances of the various sequences divided by the variance of the means yielded by each sequence. This reduction factor is sampled from exactly the second half of the series up to the iteration indicated (see below the discussion of 'burn-in time', and Fig. 1e, f). If all sequences sample the same area of the parameter space, the reduction factor will approach a value of one, values much greater than one indicate sampling of different regions around local minima by the different sequences.

The parameters that took longest to reach a common maximum, according to Gelman's criterion, were α_i for FIFE and f_{C_i} for Loobos. The evolution of the estimated mean values is shown in Fig. 1a and b, respectively, again for every 10th iteration. Also shown are one fast converging parameter, and the parameter that was most highly correlated to the first. Note that in Fig. 1b, E_{V_m} appears to be converging more slowly than f_{C_i} . The explanation is that E_{V_m} remains highly uncertain and, as we will see later, assumes an extremely non-Gaussian distribution within the posterior PDF. In general, parameters for the FIFE site seem to converge faster than for Loobos, which would be expected for an inversion with 14 instead of 23 parameters.

A more convenient way to visualize convergence of the sampling sequences is a phase diagram using the costs of the prior probability (Eqn (4b), costs of parameters) and the misfit in the Likelihood function (Eqn (3b), costs of diagnostics) as the two axes (Gelman *et al.*, 1995). As Fig. 1c and d shows for both sites and 0.25 prior uncertainty, all sequences appear to converge against a common global cost function minimum (maximum of the PDF), despite widely varying starting points. The convergence, however, is less straight for FIFE, where a local minimum with a cost of diagnostics of around 500 is initially reached by some of the simulations. Analysis of the other simulations (not shown) reveals that the sequence with 0.125 prior uncertainties remains even longer in a similar local minimum until it reaches a region with costs of diagnostics and parameters both around 200. The simulation with 0.5 prior uncertainty does not seem to find a local minimum and converges more rapidly, with costs of diagnostics around 100, and costs of parameters around 35.

The ratio of the costs of diagnostics over parameters in the region of the global minimum gives an indication of how strongly the inversions are constrained by observations. For the FIFE site, this ratio varies between around 1, 2, and 3 for 0.125, 0.25, and 0.5 prior uncertainties. For Loobos, the costs of diagnostics decrease only about 10% from 0.125 to 0.5 prior

uncertainties, and the costs of parameters all lie around 40, giving an almost constant ratio of around 10. Apparently, the more reduced model version with 14 parameters needs rather weak constraints on parameters to converge efficiently, and is still less constrained by observations than the more direct inversion against NEE and LE. Note, however, that the FIFE inversion used only 4 days and only data from daytime fluxes.

To determine a practical initial cut-off for iterations before convergence to the global PDF maximum, the so-called 'burn-in time' with length n iterations, we used again Gelman's test (Gelman & Rubin, 1992; Cowles & Carlin, 1996). It requires that the reduction factor computed for the ensuing n iterations (i.e. iterations $n + 1$ to $2n$) reaches a value of around 1.2–1.4 for all sampled quantities of interest. Figure 1e and f show this reduction factor for the same parameters as Fig. 1a and b, together with the values of the product of the slowest converging parameter with the two others, as a function of the total number of simulations, variable burn-in time plus ensuing iterations of the same length. Products are required to compute parameter covariances and appear to converge at least as rapidly as the slowest parameter. (The reduction factor is always applied to the second half of the iterations up to the number indicated. A burn-in time of $n = 50\,000$ thus corresponds to 100 000 iterations.) To be on the safe side, we chose a burn-in time of length $n = 50\,000$ iterations.

Convergence of parameters for the cases with 0.125 and 0.5 prior uncertainties was evaluated by plotting expected values of all parameters against the length of the burn-in time. The same burn-in time of 50 000 iterations was found to be sufficient for all cases except for FIFE with 0.125 prior uncertainty, where 1 000 000 iterations were chosen instead. For the case of 0.25 prior uncertainty, we carried out six sequences, for the other only one. For each sequence, we continued with another 450 000 iterations after burn-in (including the FIFE 0.125 case). Because only every 10th iteration was sampled, this yielded 45 000 parameter samplings for 0.125 and 0.5 prior uncertainties, and 270 000 samplings from six sequences for 0.25 prior uncertainties. Each sequence of 500 000 iterations took ca. 5 h central processing unit time on a Linux PC workstation with 1.9 GHz clock speed.

Parameter change and uncertainty reduction from constraining with eddy covariance data

Means and standard deviations can be estimated directly from the samplings of the posterior PDF in the space of the log-normalized parameters. As the

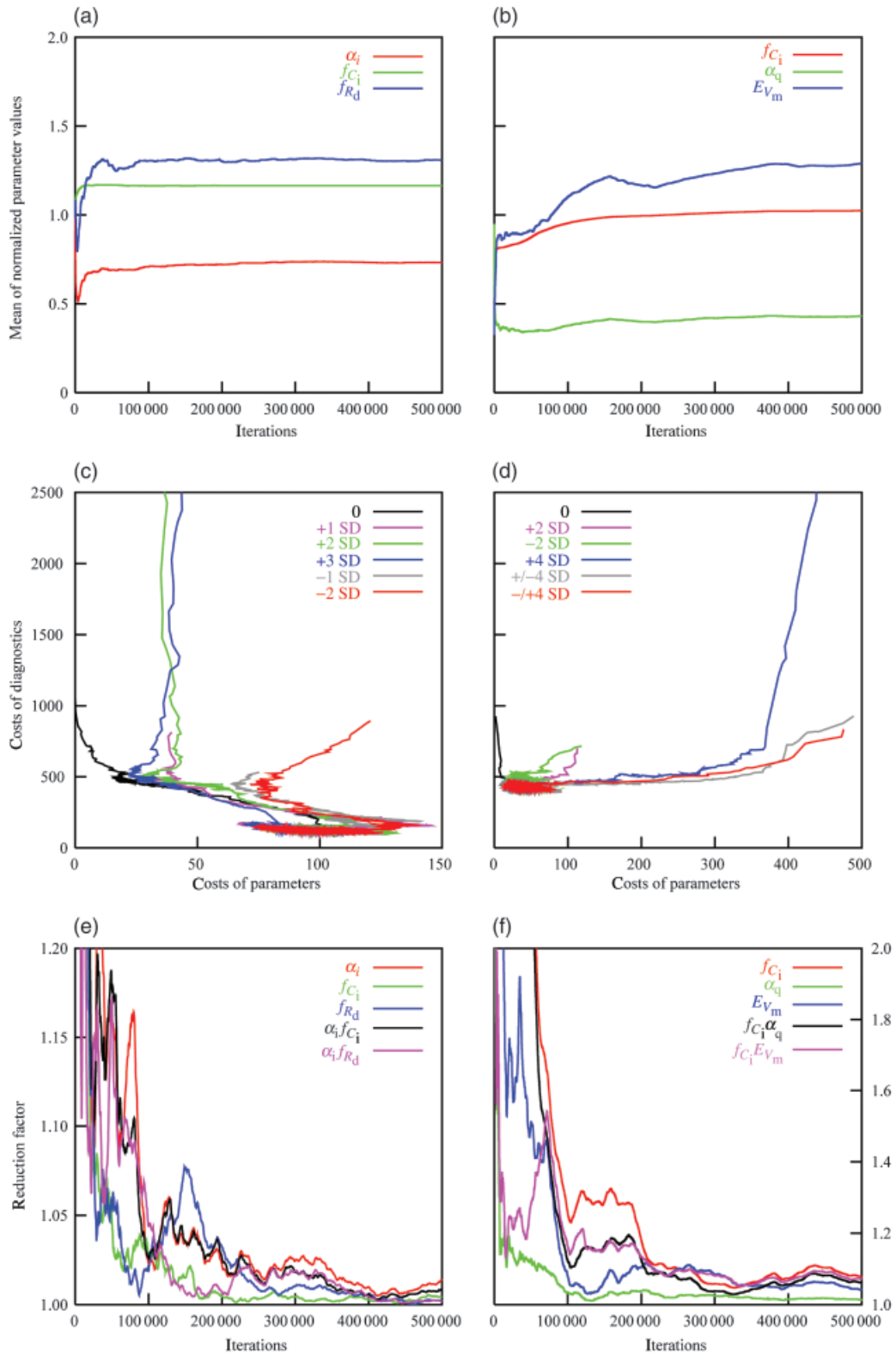


Fig. 1 Convergence of the Monte Carlo inversion for 0.25 prior uncertainty of the log-normalized parameters, for the FIFE (left: a, c, e), and the Loobos site (right: b, d, f). (a, b) Estimated mean of selected parameters depending on number of iterations; (c, d) phase diagram of the two contributions to the total cost function, measuring deviation from prior parameters and between measured and modeled diagnostics (= fluxes), for sequences with varying starting points; (e, f) Gelman's reduction factor for the same parameters as above, and for two parameter products. The selected parameters are: the slowest converging, one fast converging, and the one most highly correlated with the first.

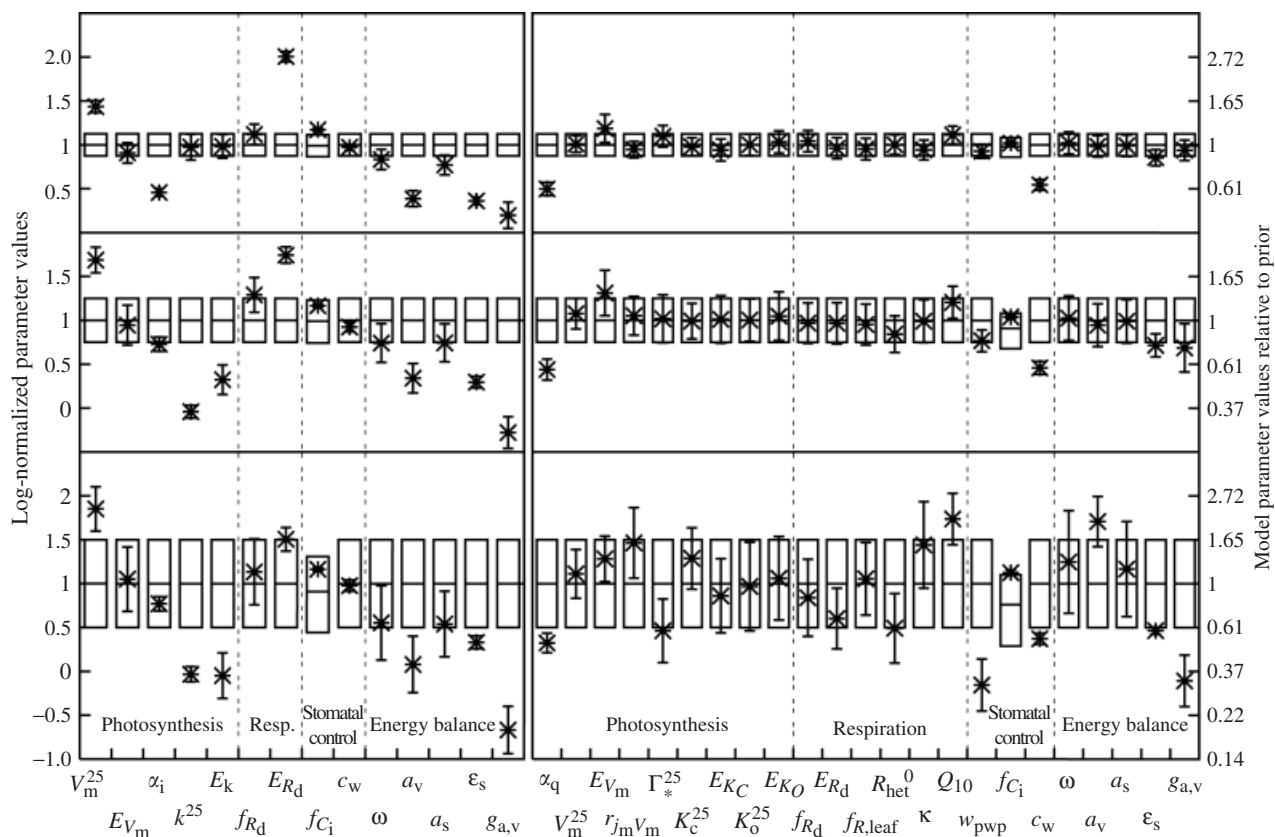


Fig. 2 Prior and posterior parameter values and uncertainties for the log-normalized parameters (transformation to model parameters see Eqn (7)). The boxes show means and one standard deviation of assumed prior parameters (SD = 0.125, 0.25, 0.5). Crosses show the posterior means, and error bars 1 SD of the posterior parameters. Left: Biosphere Energy-Transfer Hydrology (BETHY) model C4 version constrained with data from FIFE site; right: BETHY C3 version constrained with data from Loobos site. The axis on the right hand side shows the model parameter values divided by their respective priors for comparison (does not apply to parameter f_{C_i}).

parameters represent different processes, comparison with prior means and uncertainties provides valuable information on those processes about which we can learn most through the use of eddy covariance data. The means and ranges corresponding to one standard error are shown in Fig. 2 for all prior and posterior parameter values. For the non-Gaussian prior distribution of f_{C_i} , we show the corresponding percentiles.

For the C4 FIFE site, patterns of parameter change are consistent between versions 0.25 and 0.5, with version 0.125 being similar for most parameters, except for those two of the CO₂ specificity, k . The standard rate, k^{25} , and its activation energy, E_k , are decreased by a large amount when prior uncertainties are large, while they are not affected by the inversion when prior uncertainties are small. We interpret this result in the following way: both parameters describe one of three co-limiting rates that determine C4 photosynthesis (Eqn (A7)). In one case, the priors are set in such a way that the rate J_c is never limiting the actual rate A . Once prior

uncertainties are increased, the inversion gains more freedom and finds a solution where all three rates, J_c , J_e and J_v , are limiting and agreement with observations is significantly improved (see lower cost of diagnostics between the local and the global minimum in Fig. 1c).

For the Loobos C3 site, patterns of parameter changes are similar for versions 0.125 and 0.25. The pattern of version 0.5 differs from these for at least five parameters: Γ_*^{25} , K_C^{25} , E_{R_d} , k and a_v . For the photosynthesis parameters, there is a consistent pattern of lower quantum efficiency, α_q , with little change in maximum carboxylation rate, V_m^{25} , and an increase in the carboxylation rate's activation energy, E_{V_m} . For others, there is no consistency: the direction of change depends on the prior uncertainty (for $r_{j_m V_m}$, Γ_*^{25} , K_C^{25}), or changes are small overall. For the respiration parameters, there is a consistent increase in Q_{10} , and a decrease in the overall heterotrophic respiration expressed through R_{het}^0 (except for 0.125 prior uncertainty). As for FIFE, the posterior values of the stomatal parameters c_w and f_{C_i} are almost independent of the prior uncertainty ranges,

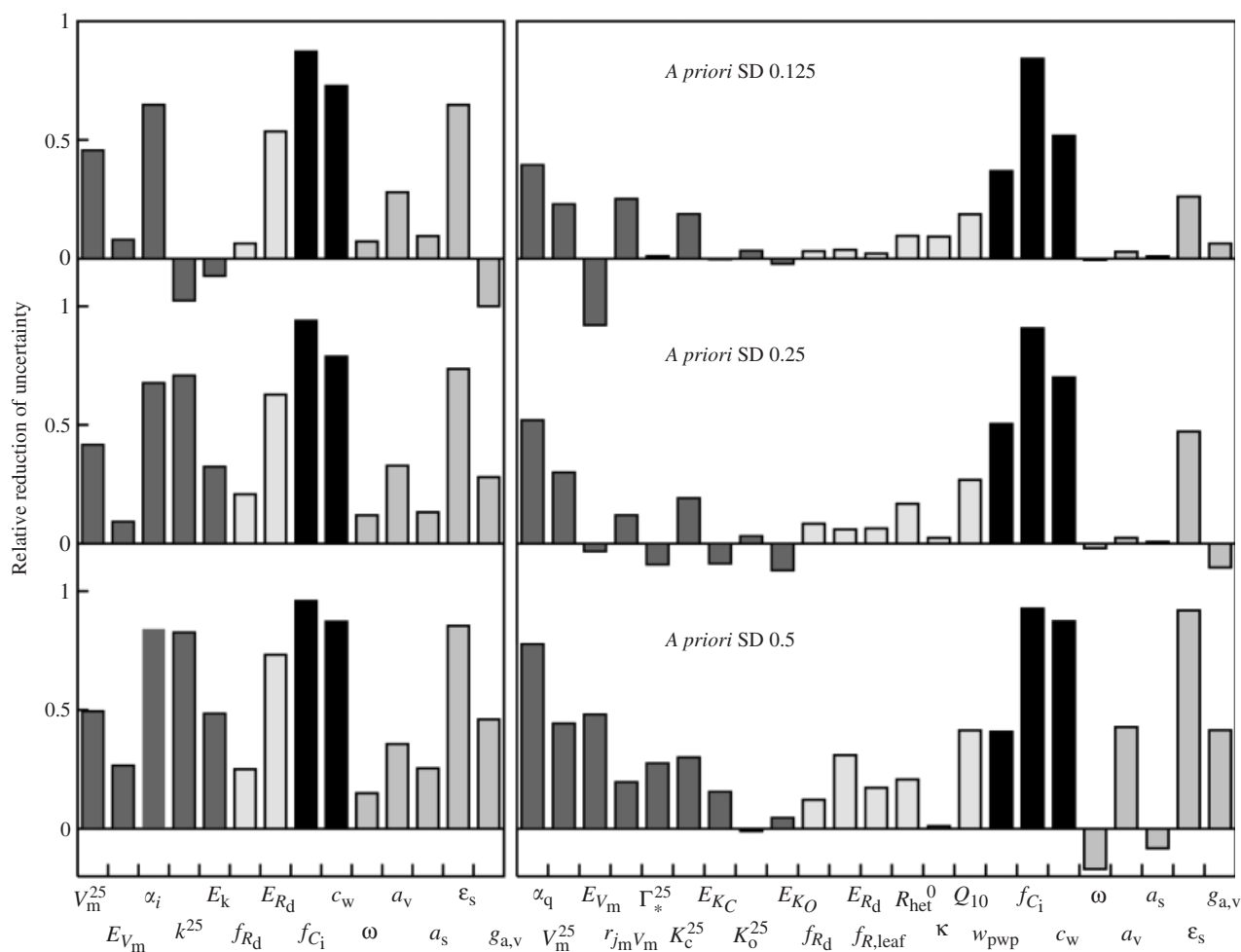


Fig. 3 Relative reduction of uncertainty of parameter values. Zero or negative relative error reduction indicates that no information about a particular parameter could be gained, one would mean perfect knowledge of the inversion. Left: Biosphere Energy-Transfer Hydrology (BETHY) model C4, FIFE site; Right: BETHY C3, Loobos site. Gray shadings denote (from left to right): photosynthesis, carbon balance, stomatal control, and energy/water balance.

and there is a universal downward adjustment of the third, w_{pwp} .

Another quantity that measures the gain in information after inversion against the eddy covariance data is the relative reduction in uncertainty, defined as $1 - \sigma_{\text{posterior}} / \sigma_{\text{prior}}$ where σ is the parameter's standard deviation. For f_{C_i} , with its non-Gaussian prior distribution, we again use the equivalent percentile range for σ_{prior} . If this value comes close to one, we have gained almost complete knowledge of the particular parameter concerned. Because σ is derived from the complete PDF, cases where this value is less than 0 are also possible. The relative reduction in uncertainty is shown in Fig. 3.

For both sites, most information is gained for the stomatal parameters, in particular f_{C_i} . This is not a great surprise, as stomata regulate water-use efficiency, (i.e.

the ratio of lost water to gained carbon dioxide molecules), and the fluxes of both (or derived quantities) are just the information that is assimilated. The next best-constrained process is photosynthesis, with most information gained for quantum efficiency (α_i or α_q for C3 or C4), maximum carboxylation rate, V_m^{25} , and for C4 the functionally similar CO_2 specificity, k^{25} (except, again, for FIFE 0.125). Within the energy and radiation balance, most information is consistently gained for the sky emissivity parameter, ϵ_s . Only in some cases, information is gained about albedo (a_v) and aerodynamic conductance ($g_{a,v}$). For FIFE, the two respiration parameters are consistently constrained, while for Loobos, only very little can be learned about either autotrophic or heterotrophic respiration. There seems to exist a principle difficulty to distinguish between autotrophic and heterotrophic respiration on

the basis of net CO₂ flux measurements. This result should caution us against the use of night-time CO₂ flux data to derive GPP from NEE, here implicit in the data from the FIFE site.

Covariances between parameters

Covariances between parameters, given in their log-normalized form in Table 2 for 0.25 prior uncertainties and both sites, can be used to find groups of parameters that tend to be constrained together. For FIFE, we rather do not find such distinct groupings of parameters. Instead, we find that 11 of the 14 parameters from different parts of the model are strongly correlated with other parameters, with a log-normalized covariance (= correlation coefficient) of up to 0.91 for the pair c_w and ε_s . Two parameters, f_{R_d} and ε_s , have a correlation of over 0.30 to four other parameters. For Loobos, however, we can identify some distinct groups of parameters for which errors are correlated. The first such emerging group consists of the six first photosynthesis parameters (α_q , V_m , E_{V_m} , $r_{J_m V_m}$, Γ_*^{25} , K_C^{25}) plus the stomatal parameter, f_{C_i} . These are linked to a second energy balance group consisting of ε_s and $g_{a,v}$ via f_{C_i} , E_{V_m} and α_q . f_{C_i} is only weakly correlated to the other, soil moisture-related stomatal parameter, c_w . This latter parameter cannot be separated from the wilting point parameter, w_{pwp} : the covariance in log-normalized space reaches 0.75, which indicates that the effect on NEE and LE of changes in one parameter is compensated by changing the other parameter in the same direction. A third group is formed by the three heterotrophic respiration parameters, R_{het}^0 , κ , and Q_{10} : these are linked to the first group by a high log-normalized covariance between Q_{10} and E_{V_m} .

There is one important difference between the two sites that affects parameter correlations: for FIFE, canopy conductance is assimilated, whereas for Loobos it is (through latent heat flux) the measured transpiration rate. For example, the correlation between c_w and ε_s is 0.91 for FIFE, but -0.12 for Loobos. Also, c_w at Loobos is highly correlated with w_{pwp} , a parameter that is absent at FIFE. For both sites, increasing c_w leads to a higher root supply rate and an increase in the canopy conductance and transpiration rate, all other parameters being equal. Increasing ε_s , through increasing net radiation, increases atmospheric demand (D in the model description, see Appendix), and through this the transpiration rate. As stomata respond to atmospheric demand by closing (Eqn (A17)), increasing ε_s leads to a decrease in canopy conductance. To match the quantity that is assimilated, both c_w (or w_{pwp}) and ε_s have either opposing effects (FIFE: on canopy conductance) and are

correlated, or have an effect that goes in the same direction (Loobos: on transpiration) and must therefore be anticorrelated to compensate each other.

Analysis of the posterior PDF

So far, we have only analyzed means and covariances derived from the PDF of the posterior parameters. Table 3 lists the prior and posterior means of both the model and the log-normalized parameters. We will now assess whether the assumption of Gaussian posterior distributions is adequate – the advantage would be easy use of the PDF in a global CCDAS (see Introduction). The analysis is based on the medium case of 0.25 prior uncertainty of log-normalized parameters. The skewness and kurtosis of the PDF projected onto each log-normalized parameter is also listed in Table 3. Skewness measures whether a distribution is 'leaning,' or skewed, towards either the left (i.e. values smaller than the mean, negative skewness), or the right (positive skewness). Kurtosis indicates 'peakedness' relative to a Gaussian distribution, where distributions flatter than Gaussian have negative values (Storch & Zwiers, 1999). Most parameters show only small deviations from a Gaussian distribution, with skewness often slightly negative.

A few parameters, however, are more negatively skewed and some have a markedly 'pointed' distribution (high positive kurtosis): ε_σ for FIFE, and E_{V_m} , $g_{a,v}$ and w_{pwp} for Loobos (see Fig. 4). E_{V_m} , $g_{a,v}$ also show an increase in the standard deviation from prior to posterior. If the distribution of a parameter is much different from Gaussian, then estimation techniques that use the gradient in parameter space to find the cost function minimum, and second derivatives of the cost function to derive parameter uncertainties, will give erroneous results. For f_{C_i} (FIFE), this would lead to a mean of 1.11 instead of 1.09, and a slight underestimate of the uncertainty. The effect would not be large for w_{pwp} (Loobos), either, and still quite acceptable for $g_{a,v}$ given the generally large uncertainties.

Extrapolation of results in time

We have obtained a constrained parameter PDF for the BETHY C4 and C3 models from 4 or 7 selected days of eddy covariance data, respectively. The question to ask now is how the gained process knowledge, expressed through reduced parameter uncertainty, translates into reduced uncertainty about the quantity of highest interest: the net sink at the site over a longer time period. For that purpose, we have computed the cumulative NEE over a period of 2 years at the Loobos site, complete with 95% confidence ranges, from the

Table 2 Elements of the reduced error covariance matrix, equal to the correlation coefficient, derived from the posterior PDF at the FIFE and Loobos sites

α_q	V_m^{25}	E_{V_m}	r_{mV_m}	Γ_*^{25}	K_C^{25}	E_{K_C}	K_O^{25}	E_{K_O}	f_{R_d}	E_{R_d}	$f_{R,leaf}$	R_{het}^0	κ	Q_{10}	w_{pwp}	f_{C_i}	c_w	ω	a_v	a_s	ε_s	$g_{a,v}$	
1.00	0.40	0.45	0.11	0.71	-0.31	-0.05	0.06	0.06	-0.02	0.11	-0.02	-0.06	-0.04	0.11	0.02	0.35	0.01	0.07	-0.05	0.00	-0.31	-0.19	
	1.00	0.51	-0.03	0.45	0.37	-0.13	-0.14	0.10	-0.18	0.20	0.06	-0.11	0.11	0.17	0.12	0.36	0.19	-0.03	-0.03	-0.02	-0.17	-0.16	
Loobos	1.00	1.00	0.41	0.47	-0.17	0.26	0.11	-0.09	-0.14	0.24	0.03	-0.18	0.03	0.31	0.09	0.62	0.15	-0.01	-0.10	0.00	-0.30	-0.33	
		1.00	1.00	0.33	-0.31	0.14	0.12	0.01	-0.07	0.12	-0.01	-0.08	-0.05	0.19	0.04	0.25	0.12	-0.01	-0.11	-0.01	-0.20	-0.24	
			1.00	1.00	-0.21	-0.05	0.04	0.06	-0.06	0.11	0.02	-0.06	0.00	-0.03	0.12	0.40	0.24	-0.01	0.00	0.02	-0.24	-0.12	
FIFE				1.00	0.05	0.18	-0.01	-0.03	-0.03	-0.04	0.05	0.03	0.08	-0.14	0.00	-0.18	0.01	-0.01	0.04	-0.02	0.12	0.07	
					1.00	0.00	-0.11	0.01	0.01	0.01	-0.01	0.01	-0.06	-0.02	-0.09	0.01	-0.04	0.02	-0.02	0.03	-0.03	0.03	
						1.00	-0.02	0.01	0.01	0.01	0.00	-0.01	-0.02	0.07	-0.04	0.04	-0.03	0.00	0.00	0.02	-0.04	-0.04	
							1.00	1.00	-0.01	0.07	-0.02	-0.04	0.01	0.03	0.03	0.06	0.00	-0.02	0.00	-0.02	-0.01	-0.01	
V_m^{25}	1.00								1.00	0.04	0.11	-0.03	0.22	-0.15	0.02	-0.02	0.00	-0.01	0.03	-0.01	0.02	0.02	
E_{V_m}	0.19	1.00							1.00	1.00	-0.08	0.05	-0.20	-0.08	0.02	0.15	0.05	0.00	-0.01	-0.01	-0.09	-0.01	
α_i	0.44	0.08	1.00						1.00	1.00	0.06	-0.22	0.15	0.15	-0.01	0.05	0.03	0.03	0.01	0.03	-0.01	-0.07	
k^{25}	0.00	-0.03	-0.25	1.00					1.00	1.00	1.00	0.06	-0.22	0.15	-0.01	-0.19	-0.04	0.01	0.02	0.00	0.06	0.03	
E_k	-0.10	-0.03	-0.30	-0.15	1.00				1.00	1.00	1.00	0.23	-0.39	0.36	-0.01	-0.19	-0.04	-0.02	0.00	-0.02	0.01	-0.05	
f_{R_d}	-0.15	-0.07	0.45	0.09	-0.08	1.00			1.00	1.00	1.00	1.00	0.36	1.00	0.00	0.07	0.04	-0.02	0.00	-0.02	0.01	-0.05	
E_{R_d}	-0.15	0.03	-0.16	-0.28	-0.05	-0.35	1.00		1.00	1.00	1.00	1.00	0.36	1.00	0.12	0.22	0.18	0.01	-0.06	-0.01	-0.12	-0.11	
f_{C_i}	-0.22	-0.06	-0.15	-0.65	0.14	0.06	0.21	1.00	1.00	1.00	1.00	0.03	0.75	0.09	1.00	0.03	0.75	0.00	0.02	0.02	0.12	0.09	
c_w	-0.26	-0.03	-0.23	-0.07	0.02	-0.42	-0.09	-0.05	1.00	1.00	1.00	1.00	0.36	1.00	1.00	1.00	1.00	-0.01	-0.06	0.02	-0.58	-0.47	
ω	0.01	0.03	0.03	0.03	0.08	-0.03	0.03	-0.03	-0.17	1.00	1.00	1.00	0.36	1.00	1.00	1.00	1.00	1.00	-0.01	-0.05	0.00	-0.12	
a_v	-0.01	-0.03	-0.01	0.08	0.13	0.02	0.13	-0.05	-0.47	-0.06	1.00	1.00	0.36	1.00	1.00	1.00	1.00	1.00	-0.01	0.01	0.02	0.01	
a_s	-0.01	-0.01	-0.01	0.04	0.06	0.01	0.02	-0.02	-0.17	-0.03	-0.13	1.00	0.36	1.00	1.00	1.00	1.00	1.00	-0.01	-0.01	0.12	0.01	
α_s	-0.32	-0.05	-0.26	-0.04	0.11	-0.50	-0.06	-0.10	0.91	-0.05	-0.17	-0.04	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.02	0.01	0.01	
$g_{a,v}$	0.16	0.08	0.15	-0.21	-0.19	0.13	0.68	0.21	-0.18	-0.10	-0.24	-0.13	-0.40	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.67	1.00	
	V_m^{25}	E_{V_m}	α_i	k^{25}	E_k	f_{R_d}	E_{R_d}	f_{C_i}	c_w	ω	a_v	a_s	ε_s	$g_{a,v}$									

Prior parameter uncertainties in log-normalized space were 0.25. Values above 0.3 or below -0.3 are considered to indicate a close correlation between parameter errors and are marked.

PDF, probability density function; FIFE, First ISLCP Field Experiment Kansas site.

Table 3 Prior and posterior parameter values in model space for FIFE (above, BETHY C4 version), and Loobos (below, BETHY C3 version); prior parameter and standard deviation (SD) in log-normalized space; posterior parameter with SD, skewness and kurtosis, also in normalized space (strongly non-Gaussian parameters highlighted)

Parameter	Model parameter		Log-normalized parameter			
	Prior p_o	Posterior	Prior SD	Posterior SD	Skewness	Kurtosis
V_m^{25}	8.00E-06	1.59E-05	0.25	0.15	0.33	0.33
E_{V_m}	5.85E+04	5.54E+04	0.25	0.23	-0.07	-0.03
α_i	4.00E-02	3.05E-02	0.25	0.08	0.30	-0.25
k^{25}	1.40E-01	4.94E-02	0.25	0.07	0.28	0.29
E_k	5.10E+04	2.59E+04	0.25	0.17	-0.04	-0.01
f_{R_d}	4.20E+02	5.62E-02	0.25	0.20	-0.01	-0.04
E_{R_d}	4.50E+04	9.47E+04	0.25	0.09	-0.17	-0.02
f_{C_i}	6.70E-01	7.82E-01	0.24*	0.02	-0.28	0.30
c_w	1.00E+00	9.26E-01	0.25	0.05	-0.33	0.10
ω	1.20E-01	9.27E-02	0.25	0.22	-0.11	-0.02
a_v	2.00E-01	1.03E-01	0.25	0.17	-0.27	0.07
a_s	5.00E-02	3.87E-02	0.25	0.22	-0.09	0.00
ϵ_s	6.40E-01	3.17E-01	0.25	0.07	-0.57	0.32
$g_{a,v}$	2.43E-02	6.43E-03	0.25	0.19	-0.10	-0.20
α_q	2.80E-01	1.60E-01	0.25	0.12	0.13	-0.14
V_m^{25}	2.90E-05	3.13E-05	0.25	0.18	-0.01	-0.12
E_{V_m}	5.85E+04	7.99E+04	0.25	0.26	-1.10	1.01
$r_{f_m}^{V_m}$	1.79E+00	1.89E+00	0.25	0.22	0.15	-0.18
Γ_*^{25}	4.25E+01	4.33E+01	0.25	0.27	0.01	-0.16
K_C^{25}	4.60E-04	4.56E-04	0.25	0.20	-0.23	0.20
E_{K_C}	5.94E+04	6.01E+04	0.25	0.27	0.06	0.03
K_O^{25}	3.30E-01	3.31E-01	0.25	0.24	0.05	0.00
E_{K_O}	3.60E+04	3.77E+04	0.25	0.28	0.14	0.16
f_{R_d}	1.00E-02	9.69E-03	0.25	0.23	-0.01	-0.05
E_{R_d}	4.50E+04	4.35E+04	0.25	0.24	0.01	0.00
$f_{R,leaf}$	5.00E-01	4.77E-01	0.25	0.23	0.01	0.11
R_{het}^0	2.07E+00	1.77E+00	0.25	0.21	0.00	0.00
κ	1.00E+00	9.91E-01	0.25	0.24	-0.01	-0.07
Q_{10}	1.72E+00	2.11E+00	0.25	0.18	-0.21	0.22
w_{pwp}	2.50E+00	1.98+00	0.25	0.12	-0.55	0.54
f_{C_i}	8.70E-01	9.05E-01	0.20*	0.02	-0.11	-0.12
c_w	1.00E+00	5.82E-01	0.25	0.08	0.02	-0.05
ω	1.20E-01	1.23E-01	0.25	0.26	0.03	0.06
a_v	2.00E-01	1.89E-01	0.25	0.24	-0.04	-0.02
a_s	5.00E-02	4.95E-02	0.25	0.25	-0.01	0.14
ϵ_s	6.40E-01	4.82E-01	0.25	0.13	0.16	0.32
$g_{a,v}$	4.00E-02	2.92E-02	0.25	0.28	-0.39	0.71

Prior distribution is Gaussian in the log-normalized space, and log-normal in model parameter space. The posterior model parameter value is the transformed log-normalized prior using Eqn (7).

*Prior distribution is a Gaussian function with a cutoff at 0 and 1 in model space. Shown is the 68.3 percentile range which is equivalent to 1 SD.

BETHY, Biosphere Energy-Transfer Hydrology; FIFE, First ISLSCP Field Experiment Kansas site.

prior, the posterior Gaussian, and the full-posterior PDF. The posterior Gaussian PDF approximates the full PDF by using only the means and the error covariance matrix. As Fig. 5 shows by the green area, prior uncertainties about parameter values of BETHY were

consistent with the Loobos site being both a strong sink (positive NEE), or a moderate source of carbon (negative NEE) over the 2 years. After constraining the model, the 95% confidence range lies outside of the median prior estimate and Loobos is now very

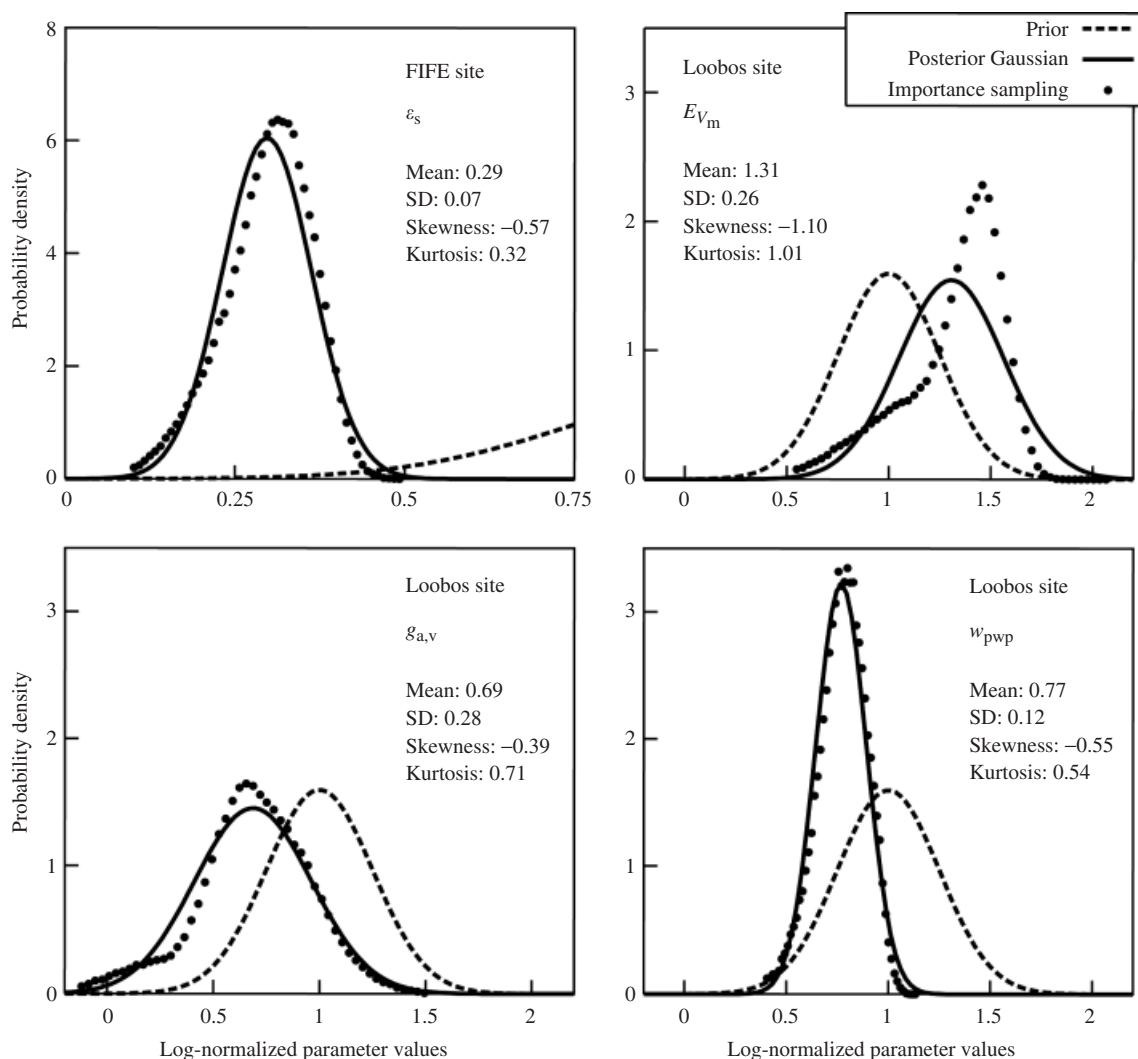


Fig. 4 Probability distributions of selected parameters from FIFE (a) and Loobos (b–d) for a prior uncertainty of 0.25 in parameter space. Comparison of importance sampling, approximating the true distribution, with the prior probability density function (PDF) and to posterior Gaussian PDF computed from mean and standard deviation. Additionally, the mean, standard deviation, skewness and kurtosis are given for the posterior distribution.

definitely identified as a CO₂ sink. This means that extrapolating 7 days of NEE and LE data through the assimilation procedure resulted in a sink estimate that was both significantly different from the best prior estimate and significantly different from zero. Further, we find that using the full PDF in parameter space results in only about half of the uncertainty in NEE over the 2 years compared with using a PDF derived from parameter means and covariances. Skewness and kurtosis of the full PDF of the cumulative NEE can also be relatively large.

Note that this result still depends on the prior uncertainty, which was only estimated in a simple and preliminary way for this study. Also, assimilating more days of flux measurements would lead to stronger

constraints of model parameters and fluxes, which would lead to even smaller uncertainties of the cumulative NEE. Here, we can instead use the measured NEE of the 2 years, with a few gaps (for which we assumed NEE = 0), to validate our time extrapolation (Fig. 5, blue line). With this additional assumption as a point of caution, we arrive at around 25 mol(CO₂)m⁻²yr⁻¹ or 300 g C m⁻²yr⁻¹ net uptake from both the observations and the model simulations. The generally good agreement between modeled (after assimilation) and measured NEE across the 2 years shows that the model is able to capture the main processes that influence this quantity. We therefore suggest, that the method shown here with all available measurements assimilated, could be a superior gap

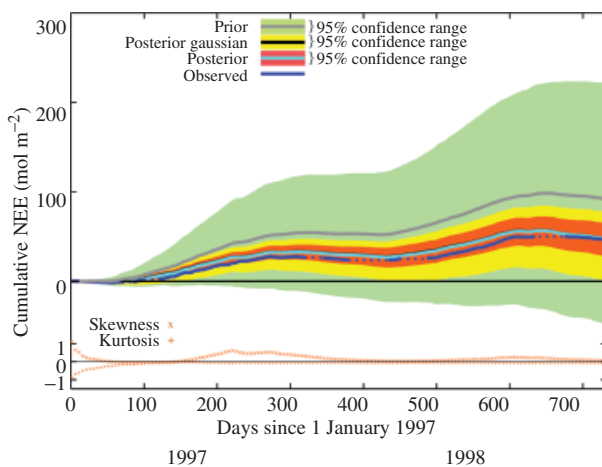


Fig. 5 Cumulative net ecosystem exchange (NEE) for 2 years, using the results from the inversion against 7 days of NEE and latent energy flux, for the Loobos site. Green: prior uncertainty range, yellow: posterior uncertainty range using posterior mean and error covariance (Gaussian posterior PDF); red: posterior uncertainty range with full PDF; blue: measurements, dashed: missing data (NEE = 0 assumed). Positive values of NEE indicate a terrestrial carbon sink.

filling method compared with the ones usually employed by the eddy covariance community.

Discussion

We have performed several Bayesian inversions of an ecosystem model, BETHY, constrained by eddy covariance data of carbon and water fluxes. There were two sites, one C3 and one C4, and three sets of assumptions about prior parameter uncertainties. We find that the method works very well, although some care has to be taken to insure algorithm convergence. Compared with non-Bayesian, standard optimization techniques (e.g. Wang *et al.*, 2001), the method treats all parameters equally and simultaneously, and is still able to distinguish between those parameters that can be constrained by the eddy covariance data, and those that cannot. With 4 or 7 days of diurnal data assimilated, the Bayesian part of the cost function in the region of the minimum was between two and 10 times the cost of the measurements, so that the inversion was found to be constrained predominantly by the flux data. Similar to Wang *et al.* (2001), who used non-Bayesian inversions, we find that typically five parameters can be effectively constrained by the method. Perhaps not surprisingly, two to three of them are stomatal parameters: stomata control the balance between carbon uptake and transpiration at the leaf level, both were assimilated, albeit at the stand level. This shows that leaf-level functional representations

can be effectively constrained. One of the parameters, ε_s , strictly speaking, belongs to the external driver of BETHY used in the computation of incoming thermal radiation. This particular result indicates that assimilating accurate radiation data (obliterating the need for ε_s) will likely improve parameter estimation further. We could also constrain parameters describing the light response (α_i, α_q), and sometimes the temperature response of photosynthesis (k^{25}). It is evident that a sufficient range of environmental conditions must be present during the period for which the data were assimilated to gain information about the dependencies on those conditions. The use of a diurnal cycle and of different dates across the seasons must have helped here, and may explain how well the eddy-flux constrained model performed against the 2-year measurements.

The method also delivers information on the error covariances of parameters. This information can be used to find out which processes can be constrained individually by the assimilation of the eddy flux data. Analysis of the full PDF, only possible by Monte Carlo methods, shows that most parameters tend to have distributions close enough to a Gaussian one for gradient and second-derivative methods to work effectively. These usually require a few orders of magnitude fewer iterations. Only one parameter was identified with a distribution so far away from a normal one that such methods would have underestimated the posterior mean and uncertainty to a large degree.

One straightforward and easy application of the method presented here would be to use the posterior means and covariances of the parameter PDF as priors in a global-scale data assimilation system (cf. Rayner *et al.*, 2005). We expect that using the Gaussian part of the complete PDF will tend to overestimate the uncertainty of the model diagnostics.

We have found that the results of our study depended on the prior uncertainty of the parameter values. This uncertainty itself will depend on the parameter in question, the scale at which the model is applied, and the amount of available information at that scale. Our preliminary results (unpublished) indicate that for V_m^{25} , 0.5 prior uncertainty at the site scale would be a realistic assumption in the log-normalized space if only the functional type of vegetation is known.

We have so far restricted our study to cases that are rather rare when considering the entire FLUXNET archive: we relied on the availability of soil moisture measurements. Applying the method for more sites, however, will be crucial for identifying representative model parameterizations by plant functional type, or some other generalization on which global models

necessarily rely. Therefore, we expect to conduct further studies using the complete BETHY model with the full water balance. If no complete data on LAI are available, a phenology scheme may also be included. LAI and soil moisture data could then also be assimilated instead of being used as input. We also suggest using more days and longer periods for assimilation, although we find that only a few days of data already deliver a strong model constraint.

Conclusions

The parameterization of global TEMs for carbon cycle studies poses great challenges. We are confronted with model errors, errors from the finite accuracy of parameter estimation, and representation errors that result from the fact that models need to work with a finite set of idealized vegetation types. This study demonstrates that inversion against eddy covariance data can be a powerful tool for using local measurements to constrain the possible range of ecosystem model parameters. Such information about parameter uncertainties is crucial for understanding to what degree of confidence we can use models to compute the global terrestrial carbon balance.

The advantage of the Monte Carlo inversion technique is that it works even for highly nonlinear models, and that it allows sampling the complete posterior PDF. This can be used to estimate how well methods will work that derive uncertainties from the curvature of the cost function at its global minimum. Because they require far fewer iterations, such methods are better suited for global applications, especially when parameters need to be inverted simultaneously.

Further use of this method will require a careful analysis of the prior uncertainties of model parameters. For the envisaged global applications, it will also be important to repeat the analysis with a sufficient number of sites per major vegetation type in order to gain an understanding of the representation error. We suggest that using such studies to determine prior parameter uncertainties for global carbon cycle data assimilation could be one of the principle applications of data from the growing network of eddy covariance measurement sites. We believe that such a method of extrapolating measurements from local sites to the global scale through the determination and spatial extrapolation of parameters would be the most promising and most adequate route to better global TEMs. These will be crucial for any application aimed at predicting the future response of the carbon cycle to climate change, including atmosphere – vegetation feedbacks.

Acknowledgments

The original eddy flux data used in this study were collected, maintained and generously provided by Eddy Moors, Jan Elbers, Wilma Jans, Han Dolman and colleagues affiliated with the Loobos Research Site in The Netherlands, run by ALTERRA, Green World Research. The authors wish to thank Bart Kruijt and Isabel van den Wyngaert for help with estimating uncertainties of eddy covariance measurements and Thomas Kaminski for useful advice and comments. This work has been financed and supported by the EU project CAMELS, contract number EVK2-CT-2002-00151, within the EU's 5th framework program for Research and Development, and the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, e.V.

References

- Collatz GJ, Ribas-Carbo M, Berry JA (1992) Coupled photosynthesis–stomatal conductance model for leaves of C4 plants. *Australian Journal of Plant Physiology*, **19**, 519–538.
- Cowles MK, Carlin BP (1996) Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, **91**, 883–904.
- Cox PM, Betts RA, Jones CD *et al.* (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**, 184–187.
- Cramer W, Kicklighter DW, Bondeau A *et al.* (1999) Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Global Change Biology*, **5** (Suppl. 1), 1–15.
- Cramer W, Bondeau A, Woodard FI *et al.* (2001) Global response of terrestrial ecosystem structure and function to CO₂ and climate change: results from six dynamic global vegetation models. *Global Change Biology*, **7**, 357–373.
- Dolman AJ, Moors EJ, Elbers JA (2002) The carbon uptake of a mid latitude pine forest growing on sandy soil. *Agricultural and Forest Meteorology*, 157–170.
- Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthesis in leaves of C3 species. *Planta*, **149**, 78–90.
- Friedlingstein P, Dufresne JL, Cox PM *et al.* (2003) How positive is the feedback between climate change and the carbon cycle? *Tellus*, **55B**, 692–700.
- Gelman A, Carlin JB, Stern HS *et al.* (1995) *Bayesian Data Analysis*. Chapman and Hall/CRC, New York.
- Gelman A, Rubin DB (1992) A single sequence from the Gibbs sampler gives a false sense of security. In: *Bayesian Statistics* (eds Bernardo JM, Berger JO, Dawid AP, Smith AFM), pp. 625–631. Oxford University Press, Oxford.
- Global Carbon Project (2003) *Science Framework and Implementation*. Global Carbon Project, Earth System Science Partnership (IGBP, IHDP, WCRP, DIVERSITAS), Canberra.
- Joos F, Plattner GK, Stocker TF *et al.* (2003) Trends in marine dissolved oxygen: implications for ocean circulation changes and the carbon budget. *EOS*, **84**, 197–201.
- Kaminski T, Knorr W, Heimann M *et al.* (2002) Assimilating atmospheric data into a terrestrial biosphere model: a case study of the seasonal cycle. *Global Biogeochemical Cycles*, **16**, 1066 (doi: 10.1029/2001GB001463).

- Kim J, Verma SB (1991) Modeling canopy photosynthesis: scaling up from a leaf to canopy in a temperate grassland ecosystem. *Agricultural and Forest Meteorology*, **57**, 149–166.
- Knorr W (1997) *Satellite Remote Sensing and Modelling of the Global CO₂ Exchange of Land Vegetation: A Synthesis Study*. Max-Planck-Institut für Meteorologie, Hamburg, Germany.
- Knorr W (2000) Annual and interannual CO₂ exchanges of the terrestrial biosphere: process-based simulations and uncertainties. *Global Ecology and Biogeography*, **9**, 225–252.
- Knorr W, Heimann M (2001a) Uncertainties in global terrestrial biosphere modeling. Part I: a comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. *Global Biogeochemical Cycles*, **15**, 207–225.
- Knorr W, Heimann M (2001b) Uncertainties in global terrestrial biosphere modeling. Part II: global constraints for a process-based vegetation model. *Global Biogeochemical Cycles*, **15**, 227–246.
- Leuning R (2002) Temperature dependence of two parameters in a photosynthesis model. *Plant, Cell and Environment*, **25**, 1205–1210.
- McGuire AD, Sitch S, Clein JS *et al.* (2001) Carbon balance of the terrestrial biosphere in the twentieth century: analyses of CO₂, climate and land use effects with four process-based ecosystem models. *Global Biogeochemical Cycles*, **15**, 183–206.
- Medlyn BE, Loustau D, Delzon S (2002) Temperature response of parameters of a biochemically based model of photosynthesis. I. Seasonal changes in mature maritime pine (*Pinus pinaster* Ait.). *Plant, Cell and Environment*, **25**, 1155–1165.
- Metropolis N, Rosenbluth AW, Rosenbluth MN *et al.* (1953) Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, **21**, 1087–1092.
- Monteith JL (1965) Evaporation and environment. *Symposium of the Society for Experimental Biology*, **19**, 205–234.
- Mosegaard K (1998) Inverse problems. *Inverse Problems*, **14**, 405–426.
- Mosegaard K, Rygaard-Hjalsted C (1999) Probabilistic analysis of implicit inverse problems. *Inverse Problems*, **15**, 572–583.
- Mosegaard K, Sambridge M (2002) Monte Carlo analysis of inverse problems. *Inverse Problems*, **18**, R29–R54.
- Mosegaard K, Tarantola A (1995) Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, **100**, 12431–12447.
- Prentice IC, Farquhar GC, Fasham MJR *et al.* (2001) The carbon cycle and atmospheric carbon dioxide. In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* (eds Houghton JT *et al.*), pp. 185–225. Cambridge University Press, Cambridge.
- Raich JW, Potter CS, Bhagawati D (2002) Interannual variability in global soil respiration, 1980–94. *Global Change Biology*, **8**, 800–812.
- Rayner PJ, Scholze M, Knorr W *et al.* (2005) Two decades of terrestrial carbon fluxes from a Carbon Cycle Data Assimilation System (CCDAS). *Global Biogeochemical Cycles* (in press).
- Reichstein M, Rey A, Freibauer A (2003) Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices. *Global Biogeochemical Cycles*, **17**, 1104 (doi:10.1029/2003GB002035).
- Schachtschabel P, Blume HP, Brummer G *et al.* (1992) *Lehrbuch der Bodenkunde*. Ferdinand Enke Verlag, Stuttgart.
- Sellers PJ (1985) Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, **6**, 1335–1372.
- v. Storch H, Zwiers FW (1999) *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, UK.
- Wang YP, Leuning R, Cleugh H *et al.* (2001) Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful? *Global Change Biology*, **7**, 495–510.
- Weiss A, Norman JA (1985) Partitioning solar radiation into direct and diffuse, visible and near-infrared components. *Agricultural and Forestry Meteorology*, **34**, 205–213.
- White MA, Thornton PE, Running SW *et al.* (2000) Parameterization and sensitivity analysis of the BIOME–BGC terrestrial ecosystem model: net primary production controls. *Earth Interactions*, **4**, 1–85.
- Wullschlegler SD (1993) Biochemical limitations to carbon assimilation in C₃ plants – a retrospective analysis of the A/c_i curves from 109 species. *Journal of Experimental Botany*, **44**, 907–920.

Appendix: The BETHY model

Overview

We use a process-based model of the coupled photosynthesis and energy balance system, the BETHY scheme, to simulate the exchange of CO₂, water and energy between the plant canopy and the atmosphere. BETHY computes absorption of PAR in three layers, while the canopy air space is treated as a single, well mixed air mass with a single temperature. Evapotranspiration and sensible heat fluxes are calculated from the Penman–Monteith equation (Monteith, 1965). Carbon uptake is computed with the model by Farquhar *et al.* (1980) for C₃, and the one by Collatz *et al.*, (1992) for C₄ plants. The stomata and canopy model of Knorr (2000) simulates canopy conductance in response to PAR; in the absence of water stress in such a way as to satisfy the demand for CO₂. In water-stressed situations, stomata are further closed until transpiration reaches a specific root supply rate that depends on soil moisture. The carbon balance is computed as plant and soil respiration subtracted from the photosynthesis rate to yield net CO₂ fluxes. The full version of BETHY, described in Knorr (2000) and Knorr and Heimann (2001a), also contains submodels for soil water balance, snow, canopy and soil evaporation, and phenology, which are not used here. Instead, LAI and soil moisture are treated as external forcing data (elements of *s* in Eqn (1)). The version of BETHY for C₃ vegetation used here has 23 free parameters, while the C₄ version has 14. Following is a description of all free model parameters

and their meaning in the context of the model. Parameters have been marked as underlined mathematical symbols and are listed in Table 1, complete with their prior values. (Those that do not appear in one of the equations appear underlined in the text.)

Photosynthesis

For C3 vegetation, gross leaf CO₂ uptake, A , is calculated as (cf. Farquhar *et al.*, 1980):

$$A = \min[J_c, J_e] \quad (\text{A1a})$$

with

$$J_c = V_m \frac{C_i - \Gamma_*}{C_i + K_C(1 + O_x/K_O)}, \quad (\text{A1b})$$

$$J_e = \frac{\alpha_q J_m}{\sqrt{J_m^2 + \alpha_q^2 I^2}} \frac{C_i - \Gamma_*}{4(C_i + 2\Gamma_*)} \quad (\text{A1c})$$

A is gross photosynthesis, or GPP, I is absorbed PAR, C_i the leaf-internal CO₂, and O_x the oxygen concentration (= 0.21 mol(O₂)/mol(air)). α_q is the quantum efficiency of photon capture (mol(e⁻)/mol(photons)⁻¹) and V_m the maximum carboxylation rate (in mol(CO₂)m⁻²s⁻¹), expressed as

$$V_m = \underline{V_m}^{25} \exp\left\{\frac{E_{V_m}(T_K - 298 \text{ K})}{RT_K 298 \text{ K}}\right\} \exp\left\{k_{12} \frac{\Lambda_u + \Lambda_l}{2}\right\} \quad (\text{A2})$$

with the activation energy E_{V_m} (in J mol⁻¹). Λ_u and Λ_l are the LAI of the upper and lower bounds of the specific canopy layer under consideration, and k_{12} a leaf nitrogen scaling parameter set to 0.5/cos(θ_{12}). θ_{12} is the solar zenith angle at noon.

Further, J_m , the maximum electron transport rate (same units as V_m), is expressed as

$$J_m = \underline{r_{J_m V_m}}^{25} \frac{V_m^{25} T_C}{25^\circ \text{C}}. \quad (\text{A3})$$

This rate, at standard temperature, is assumed proportional to V_m with an additional proportionality constant (e.g. Wullschlegel, 1993). T_c is the canopy temperature in °C, T_k the canopy temperature in Kelvin, and R the universal gas constant (8.314 J K⁻¹ mol⁻¹). The CO₂ compensation point without dark respiration, Γ_* , follows from:

$$\Gamma_* = \underline{\Gamma_*}^{25} \frac{T_C}{25^\circ \text{C}}. \quad (\text{A4})$$

The two Michaelis–Menten constants for carboxylation and oxygenation, K_C and K_O , respectively (in J mol⁻¹), have a temperature dependence based on Arrhenius'

equation similar to V_m :

$$K_C = \underline{K_C}^{25} \exp\left\{\frac{E_{K_C}(T_K - 298 \text{ K})}{RT_K 298 \text{ K}}\right\} \quad (\text{A5})$$

and

$$K_O = \underline{K_O}^{25} \exp\left\{\frac{E_{K_O}(T_K - 298 \text{ K})}{RT_K 298 \text{ K}}\right\}. \quad (\text{A6})$$

For C4 photosynthesis, the model of Collatz *et al.* (1992) is used with:

$$A = \min[J_e, J_c, J_i] \quad (\text{A7a})$$

with

$$J_e = V_m, \quad (\text{A7b})$$

$$J_c = kC_i, \quad (\text{A7c})$$

$$J_i = \alpha_i I \quad (\text{A7d})$$

V_m is computed from Eqn (A2), α_i is the C4 quantum efficiency (in mol(CO₂)/mol(photons)), and k is the C4 CO₂ specificity (in mol(CO₂)m⁻²s⁻¹), with

$$k = \underline{k}^{25} \exp\left\{\frac{E_k(T_K - 298 \text{ K})}{RT_K 298 \text{ K}}\right\}. \quad (\text{A8})$$

Photosynthesis rates are computed across three different layers of the canopy, each with its own value for I and V_m , and thus A . The sum over the three layers yields A_c , the canopy gross photosynthesis.

Carbon balance

To compute net CO₂ uptake by the leaves, leaf or 'dark' respiration, R_d , is subtracted from A to yield net leaf CO₂ uptake, A_n , with

$$A_n = A - R_d \quad (\text{A9})$$

and

$$R_d = \underline{f_{R_d}}^{25} V_m^{25} \exp\left\{\frac{E_{R_d}(T_K - 298 \text{ K})}{RT_K 298 \text{ K}}\right\}. \quad (\text{A10})$$

Summation of R_d across canopy layers yields $R_{d,c}$, the canopy dark respiration.

There are two standard values for f_{R_d} , one for C3 and one for C4 vegetation. Dark respiration is assumed to be a constant fraction of total plant, or autotroph respiration, R_{aut} , such that

$$R_{aut} = \frac{R_{d,c}}{f_{R,\text{leaf}}}. \quad (\text{A11})$$

$f_{R,\text{leaf}}$ stands for the fraction of total plant respiration contributed by the leaves. This formulation differs from the form chosen in the original description of BETHY, which contains an additional term for 'growth respira-

tion' assumed proportional to net primary productivity. Such an implicit formulation yields a sum of two terms, one of which is proportional to R_{aut} of Eqn (A10), the other to GPP. The above formulation was chosen for simplicity in order to avoid unnecessary co-dependence of parameters.

For soil (excluding root), or heterotroph respiration, R_{het} , we use an exponential temperature dependence on air temperature (T_a , in °C) times a soil water factor (with zero respiration at zero plant-available soil moisture):

$$R_{\text{het}} = R_{\text{het}}^0 \left(\frac{w}{w_m} \right)^k Q_{10}^{T_a/10^\circ\text{C}}. \quad (\text{A12})$$

w_m is the plant available soil water content at field capacity (% volume), and R_{het}^0 soil respiration at 0 °C and with soil water content at field capacity. w is the plant-available soil water content (% volume) and is computed from total soil moisture, w_{tot} as $w = \max\{w_{\text{tot}}, w_m\} - \underline{w}_{\text{pwp}}$, is the soil water content at the permanent wilting point, which is used as another free parameter in the case that total soil water content is used as input.

Finally, the net carbon flux of the site is given by

$$F_{\text{CO}_2} = A - R_{\text{aut}} - R_{\text{het}}. \quad (\text{A13})$$

Stomatal control

The model of stomatal control follows the assumption that, in the absence of water stress, leaf-level photosynthesis operates at a standard ratio between the leaf-internal CO_2 concentration, C_i , and the CO_2 concentration of free air, C_a . This value is given by

$$C_{i,0} = f_{C_i} C_a \quad (\text{A14})$$

with two values for f_{C_i} , one for C3 and one for C4 vegetation. In order to determine the demand for CO_2 uptake, A_n is first calculated as $A_{n,0}$ for $C_i = C_{i,0}$, and $T_c = T_a$. Inversion of the diffusion equation for CO_2 at the stomatal boundary is then used to compute stomatal conductance in the absence of water stress at each canopy layer (in m s^{-1}):

$$g_{s,0} = \frac{1.6A_{n,0}}{C_a - C_{i,0}} \frac{RT_k}{p}. \quad (\text{A15})$$

p is air pressure (in Pa). If at the time of highest demand, D , transpiration rates exceed a root water supply rate, S , stomatal conductance at each canopy layer is reduced according to

$$g_s = \frac{g_{s,0}}{1 + b_e D_a} \quad (\text{A16})$$

by adjusting b_e such that $S = D$. D_a is the VPD of the free air. This supply rate is taken as

$$S = \frac{c_w}{w_m} \frac{\omega}{w_m}. \quad (\text{A17})$$

Next, the canopy temperature, T_c , is computed consistent with the energy balance after integrating g_s over the canopy to obtain the canopy conductance used in the Penman–Monteith equation. Then, the photosynthesis model is run again, but at a fixed stomatal conductance, g_s , obtained from Eqn (A16), which yields the final gross, A , and net photosynthesis rate, A_n .

Energy and radiation balance

PAR absorption is calculated according to the two-flux scheme by Sellers (1985) with three vertical layers of equal LAI. The diffuse fraction of PAR is calculated according to a procedure by Weiss and Norman (1985). Leaf-angle distribution is assumed to be uniform, and the only free parameters for this scheme is $\underline{\omega}$, the leaf single-scattering albedo.

To determine evapotranspiration rates from the Penman–Monteith formula, BETHY computes net radiation balance of the canopy, $R_{n,c}$, according to the following equation:

$$R_{n,c} = (1 - t_{l,v}) \left[(\varepsilon_a - \varepsilon_{\text{sfc}}) \sigma T_{K,a}^4 - G \right] - (1 - \underline{a}_v - \underline{a}_s) f_{\text{PAR}} R_S. \quad (\text{A18})$$

ε_a and $\varepsilon_{\text{sfc}} = 0.97$ are sky and surface emissivity, respectively, $T_{K,a}$ air temperature in Kelvin, $\sigma = 5.6703 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ the Stefan–Boltzmann constant, and $t_{l,v}$ the longwave transmissivity of the vegetation, assumed $t_{l,v} = f_c \exp(-0.5\Lambda/f_c) + (1-f_c)$. f_c is the fraction of soil covered by vegetation. For the shortwave part, R_S is incoming solar radiation (W m^{-2}), f_{PAR} is the fraction of PAR absorbed by the vegetation and computed by the two-flux scheme, \underline{a}_v the albedo of the vegetation at the limit of high LAI and closed canopy, and \underline{a}_s the amount of solar radiation absorbed by the soil under the canopy at the same limit. G is the ground heat flux, assumed to be a fixed fraction of total net radiation. (Contrary to $R_{n,c}$, total net radiation does not depend on G , so there is no implicit equation to be solved.) The sky emissivity is computed from

$$\varepsilon_a = \underline{\varepsilon}_s \left(\frac{\varepsilon_a}{T_{K,a}} \right)^{1/7} (1 + 0.22n_c^2) \quad (\text{A19})$$

with the cloud cover fraction n_c . If no separate radiation data for PAR and solar radiation are available, R_S is calculated from PAR according to Weiss and Norman (1985).

The aerodynamic exchange between the canopy and the free air is described as

$$G_a = \underline{g}_{a,v} u \quad (\text{A20})$$

with wind speed, u , and a proportionality factor serving as a free model parameter. Wind speeds below 1 m s^{-1} are uniformly set to 1 m s^{-1} to avoid unrealistically high canopy temperature under conditions of extremely still air and high incoming radiation. The prior value of $g_{a,v}$ is determined from the following formula:

$$\underline{g_{a,v}} = \frac{k^2}{[\ln(h_{\text{ref}}/r_z h_c + a_z)]^2}. \quad (\text{A21})$$

h_{ref} is the reference height above canopy (10 m), h_c the canopy height, $k = 0.41$, $r_z = 0.1$ and $a_z = 1$.