Inversion of Terrestrial Ecosystem Model Parameter Values against Eddy Covariance Measurements by Monte Carlo Sampling

Wolfgang Knorr and Jens Kattge

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

Date of receipt:

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

Corresponding author:
Wolfgang Knorr
Max-Planck Institute for Biogeochemistry
PO Box 10 01 64
07701 Jena, Germany
Tel.: +49 40 41173 282
Fax: +49 40 41173 298
email: wknorr@bgc-jena.mpg.de

Running Title:
Monte Carlo Inversion of Ecosystem Model

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 (PDF) of parameters of a terrestrial ecosystem model. By thus assimilating half-hourly eddy covariance measurements of CO_2 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 seven days of measurements, uncertainties for net carbon uptake over two years for the forest site can be substantially reduced, with the median estimate in excellent agreement with measurements.

Introduction

Only about half of the increasing emissions of CO_2 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 CO_2 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 (Prentice et al., 2001; McGuire et al., 2001). However, results still vary significantly due to 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 and Heimann (2001a, b) have shown that uncertainties of TEM process parameters lead at least to the same spread of simulated atmosphere–vegetation carbon fluxes than inter-model 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. (2004). CCDAS requires to specify prior means and error covariance matrices of model parameters, as an approximation of the prior probability density function (PDF) of parameter. To generate and analyse such a PDF is one purpose of the present study.

Few attempts exist at quantifying uncertainty ranges based directly on experimental data (White et al., 2000; Knorr, 2000; Knorr and Heimann, 2001a). It is therefore of general interest to utilize the still growing amount of eddy covariance measurements of CO_2 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 CO_2, 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 thus 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., 2004), 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 TEMs, 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 CO_2 and water fluxes for a given set of prior parameter uncertainties and for a given error margin of measurements, using a particular TEM. The analysis of covariances is then used to determine which parameter values cannot be determined independently by the method. 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 CO_2 fluxes. The method is thus presented as a prototype for an initial step of CCDAS that allows the exploitation of widely availability 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 **p**, given a set of measurements **f**, in this case fluxes. Whether a given vector **p** agrees with **f** is determined by running the model **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 **c** and **s** vectors of environmental boundary conditions and model state variables, respectively. **f**, $\mathbf{f}_M$, **c** and **s** contain values across both time and types of data (CO_2, water and heat fluxes; temperature, solar radiation, humidity; soil moisture and leaf area index), while **p** is assumed invariable in time. For a process-based TEM, **M** is usually non-linear 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 savely be solved by direct sampling of the probability density function (PDF) in parameter space using Monte Carlo techniques. Developed for applications in nuclear (Metropolis et al., 1953), and later geophysics (Mosegaard and Tarantola, 1995; Mosegaard, 1998; Mosegaard and 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 $\mathbf{p}^1 ... \mathbf{p}^N$ 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 and 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 normalized parameters (see below). 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, $J_{\mathbf{f}}(\mathbf{p})$, such that

$$L(\mathbf{p}) = \exp\{-J_{\mathbf{f}}(\mathbf{p})\} \tag{3a}$$

with

$$J_{\mathbf{f}}(\mathbf{p}) = \frac{1}{2}\left(M(\mathbf{p}) - \mathbf{f}\right)^T \mathbf{C}_{\mathbf{f}}^{-1}\left(M(\mathbf{p}) - \mathbf{f}\right) \tag{3b}$$

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

$$\rho(\mathbf{p}) = \exp\{-J_{\mathbf{p}}(\mathbf{p})\} \tag{4a}$$

and

$$J_{\mathbf{p}}(\mathbf{p}) = \frac{1}{2}\left(\mathbf{p} - \mathbf{p}_0\right)^T \mathbf{C}_{\mathbf{p}}^{-1}\left(\mathbf{p} - \mathbf{p}_0\right) \tag{4b}$$

with $\mathbf{p}_0$, the vector of prior (normalized) parameter values, and $\mathbf{C}_{\mathbf{p}}$, the error covariance matrix of the priors.

In standard inversion techniques, the inversion problem consists of finding 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, $\mathbf{p}^1 \ldots \mathbf{p}^N$, simply has a distribution with its highest density in the vicinity of the maximum of $f(\mathbf{p})$. If the objective is less to find the exact optimum but more to gain understanding of the probability distribution of parameters, this technique has obvious advantages. The sampled distribution can subsequently be used to compute the expected values 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 multi-dimensional PDF onto the dimension of a single parameter (importance sampling) from:

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

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

The complete method of Monte Carlo inversion is described in detail by Mosegaard and Tarantola (1995) and reviewed by Mosegaard and 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 $\mathbf{p}^2$, $\mathbf{p}^3$,... in the 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 $\rho(\mathbf{p}^i+\Delta\mathbf{p})/\rho(\mathbf{p})<1$. The second step is assessed in the same way as the first, only 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 $\mathbf{p}^{i+1}=\mathbf{p}^i$. 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).

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 BETHY. The reduced version of BETHY is used together with the C4 photosynthesis model and excludes the heterotrophic respiration part. Compared to 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 and Verma (1991) from the 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 northeastern 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 day-time data of net canopy assimilation (GPP minus total-canopy leaf respiration) derived from eddy covariance measurements of NEE by subtracting soil and plant, excluding leaf, respiration rates derived from night-time CO_2 fluxes. We also assimilated day-time canopy conductance values that were obtained through inversion of the Penman-Monteith equation against day-time latent energy flux measurements. PAR, air temperature, VPD, and relative plant available soil moisture (w/w_m, Equ. A12, A17) were used as input data. All data, for four different days between June and August 1987, were taken from Kim and Verma (1991). Global radiation was computed from Julian day, longitude and latitude, while wind speed and free-air CO_2 concentration were left constant at 3 m/s 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 understorey of the grass *Deschampsia flexuosa* (Dolman et al. 2002). Global radiation, photosynthetically active radiation (PAR), air temperature,

ambient CO_2 concentration, wind speed, vapor pressure deficit (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 net ecosystem exchange (NEE) and latent energy flux (LE) from seven days in 1997 and 1998. 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 Wm^{-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_{JmVm} was derived from data by Wullschläger et al. (1993), Medlyn et al. (2002) and Leuning (2002); k^{25} and E_k follow Knorr (1997); E_{Rd} was set to the value cited by Kim and 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).

For the prior parameter uncertainty, we chose 0.125, 0.25, and 0.5, respectively, as one standard error in the space of normalized parameter values, p_i. These values were the same as those sampled by the Metropolis algorithm (see above), and were uniformly set to a prior value of 1. Those normalized parameter values are translated into model parameters, $\underline{p}_i$ (see previous section) through a logarithmic transformation, given by:

$$\underline{p_i} = \ln(p_i - 1)\underline{p_{i,0}} \quad (7)$$

$\underline{p_{i,0}}$ is the prior value in model parameter space as listed in Table 1. We chose a log-normal distribution because of large prior uncertainties with all parameters positive by definition. For f_{Ci}, however, we require $0 \leq f_C \leq 1$, and so that we chose to use a normal distribution cut off at 0 and 1. $f_{Ci,0}$ is the prior estimate of f_{Ci}, and $f_{Ci}=p_k f_{Ci,0}$ replaces Equ. 7, where k is the parameter index for f_{Ci}. The vector of prior normalized parameters is thus $\mathbf{p}_0=\{1,...,1\}$, and $\mathbf{C_p}$, the error covariance matrix of the priors:

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

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, $\rho(\mathbf{p})$ (Equ. 4), we have the additional condition

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

291 **Results**

292

293 We will first show results to demonstrate convergence of the algorithm. Next,
294 optimized parameter values will be described by their means, standard errors, and
295 covariances, all in the space of normalized parameters (cf. Equ. 7). Comparison with
296 prior means and errors indicates about how many parameters we have learned
297 something through the assimilation of the eddy covariance data. We also assess for
298 which parameters the posterior PDF differs from the Gaussian distribution assumed
299 for the prior PDF. For the Loobos site, we eventually compute the cumulative NEE
300 with and without optimized parameters over a period of two years to test the validity
301 of parameterizations across time and to assess to what degree the inversion has lead to
302 a constraint on the modeled ecosystem carbon balance.

303

304 *Convergence of the algorithm*

305

306 To insure convergence, we performed rather long integrations with 500,000 iterations
307 (and more in one case). For the two cases with 0.25 prior uncertainty, we produced a
308 series of six independent simulations starting from different points in parameters
309 space: the prior parameter vector, $\mathbf{p}_0=\{1,...,1\}$ in the space of normalized parameters,
310 and points shifted away from the estimated posterior optimum, **p',** by one to several
311 times the posterior standard deviations, $\boldsymbol{\sigma}'=\{\sigma'_1,...,\sigma'_n\}$ estimated from preliminary
312 simulations. For FIFE, the starting points were $\mathbf{p}_0$, $\mathbf{p}'+\boldsymbol{\sigma}'$, $\mathbf{p}'+2\boldsymbol{\sigma}'$, $\mathbf{p}'+3\boldsymbol{\sigma}'$, $\mathbf{p}'-\boldsymbol{\sigma}'$,
313 $\mathbf{p}'-2\boldsymbol{\sigma}'$, for Loobos $\mathbf{p}_0$, $\mathbf{p}'+2\boldsymbol{\sigma}'$, $\mathbf{p}'-2\boldsymbol{\sigma}'$, $\mathbf{p}'+4\boldsymbol{\sigma}'$, $\mathbf{p}'+4\{+\sigma'_1,-\sigma'_2,+\sigma'_3,...\}$ and
314 $\mathbf{p}'-4\{+\sigma'_1,-\sigma'_2,+\sigma'_3,...\}$. To determine at which iteration the sequences have
315 converged to a common maximum, as opposed to sampling local maxima, we applied
316 Gelman's criterion of convergence (Gelman et al., 1995) for all parameters. This test
317 of convergence, designed for practical purposes, yields a reduction factor that
318 measures both the variance within each sequence of the series, and the variance of
319 means across sequences for exactly the second half of the series up to the iteration
320 indicated.

321

322 The parameters that took longest to reach a common maximum, according to
323 Gelman's criterion, were α_i for FIFE and f_{Ci} for Loobos. The evolution of the

estimated mean values are shown in Figs. 1a and 1b, respectively, for every tenth iteration. Also shown are one fast converging parameter, and the parameter that was most highly correlated to the first. Note that for in Fig. 1b, E_{Vm} appears to be converging more slowly than f_{Ci}. The explanation is that E_{Vm} 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 (Equ. 4b, costs of parameters) and the misfit in the Likelihood-function (Equ. 3b, costs of diagnostics) as the two axes (Gelman et al., 1995). As Figs. 1c and 1d show 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 week 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 day-time fluxes.

357

358 To determine a practical initial cut-off for iterations before convergence to the global
359 PDF maximum, the so-called "burn-in time" with length n iterations, we used again
360 Gelman's test (Gelman, 1992; Cowles and Carlin, 1996). It requires that the reduction
361 factor computed for iterations n+1 to 2n reaches a value of around 1.2 to 1.4 for all
362 sampled quantities of interest. To be on the safe side, we chose 2n=100,000. Fig. 1e
363 and 1f show this reduction factor for the same parameters as Fig. 1a and 1b, together
364 with the values of the product of the fastest converging parameter with the two others.
365 Such products are required to compute parameter covariances and appear to converge
366 at least as rapidly as the slowest parameter.

367

368 Convergence of parameters for the cases with 0.125 and 0.5 prior uncertainties was
369 evaluated by plotting expected values of all parameters against the length of the burn-
370 in time. A burn-in time of 50,000 iterations was found to be sufficient for all cases but
371 FIFE 0.125, where 1,000,000 iterations were chosen instead. Of the following
372 450,000 iterations, we used every tenth step for parameter sampling to avoid
373 correlations between subsequent samplings. The following analysis was thus based on
374 45,000 parameter samplings for 0.125 and 0.5 prior uncertainties, and 270,000
375 samplings from six sequences for 0.25 prior uncertainties. Each sequence of 500,000
376 iterations took ca. 5 hours CPU time on a Linux PC workstation.

377

378 *Parameter change and uncertainty reduction from constraining with eddy covariance*
379 *data*

380

381 Means and standard deviations can be estimated directly from the samplings of the
382 posterior PDF in the space of the normalized parameters. Since the parameters
383 represent different processes, comparison with prior means and uncertainties provides
384 valuable information on those processes about which we can learn most through the
385 use of eddy covariance data. The means and ranges corresponding to one standard
386 error are shown in Fig. 2 for all prior and posterior parameter values. For the non-
387 Gaussian prior distribution of f_{Ci}, we show the corresponding percentiles.

388

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 CO_2 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 (Equ. 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_e, J_c, and J_i, 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 5 parameters: Γ_*^{25}, K_C^{25}, E_{Rd}, 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_{Vm}. For others, there is no consistency: the direction of change depends on the prior uncertainty (for r_{JmVm}, 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_{Ci} are almost independent of the prior uncertainty ranges, and there is a universal downward adjustment of the third.

Another quantity that measures the gain in information after inversion against the eddy covariance data is the relative reduction in uncertainty, defined as $1-\sigma_{prior}/\sigma_{posterior}$, where σ is the parameter's standard deviation. For f_{Ci} 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_{Ci}. This is not a great surprise, since 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 CO_2 flux measurements. This results should caution us against the use of night-time CO_2 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 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 normalized covariance (=correlation coefficient) of up to 0.91 for the pair c_w and ε_s. Two parameters, f_{Rd} 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_{Vm}, r_{JmVm}, Γ_*^{25}, K_C^{25}) plus the stomatal parameter f_{Ci}. These are linked to a second energy balance group consisting of ε_s and $g_{a,v}$ via f_{Ci}, E_{Vm} and α_q. f_{Ci} 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 normalized covariance 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 normalized covariance between Q_{10} and E_{Vm}.

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, here of the model parameters instead of the 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 carbon cycle data assimilation system (see Introduction). The analysis is based on the medium case of 0.25 prior uncertainty of normalized parameters. The skewness and kurtosis of the PDF projected onto each normalized parameter, also listed in Table 3, show only small deviations from Gaussian distributions. Exceptions are ε_s for FIFE, and E_{Vm}, $g_{a,v}$ and w_{pwp} for Loobos (see Fig. 4). E_{Vm}, $g_{a,v}$ 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_{Ci} (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 now obtained a constrained parameter PDF for the BETHY C4 and C3 models from four or seven 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

488 have computed the cumulative NEE over a period of two years at the Loobos site,
489 complete with 95% confidence ranges, from the prior, the posterior Gaussian, and the
490 full posterior PDF. The posterior Gaussian PDF approximates the full PDF by using
491 only the means and the error covariance matrix. As Fig. 5 shows by the green area,
492 prior uncertainties about parameter values of BETHY were consistent with the
493 Loobos site being both a strong sink (positive NEE), or a moderate source of carbon
494 (negative NEE) over the two years. After constraining the model, the 95% confidence
495 range lies outside of the median prior estimate. This means that extrapolating seven
496 days of NEE and LE data through the assimilation procedure resulted in a sink
497 estimate that was significantly different from the best prior estimate. Further, we find
498 that using the full PDF in parameter space results in only about half of the uncertainty
499 in NEE over the two years compared to using a PDF derived from parameter means
500 and covariances. Skewness and kurtosis of the full PDF of the cumulative NEE can
501 also be relatively large.
502
503 Note that this result still depends on the prior uncertainty, which was only estimated
504 in a simple and preliminary way for this study. Also, assimilating more days of flux
505 measurements would lead to stronger constraints of model parameters and fluxes,
506 which would lead to even smaller uncertainties of the cumulative NEE. Here, we can
507 instead use the measured NEE of the two years, with a few gaps (for which we
508 assumed NEE=0), to validate our time extrapolation (Fig. 5, blue line). With this
509 additional assumption as a point of caution, we arrive at around 25 mol(CO_2)/m^2/yr or
510 300 gC/m^2/yr net uptake from both the observations and the model simulations. The
511 generally good agreement between modeled (after assimilation) and measured NEE
512 across the two years shows that the model is able to capture the main processes that
513 influence this quantity. We therefore suggest that the method shown here, with all
514 available measurements assimilated, could be a superior gap filling method compared
515 to 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 to 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 can not. With four or seven days of diurnal data assimilated, the Bayesian part of the cost function in the region of the minimum was between two and ten 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. Even though this depends on somewhat subjective assumptions about the prior uncertainty and what degree or relative error reduction can be considered as “effectively constrained”, this particular result is rather robust.

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 be distributed 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., 2004). We expect that using the Gaussian part of the complete PDF will tend to overestimate the uncertainty of the model diagnostics.

549

550 We have so far restricted our study to cases that are rather rare when considering the
551 entire FLUXNET archive: we relied on the availability of soil moisture
552 measurements. Applying the method for more sites, however, will be crucial for
553 identifying representative model parameterizations by plant functional type, or some
554 other generalization on which global models necessary rely on. Therefore, we expect
555 to conduct further studies with the complete BETHY model with the full water
556 balance. If no complete data on LAI are available, a phenology scheme may also be
557 included. LAI and soil moisture data could then also be assimilated instead of being
558 used as input. We also suggest to use more days and longer periods for assimilation,
559 although we find that only a few days of data already deliver a strong model
560 constraint.

561

562 Finally, we have considered only some first approximations for prior parameter
563 uncertainties, and for uncertainties in eddy flux measurements. We found that in some
564 cases (for the FIFE site, in particular), the results depended rather strongly on the
565 choice of the prior parameter uncertainties. This means that a careful choice of prior
566 parameter values and their uncertainties can be important for the optimization.

567

568

Conclusions

The parameterization of global terrestrial ecosystem models 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 non-linear models, and that it allows to sample the complete posterior probability density function. 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 at the global scale.

Further application 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 application 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 terrestrial ecosystem models. These will be crucial for any application aimed at predicting the future response of the carbon cycle to climate change, including atmosphere–vegetation feedbacks.

602 **Acknowledgments**

603

604 The authors wish to thank Isabel van den Wyngaert and Bart Kruijt for most
605 cooperative preparation of data and help with estimating uncertainties of
606 measurements at the Lobos site, and Thomas Kaminski for useful advice and
607 comments. This work has been financed and supported by the EU project CAMELS,
608 contract number EVK2-CT-2002-00261, within the EU's 5th framework program for
609 Research and Development, and the Max-Planck-Gesellschaft zur Förderung der
610 Wissenschaften, e.V.

611

612 **Literature**

613

614 Collatz GJ, Ribas-Carbo M, Berry JA (1992) Coupled photosynthesis-stomatal
615 conductance model for leaves of C4 plants. *Australian Journal of Plant*
616 *Physiology*, **19**, 519-538.
617 Cowles MK, Carlin BP (1996) Markov chain Monte Carlo convergence diagnostics:
618 A comparative review. *Journal of the American Statistical Association*, **91**, 883-
619 904.
620 Cox, PM, Betts RA, Jones CD, Spall SA, and Totterdell IJ (2000) Acceleration of
621 global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*,
622 **408**, 184-187.
623 Cramer W, Kicklighter DW, Bondeau A *et al*. (1999) Comparing global models of
624 terrestrial net primary productivity (NPP): overview and key results. *Global*
625 *Change Biology*, **5, Suppl. 1**, 1-15.
626 Cramer W, Bondeau A, Woodard FI *et al*. (2001) Global response of terrestrial
627 ecosystem structure and function to CO2 and climate change: results from six
628 dynamic global vegetation models. *Global Change Biology*, **7**, 357-373.
629 Dolman AJ, Moors EJ, Elbers JA (2002) The carbon uptake of a mid latitude pine
630 forest growing on sandy soil. *Agricultural and Forest Meteorology*, 157–170.
631 Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of
632 photosynthesis in leaves of C3 species. *Planta*, **149**, 78-90.
633 Friedlingstein P, Dufresne JL, Cox PM, Rayner PM (2003) How positive is the
634 feedback between climate change and the carbon cycle? *Tellus*, **55B**, 692–700.
635 Gelman A, Carlin JB, Stern HS, Rubin DB (1995) *Bayesian Data Analysis*, Chapman
636 and Hall/CRC, New York.
637 Gelman A, Rubin DB (1992) A single squence from the Gibbs sampler gives a false
638 sense of security. In: *Bayesian Statistics* (eds Bernardo JM, Berger JO, Dawid AP,
639 Smith AFM), pp. 625-631. Oxford University Press, Oxford.
640 Global Carbon Project (2003) *Science Framework and Implementation*, Global
641 Carbon Project, Earth System Science Partnership (IGBP, IHDP, WCRP,
642 DIVERSITAS), Canberra.
643 Joos F, Plattner GK, Stocker TF, Körtzinger A,Wallace DWR (2003) Trends in
644 Marine Dissolved Oxygen: Implications for Ocean Circulation Changes and the

645 Carbon Budget. *EOS*, **84**, 197-201.
646 Kaminski T, Knorr W, Heimann M, Rayner P (2002) Assimilating Atmospheric data
647 into a Terrestrial Biosphere Model: A case study of the seasonal cycle. *Global*
648 *Biogeochemical Cycles*, **16**, 1066, doi:10.1029/2001GB001463.
649 Kim J, Verma SB (1991) Modeling canopy photosynthesis: scaling up from a leaf to
650 canopy in a temperate grassland ecosystem. *Agricultural and Forest Meteorology*,
651 **57**, 149-166.
652 Knorr W (1997) *Satellite remote sensing and modelling of the global CO_2 exchange of*
653 *land vegetation: a synthesis study*, Max-Planck-Institut für Meteorologie,
654 Hamburg, Germany,185 pp.
655 Knorr W (2000) Annual and interannual CO_2 exchanges of the terrestrial biosphere:
656 process-based simulations and uncertainties. *Global Ecology and Biogeography*, **9**,
657 225-252.
658 Knorr W, Heimann M (2001a) Uncertainties in global terrestrial biosphere modeling.
659 Part I: A comprehensive sensitivity analysis with a new photosynthesis and energy
660 balance scheme. *Global Biogeochemical Cycles*, **15**, 207-225.
661 Knorr W, Heimann M (2001b) Uncertainties in global terrestrial biosphere modeling.
662 Part II: global constraints for a process-based vegetation model. *Global*
663 *Biogeochemical Cycles*, **15**, 227-246.
664 Leuning R (2002) Temperature dependence of two parameters in a photosynthesis
665 model. *Plant, Cell and Environment*, **25**, 1205-1210.
666 McGuire AD, Sitch S, Clein JS *et al*. (2001) Carbon balance of the terrestrial
667 biosphere in the twentieth century: Analyses of CO2, climate and land use effects
668 with four process-based ecosystem models. *Global Biogeochemical Cycles*, **15**,
669 183-206.
670 Medlyn BE, Loustau D, Delzon S (2002) Temperature response of parameters of a
671 biochemically based model of photosynthesis. I. Seasonal changes in mature
672 maritime pine (Pinus pinaster Ait.). *Plant, Cell and Environment*, **25**, 1155-1165.
673 Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation
674 of state calculations by fast computing machines. *Journal of Chemical Physics*, **21**,
675 1087-1092.
676 Monteith JL (1965) Evaporation and environment. *Symposium of the Society for*
677 *Experimental Biology*, **19**, 205-234.

678 Mosegaard K (1998) Inverse Problems. *Inverse Problems*, **14**, 405-426.
679 Mosegaard K, Rygaard-Hjalsted C (1999) Probabilistic analysis of implicit inverse
680 problems. *Inverse Problems*, **15**, 572-583.
681 Mosegaard K, Sambridge M (2002) Monte Carlo analysis of inverse problems.
682 *Inverse Problems*, **18**, R29-R54.
683 Mosegaard K, Tarantola A (1995) Monte Carlo sampling of solutions to inverse
684 problems. *Journal of Geophysical Research*, **100**, 12431-12447.
685 Prentice IC, Farquhar GC, Fasham MJR *et al.* (2001) The Carbon Cycle and
686 Atmospheric Carbon Dioxide. In: *Climate Change 2001: The Scientific Basis.*
687 *Contribution of Working Group I to the Third Assessment Report of the*
688 *Intergovernmental Panel on Climate Change* (eds Houghton JT *et al.*), pp. 185-
689 225, Cambridge University Press, Cambridge.
690 Raich JW, Potter CS, Bhagawati D (2002) Interannual variability in global soil
691 respiration, 1980-94. *Global Change Biology*, **8**, 800-812.
692 Rayner, PJ, Scholze M, Knorr W, Kaminski T, Giering R, Widmann H (2004) Two
693 decades of terrestrial carbon fluxes from a Carbon Cycle Data Assimilation System
694 (CCDAS). *Global Biogeochemical Cycles*, submitted.
695 Reichstein M, Rey A, Freibauer A (2003) Modeling temporal and large-scale spatial
696 variability of soil respiration from soil water availability, temperature and
697 vegetation productivity indices. *Global Biogeochemical Cycles*, **17**, 1104,
698 doi:10.1029/2003GB002035.
699 Schachtschabel P, Blume HP, Bruemmer G, Hartge KH, and Schwertmann U (1992)
700 *Lehrbuch der Bodenkunde*, Ferdinand Enke Verlag, Stuttgart, 178pp.
701 Sellers PJ (1985) Canopy reflectance, photosynthesis and transpiration. *International*
702 *Journal of Remote Sensing*, **6**, 1335-1372.
703 Wang YP, Leuning R, Cleugh H, Coppin PA (2001) Parameter estimation in surface
704 exchange models using nonlinear inversion: how many parameters can we estimate
705 and which measurements are most useful? *Global Change Biology*, **7**, 495-510.
706 Weiss A, Norman JA (1985) Partitioning solar radiation into direct and diffuse,
707 visible and near-infrared components. *Agricultural and Forestry Meteorology*, **34**,
708 205-213.
709 White MA, Thornton PE, Running SW, Nemani RR (2000) Parameterization and
710 Sensitivity Analysis of the BIOME–BGC Terrestrial Ecosystem Model: Net

711 Primary Production Controls. *Earth Interactions*, **4**, 1-85.
712 Wullschleger SD (1993) Biochemical limitations to carbon assimilation in C3 plants -
713 a retrospective analysis of the A / c_i curves from 109 species. *Journal of*
714 *Experimental Botany*, **44**, 907-920.
715

716 Table 1: Parameters of BETHY that were used in the inversion against eddy
717 covariance measurements.

symbol	description	value	unit	Equ.	C3	C4
photosynthesis						
α_q	quantum efficiency of photon capture (C3)	0.28	mol(e^-)/mol	A1c	X	
V_m^{25}	maximum carboxylation rate at 25 °C (C3)	29	μmol/m^2/s^1	A2	X	
V_m^{25}	maximum carboxylation rate at 25 °C (C4)	8	μmol/m^2/s^1	A2		X
E_{Vm}	activation energy of V_m	58520	J/mol	A2	X	X
r_{JmVm}	ratio of J_m to V_m at 25 °C	1.79	-	A3	X	
Γ_*^{25}	CO_2 compensation point without dark resp. at 25 °C	42.5	μmol/mol	A4	X	
K_C^{25}	Michaelis Menten constant for carboxylation at 25 °C	460	μmol/mol	A5	X	
E_{Kc}	activation energy of K_C	59356	J/mol	A5	X	
K_O^{25}	Michaelis Menten constant for oxygenation at 25 °C	0.33	mol/mol	A6	X	
E_{Ko}	activation energy of K_O	35948	J/mol	A6	X	
α_i	quantum efficiency of photon capture (C4)	0.04	mol(CO_2)/mol	A7		X
k^{25}	CO_2 specificity at 25 °C	0.14	mol/m^2/s^1	A8		X
E_k	activation energy of k	50967	J/mol	A8		X
carbon balance						
f_{Rd}	ratio of leaf dark respiration at 25 °C and V_m^{25} (C3)	0.011	-	A10	X	
f_{Rd}	ratio of leaf dark respiration at 25 °C and V_m^{25} (C4)	0.042	-	A10		X
E_{Rd}	activation energy of leaf dark respiration	45000	J/mol	A10	X	X
$f_{R,leaf}$	ratio of canopy to total plant respiration	0.5	-	A11	X	
R_{het}^0	heterotrophic respiration at 0 °C and field capacity	2.07	μmol/m^2/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	
w_{pwp}	soil water content at permanent wilting point	2.5	vol%	-	X	
stomatal control						
f_{Ci}	non water limited ratio of $C_{i,0}$ and C_a (C3)	0.87	-	A14	X	
f_{Ci}	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/hour	A17	X	X
energy and radiation balance						
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 rad. abs. 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	

718

719 Table 2: Elements of the reduced error covariance matrix, equal to the correlation
720 coefficient, derived from the posterior PDF the FIFE and Loobos sites. Prior
721 parameter uncertainties in normalized space were 0.25. Values above 0.3 or below -
722 0.3 are considered to indicate a close correlation between parameter errors and are
723 marked.

α_q	V_m^{25}	E_{Vm}	r_{JmVm}	Γ_*^{25}	K_c^{25}	E_{Kc}	K_O^{25}	E_{Ko}	f_{Rd}	E_{Rd}	$f_{R,leaf}$	R_{het}^0	κ	Q_{10}	w_{pwp}	f_{Ci}	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	α_q
	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	V_m^{25}
Loobos		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	E_{Vm}
			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	r_{JmVm}
				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	Γ_*^{25}
FIFE					1.00	0.05	0.18	-0.01	-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	K_c^{25}
						1.00	0.00	-0.11	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	E_{Kc}
							1.00	-0.02	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	K_O^{25}
V_m^{25}	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	E_{Ko}
E_{Vm}	0.19	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	f_{Rd}
α_i	0.44	0.08	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	E_{Rd}
k^{25}	0.00	-0.03	-0.25	1.00							1.00	0.06	-0.22	0.15	-0.01	0.05	0.03	0.03	0.01	0.03	-0.01	-0.07	$f_{R,leaf}$
E_k	-0.10	-0.03	-0.30	-0.15	1.00							1.00	0.23	-0.39	-0.01	-0.19	-0.04	0.01	0.02	0.00	0.06	0.03	R_{het}^0
f_{Rd}	-0.15	-0.07	0.45	0.09	-0.08	1.00							1.00	0.36	0.00	0.07	0.04	-0.02	0.00	-0.02	0.01	-0.05	κ
E_{Rd}	-0.15	0.03	-0.16	-0.28	-0.05	-0.35	1.00							1.00	0.12	0.22	0.18	0.01	-0.06	-0.01	-0.12	-0.11	Q_{10}
f_{Ci}	-0.22	-0.06	-0.15	-0.65	0.14	0.06	0.21	1.00							1.00	0.03	0.75	0.00	0.00	0.02	0.12	0.09	w_{pwp}
c_w	-0.26	-0.03	-0.23	-0.07	0.02	-0.42	-0.09	-0.05	1.00							1.00	0.17	-0.01	-0.06	0.02	-0.58	-0.47	f_{Ci}
ω	0.01	0.03	0.03	0.03	0.08	-0.03	0.03	-0.03	-0.17	1.00							1.00	-0.01	-0.05	0.00	-0.12	-0.10	c_w
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	-0.01	0.01	0.02	0.01	ω
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							1.00	-0.01	0.12	0.01	a_v
ε_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	0.02	0.01	a_s
$g_{a,v}$	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	0.67	ε_s
	V_m^{25}	E_{Vm}	α_i	k^{25}	E_k	f_{Rd}	E_{Rd}	f_{Ci}	c_w	ω	a_v	a_s	ε_s	$g_{a,v}$								1.00	$g_{a,v}$

724

Table 3: Prior and posterior parameter values in model space for FIFE (above, BETHY C4 version), and Loobos (below, BETHY C3 version); standard deviation (SD) of the prior parameters, as well as SD, skewness and kurtosis of the posterior parameters, in normalized space. In the normalized parameter space, prior distributions are Gaussian.

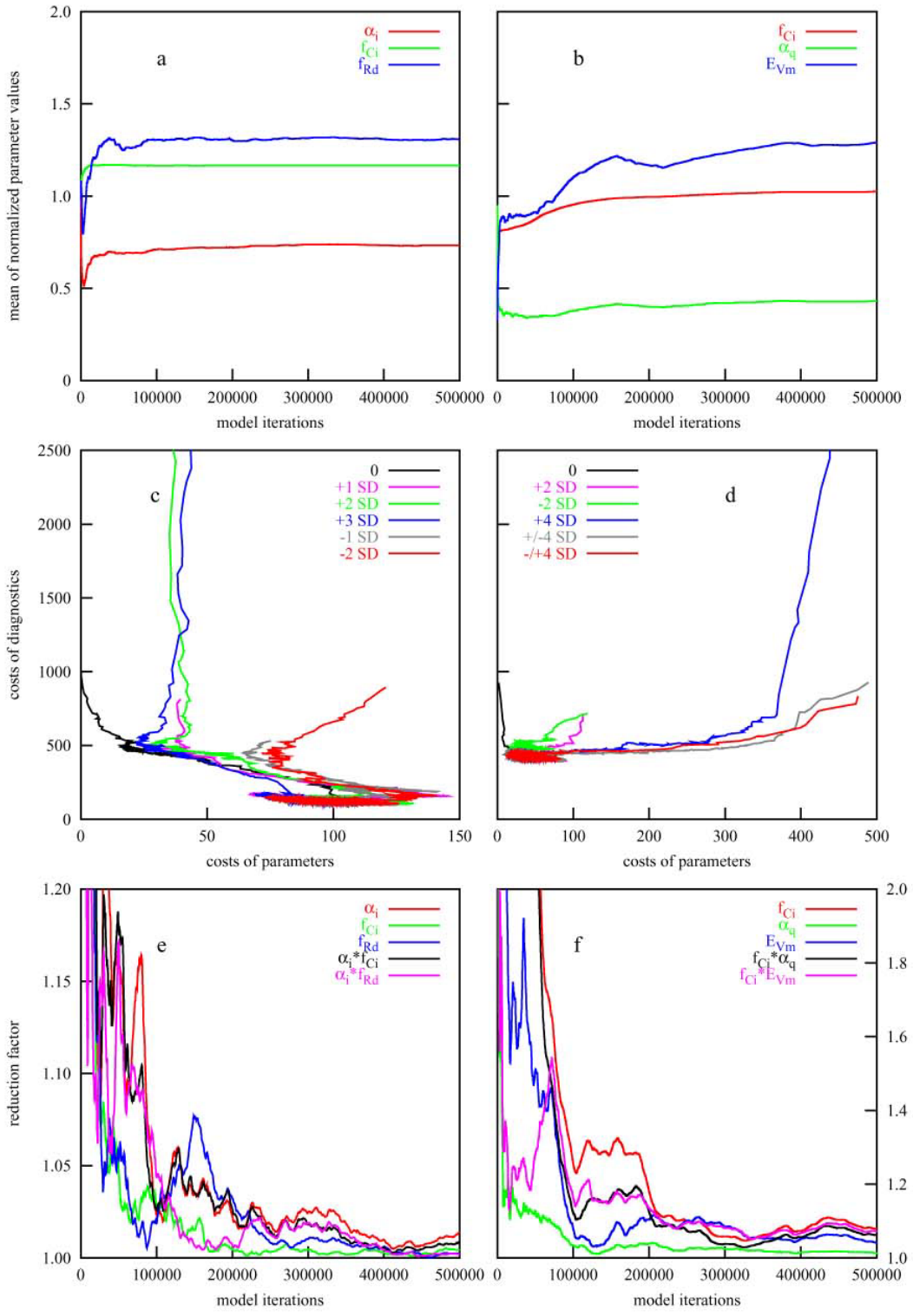
[1)]Prior distribution is Gaussian with a cutoff at 0 and 1 in model space. Shown is the 68.3 percentile range which is equivalent to 1 SD.

	model parameter		normalized parameter			
parameter	**prior p_0**	**posterior mean**	**prior SD**	**posterior SD**	**skewness**	**kurtosis**
V_m^{25}	8.00E-06	1.59E-05	0.25	0.15	0.33	0.33
E_{Vm}	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_{Rd}	4.20E-02	5.62E-02	0.25	0.20	-0.01	-0.04
E_{Rd}	4.50E+04	9.47E+04	0.25	0.09	-0.17	-0.02
f_{Ci}	6.70E-01	7.82E-01	0.24[1)]	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_{Vm}	5.85E+04	7.99E+04	0.25	0.26	-1.10	1.01
r_{JmVm}	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_{Kc}	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_{Ko}	3.60E+04	3.77E+04	0.25	0.28	0.14	0.16
f_{Rd}	1.00E-02	9.69E-03	0.25	0.23	-0.01	-0.05
E_{Rd}	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	1.98E+00	0.25	0.12	-0.55	0.54
f_{Ci}	8.70E-01	9.05E-01	0.20[1)]	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

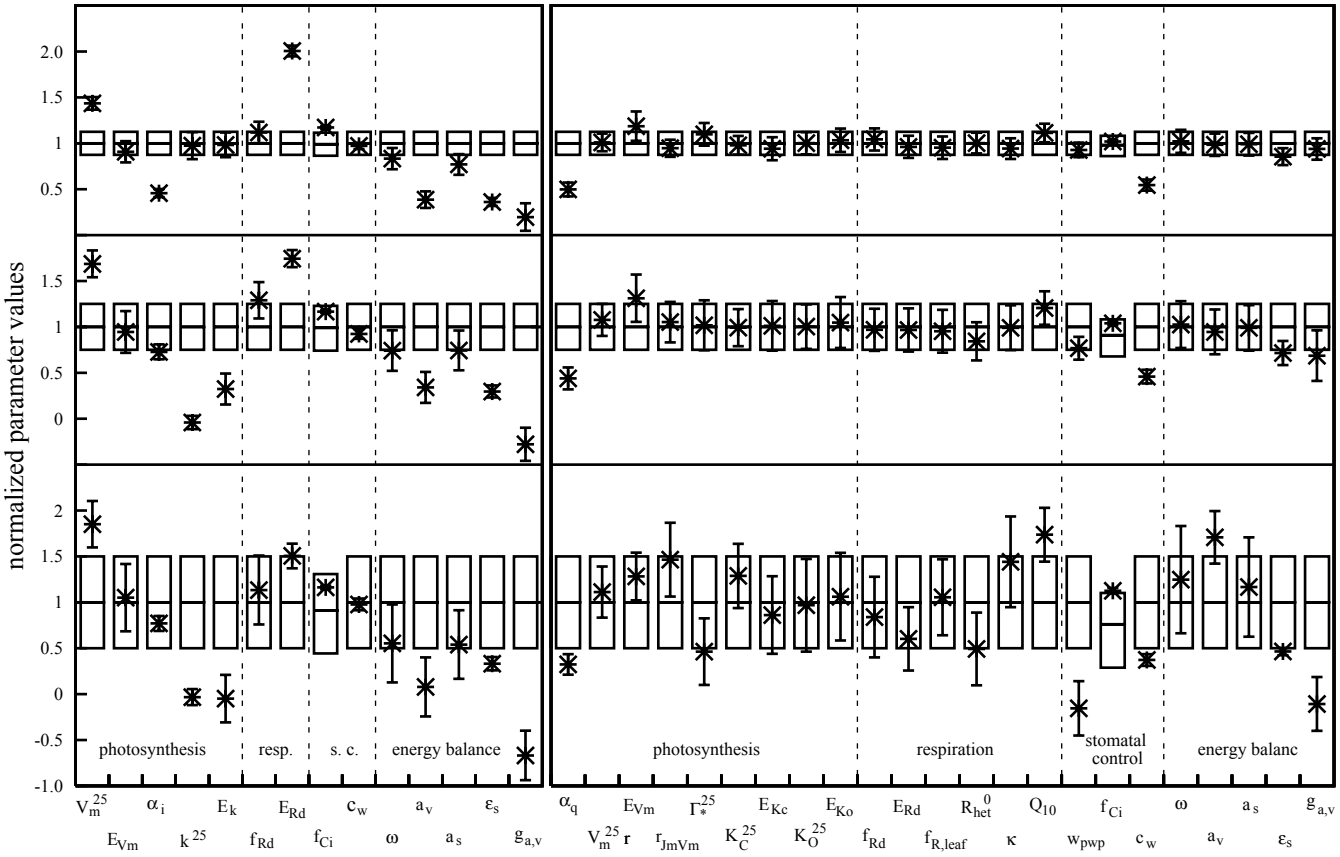
737 Figure 1: Convergence of the Monte Carlo inversion, for the FIFE (left: a, c, e), and
738 the Loobos site (right: b, d, e). a, b: Estimated mean of selected parameters depending
739 on number of iterations; c, d: phase diagram of the two contributions to the total cost
740 function, measuring deviation from prior parameters and between measured and
741 modeled diagnostics (=fluxes), for sequences with varying starting points; e, f:
742 Gelman's reduction factor for the same parameters as above, and for two parameter
743 products. The selected parameters are: the slowest converging, one fast converging,
744 and the one most highly correlated with the first. Prior uncertainty of normalized
745 parameters was 0.25.
746
747
748
749
750
751 Figure 2: Prior and a posterior parameter values and uncertainties in normalized
752 space. The boxes show means and one standard deviation of assumed prior parameters
753 (SD = 0.1, 0.25, 0.5). Crosses show the posterior means, and error bars one standard
754 deviation of the posterior parameters. Left: BETHY C4 version constrained with data
755 from FIFE site; right: BETHY C3 version constrained with data from Loobos site.
756
757
758
759
760
761 Figure 3: Relative reduction of uncertainty of parameter values. Zero or negative
762 relative error reduction indicates that no information about a particular parameter
763 could be gained, one would mean perfect knowledge of the inversion. Left: BETHY
764 C4, FIFE site; Right: BETHY C3, Loobos site.
765
766
767
768
769

770 Figure 4: Probability distributions of selected parameters from FIFE (a) and Loobos
771 (b–d). Comparison of importance sampling, approximating the true distribution, to the
772 prior PDF and to posterior Gaussian PDF computed from mean and standard
773 deviation. Additionally, the mean, standard deviation, skewness and kurtosis are given
774 for the posterior distribution. Prior uncertainty: 0.25 in normalized parameter space.
775
776
777 Figure 5: Cumulative NEE for two years, using the results from the inversion against
778 seven days of NEE and LE, for the Loobos site. Green: prior uncertainty range,
779 yellow: posterior uncertainty range using posterior mean and error covariance
780 (Gaussian posterior PDF);red: posterior uncertainty range with full PDF; blue:
781 measurements, dashed: missing data (NEE=0 assumed).

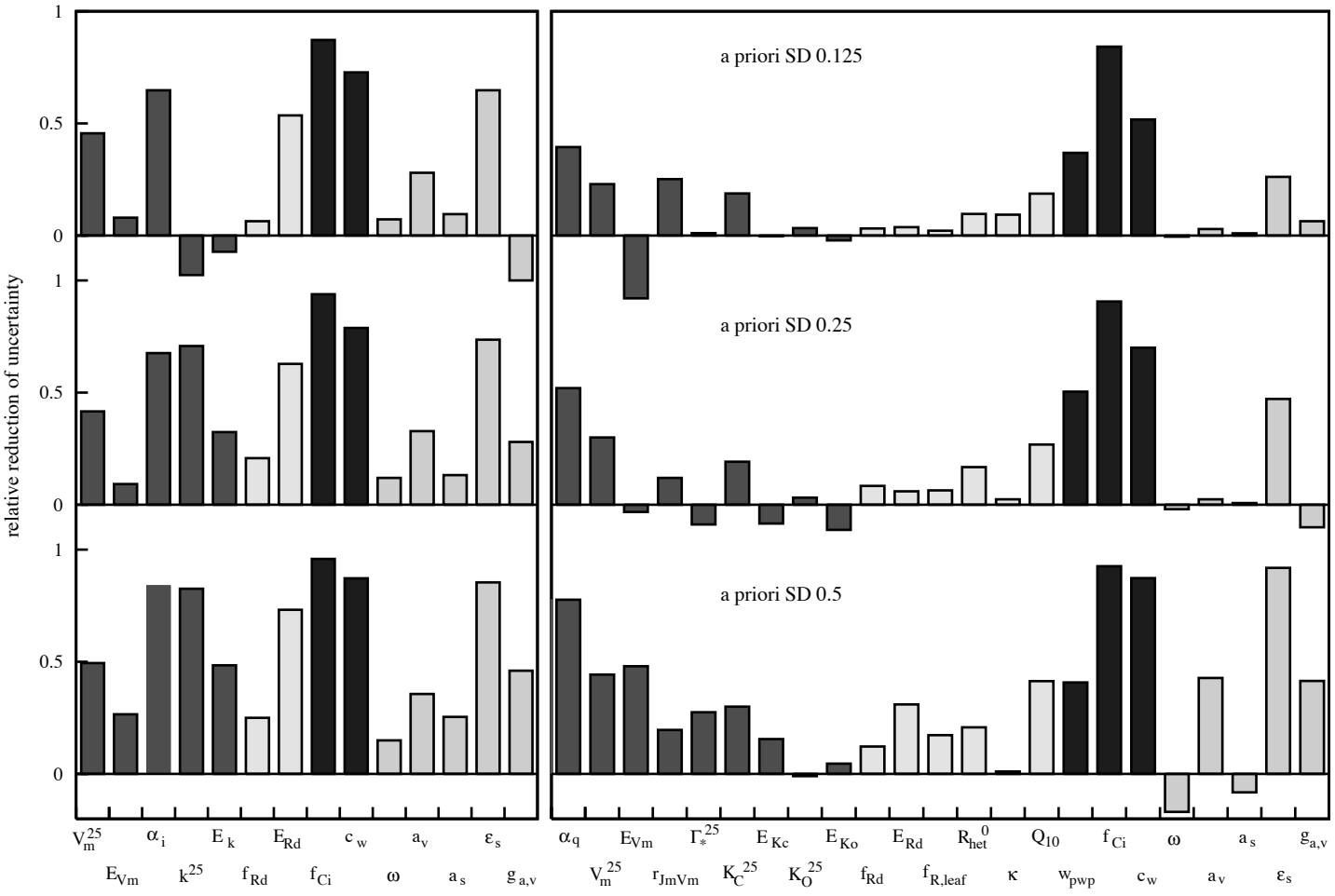
782 Figure 1:


a
b
c
d
e
f
mean of normalized parameter values
model iterations
costs of diagnostics
costs of parameters
reduction factor
α_i
f_{Ci}
f_{Rd}
α_q
E_{Vm}
α_i*f_{Ci}
α_i*f_{Rd}
f_{Ci}*α_q
f_{Ci}*E_{Vm}
0
+1 SD
+2 SD
+3 SD
-1 SD
-2 SD
+4 SD
+/-4 SD
-/+4 SD

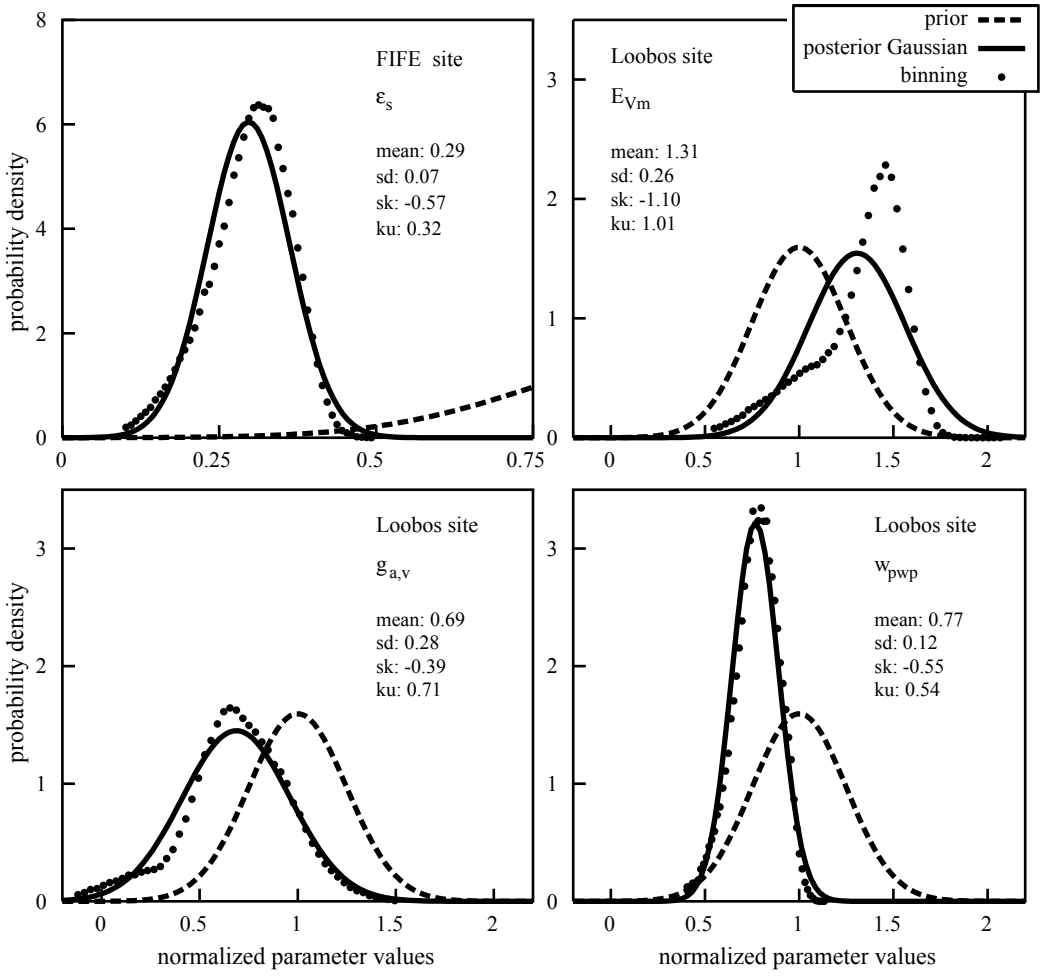
783 Figure 2:


normalized parameter values
2.0
1.5
1
0.5
1.5
1
0.5
0
2
1.5
1
0.5
0
-0.5
-1.0
photosynthesis
resp.
s. c.
energy balance
photosynthesis
respiration
stomatal control
energy balanc
V_m^{25} α_i E_k E_{Rd} c_w a_v ε_s
E_{Vm} k^{25} f_{Rd} f_{Ci} ω a_s $g_{a,v}$
α_q E_{Vm} Γ_*^{25} E_{Kc} E_{Ko} E_{Rd} R_{het}^0 Q_{10} f_{Ci} ω a_s $g_{a,v}$
V_m^{25} r r_{JmVm} K_C^{25} K_O^{25} f_{Rd} $f_{R,leaf}$ κ w_{pwp} c_w a_v ε_s

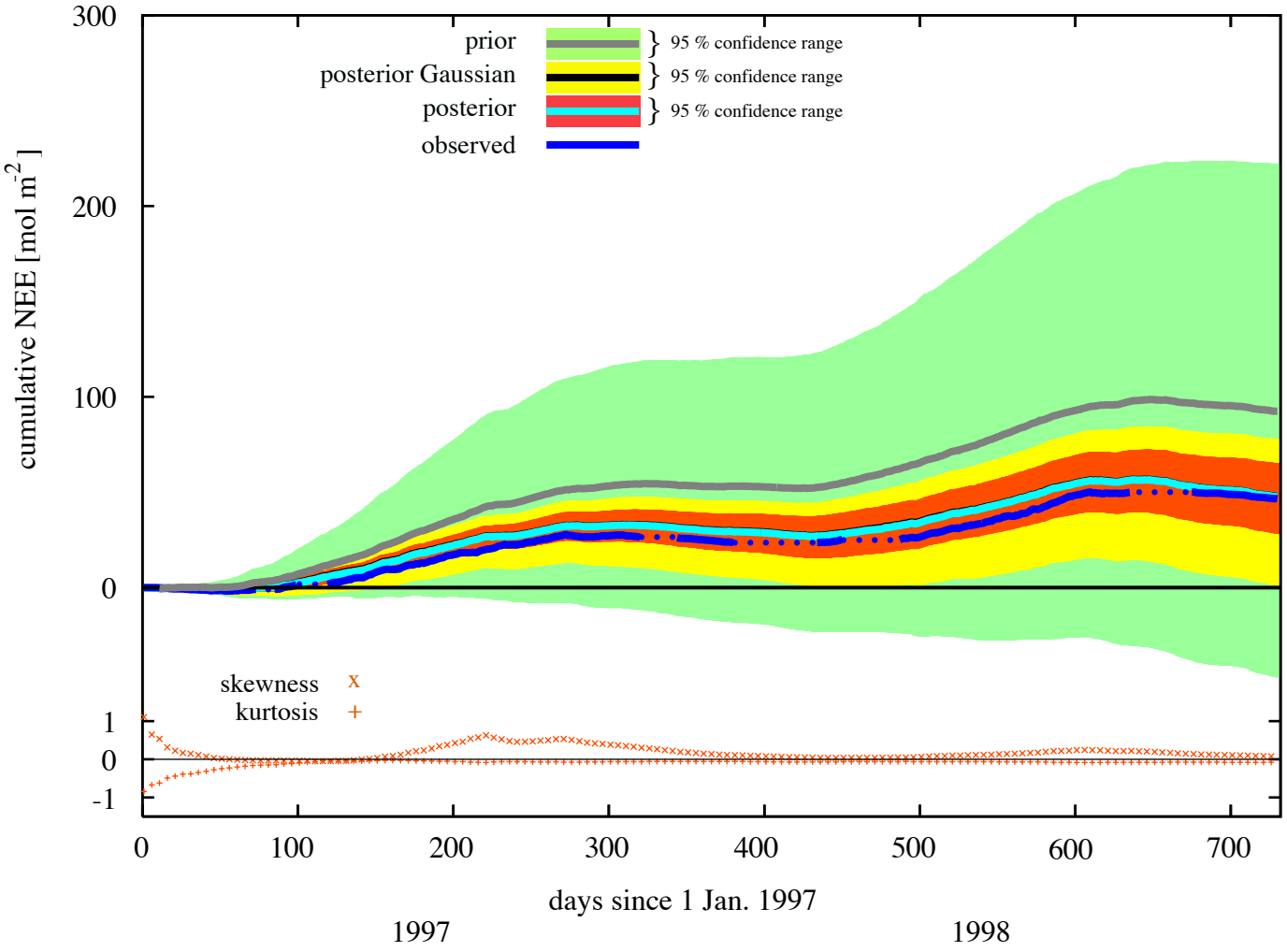
784 Figure 3:


relative reduction of uncertainty
a priori SD 0.125
a priori SD 0.25
a priori SD 0.5
1
0.5
0
V_m^{25} α_i E_k E_{Rd} c_w a_v ε_s
E_{Vm} k^{25} f_{Rd} f_{Ci} ω a_s $g_{a,v}$
α_q E_{Vm} Γ_*^{25} E_{Kc} E_{Ko} E_{Rd} R_{het}^0 Q_{10} f_{Ci} ω a_s $g_{a,v}$
V_m^{25} r_{JmVm} K_C^{25} K_O^{25} f_{Rd} $f_{R,leaf}$ κ w_{pwp} c_w a_v ε_s

785 Figure 4:


prior
posterior Gaussian
binning
FIFE site
ε_s
mean: 0.29
sd: 0.07
sk: -0.57
ku: 0.32
Loobos site
E_{Vm}
mean: 1.31
sd: 0.26
sk: -1.10
ku: 1.01
Loobos site
$g_{a,v}$
mean: 0.69
sd: 0.28
sk: -0.39
ku: 0.71
Loobos site
w_{pwp}
mean: 0.77
sd: 0.12
sk: -0.55
ku: 0.54
probability density
probability density
normalized parameter values
normalized parameter values

786 Figure 5:


300
200
100
0
1
0
-1
cumulative NEE [mol m-2]
prior
posterior Gaussian
posterior
observed
95 % confidence range
95 % confidence range
95 % confidence range
skewness x
kurtosis +
0
100
200
300
400
500
600
700
days since 1 Jan. 1997
1997
1998

Appendix: The Biosphere Energy-Transfer Hydrology model (BETHY)

Overview

We use a process-based model of the coupled photosynthesis and energy balance system, the Biosphere Energy-Transfer Hydrology (BETHY) scheme, to simulate the exchange of CO_2, water and energy between the plant canopy and the atmosphere. BETHY computes absorption of photosynthetically active radiation (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 CO_2. 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 CO_2 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, leaf area index and soil moisture are treated as external forcing data (elements of $\mathbf{s}$ in Equ. 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 CO_2 uptake, A, is calculated as (cf. Farquhar 1980):

819 $$A = \min[J_C, J_E] \quad \text{(A1a)}$$

820

821 with

822

823 $$J_C = V_m \frac{C_i - \Gamma_*}{C_i + K_C(1 + O_x / K_O)} \quad \text{(A1b)}$$

824

825 $$J_E = \frac{\underline{\alpha_q} I J_m}{\sqrt{J_m^2 + \underline{\alpha_q}^2}} \frac{C_i - \Gamma_*}{4(C_i - 2\Gamma_*)} \quad \text{(A1c)}$$

826

827 A is gross photosynthesis, or gross primary productivity (GPP), I is absorbed PAR, C_i
828 the leaf-internal CO_2, and O_x the oxygen concentration (=0.21 mol(O_2)/mol(air)). α_q is
829 the quantum efficiency of photon capture (mol(e^-)/mol(photons)) and V_m the
830 maximum carboxylation rate (in mol(CO_2)$m^{-2}s^{-1}$), expressed as

831

832 $$V_m = \underline{V_m^{25}} \exp\left\{\frac{\underline{E_{Vm}}(T_k - 298\,\mathrm{K})}{RT_k 298\,\mathrm{K}}\right\} \exp\left\{k_{12}\frac{\Lambda_u + \Lambda_l}{2}\right\} \quad \text{(A2)}$$

833

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

837

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

839

840 $$J_m = \underline{r_{JmVm}} * \underline{V_m^{25}} * \frac{T_c}{25°\mathrm{C}} \quad \text{(A3)}$$

841

842 This rate, at standard temperature, is assumed proportional to V_m with an additional
843 proportionality constant (e.g. Wullschläger et al. 1993). T_c is the canopy temperature
844 in °C, T_k the canopy temperature in Kelvin, and R the universal gas constant (8.314
845 $JK^{-1}mol^{-1}$). The CO_2 compensation point without dark respiration, Γ_*, follows from:

846

847 $$\Gamma_* = \underline{\Gamma_*^{25}} * \frac{T_c}{25°\mathrm{C}} \quad \text{(A4)}$$

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:

$$K_c = \underline{K_C^{25}} \exp\left\{ \frac{\underline{E_{Kc}}(T_k - 298\,\mathrm{K})}{RT_k 298\,\mathrm{K}} \right\} \quad \text{(A5)}$$

and

$$K_o = \underline{K_O^{25}} \exp\left\{ \frac{\underline{E_{Ko}}(T_k - 298\,\mathrm{K})}{RT_k 298\,\mathrm{K}} \right\} \quad \text{(A6)}$$

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

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

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

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

V_m is computed from Equ. A2, α_i is the C4 quantum efficiency (in mol(CO2)/ mol(photons)), and k is the C4 CO_2 specificity (in mol(CO2)m^{-2}s^{-1}), with

$$k = \underline{k^{25}} \exp\left\{ \frac{\underline{E_k}(T_k - 298\,\mathrm{K})}{RT_k 298\,\mathrm{K}} \right\} \quad \text{(A8)}$$

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.

878 *Carbon balance*

879

880 To compute net CO_2 uptake by the leaves, leaf or "dark" respiration, R_d, is subtracted
881 from A to yield net leaf CO_2 uptake, A_n, with

882

883 $$A_n = A - R_d \tag{A9}$$

884

885 and

886 $$R_d = \underline{f_{Rd}} \underline{V_m^{25}} \exp\left\{ \frac{\underline{E_{Rd}}(T_k - 298\,\mathrm{K})}{RT_k 298\,\mathrm{K}} \right\} \tag{A10}$$

887

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

889

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

893 $$R_{aut} = \frac{R_{d,c}}{\underline{f_{R,leaf}}} \tag{A11}$$

894

895 $f_{R,leaf}$ stands for the fraction of total plant respiration contributed by the leaves. This
896 formulation differs from the form chosen in the original description of BETHY, which
897 contains an additional term for "growth respiration" assumed proportional to net
898 primary productivity. Such an implicit formulation yields a sum of two terms, one of
899 which is proportional to R_{aut} of Equ. A10, the other to GPP. The above formulation
900 was chosen for simplicity in order to avoid unnecessary co-dependence of parameters.

901

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

905

906 $$R_{het} = \underline{R_{het}^0} \left(\frac{w}{w_m} \right)^{\kappa} \underline{Q_{10}}^{T_a/10°C} \tag{A12}$$

907

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_{tot}, w_m}-w_{pwp}. w_{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

$$F_{CO2} = A - R_{aut} - R_{het} \tag{A13}$$

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

$$C_{i,0} = f_{Ci} C_a \tag{A14}$$

with two values for f_{Ci}, 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):

$$g_{s,0} = \frac{1.6 A_{n,0}}{C_a - C_{i,0}} \frac{RT_k}{p} \tag{A15}$$

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

$$g_s = \frac{g_{s,0}}{1 + b_e D_a} \tag{A16}$$

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

$$S = \underline{c_w} \frac{w}{w_m} \tag{A17}$$

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 (see below). Then, the photosynthesis model is run again, but at a fixed stomatal conductance, g_s, obtained from Equ. 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 leaf area index (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:

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

ε_a and ε_{sfc}=0.97 are sky and surface emissivity, respectively, $T_{K,a}$ air temperature in Kelvin, σ=5.6703×10^{-8} Wm^{-2}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

971 of soil covered by vegetation. For the shortwave part, R_S is incoming solar radiation
972 (Wm^{-2}), f_{PAR} is the fraction of PAR absorbed by the vegetation and computed by the
973 two-flux scheme, a_v the albedo of the vegetation at the limit of high LAI and closed
974 canopy, and a_s the amount of solar radiation absorbed by the soil under the canopy at
975 the same limit. The sky emissivity is computed from
976

977 $$\varepsilon_a = \underline{\varepsilon_s}\left(\frac{e_a}{T_{K,a}}\right)^{1/7}(1+0.22n_c^2) \qquad \text{(A19)}$$

978

979 with the cloud cover fraction n_c. If no separate radiation data for PAR and solar
980 radiation are available, R_S is calculated from PAR according to Weiss and Norman
981 (1985).
982

983 The aerodynamic exchange between the canopy and the free air is described as
984

985 $$G_a = \underline{g_{a,v}} * u \qquad \text{(A20)}$$

986

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

991 $$\underline{g_{a,v}} = \frac{k^2}{\left[\ln\left(\frac{h_{ref}}{r_z h_c} + a_z\right)\right]^2} \qquad \text{(A21)}$$

992

993 h_{ref} is the reference height above canopy (10 m), h_c the canopy height, k=0.41, r_z=0.1
994 and a_z =1.
995
996