

Limiting the parameter space in the Carbon Cycle Data Assimilation System (CCDAS)

S. Kemp¹, M. Scholze², T. Ziehn³ and T. Kaminski⁴

[1]{School of Earth Sciences, University of Bristol, Bristol, UK}

[2]{Department of Physical Geography and Ecosystem Science, Lund University}

[3]{Centre for Australian Weather and Climate Research, CSIRO Marine and Atmospheric Research, Aspendale, VIC 3195, Australia}

[4] {FastOpt GmbH, Hamburg, Germany}

Correspondence to: S. Kemp (sarah.kemp@bristol.ac.uk)

Abstract

Terrestrial ecosystem models are employed to calculate the sources and sinks of carbon dioxide between land and atmosphere. These models may be heavily parameterised. Where reliable estimates are unavailable for a parameter, it remains highly uncertain; uncertainty of parameters can substantially contribute to overall model output uncertainty. This paper builds on the work of the terrestrial Carbon Cycle Data Assimilation System (CCDAS), which, here, assimilates atmospheric CO₂ concentrations to optimise 19 parameters of the underlying terrestrial ecosystem model (Biosphere Energy Transfer and Hydrology scheme, BETHY). Previous experiments have shown that the identified minimum may contain non-physical parameter values. One way to combat this problem is to use constrained optimisation and so avoid the optimiser searching non-physical regions. Another technique is to use penalty terms in the cost function, which are added when the optimisation searches outside of a specified region. The use of parameter transformations is a further method of avoiding this problem, where the optimisation is carried out in a transformed parameter space, thus ensuring that the optimal parameters at the minimum are in the physical domain. We compare these different methods of achieving meaningful parameter values, finding that the parameter transformation method shows consistent results and the other two provide no useful results.

1 **1 Introduction**

2 The response of the global carbon cycle to future changes in climate is highly uncertain. It has
3 been proposed that there is a positive climate-carbon cycle feedback that might significantly
4 accelerate climate change; the study of Friedlingstein et al. (2006) used eleven Earth System
5 models with an interactive carbon cycle and two simulations with each model, to isolate the
6 feedback between climate change and the carbon cycle. All of the models showed that future
7 climate change would reduce the efficiency of the Earth system and in particular the land
8 biosphere to absorb the anthropogenic carbon perturbation, with an additional CO₂ of between 20
9 and 200ppm between the two most extreme models by 2100. Friedlingstein et al. (2006) estimated
10 that this rise in CO₂ would lead to a further warming of 0.1°C to 1.5°C.

11
12 The sources and sinks of carbon dioxide between land and atmosphere can be calculated using
13 terrestrial ecosystem models (TEMs). State of the art TEMs, such as the Biosphere Energy
14 Transfer and Hydrology (BETHY) scheme (Knorr, 2000), encapsulate large numbers of
15 biogeochemical processes and hence involve a large number of parameters. Results from TEMs
16 can diverge markedly, indicating limited understanding and representation of the processes
17 involved. The study of Sitch et al. (2008) used five Dynamic Global Vegetation Models (DGVMs)
18 to model the contemporary terrestrial carbon cycling. They coupled the DGVMs to a fast ‘climate
19 analogue model’ based on the Hadley Centre General Circulation Model, and ran the coupled
20 models to the year 2100 using four Special Report Emissions Scenarios. The most extreme
21 projections differed by up to 494 PgC of cumulative land uptake across the DGVMs over the 21st
22 Century (over 50 years of anthropogenic emissions at current levels; Sitch et al., 2008), although
23 they remained consistent with the contemporary global land carbon budget. Furthermore,
24 Huntingford et al. (2013) explored uncertainties of potential future carbon loss from tropical
25 forests. They found that the DGVM response uncertainty dominated over variation between
26 emission scenarios and climate models.

27
28 There are various sources of uncertainty within the model, for example structural uncertainty,
29 which depends on the formulation of individual processes and their numerical representation.
30 Another source of uncertainty is parametric uncertainty, which results from the uncertainty of the
31 process parameter values used in the models' parameterisation, either due to a lack of knowledge
32 or to upscaling to larger spatial domains. Model parameter values are commonly based on “expert

1 knowledge”. Where little information is known, this can be just an educated guess. If estimates are
2 unavailable for a parameter, it remains highly uncertain. Uncertainty of parameters can
3 substantially contribute to overall model output uncertainty. In this case, parameter estimation to
4 constrain the model against observations can be very useful.

5

6 Many parameter estimation methods, such as gradient-based, Kalman Filter, Monte Carlo
7 inversion, Levenberg-Marquardt and genetic algorithm, use the Bayesian approach [Tarantola,
8 1987, 2005], which combines probability density functions (PDFs) of observational information,
9 prior information and the model dynamics. Four-dimensional variational (4D-Var) schemes use
10 the gradient of the model for the optimisation of parameters; this is usually provided by the
11 adjoint. These approaches are generally computationally efficient but unlike some other
12 variational data assimilation methods, for example the Markov Chain Monte Carlo method, it is
13 possible to identify only a local minimum. Another weakness of 4D-Var schemes is that they
14 concentrate only on the optimal solution without considering uncertainties. However, there are
15 some 4D-Var schemes, such as the one used in the Carbon Cycle Data Assimilation System
16 (CCDAS) (Rayner et al., 2005), which are able to approximate posterior parameter uncertainties
17 using the inverse of the second order derivative of the cost function with respect to the parameters
18 (Hessian) at the global minimum.

19 Generally, Gaussian distributions are assumed for the prior probability distributions of the
20 parameters. This is not always a good assumption as sometimes parameters are restricted to
21 certain values; many are positive, for example and some are restricted between two values, such as
22 a fraction between 0 and 1. Another example, is the terrestrial carbon parameter Q10, which
23 regulates the response of the decomposition rate of organic material to changes in temperature and
24 is known to be greater than 1 (“A rule of thumb widely accepted in the biological research
25 community is that... the Q10 of decomposition is two” Davidson et al., 2006). Where parameters
26 are limited to certain values, optimal solutions can contain non-physical parameter values, as has
27 been seen in Koffi et al. (2012) when using CCDAS (Rayner et al., 2005) without attempting to
28 limit the parameter space. Here, the optimal value of one of the parameters in the photosynthesis
29 scheme was negative, which is unrealistic and would lead to a reversed photosynthesis. Kaminski
30 et al. (2012) used, in addition, quadratic and double bounded transformations to achieve a limited
31 parameter space. Further, in Trudinger et al. (2007), an optimisation inter-comparison study of
32 parameter estimation methods in terrestrial biogeochemical models, and in Fox et al. (2009),

1 another inter-comparison project, the parameter space needed to be limited to avoid non-physical
2 values.

3

4 A simple method of avoiding these non-physical values would be to place hard constraints within
5 the search algorithm. Byrd et al. (1995) described a limited memory quasi-Newton algorithm for
6 solving large nonlinear optimisation problems, which can be applied to parameter estimation.

7

8 Alternatively, it is possible to modify the cost function formulation by adding a so-called penalty
9 term associated with some of the parameters. The penalty term is zero when the parameter is
10 within its specified limits and increases as the parameter goes further away from these limits. This
11 has been implemented in a study to estimate the turnover time of terrestrial carbon (Barrett, 2002).
12 A genetic algorithm was used to improve consistency between estimated model parameters and
13 data. All of the parameters were limited between two values and a penalty term was added
14 whenever they violated these constraints.

15

16 A further option to avoid these non-physical values would be to alter the estimation problem by
17 using a parameter transformation (i.e. a nonlinear change of parameters' PDFs) so that the
18 parameter limits can never be reached. Simon and Bertino (2009) performed a twin experiment
19 with a coupled ocean ecosystem model (HYCOM-NORWECOM) with an ensemble Kalman filter
20 (EnKf), with and without parameter transformations to limit parameters to positive values. The
21 study compared EnKF with parameter transformations and the plain EnKf with post-processing of
22 results, where negative values are increased to zero. These two methods led to similar results,
23 however, the parameter transformations had an advantage in efficiency. In this work they use the
24 term "Gaussian anamorphosis", however, we will continue to use the term "parameter
25 transformation".

26

27 Within CCDAS, a parameter transformation from a Gaussian prior parameter distribution to a log-
28 normal prior parameter distribution is already routinely in use for some selected parameters such
29 as the Q10 parameters. Koffi et al. (2012) showed that the choice of prior parameter distribution
30 can have a great effect on the parameter's uncertainty and the resulting flux field. In their

experiments a log-normal PDF on prior parameters reduced the sensitivity of net CO₂ exchange flux (net ecosystem productivity, NEP) to the observational network as well as the transport model. In the study, the differences in NEP between two configurations are quantified by calculating the root mean square difference (rmsd) over all the grid cells and all months in the study period. After applying the log-normal PDF, the rmsd between the observational networks went from 42 gCm²/yr to 16 gCm²/yr.

This paper builds upon the findings of Koffi et al. (2012) and systematically investigates the ability of the above mentioned three different methods to limit the parameter space within CCDAS.

The outline of the paper is as follows:

First, we give an overview of the data assimilation system and the model, going on to describe the parameter limiting methods and the experiments (Sect. 2). Section 3 describes the results and discussion. We finish with Sect. 4, providing some concluding remarks.

2 Methodology

CCDAS employs a terrestrial ecosystem model BETHY (Knorr, 2000) and an atmospheric tracer transport model TM2 (Heimann, 1995), along with prescribed CO₂ fluxes constituting land-use change, sea surface-atmosphere exchange flux and fossil fuel emission (Rayner et al., 2005; Scholze et al., 2007) that are not calculated by the BETHY model. The biosphere model parameters are estimated using the variational approach. The configuration of CCDAS has been comprehensively described by Scholze et al. (2003), Rayner et al. (2005) and Ziehn et al. (2011b). Here, we provide a brief summary and an explanation of the points where we differ.

2.1 Data assimilation system

There are two steps to the data assimilation in CCDAS as can be seen in Fig. 1. The first uses the full version of BETHY to assimilate space-borne remote sensing data of vegetation activity to optimise the model's phenology and hydrology. The second is a simplified form of BETHY and uses the optimised leaf area index (LAI) and soil moisture fields from the full version as input.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

This paper focuses on the soil carbon balance, a simplified part of the second step. This simplification of the model keeps parameters that control net primary productivity (NPP) fixed; previous studies (Rayner et al., 2005; Scholze et al., 2007) have demonstrated that atmospheric CO₂ data constrain these parameters only moderately. The NPP parameters are calculated by an additional forward simulation covering the full 25-year simulation period after the first step. They are then used as input, similar to soil moisture from the first step.

Posterior parameter values are obtained via iterative minimisation of a cost function $J(\mathbf{x})$. The cost function yields the mismatch between the parameter vector \mathbf{x} and their priors \mathbf{x}_θ and modelled concentrations $M(\mathbf{x})$ and observations \mathbf{c} , where each is weighted by the uncertainties \mathbf{C}_{x_θ} and \mathbf{C}_c of the prior and the observations, respectively (Rayner et al., 2005):

$$J(\mathbf{x}) = \frac{1}{2} \left((\mathbf{x} - \mathbf{x}_\theta)^T \mathbf{C}_{x_\theta}^{-1} (\mathbf{x} - \mathbf{x}_\theta) + (M(\mathbf{x}) - \mathbf{c})^T \mathbf{C}_c^{-1} (M(\mathbf{x}) - \mathbf{c}) \right) \quad (1)$$

The formulation of the cost function uses a Bayesian approach (Tarantola, 1987, 2005) and reflects an assumption of Gaussian probability distributions on the observed concentrations and the prior information on the parameters (explained further in Ziehn et al. (2012)). Minimisation of the cost function uses the gradient of J with respect to the parameters \mathbf{x} at each iteration. Transformation of Algorithms in Fortran (TAF) (Giering and Kaminski, 1998; Kaminski et al., 2003) is used to generate derivative code from the model's source code.

At the minimum in the cost function, the Hessian approximates the inverse covariance of the parameter uncertainties (Tarantola, 1987) and can therefore be used to estimate the posterior uncertainties in the process parameters. Calculation of the Hessian is done by using TAF once more to differentiate the gradient vector in forward mode with respect to the process parameters. Although there is a significant reduction of the cost function within a few tens of iterations, for the Hessian assumptions to hold, many more iterations are required to achieve the near zero gradient of a cost function minimum.

1 When using the gradient-based approach, it is possible that only a local minimum is identified.
2 Therefore, an ensemble of optimisations is performed, with each optimisation starting in slightly
3 varied points in parameter space. In this way, if they all converge to the same minimum, we have
4 confidence that we have found a minimum that is more likely to be a global minimum within the
5 physical parameter space.

6

7 Using the atmospheric tracer transport model TM2, calculated fluxes from BETHY are mapped
8 onto atmospheric concentrations for comparison with measurements of observations of CO₂
9 obtained from the GLOBALVIEW database (GLOBALVIEW-CO₂, 2008). As in previous studies
10 (Rayner et al., 2005), we are using global monthly mean atmospheric CO₂ concentration data from
11 41 sites but here, we use data from over 25 years (1979-2003).

12

13 As the interest of this study is the natural CO₂ exchange flux between land-atmosphere, the
14 remaining fluxes contributing to the atmospheric CO₂ content are added separately. We use the
15 estimates of Houghton (2008) for the land-use flux, without seasonality or interannual variability,
16 following the procedure of Rayner et al. (2005). The flux pattern and magnitude of ocean CO₂
17 exchange is taken from Takahashi et al. (1999) and estimations of inter annual variability from Le
18 Quéré et al. (2007). Background fluxes for fossil fuel emissions, based on the flux magnitudes
19 from Boden et al. (2009), are described by the method of Scholze et al. (2007).

20

21 **2.2 Terrestrial biosphere model and parameters**

22 BETHY, a process-based model of the terrestrial biosphere (Knorr, 2000), simulates carbon
23 uptake and soil respiration within a full energy and water balance and phenology scheme. The grid
24 resolution of BETHY in this study is 2° × 2° with the global vegetation mapped onto 13 Plant
25 Functional Types (PFT) based on Wilson and Henderson-Sellers (1985). Each grid cell can
26 contain up to three PFTs. The amount of present PFTs within a grid cell is specified by their
27 fractional coverage.

28

29 In BETHY, NEP is defined as

$$\text{NEP} = \text{NPP} - R_s = \text{NPP} - (R_{s,s} + R_{s,f}) \quad (2)$$

Where $R_{s,s}$ and $R_{s,f}$ are respiration fluxes from the slowly and rapidly decomposing soil carbon pools. Input to the fast pool is parameterised by the annual course of LAI for deciduous PFTs and the constant fraction of the leaf carbon pool for evergreen PFTs. Soil respiration is simulated to be soil moisture and temperature dependent assuming the following functional dependencies:

$$R_{s,f} = (1 - f_s)k_f C_f \quad (3)$$

$$R_{s,s} = k_s C_s \quad (4)$$

where C_f and C_s represent sizes of the fast and slow carbon pool, respectively, and f_s the fraction of decomposition from the fast pool to the long-lived soil carbon pool. The rate constants are

$$k_f = \omega^\kappa Q_{10,f}^{T_a/10} / \tau_f \quad (5)$$

$$k_s = \omega^\kappa Q_{10,s}^{T_a/10} / \tau_s \quad (6)$$

where ω is the dimensionless plant available soil moisture, i.e. divided by the field capacity of the soil in the respective grid cell (a value between 0 and 1), T_a air temperature, κ a soil moisture dependence parameter, $Q_{10,f}$ and $Q_{10,s}$, temperature dependence parameters for the fast and slow pool, τ_f and τ_s the pool turnover times at 25 °C.

A parameter can either be global or differentiated by certain criteria (in this study, PFT). In this simplified version with NPP kept fixed, there are 6 controlling parameters; five are global and one, the β parameter, is PFT dependent. There is an additional parameter, the offset, representing the carbon dioxide concentration at the beginning of the optimisation, giving 19 process parameters, as can be seen in Table 3. Also shown in Table 3 are the prior uncertainties. As little information is known about some of the parameters, we have chosen to start with larger, realistic uncertainties. The five global parameters are $Q_{10,f}$ and $Q_{10,s}$, the temperature dependence parameters for the fast and slow pool, τ_f the fast pool turnover time at 25°C, f_s the fraction of decomposition from the fast pool to the slow decomposing soil carbon pool and κ the soil moisture dependence parameter. The PFT dependent parameter β , described in Eq. (7) and in more detail in Ziehn et al. (2012), is the carbon balance parameter and determines whether a PFT is a long-term source ($\beta > 1$) or sink ($0 < \beta < 1$):

$$\overline{\text{NEP}} = \overline{\text{NPP}}(1 - \beta) \quad (7)$$

note that the vertical lines above denote the temporal average value over the full 25 year simulation period at each subgrid cell. The β parameter is strictly positive and, whilst it has no physical upper bound, it shouldn't be unrealistically large; a value of 10, for example, would indicate that locations covered by this PFT have a net flux, NEP, 9 times that of NPP as described in Ziehn et al. (2011a). Therefore, an upper bound of 2 is a reasonable selection and is the value we have chosen when bounding this parameter.

We distinguish between the physical model parameters p_i and the parameters as seen by the optimisation routine, the control variables x_i . Control variables have variance 1, in this sense, all the parameters are on the same dimensionless scale and so a change of 1 in that scale to the value of each parameter contributes equally to the value of the cost function. Furthermore, the control variables have PDFs assuming a Gaussian distribution, as mentioned above. To obtain the control variables and to achieve the unit uncertainty, physical parameter values are divided by their prior standard deviation.

$$x_i = \frac{p_i}{\sigma_{p_{0i}}} \quad (8)$$

2.3 Limiting the parameter space

It is not always the case that the physical parameters are distributed in a Gaussian way. For example, this gives positive probability of negative values and as mentioned, some model parameters are only physically meaningful with strictly positive values. Three methods of avoiding these non-physical parameter values are examined in this paper. Two of the methods incorporate the bounding directly into the optimisation. The first, constrained optimisation, seeks a solution within the physically meaningful parameter space. The second adds a penalty term to the cost function when the optimiser begins to search the non-physical domain. This encourages it to stay within the physically meaningful parameter space. The final method investigated in this paper, parameter transformations, performs the optimisation in a transformed parameter space, which ensures that, when back-transformed, the minimum is always in the physically meaningful parameter space. In addition to testing these three methods, we go on to investigate the effect of

1 different parameter transformations on the inferred target quantities and their posterior
2 uncertainties. One particular transformation, the log transformation has already been used in
3 CCDAS and found to have a large impact on the optimised parameter values and also the resulting
4 flux fields as explained above (Koffi et al., 2012). In addition to this log-normal transformation we
5 propose two other transformations: quadratic and double bounded log. The quadratic and log
6 transformations are used to provide a lower bound on a parameter and the double bounded log can
7 be used to provide an upper and lower bound on a parameter. We will examine the effect of using
8 these different parameter transformations on parameter values and their uncertainties.

9

10 The essential difference between the three approaches is the form of the prior (and thus posterior)
11 PDFs in (physical) parameter space. Both the constrained and the penalty function approaches
12 produce a prior of Gaussian shape inside the bounds/the non-penalised region. Outside, the
13 constrained approach produces a zero probability while the penalty approach produces a non-zero
14 probability consisting of a gradual reduction of the Gaussian probability with increasing distance
15 from the bounds. The parameter transformation approach produces a zero probability outside the
16 bounds and a non- zero but non-Gaussian probability within the bounds.

17

18

19 **2.3.1 Constrained optimiser**

20 When using the constrained optimisation, the optimiser can only choose from amongst a restricted
21 well-defined set. Minimisation of the cost function is done via a gradient based algorithm updating
22 an approximation of the Hessian through the L-BFGS-B method [Byrd et al., 1995; Zhu et al.,
23 1997], which limits the control parameter space to the restricted set. This is a variant of the
24 Davidon-Fletcher-Powell (DFP) formula (Fletcher and Powell, 1963; Press et al., 1996).

25

26 **2.3.2 Penalty term in the cost function**

27 For the penalty term optimisation, we use BFGS but add a penalty term to the cost function when
28 the optimiser begins to search a non-physical region in the form of

$$J(\mathbf{x}) = \frac{1}{2} \left((\mathbf{x} - \mathbf{x}_0)^T \mathbf{C}_{\mathbf{x}_0}^{-1} (\mathbf{x} - \mathbf{x}_0) + (\mathbf{M}(\mathbf{x}) - \mathbf{c})^T \mathbf{C}_c^{-1} (\mathbf{M}(\mathbf{x}) - \mathbf{c}) \right) + \sum_{r=1}^R P_r(D_r g_r \delta_r \mu_r) ,$$

$$r=1, \dots, R=19 \text{ (i.e. the number of parameters)} \quad (9)$$

where D_r is a penalty factor that scales the penalty function, g_r is the threshold function, δ_r invokes the penalty when the threshold is violated and μ_r determines the sensitivity of the penalty function to threshold violation (with even, integer values).

$$P_r = D_r \delta_r g_r^{\mu_r} \quad (10)$$

$$g_r(\alpha_r) = (\alpha_r^* - \alpha_r) \quad (11)$$

Where α_r is the current value of the r^{th} parameter and α_r^* is the threshold value, the value beyond which the threshold is violated and the penalty imposed.

10

$$\delta_r = \begin{cases} 1, & \text{if threshold violated} \\ 0, & \text{if threshold not violated} \end{cases} \quad (12)$$

12

2.3.3 Parameter transformations

Depending on the transformation used on the parameter, different equations are used to convert them from the model parameters p_i into the control variables x_i . The equations give control variables with a variance of 1. Where no transformation is used (i.e. the prior is assumed to have a Gaussian distribution), the parameters are just normalised using Eq. (8) as mentioned above. p_{0i} is the prior value of the i^{th} model parameter and $\sigma_{p_{0i}}$ is its prior uncertainty. As further options we have a double-bounded log, a lower-bounded log, and a quadratic transformation.

Where a parameter has a lower and upper bound, a and b , the parameter transformation from optimisation space to physical space is given by an equation of the form:

$$p(x) = \frac{(b-a)}{(1+e^{-x})} + a \quad (13)$$

For the log transformation with only a lower bound of a , this simplifies to an equation of the form:

$$p(x) = e^x + a \quad (14)$$

The quadratic transformation with lower bound a is computed by a function like this:

$$p(x) = x^2 + a \quad (15)$$

Minimisation of the cost function is achieved via a gradient-based algorithm updating an approximation of the Hessian through the Broyden-Fletcher-Goldfarb-Shannon formula (Fletcher and Powell, 1963; Press et al., 1996), a Quasi-Newton method.

2.4 Experiments

We performed a total of five experiments investigating the impact of the three parameter space restriction methods on the results of the optimisation. Table 1 provides an overview of the experiments and how they differ.

Previous experiments with this reduced version of CCDAS using no parameter transformations indicated three β parameters (8, 13, 18) that were either negative or extremely high, so in this paper, unless otherwise stated, they have all been limited between 0 and 2 using a double bounded log transformation. The rest of the β parameters have been left untransformed (i.e., assumed Gaussian), as they didn't require any bounding since their posterior values already lied between 0 and 2.

To explore the effect that parameter transformations have in the model, the Q_{10f} parameter's treatment was varied between Gaussian, Log and Quadratic, whilst keeping all but the three β parameters' (8, 13 and 18) treatments Gaussian. (Experiments PTG, PTL and PTQ.)

For the default penalty term optimisation, we only added a penalty term when the β parameter for crops (parameter 18) became negative. In Eq. (10), we chose $D_{18}=10^4$ as the penalty factor, since this is on the same order of magnitude as the cost function minimum from previous experiments, and $\mu_{18}=4$ for the sensitivity value, as it has to be positive but we do not use 2 since we use the second derivative to calculate posterior uncertainties. The other two β parameters (8 and 13) were still transformed using the double bounded log transformation. (Experiment PEN.)

1 In the default constrained optimisation, the three β parameters (8, 13 and 18) were restricted
2 between 0 and 2 by the hard limits imposed by the constrained optimiser. (Experiment CONS.)

3

4 For each of the experiments above, four extra optimisations (building together an ensemble of five
5 optimisations) were performed with the default prior parameter values randomly perturbed by up
6 to 10%. This ensures that if most of the optimisations converge to the same minimum we have
7 found a robust solution.

8

9

10 **3 Results and discussion**

11 We present the results of the different experiments, with a focus on the parameter transformations,
12 as these are the experiments that successfully located a minimum within the physical parameter
13 space. The other two methods were not successful and so are of limited use. We commence with
14 the constrained optimisation (CONS), then briefly discuss the penalty term experiment (PEN) and
15 finish with the results from the parameter transformations (PTG, PTL, PTQ). An overview of the
16 optimisation results is presented in Table 2.

17

18 CONS: Here, for the default prior parameter values, the optimisation did not converge and reached
19 the preset maximum number of iterations (5000). We did not continue this optimisation as the
20 number of iterations was already about 10 times more than the average number of iterations for
21 the parameter transformation experiments. At this point there had been a significant reduction in
22 the cost function by around a factor of 550 but the bounded parameters (8,13,18) were exactly at
23 their bounds of 0 or 2, which at least in the case of 0 for a beta parameter does not make sense.
24 Furthermore, there was not a near zero gradient. The other four ensemble members terminated
25 after fewer iterations (10-382), without finding a minimum because of internal numerical
26 problems within the optimiser. There was some reduction in the cost function by between a factor
27 of around 20 and 400 but there was still a very large gradient of at least 4000 and all of these four
28 ensembles had negative values for the soil moisture dependence parameter, κ .

29 Since this method of limiting the parameter space was unsuccessful for this problem, uncertainties
30 have not been calculated.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

PEN: All of the five optimisations converged and found a minimum in a mathematical sense (zero or at least close to zero gradient). However, they did not achieve the parameter bounding, as the limited β parameter (parameter 18) was slightly negative (-0.024), contributing a penalty term of 0.8 to the cost function. As the penalty was non-zero, the experiment was not successful in our aim of limiting the parameter. The optimisation is able to offset this small negative penalty contribution by achieving a smaller input to the cost function from the data and the parameters. We performed further experiments adding a penalty to parameter 8 as well, with no successful bounding of these parameters. We also increased the penalty term by a factor of 100 but still the β parameter was slightly negative. Again, uncertainties have not been calculated due to the unsuccessful optimisation of the parameters to physically meaningful values.

PTG, PTL, PTQ: Using the parameter transformations we were able to successfully limit the parameter space. In this case, the transformation of $Q_{10,f}$ did not seem to have an effect on the final value of the cost function. Of the 15 (3 x 5) optimisations, 12 converged to the same minimum in the cost function of $J = 9667$ (reduced from an initial value of 5294051 when using the prior parameter values) and took between 174 and 876 iterations. The other three (one from each of the Gaussian, Log and Quadratic) converged to a different value of 9515, but were outside of the physical parameter space since another of the β parameters (parameter 9) was negative (-0.057), and are therefore not relevant. Having been reduced from over 10^7 to 10^{-3} , the gradient of the minimum in the cost function can be considered to be sufficiently small enough to indicate that a minimum has been located for all three parameter transformation experiments. We calculate posterior parameter uncertainties and also propagate these uncertainties onto the net carbon flux using a linearisation of the model (Kaminski et al., 2003).

Prior and optimised parameter values for all the parameter transformation experiments are shown in Table 3. Also shown are prior and posterior uncertainties and percentage reduction in uncertainty. The 3 β parameters that were double bounded show their upper and lower percentiles, equivalent to one standard deviation. The global parameters behave in a consistent way to previous studies (Ziehn et al., 2011a). The temperature dependence parameter of the fast carbon pool, $Q_{10,f}$, is somewhat reduced to 1.07 compared to its initial value of 1.5, although this change is within the range of the prior parameter uncertainty. The temperature dependence parameter of

1 the slow pool, $Q_{10,s}$ is increased from its initial value of 1.5 to 1.82, which again lies within the
2 one sigma range of the parameter's prior uncertainty. The two Q_{10} parameters posterior
3 uncertainties are lowered by more than one order of magnitude, which confirms the result of
4 Scholze et al. (2007) that atmospheric CO₂ data constrain the parameters of soil respiration
5 relatively well.

6 The fast pool turnover time τ_f is also within the prior uncertainty range of one standard deviation
7 increasing from 1.5 to 3.46. As is the soil moisture dependence parameter, κ , which is reduced
8 from 1 to 0.57. The small posterior uncertainty of this parameter indicates that it is also well
9 constrained by the data. The optimised parameter value of the fraction, f_s , however is outside of
10 the prior uncertainty range, increasing from 0.2 to 0.74. It behaves similarly to previous studies
11 (Ziehn et al., 2011a). Again, the posterior uncertainty is very small. Lastly, the offset parameter
12 also behaves in a consistent way to Ziehn et al. (2011a). The posterior uncertainties for all the
13 global parameters are reduced by over 95% compared to their prior uncertainty. This is due in part
14 to the fact that the global atmospheric CO₂ network strongly observes those parameters that act
15 globally at all subgrid cells and is further explained by the fact that moderately large prior
16 parameter uncertainty values are used.

17

18 For the PFT-dependent β parameters, the uncertainty reduction varies between 5-90% and so is
19 clearly less than for the global parameters. This is partly due to the β parameter being
20 differentiated by PFT, which means each PFT is less well observed by the atmospheric network.

21

22 The cost function reduction of all of the five experiments is shown in Fig. 2 up to the first 400
23 iterations on a log scale. By 400 iterations, all of the parameter transformations (PTG, PTL, PTQ)
24 and the penalty term optimisation (PEN) had converged, the constrained optimisation (CONS) had
25 not. The rest of the constrained optimisation's performance is shown inlaid in Fig. 2 on a linear
26 scale. The majority of the cost function reduction (around 2 orders of magnitude) is within the first
27 30 iterations. After this, convergence is slower but the optimisation continues until a near zero
28 gradient is achieved, which indicates that we have found a minimum. Only at the minimum can
29 the inverse of the hessian be used to estimate the posterior parameter uncertainty.

30

The gradient value for all five optimisations is shown in Fig. 3, again up to 400 iterations, where we can see that the parameter transformations and the penalty term experiments have converged. The constrained optimisation has not been included up to its full 5000 iterations; it continues in much the same way after this and does not achieve a near-zero gradient. The different methods and the minimisations from different starting points converge differently as they are solving different problems. Each change in the formulation of the cost function results in a different optimisation problem. When an optimisation starts at a different point in the parameter space, it follows a different trajectory to find a minimum.

Figure 4 shows a time-series of our target output quantity of global mean NEP, along with uncertainties. We calculated the values for all three of the parameter transformation experiments but as they are all within the same numerical limits only the Gaussian case has been shown. The global mean NEP time-series and their uncertainties resemble that of Ziehn et al. (2011b). This is because we are using exactly the same set-up with identical forcing and assimilation data. We also calculated NEP using the parameter values obtained from the constrained and the penalty experiments. Per year, NEP from these two cases does not differ much from the parameter transformations (between 2-4%), so we have not added this to Fig. 4. The resulting NEP fields look very similar to each other as do the posterior parameter values but since each of the constrained and penalty experiments yielded at least one unphysical value, we shall not consider these any further. The effect of these unphysical parameters does not show up in aggregated quantities such as annual global values or even annual grid cell values since the NEP of a grid cell is also the sum of the NEP of the individual PFTs within that grid cell.

We also show the covariance between the flux uncertainties, which, as in Ziehn et al. (2011b), we express using the uncertainty correlation matrix of diagnostics, \mathbf{R}_d , defined as follows:

$$R_d^{i,j} = \frac{C_d^{i,j}}{\sigma_i \sigma_j} \quad (16)$$

Where C_d^{ij} is element i,j of the error covariance matrix of global net CO₂ exchange flux (NEP) per year and σ_i the posterior uncertainty of parameter i obtained from the diagonal element $C_d^{i,i}$ of the matrix \mathbf{C}_d . For mean global NEP for the Gaussian case this uncertainty correlation matrix is shown in Fig. 5. There are a large number of negative correlations, which is the reason for a relatively small overall uncertainty for the global mean NEP over the whole period 1979 to 2003. However, the uncertainty for the global mean NEP for a single year (as shown in Fig 4) is substantially larger. It is worthwhile to note that, between the different parameter transformations, there is no difference in the uncertainty correlation matrix and thus also in the posterior parameter covariances.

It seems in general, that a lower cost function value can be achieved when the optimiser is allowed to search the whole space. For example, in one of each of the parameter transformation experiments the cost function at the minimum was 9515 but one of the β parameters (parameter 9) was negative (-0.057). It is possible that the global minimum is within the non-physical space because the model is highly non-linear with a complex, 19-dimensional parameter space and from a purely mathematical point of view, a smaller minimum can be found outside the physically meaningful parameter space. However, this does not constitute a solution for our optimisation problem. Another possible reason for finding a minimum outside the physically valid parameter space is that the model, as it stands, is missing or does not fully describe a relevant process and therefore the optimisation has to compensate for this missing process by choosing non-physical parameter values. The analysis of such non-physical parameter values can provide useful information for further model development. However, this is not always feasible and therefore limiting the parameter space with parameter transformations seems to be the most effective way to ensure physically meaningful parameter values.

4 Conclusions

We systematically investigated the effects of different methods of limiting the parameter space, which is an emerging issue in parameter optimisation studies. In our simplified set-up of CCDAS, we saw that two of the methods were not successful; both the constrained and the penalty term optimisation had values outside of the physically meaningful parameter space and in fact, the former did not converge to a minimum at all. Parameter transformations however, were successful in locating an optimal solution within the limits. All of the physically meaningful ensembles

converged to the same minimum, so we can be confident that this is the global minimum. We tested two parameter transformations against standard scaling and found that these three experiments all reached the same minimum, indicating that the transformation does not alter the optimisation problem. This is in contrast to the study of Koffi et al. (2012) however, we note that in this study a more complex system was used that involved 57 parameters compared to our 19. Furthermore, in Koffi et al. (2012), the optimisations do not converge to a minimum and have been stopped after a certain reduction in the cost function value, without obtaining a near zero gradient. In our experiments, for this manuscript, all the optimisations with parameter transformations have converged to a minimum with a final gradient approaching zero. A future experiment of interest may involve systematically investigating parameter transformations within the fully complex model.

In our experience, we would therefore recommend the parameter transformations as the most suitable solution to the problem of limiting parameter spaces. As the parameter transformations are applied outside of the optimisation routine this would be a good general method for any problem involving restricted parameter sets. As for CCDAS, the quadratic transformation is slightly preferred to the log as it may have a lower range for the number of iterations required to achieve convergence.

5 Code availability

For obtaining the code, please contact M. Scholze (marko.scholze@nateko.lu.se).

1 **References**

- 2 Barrett, D. J.: Steady state turnover time of carbon in the Australian terrestrial biosphere, *Global*
3 *Biogeochem. Cycles*, 16(4), 1108, doi:10.1029/2002GB001860, 2002.
- 4 Boden, T. A., G. Marland, and R. J. Andres: Global, regional, and national fossil-fuel CO₂
5 emissions, *Carbon Dioxide Inf. Anal. Cent.*, Oak Ridge Natl. Lab., Oak Ridge, Tenn.,
6 doi:10.3334/CDIAC/00001, (2009).
- 7 Byrd, R. H., LU, P., Nocedal, J., and Zhu, C.: A limited memory algorithm for bound constrained
8 optimization. *SIAM J. Sci. Comput.* 16, 5 (Sept.), 1190–1208, 1995.
- 9 Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and
10 feedbacks to climate change. *Nature* 440:165-173, 2006.
- 11 Fletcher, R., and Powell, M. J. D.: A rapidly convergent descent method for minimization,
12 *Comput. J.*, 6, 163–168, 1963.
- 13 Fox, A., Williams, M., Richardson, A. D., Cameron, D., Gove, J. H., Quaife, T., Ricciuto, D.,
14 Reichstein, M., Tomelleri, E., Trudinger, C. M., and Van Wijk, M. T.: The REFLEX project:
15 Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem
16 model against eddy covariance data, *Agricultural and Forest Meteorology*, Volume 149, Issue 10,
17 1 October 2009, Pages 1597-1615, ISSN 0168-1923,
18 <http://dx.doi.org/10.1016/j.agrformet.2009.05.002>, 2009.
- 19 Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S.,
20 Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W.,
21 Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G.,
22 Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zengq, N.: Climate-carbon cycle
23 feedback analysis: results from the C4MIP model intercomparison, *J. Climate*, 19, 3337–3353,
24 2006.
- 25 Giering, R., and Kaminski, T.: Recipes for adjoint code construction, *ACM Trans. Math.*
26 *Software*, 24, 437–474, 1998.
- 27 Heimann, M.: The Global Atmospheric Tracer Model TM2, Technical Report, 10, Deutsches
28 Klimarechenzentrum, Hamburg, Germany, 1995.

1 Houghton, R. A.: Carbon flux to the atmosphere from land-use changes: 1850–2005, in TRENDS:
2 A Compendium of Data on Global Change, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl.
3 Lab., Oak Ridge, Tenn, 2008.

4 Huntingford C., Mercado LM., Harris PP., Zelazowski P., Galbraith D., Malhi Y., Sitch S., Lewis
5 SL., Phillips OL., Lloyd J., Gloor E., Fisher R., Lomas M., Walker AP., Jones CD., Booth BBB.,
6 Hemming D., Kay G., Good P., Betts R., Atkin OK., Meir P., Zaragoza-Castells J., Nobre C.,
7 Marengo J. and Cox PM.: Simulated resilience of tropical rainforests to CO₂ -induced climate
8 change, *Nature Geoscience*, **6**, pp.268-273. doi: 10.1038/ngeo1741, 2013.

9 Kaminski, T., Giering, R., Scholze, M., Rayner, P. and Knorr, W.: An example of an automatic
10 differentiation-based modelling system, in Computational Science, ICCSA 2003, International
11 Conference Montreal, Canada, Lecture Notes Comput. Sci., vol. 2668, edited by Kumar, V.,
12 Gavrilova, L., Tan, C. J. K., and L'Ecuyer, P., pp. 95–104, Springer, New York, 2003.

13 Kaminski, T., Rayner, P. J., Voßbeck, M., Scholze, M., and Koffi, E.: Observing the continental-
14 scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial
15 assimilation system by means of quantitative network design. *Atmos. Chem. Phys.*, **12**(16), 7867-
16 7879, doi:10.5194/acp-12-7867-2012, 2012.

17 Knorr, W.: Annual and interannual CO₂ exchanges of the terrestrial biosphere: Process based
18 simulations and uncertainties, *Global Ecol. Biogeogr.*, **9**, 225–252, 2000.

19 Knorr, W., and Heimann, M.: Uncertainties in global terrestrial biosphere modeling: 1. A
20 comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme, *Global*
21 *Biogeochem. Cycles*, **15**, 207–225, 2001.

22 Koffi, E. N., Rayner, P. J., Scholze, M., & Beer, C.: Atmospheric constraints on gross primary
23 productivity and net ecosystem productivity: Results from a carbon-cycle data assimilation
24 system. *Global Biogeochem. Cycles*, **26**(1), GB1024. AGU. doi:10.1029/2010GB003900, 2012.

25 Le Quéré, C., **Rödenbeck, C.**, Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A.,
26 Labuschagne, C., Ramonet, M., Nakaza, T., Metzl, N., Gillett N., and Heimann, M.: Saturation of
27 the Southern Ocean CO₂ sink due to recent climate change, *Science*, **316**, 1735–1738,
28 doi:10.1126/science.1136188, 2007.

29 Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numerical Recipes in
30 FORTRAN 77: The Art of Scientific Computing, Cambridge Univ. Press, New York, 1996.

1 Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., & Widmann, H.: Two decades of
2 terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Global*
3 *Biogeochem. Cycles*, 19(2), GB2026. AGU. doi:10.1029/2004GB002254, 2005.

4 Scholze, M.: Model studies on the response of the terrestrial carbon cycle on climate change and
5 variability, Ph.D. thesis, Max Planck Inst. für Meteorol., Hamburg, Germany, 2003.

6 Scholze, M., Kaminski, T., Rayner, P., Knorr, W., and Giering, R.: Propagating uncertainty
7 through prognostic carbon cycle data assimilation system simulations, *J. Geophys. Res.*, 112,
8 D17305, doi:10.1029/2007JD008642, 2007.

9 Simon, E. and Bertino, L.: Application of the Gaussian anamorphosis to assimilation in a 3D
10 coupled physical–ecosystem model of the North Atlantic with the EnKF: a twin experiment.
11 *Ocean Science* 5, 495–510, 2009.

12 Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P.,
13 Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C. and Woodward, F. I.: Evaluation of the
14 terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five
15 Dynamic Global Vegetation Models (DGVMs). *Global Change Biology*, 14: 2015–2039.
16 doi: 10.1111/j.1365-2486.2008.01626.x, 2008.

17 Takahashi, T., Wanninkhof, R. H., Feely, R. A., Weiss, R. F., Chipman, D. W., Bates, N.,
18 Olafsson, J., Sabine, C., and Sutherland, S. C.: Net sea-air CO₂ flux over the global oceans: an
19 improved estimate based on the sea-air pCO₂ difference, Paper presented at 2nd International CO₂
20 in the Oceans Symposium, Center for Global Environmental Research, National Institute for
21 Environmental Studies, Tsukuba, Japan, January 18–23, 1999.

22 Tarantola, A.: *Inverse Problem Theory: Methods for Data Fitting and Model Parameter*
23 *Estimation*, Elsevier, New York, 1987.

24 Tarantola, A.: *Inverse Problem Theory and Methods for Model Parameter Estimation*, Soc. for
25 *Ind. and Appl. Math.*, Philadelphia, Pa, 2005.

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

1 Wilson, M. F., and Henderson-Sellers, A.: A global archive of land cover and soil data for use in
2 general circulation climate models, *J. Climatol.*, 5, 119 – 143, 1985.

3 Ziehn, T., Knorr, W., and Scholze, M.: Investigating spatial differentiation of model parameters in
4 a carbon cycle data assimilation system. *Global Biogeochem. Cycles*, 25(2), GB2021. AGU.
5 doi:10.1029/2010GB003886, 2011a.

6 Ziehn, T., Scholze, M., and Knorr, W.: Development of an ensemble-adjoint optimization
7 approach to derive uncertainties in net carbon fluxes, *Geosci. Model Dev.*, 4, 1011-1018,
8 doi:10.5194/gmd-4-1011-2011, 2011b.

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

12 Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines
13 for large scale bound constrained optimization. *ACM Trans. Math. Softw.* 23, 4, 550–560, 1997.

14

1

2

Table 1: Experiments performed to investigate the impact of different methods of limiting the parameter space on the optimisation.

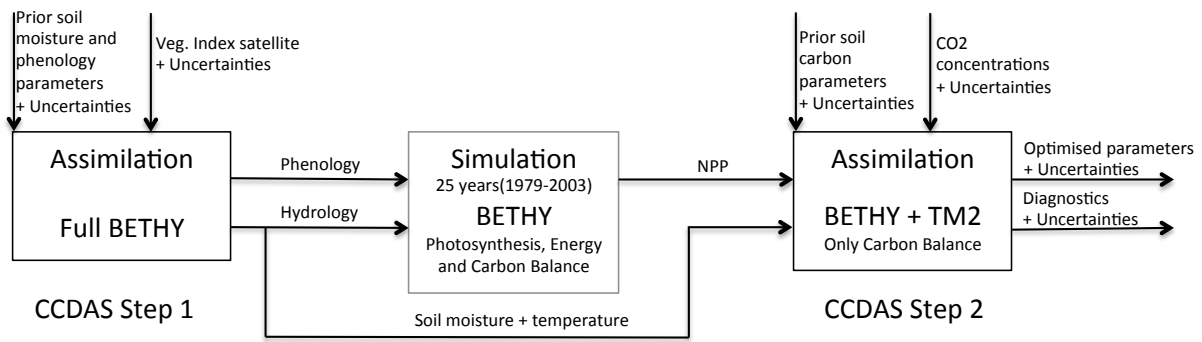
Experiment	Default set-up: No Parameter Transformation of $Q_{10,f}$ (Gaussian)	Parameter Transformation of $Q_{10,f}$ (Log)	Parameter Transformation of $Q_{10,f}$ (Quadratic)	Constrained	Penalty Term
Optimiser used	BFGS	BFGS	BFGS	L-BFGS-B	BFGS
Parameters treated with a double bounded parameter transformation	8,13,18	8,13,18	8,13,18	none	8,13
Treatment of $Q_{10,f}$ in parameter transformation experiments	No transformation (assuming Gaussian PDF)	Logarithmic transformation	Quadratic transformation	none	none
Parameters constrained	none	none	none	8,13,18: constrained between 0 and 2	none
Parameters with a penalty term added	none	none	none	none	18: penalty term added when negative
Abbreviation	PTG	PTL	PTQ	CONS	PEN

Table 2. Values of the cost function, the contributions from data and parameters, the gradient, the number of iterations to achieve convergence and how many optimisations from the ensemble of 5 converged to this value for the different parameter transformations of $Q_{10,f}$ and the constrained and penalty term experiments.

	<i>Parameter Treatment Used for $Q_{10,f}$</i>			<i>Constrained</i>	<i>Penalty term</i>
	<i>Gaussian</i>	<i>Log</i>	<i>Quadratic</i>		
Optimised cost function value	9666.8	9667.0	9666.9	9613.6	9639.0
Data contribution	9584.3	9584.3	9584.3	9552.8	9571.0
Parameter contribution	82.5	82.7	82.6	60.8	67.1
Final gradient value	2.3×10^{-3}	1.6×10^{-4}	1.2×10^{-3}	49.8	5.2×10^2
Number of iterations	224	319	365	5000	154
Range of number of iterations	174-876	319-595	232-478	10-5000	154-208
Number of optimisations that successfully converged	4	4	4	0	0

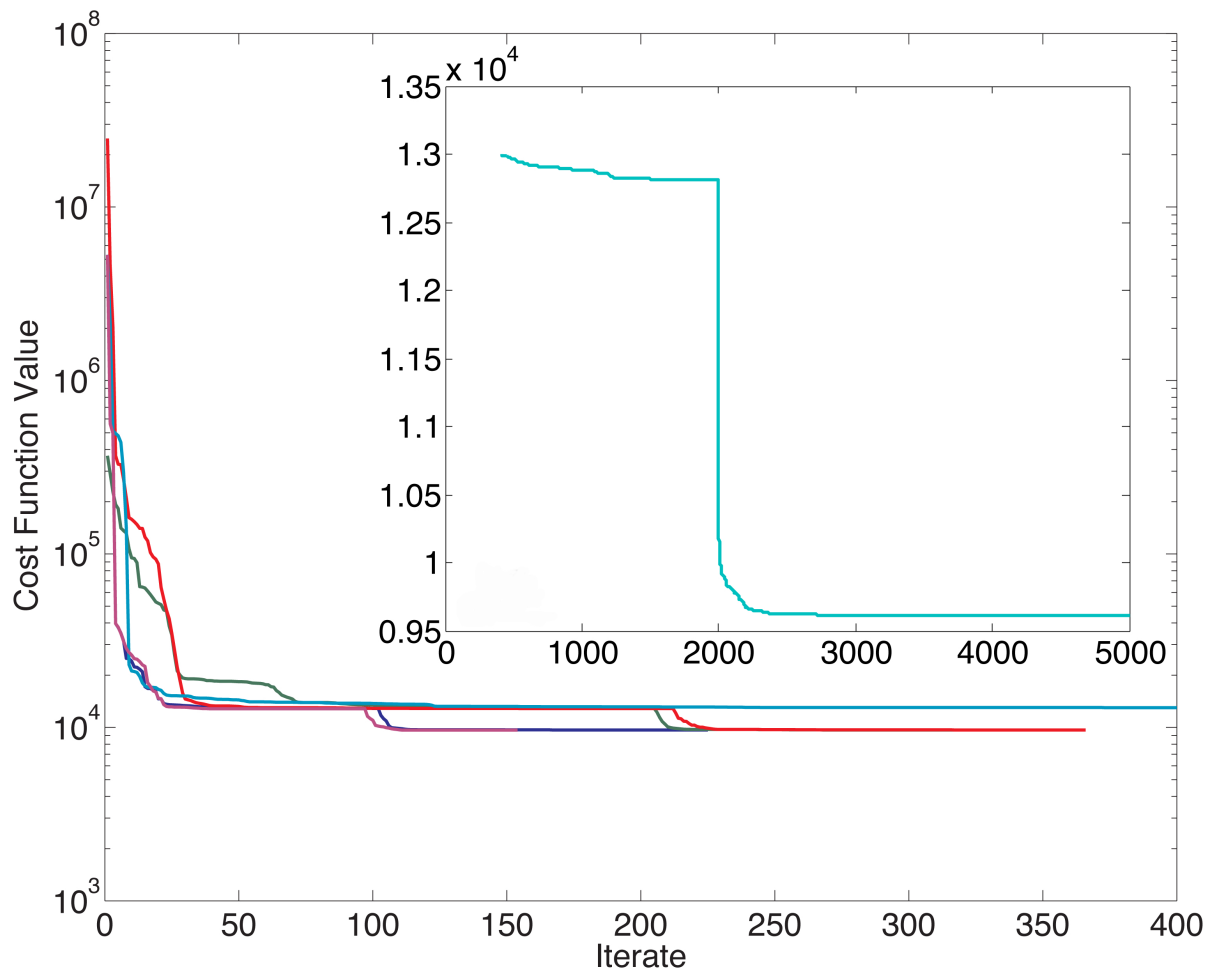
Table 3. Controlling parameters and their initial and optimised values for each experiment and parameters' initial and posterior uncertainty (equivalent to one standard deviation) and percentage reduction in uncertainty (relative to the upper standard deviation) after the optimisation for the parameter transformation experiments. For the three β parameters that were transformed using the double-bounded log transformation, upper and lower percentiles, equivalent to one standard deviation, are shown.

<i>Parameter</i>	<i>Initial Value</i>	<i>Optimised Value</i>					<i>Initial Uncertainty</i>	<i>Optimised Uncertainty</i>	<i>Percentage Reduction in Uncertainty</i>
		<i>CONS</i>	<i>PEN</i>	<i>PTG</i>	<i>PTL</i>	<i>PTQ</i>			
1 $Q_{10,f}$	1.5	1.071	1.068	1.069	1.069	1.069	0.75	0.016	97.899
2 $Q_{10,s}$	1.5	1.81	1.811	1.817	1.816	1.816	0.75	0.019	97.492
3 τ_f	1.5	3.456	3.417	3.435	3.436	3.435	3.0	0.12	95.993
4 κ	1	0.574	0.574	0.571	0.571	0.571	9.0	0.011	99.877
5 f_s	0.2	0.734	0.735	0.735	0.735	0.735	0.2	0.004	98.073
6 $\beta(\text{TrEv})$	1	0.803	0.807	0.796	0.796	0.796	0.25	0.017	93.387
7 $\beta(\text{TrDec})$	1	0.889	0.896	0.926	0.926	0.926	0.25	0.043	82.999
8 $\beta(\text{TmpEv})$	1	0	0.284	0.29	0.29	0.29	0.25	0.071/0.089	37.231
9 $\beta(\text{TmpDec})$	1	0.075	0.101	0.037	0.037	0.037	0.25	0.05	79.889
10 $\beta(\text{EvCn})$	1	1.274	1.278	1.272	1.272	1.272	0.25	0.023	90.774
11 $\beta(\text{DecCn})$	1	0.36	0.398	0.325	0.325	0.325	0.25	0.121	51.785
12 $\beta(\text{EvShr})$	1	0.151	0.171	0.196	0.196	0.196	0.25	0.095	61.899
13 $\beta(\text{DecShr})$	1	2	1.912	1.913	1.913	1.913	0.25	0.024/0.019	48.949
14 $\beta(\text{C3Gr})$	1	1.508	1.509	1.485	1.485	1.485	0.25	0.029	88.468
15 $\beta(\text{C4Gr})$	1	1.138	1.131	1.129	1.129	1.129	0.25	0.024	90.508
16 $\beta(\text{Tund})$	1	0.866	0.866	0.876	0.876	0.876	0.25	0.05	79.816
17 $\beta(\text{Wetl})$	1	2.241	2.165	2.211	2.211	2.211	0.25	0.238	4.9
18 $\beta(\text{Crop})$	1	0	-0.024	0.066	0.066	0.066	0.25	0.013/0.017	54.095
19 offset	338	336.423	336.421	336.421	336.421	336.421	1.0	0.049	95.099



1
2
3
4
5
6

Figure 1. CCDAS structure. Top arrows indicate the parameters to be optimised and the observational data used in the various steps.

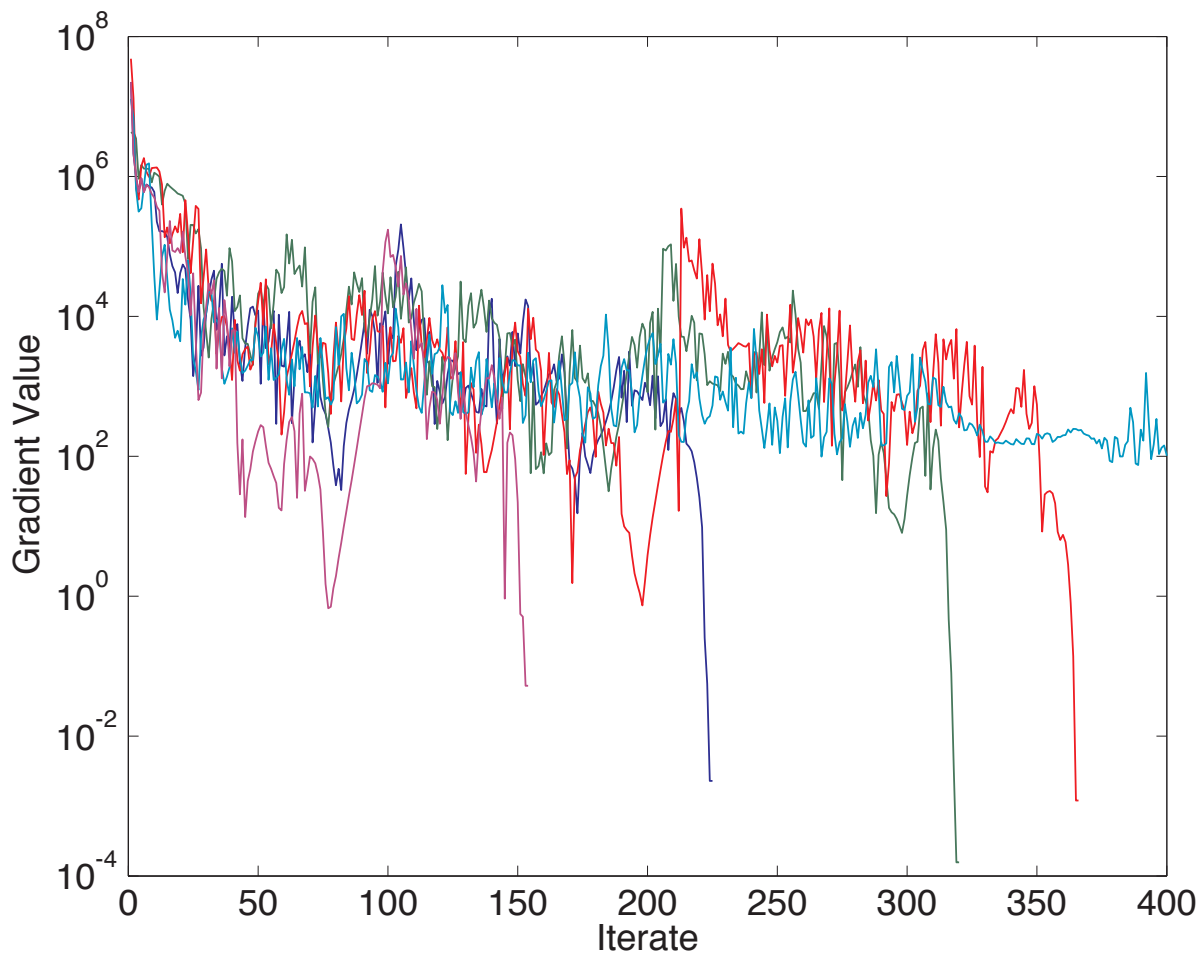


1

2

3 Figure 2. Cost function value for all five experiments in log scale for the first 400 iterations and
 4 cost function value for constrained optimisation from 400 to 5000 in linear scale (dark blue:
 5 Gaussian, green: log, red: quadratic, light blue: constrained, pink: penalty).

6

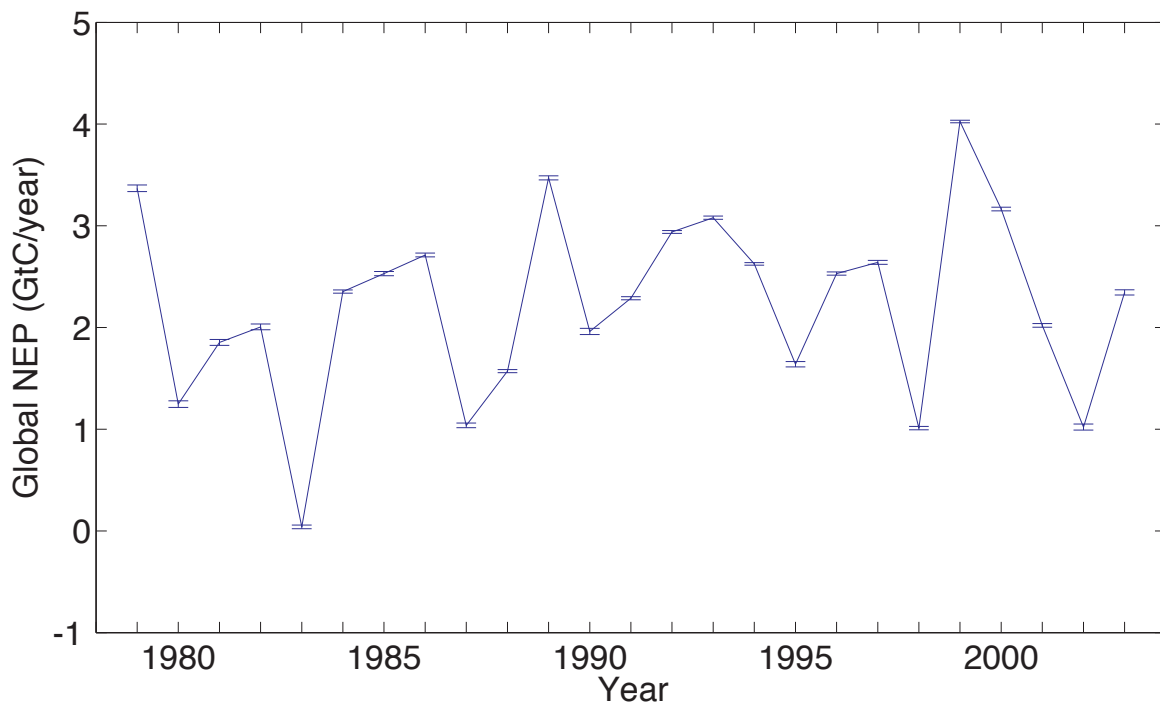


1

2

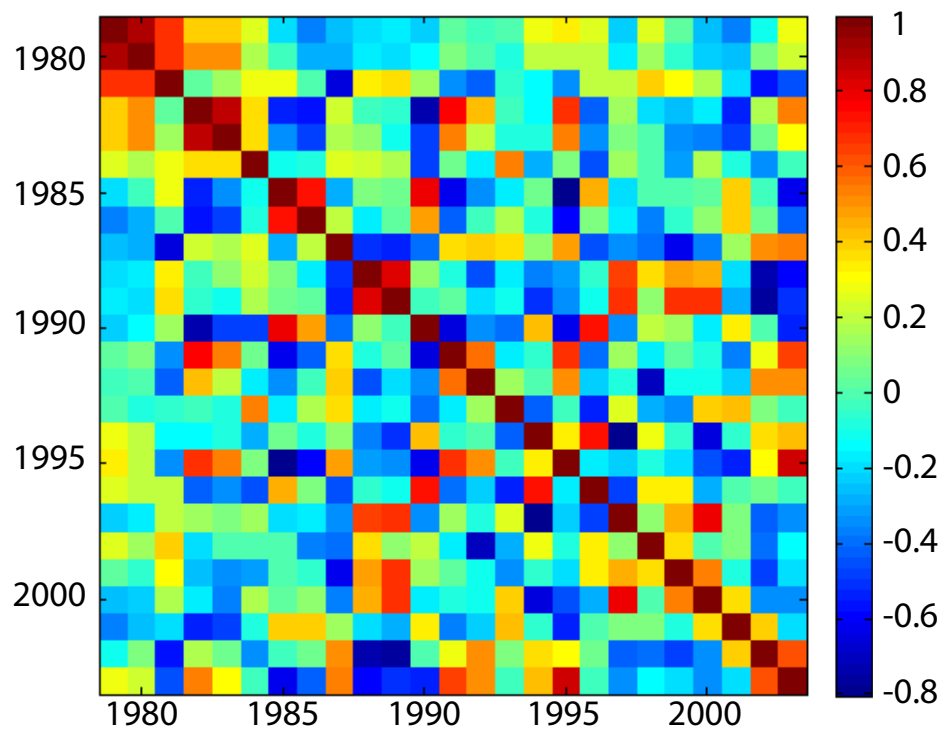
3 Figure 3. Gradient value for all five experiments in log scale up to 400 iterations (dark blue:
 4 Gaussian, green: log, red: quadratic, light blue: constrained, pink: penalty).

5



1
2
3
4
5

Figure 4. Time series of annual global mean net ecosystem productivity (NEP) with posterior uncertainty.



1
2
3 Figure 5. Uncertainty correlation matrix of global mean NEP.
4