Geosci. Model Dev. Discuss., 4, 1513–1531, 2011
www.geosci-model-dev-discuss.net/4/1513/2011/
doi:10.5194/gmdd-4-1513-2011

Geoscientific Model Development Discussions

Development of an ensemble-adjoint optimization approach to derive uncertainties in net carbon fluxes

T. Ziehn[1], **M. Scholze**[1], **and W. Knorr**[1,2]

[1]Department of Earth Sciences, University of Bristol, Wills Memorial Building, Queen's Road, Bristol, BS8 1RJ, UK
[2]Department of Meteorology and Climatology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

Received: 9 June 2011 – Accepted: 23 June 2011 – Published: 8 July 2011

Correspondence to: T. Ziehn (tilo.ziehn@bristol.ac.uk)

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

Abstract

Accurate modelling of the carbon cycle strongly depends on the parametrization of its underlying processes. The Carbon Cycle Data Assimilation System (CCDAS) can be used as an estimator algorithm to derive posterior parameter values and uncertainties for the Biosphere Energy Transfer and Hydrology scheme (BETHY). However, the simultaneous optimization of all process parameters can be quite challenging, due to the complexity and non-linearity of the BETHY model. Therefore, we propose a new overall concept, which uses ensemble runs and the adjoint optimization approach of CCDAS to derive the full probability density function (PDF) for posterior soil carbon parameters and the net carbon flux at the global scale. In this way, we only optimize those parameters, which can be constrained best by atmospheric carbon dioxide (CO_2) data. The prior uncertainties of the remaining parameters are included in a consistent way through ensemble runs, but are not constrained by data. The final PDF for the optimized parameters and the net carbon flux are then derived by superimposing the individual PDFs for each ensemble member. We find that the optimization with CCDAS converges much faster, due to the smaller number of processes involved. Moreover, it is more likely that we find the global minimum in the reduced parameter space.

1 Introduction

The terrestrial biosphere plays an important role in the global carbon cycle and has a great impact on the accumulation of carbon dioxide (CO_2) in the atmosphere. Feedbacks between the carbon cycle and climate change, generally known as carbon-climate feedbacks, have the potential to accelerate the rise in atmospheric CO_2 which causes further global warming (Matthews et al., 2007). The quantification of the carbon cycle-climate feedback is therefore important with respect to determining the magnitude of future climate change. However, large uncertainties exists in the natural sinks of the terrestrial carbon cycle which in turn have a major effect on the uncertainty in

GMDD
4, 1513–1531, 2011

climate predictions (Zaehle et al., 2005; Denman et al., 2007). The large variations in the prediction of the future atmospheric CO_2 load result from differences in the models (Cramer et al., 1999; Friedlingstein et al., 2006) but also from uncertainties of the process parameters of the terrestrial ecosystem models (TEMs) (Knorr and Heimann, 2001).

The increase in the complexity of TEMs over the recent years has also led to an increase in the number of their parameters. Prior parameter values are usually based on "expert knowledge", which in some cases is little more than an informed guess. In addition the abstraction of the model means, that even those parameters that have system analogues are not directly analogous to the values derived from laboratory experiments or site-scale experiments. Parameter optimization methods are very useful in this context, because they provide a way of constraining the model parameters against observations and in this way are able to reduce the parameter uncertainties.

The Bayesian approach has been proven to provide an powerful and convenient framework for combining prior knowledge about parameters with additional information such as observations (Rayner et al., 2005). The resulting inverse problem expressed by Bayes' theorem can be solved in different ways, for example through Monte Carlo inversion (Sambridge and Mosegaard, 2002) or variational data assimilation. Monte Carlo inversion methods such as the Markov Chain Monte Carlo (MCMC) method are able to find an optimal solution by sampling the posterior probability density function (PDF) of the parameters directly. They are easy to implement and require no assumptions about the model (i.e. continuity), however they may require a very large sample size which is not always feasible to simulate due to computational limitations (i.e. long computing time of TEMs). Variational data assimilation such as the four dimensional variational scheme (4-D-Var) is the most advanced approach to accumulate observed information into a model. It uses derivative code (i.e. the adjoint of the model) for the optimization of the parameters and therefore requires the model to be differentiable with respect to all parameters. Although, the 4-D-Var approach is very efficient in most cases, the optimization might not always converge in time or might only identify a local

GMDD
4, 1513–1531, 2011

minimum. These issues arise due to the complexity and non-linearity of state-of-the-art TEMs and the potentially high-dimensional parameter space.

In this contribution we address the convergence issue of the optimization scheme in the 4-D-Var approach as used in the Carbon Cycle Data Assimilation System (CCDAS) (Rayner et al., 2005) and propose a new concept for deriving the posterior PDF for parameters and target quantities in a global TEM. Although, we focus only on one TEM, the Biosphere Energy Transfer and Hydrology scheme (BETHY) (Knorr, 2000), the approach is universal and could potentially be applied to any other model. The main idea is, that we only optimize parameters controlling the heterotrophic respiration (in the follwing called soil carbon parameters) in BETHY, because they are best constrained by atmospheric CO_2 concentration observations as demonstrated by Rayner et al. (2005) and Scholze et al. (2007). The prior uncertainties of the remaining parameters, parameters related to photosynthesis and autotrophic respiration, i.e. related to the net primary productivity (NPP), are included through ensemble runs and are therefore not constrained by the observations. The new concept allows to treat all parameter uncertainties in a consistent way.

2 Materials and methods

The 4-D-Var data assimilation scheme is the most advanced method in the field of parameter optimization and has been successfully applied within CCDAS to constrain process parameters in a TEM. CCDAS can be used in various modes. For example, in calibration mode it serves as an estimator algorithm for a set of photosynthesis, autotrophic and heterotrophic respiration process parameters by using automatically generated adjoint code (first derivative) for parameter optimization. In Hessian mode, the Hessian model code (second order derivative) is used for estimating posterior parameter uncertainties. As its ecosystem model, CCDAS uses the BETHY model, which simulates carbon assimilation and soil respiration within a full energy and water balance and phenology scheme. Calculated fluxes are then mapped to atmospheric concentrations using the atmospheric transport model TM2 (Heimann, 1995).

GMDD
4, 1513–1531, 2011

The CCDAS framework has been previously described in detail by Scholze (2003) and Rayner et al. (2005). Therefore, we provide only a brief summary. The data assimilation in CCDAS is performed in two steps: In the first step, the full BETHY model is used to assimilate global monthly fields of the fraction of Absorbed Photosynthetically Active Radiation (fAPAR) for optimizing parameters controlling soil moisture and phenology (Knorr and Schulz, 2001). In the second step, soil moisture and leaf area index (LAI) fields are provided as inputs for a reduced version of BETHY, in the following referred to as CarbonBETHY. This version is used to assimilate atmospheric CO_2 concentration observations from a large number of observation stations for optimizing photosynthesis and soil carbon parameters and to derive their posterior uncertainties (Rayner et al., 2005; Scholze et al., 2007).

2.1 Data assimilation

The Bayesian approach (Tarantola, 1987, 2005) provides a consistent framework for constraining model parameters $\boldsymbol{x}$ against observations $\boldsymbol{c}$. This framework enables us to combine the prior probability distribution of the parameters $P(\boldsymbol{x})$ with the probability distribution of the observations given the parameters $P(\boldsymbol{c}|\boldsymbol{x})$ in order to determine the inverse (posterior) probability distribution of the parameters given the observations $P(\boldsymbol{x}|\boldsymbol{c})$:

$$P(\boldsymbol{x}|\boldsymbol{c}) = \frac{1}{A} P(\boldsymbol{c}|\boldsymbol{x}) P(\boldsymbol{x}) . \quad (1)$$

The factor $1/A$ is a normalisation constant and independent of the parameters $\boldsymbol{x}$. In many cases a normal distribution is assumed for the prior parameter values and the observations. This Gaussian assumption has the advantage that only the mean and covariance have to be provided in order to describe the prior probability distribution of each variable. Applying a normal distribution to Eq. (1) leads to the following expression:

GMDD
4, 1513–1531, 2011

$$P(\boldsymbol{x}|\boldsymbol{c}) = \frac{1}{A}\exp\left(-\frac{1}{2}(\boldsymbol{c}_{\mathrm{M}}-\boldsymbol{c})^T\mathbf{C}_{\boldsymbol{c}}^{-1}(\boldsymbol{c}_{\mathrm{M}}-\boldsymbol{c})\right)\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)^T\mathbf{C}_{\boldsymbol{x}_0}{}^{-1}(\boldsymbol{x}-\boldsymbol{x}_0)\right), \quad (2)$$

where $\boldsymbol{c}_{\mathrm{M}} = M(\boldsymbol{x})$ are the modelled observations. The covariance matrices $\mathbf{C}_{\boldsymbol{c}}$ and $\mathbf{C}_{\boldsymbol{x}_0}$ express the uncertainty for the observations $\boldsymbol{c}$ and for the model parameter priors $\boldsymbol{x}_0$, respectively. We would like to find the peak in $P(\boldsymbol{x}|\boldsymbol{c})$, which will give us the most likely parameter value. This can be achieved in two ways: firstly, we can maximise Eq. (2) using Monte Carlo inversion or secondly, we can minimise the negative exponent of Eq. (2) using variational data assimilation. The cost function $J(\boldsymbol{x})$

$$P(\boldsymbol{x}|\boldsymbol{c}) = \frac{1}{A}\exp(-J(\boldsymbol{x})) \quad (3)$$

$$J(\boldsymbol{x}) = \frac{1}{2}\left((\boldsymbol{c}_{\mathrm{M}}-\boldsymbol{c})^T\mathbf{C}_{\boldsymbol{c}}{}^{-1}(\boldsymbol{c}_{\mathrm{M}}-\boldsymbol{c}) + (\boldsymbol{x}-\boldsymbol{x}_0)^T\mathbf{C}_{\boldsymbol{x}_0}{}^{-1}(\boldsymbol{x}-\boldsymbol{x}_0)\right) \quad (4)$$

describes the mismatch between the observations and their modelled equivalents and the mismatch between the parameters and their priors.

The data assimilation in CCDAS is based on the 4-D-Var scheme. In our case the observations $\boldsymbol{c}$ are the atmospheric CO_2 concentrations measured at 41 remote monitoring stations (GLOBALVIEW-CO_2, 2004). CarbonBETHY is used to calculate surface fluxes which are then mapped via the atmospheric transport model TM2 to atmospheric concentrations $\boldsymbol{c}_{\mathrm{M}}$. Since BETHY calculates only the natural land-atmosphere fluxes, we have to add land use change as an external flux as described in Rayner et al. (2005). Background fluxes for fossil fuel emissions are based on the flux magnitudes from Boden et al. (2009) as described in Scholze et al. (2007). The flux pattern and magnitude of ocean CO_2 exchange is taken from Takahashi et al. (1999) with estimates of inter annual variability taken from Le Quéré et al. (2007).

The parameter vector $\boldsymbol{x}$ contains the photosynthesis and soil carbon parameters in CarbonBETHY with their prior values represented by $\boldsymbol{x}_0$ (see Tables S1 and S2). A quasi-Newton method, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) variant of the

GMDD
4, 1513–1531, 2011

Davidon-Fletcher-Powell (DFP) formula (Fletcher and Powell, 1963; Press et al., 1996), is used for the minimization of the cost function, which requires the calculation of the gradient of J with respect to the control parameters $\boldsymbol{x}$ in each iteration. All derivative code is generated from the model's source code using the tool Transformation of Algorithms in Fortran (TAF) (Giering and Kaminski, 1998; Kaminski et al., 2003).

2.2 Challenges

The simultaneous optimization of the photosynthesis and soil carbon parameters in CCDAS as described in the previous section can be challenging due to slow convergence or failure of convergence. Even if a convergence requirement is fulfilled (i.e. test for convergence on Δx), the gradient of the cost function may not be sufficiently close to zero at the final convergence point in parameter space. As a consequence the Hessian is not positive definite (i.e. contains negative eigen-values), which indicates that an exact minimum has not been found. This has been noted by Rayner et al. (2005) where, in order to derive the posterior parameter uncertainties, the Hessian had to be modified manually. Another concern is that due to the large input space dimension and the fact that the BETHY model is highly non-linear, it is likely that we only identify a local minimum with CCDAS.

The study by Ziehn et al. (2011a) has revealed that the performance of the optimization in CCDAS can be significantly improved if only the soil carbon parameters are constrained with atmospheric CO_2 concentration data, while all parameters controlling net primary productivity (NPP) were kept fixed. Earlier studies with CCDAS (Rayner et al., 2005; Scholze et al., 2007) confirmed that NPP-related parameters (i.e. photosynthesis parameters) are constrained relatively little by the assimilation of CO_2 concentration observations. Most importantly it could be shown in the study by Ziehn et al. (2011a) that an ensemble run of optimizations, starting each of them in a different point, identified only one minimum in the physical parameter space, which makes it more likely that a global minimum has been found. Additionally, the gradient in the cost function minimum was almost zero and the Hessian was also positive definite, so

GMDD
4, 1513–1531, 2011

that no manual modification was required. Although the overall performance of the optimization has improved significantly, there is one drawback of the technique presented in their work: the uncertainties in the NPP-related parameters have not been included, which means that estimated uncertainties of the soil carbon parameters and diagnostics were only a lower bound. Therefore, we propose a new concept, that treats the uncertainties in the photosynthesis parameters via ensemble runs and optimizes the soil carbon parameters using the adjoint optimization approach within CCDAS.

2.3 New concept and test case

A flow chart of the concept developed in this contribution is presented in Fig. 1. In a first stage, ensemble runs are performed using CarbonBETHY by varying the NPP-related parameters randomly according to a normal distribution defined by their prior mean and standard deviation (see Table S1). All other parameters (i.e. soil carbon parameters) are kept fixed. Here, we use a sample size of $N = 200$. We also perform one additional forward run, referred to as base case, where all NPP controlling parameters are set to their prior mean.

CarbonBETHY is driven by observed climate data over 25 yr for the period 1979 to 2003. Global vegetation is mapped onto 13 different PFTs and each grid cell can contain sub-areas (sub-grid cells) with up to three different PFTs with their amount specified by each PFT's fractional cover. CarbonBETHY is run on a $2^{\circ} \times 2^{\circ}$ grid with 3462 land grid cells (excl. Antarctica).

Each ensemble run (including the base case) provides a monthly field of NPP, which is used as an input field in the second stage. Here, we apply CCDAS to optimize only the soil carbon parameters using atmospheric CO_2 concentration observations. Most of the soil carbon parameters are globally valid (i.e. they have the same value in each of the grid cells), only the carbon balance parameter β is differentiated by PFT and region (Ziehn et al., 2011a). In addition to the 13 PFTs (Fig. S1 and Table S3) we also consider 6 different regions (Fig. S2 and Table S4), which results in a set of 73 parameters (67 βs + 5 global parameters + 1 offset). For each NPP input field we obtain

GMDD
4, 1513–1531, 2011

a different set of optimal soil carbon parameters including their uncertainties. We can then propagate the posterior uncertainties for those parameters to any output target quantity of interest, for example the net ecosystem productivity (NEP) calculated as

$$\mathrm{NEP} = \mathrm{NPP} - R_\mathrm{S} = \mathrm{NPP} - (R_{\mathrm{S,s}} + R_{\mathrm{S,f}}), \tag{5}$$

where $R_{\mathrm{S,s}}$ and $R_{\mathrm{S,f}}$ are the respiration fluxes from the slowly and rapidly decomposing soil carbon pools, respectively.

In a third stage, we superimpose the posterior PDFs for the soil carbon parameters and the output target quantity in order to obtain their final PDF, which then also accounts for the prior uncertainties in the NPP-related parameters. The calculation of the final PDF $p(y)$ for the output target quantity y is given by the following equations:

$$p(y) = \frac{p'(y)}{N} \tag{6}$$

$$p'(y) = \sum_{i=1}^{N} \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{(y_i - \mu_i)^2}{2\sigma_i^2}\right), \tag{7}$$

where N is the ensemble size and μ_i and σ_i are the mean and standard deviation for each individual output y_i. The calculation of the final (superimposed) soil carbon parameter PDFs is performed in a similar way.

3 Results and discussion

The optimization within CCDAS (data assimilation in stage 2, see Fig. 1) reached convergence for 198 out of the 200 ensemble members and required about 1700 iterations on average. However, from the 198 successful optimization runs we had to exclude further 28 runs, where either the gradient in the cost function minimum was not sufficiently small enough or the optimal (posterior) parameter set contained non-physical parameter values. This leaves us with 170 sets of optimal soil carbon parameters.

GMDD
4, 1513–1531, 2011

A list of the posterior parameter values for the five global parameters including their uncertainties is presented in Table 1, the values for the parameter β for each PFT and region and the offset (global atmospheric CO_2 concentration at the beginning of the optimization period) are presented in Table S2. In the following, we focus only on the five global parameters.

The optimal values for the temperature sensitivity of the fast and slow carbon pool respiration ($Q_{10,f}$ and $Q_{10,s}$) for the base case are close to their prior values and within the prior uncertainty range. The soil moisture dependence parameter κ is reduced from its initial value, but is also within its prior uncertainty range. The optimized parameter values for the fast pool turnover time, τ_f, and the fraction f_s of the decomposition flux going from the fast to the long-lived soil carbon pool are much larger than their priors and both outside the prior uncertainty range. All five global parameters are well constrained by the CO_2 data, shown by the small posterior uncertainty in the base case. The posterior mean values for all soil carbon parameters are very similar in both cases (base case and superimposed case), showing that the mean values are not heavily effected by changes in the NPP-related parameters.

Our target quantities are global mean NEP for the 1980s and 1990s. The PDFs for those quantities are presented in Fig. 2. We obtain the final PDF by superimposing the 170 individual PDFs (Eq. 7) from each optimization run. In this way the final PDF accounts for both, the posterior soil carbon parameter uncertainties and the prior uncertainties in the NPP-related parameters. The superimposed PDF is non-Gaussian. For the 1990s the PDF can be approximated very well by a normal distribution, as indicated by skewness and kurtosis (Fig. 2b). The mean values for our target quantities are nearly identical for the base case and the superimposed case, showing again that the mean values are not heavily effected by changes in the NPP-related parameters. The uncertainties for the target quantities however increase by more than 50 % for the 1980s and by more than 100 % for the 1990s using the ensemble-adjoint method.

According to Denman et al. (2007) the terrestrial carbon sink removed $-1.7\,\mathrm{Pg\,C\,yr^{-1}}$ (range: -3.4 to $+0.2\,\mathrm{Pg\,C\,yr^{-1}}$) during the 1980s and $-2.6\,\mathrm{Pg\,C\,yr^{-1}}$ (range: -4.3 to

GMDD
4, 1513–1531, 2011

$-0.9\,\mathrm{Pg\,C\,yr^{-1}}$) during the 1990s from the atmosphere. The results from our study match the mean values well, with a carbon flux of $-1.83\,\mathrm{Pg\,C\,yr^{-1}}$ (range: -1.84 to $-1.82\,\mathrm{Pg\,C\,yr^{-1}}$) for the decade of the 1980s and $-2.55\,\mathrm{Pg\,C\,yr^{-1}}$ (range: -2.57 to $-2.54\,\mathrm{Pg\,C\,yr^{-1}}$) for the decade of the 1990s. However, the uncertainties of our results are small in comparison to those from Denman et al. (2007). One reason for this is the large number of negative entries in the error covariance matrix of global mean NEP for the 1980s and 1990s. The covariance between the flux uncertainties can be expressed via the uncertainty correlation matrix of diagnostics, $\mathbf{R_d}$, which is defined as follows:

$$R_{\mathrm{d}}^{\,i,j} = \frac{C_{\mathrm{d}}^{\,i,j}}{\sigma_i \sigma_j}, \tag{8}$$

where $C_{\mathrm{d}}^{\,i,j}$ is element i,j of the uncertainty covariance matrix of the diagnostics (global NEP per year), and σ_i the posterior uncertainty of parameter i derived from the diagonal elements $C_{\mathrm{d}}^{\,i,i}$ of the matrix $\mathbf{C_d}$. Figure 3 shows the correlation matrix for global mean NEP for the 1980s and the 1990s. Due to the large number of negative correlations the overall uncertainty for global mean NEP over the 10 yr period (1980s and 1990s) is rather small. However, the uncertainty for global mean NEP for a single year, for example the year 1990, is by at least a factor of two larger then global mean NEP for the 1990s in the base case and increases by the ensemble-adjoint method by more than a factor of four.

4 Conclusions

The developed ensemble-adjoint optimization approach allows us to treat all parameter uncertainties in a TEM in a consistent way. Some parameters are constrained against data using the 4-D-Var data assimilation scheme, whereas the uncertainties of the remaining parameters are included via ensemble runs. In this way we optimize only those parameters which are constrained best by the kind of observations used in

GMDD
4, 1513–1531, 2011

the 4-D-Var step. This has the advantage, that fewer parameters and processes are involved within the optimization process, which, in turn, speeds up the convergence of the optimization. We are also more likely to identify the global minimum in the reduced parameter space.

In this study we have illustrated the usefulness of the ensemble-adjoint optimization approach by including prior uncertainties of unconstrained model process parameters (here the NPP-related parameters, which have not been constrained by the atmospheric CO_2 data) to derive a full probability density function on the model's target output quantities. For future applications the proposed concept also allows the inclusion of posterior uncertainties for the remaining, yet unconstrained parameters. Ziehn et al. (2011b) have demonstrated how to constrain the parameters of the Farquhar et al. (1980) photosynthesis model using an extensive set of plant traits and therefore provide a way on how to derive the posterior PDF for the NPP-related parameters. Those results could potentially been used within the same ensemble-adjoint optimization framework, but in this case all parameters would be constrained by observational data.

Supplementary material related to this article is available online at: http://www.geosci-model-dev-discuss.net/4/1513/2011/gmdd-4-1513-2011-supplement.pdf.

Acknowledgements. Stimulating discussions with Peter Jan van Leeuwen are gratefully acknowledged. This work was supported by the NERC National Centre for Earth Observation.

References

Boden, T. A., Marland, G., and Andres, R. J.: Global, Regional, and National Fossil-Fuel CO_2 Emissions, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tenn., USA, doi:10.3334/CDIAC/00001, 2009. 1518

GMDD
4, 1513–1531, 2011

Cramer, W., Kicklighter, D. W., Bondeau, A., Moore, B., Churkina, C., Nemry, B., Ruimy, A., and Schloss, A. L.: Comparing global models of terrestrial net primary productivity (NPP): overviewand key results, Glob. Change Biol., 5, 1–15, 1999. 1515

Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P. M., Dickinson, R. E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., da Silva Dias, P. L., Wofsy, S. C., and Zhang, X.: Couplings between changes in the climate system and biogeochemistry, in: Climate Change 2007: The Physical Science Basis, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M. M. B., and Miller, H.L., Working Group 1 Contribution to the Fourth Assessment Report of the Intergovernmental Panell on Climate Change (IPCC), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 499–588, 2007. 1515, 1522, 1523

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species, Planta, 149, 78–90, 1980. 1524

Fletcher, R. and Powell, M. J. D.: A rapidly convergent descent method for minimization, Comput. J., 6, 163–168, 1963. 1519

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zengq, N.: Climate-carbon cycle feedback analysis: results from the C4MIP model intercomparison, J. Climate, 19, 3337–3353, 2006. 1515

Giering, R. and Kaminski, T.: Recipes for adjoint code construction, ACM T. Math. Software, 24, 437–474, 1998. 1519

GLOBALVIEW-CO_2: Cooperative Atmospheric Data Integration Project – Carbon Dioxide, CD-ROM, NOAA CMDL, Boulder, Colorado, available at: ftp.cmdl.noaa.gov/ccg/co2/GLOBALVIEW (last access: 6 July 2011), 2004. 1518

Heimann, M.: The global atmospheric tracer model TM2, Technical Report, 10, Deutsches Klimarechenzentrum, Hamburg, Germany, 1995. 1516

Kaminski, T., Giering, R., Scholze, M., Rayner, P., and Knorr. W.: An example of an automatic differentiation-based modelling system, in: Computational Science – ICCSA 2003, International Conference Montreal, Canada, Lecture Notes in Computer Science, edited by: Kumar, V., Gavrilova, M. L., Tan, C. J. K., and L'Ecuyer, P., Springer, New York, NY, vol. 2668, 95–104, 2003. 1519

GMDD
4, 1513–1531, 2011

Knorr, W.: Annual and interannual CO_2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Global Ecol. Biogeogr., 9, 225–252, 2000. 1516

Knorr, W. and Heimann, M.: Uncertainties in global terrestrial biosphere modeling, 1. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme, Global Biogeochem. Cy., 15, 207–225, 2001. 1515

Knorr, W. and Schulz, J. P.: Using satellite data assimilation to infer global soil moisture an vegetation feedback to climate, in: Remote Sensing and Climate Modelling: Synergies and Limitations, edited by: Beniston, M. and Verstraete, M., Springer, New York, 207–225, 2001. 1517

Le Quéré, C., Rödenbeck, C., Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A., Labuschagne, C., Ramonet, M., Nakazawa, T., Metzl, N., Gillett, N., and Heimann, M.: Saturation of the southern ocean CO_2 sink due to recent climate change, Science, 316, 1735–1738, doi:10.1126/science.1136188, 2007. 1518

Matthews, H. D., Eby, M., Ewen, T., Friedlingstein, P., and Hawkins, B. J.: What determines the magnitude of carbon-climate feedbacks?, Global Biogeochem. Cy., 21, GB2012, doi:10.1029/2006GB002733, 2007. 1514

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numerical Recipes in FORTRAN 77: the Art of Scientific Computing, Cambridge University Press, New York, NY, 1996. 1519

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), Global Biogeochem. Cy., 19, GB2026, doi:10.1029/2004GB002254, 2005. 1515, 1516, 1517, 1518, 1519

Sambridge, M. and Mosegaard, K.: Monte Carlo methods in geophysical inverse problems, Rev. Geophys., 40, 1009, doi:10.1029/2000RG000089, 2002. 1515

Scholze, M.: Model studies on the response of the terrestrial carbon cycle on climate change and variability, Ph.D. thesis, Max-Planck-Institute für Meteorologie, Hamburg, Germany, 2003. 1517

Scholze, M., Kaminski, T., Rayner, P., Knorr, W., and Giering, R.: Propagating uncertainty through prognostic carbon cycle data assimilation system simulations, J. Geophys. Res., 112, D17305, doi:10.1029/2007JD008642, 2007. 1516, 1517, 1518, 1519

Takahashi, T., Wanninkhof, R. H., Feely, R. A., Weiss, R. F., Chipman, D. W., Bates, N., Olafsson, J., Sabine, C., and Sutherland, S. C.: Net sea-air CO_2 flux over the global oceans: an

GMDD

4, 1513–1531, 2011

improved estimate based on the sea-air pCO_2 difference, Paper presented at 2nd International CO_2 in the Oceans Symposium, Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan, January 18–23, 1999. 1518

Tarantola, A.: Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation, Elsevier, New York, NY, 1987. 1517

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, Philadelphia, PA, 2005 1517

Zaehle, S., Sitch, S., Smith, B., and Hatterman, F.: Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics, Global Biogeochem. Cy., 19, GB3020, doi:10.1029/2004GB002395, 2005. 1515

Ziehn, T., Knorr, W., and Scholze, M.: Investigating spatial differentiation of model parameters in a carbon cycle data assimilation system (CCDAS), Global Biogeochem. Cy., 25, GB2021, doi:10.1029/210GB003886, 2011a. 1519, 1520

Ziehn, T., Kattge, J., Knorr, W., and Scholze, M.: Improving the predictability of global CO_2 assimilation rates under climate change, Geophys. Res. Lett., 38, L10404, doi:10.1029/2011GL047182, 2011b. 1524

GMDD

4, 1513–1531, 2011

Table 1. Prior and posterior parameter values including their uncertainties for five global parameters. Upper and lower percentiles equivalent to one standard deviation are given. Prior and posterior parameter values for the β parameters are provided in Table S2. Units: τ_f, years; all others unitless.

	Prior			Posterior					
				Base case			Superimposed		
Parameter	μ	$\mu-\sigma$	$\mu+\sigma$	μ	$\mu-\sigma$	$\mu+\sigma$	μ	$\mu-\sigma$	$\mu+\sigma$
$Q_{10,f}$	1.50	1.00	2.25	1.24	1.20	1.27	1.22	1.10	1.33
$Q_{10,s}$	1.50	1.00	2.25	1.65	1.63	1.67	1.65	1.53	1.73
τ_f	1.50	0.50	4.50	4.55	4.27	4.85	4.46	3.50	5.47
κ	1.00	0.10	10.00	0.60	0.59	0.61	0.60	0.55	0.71
f_s	0.20	0.10	0.40	0.80	0.79	0.80	0.79	0.75	0.82

GMDD
4, 1513–1531, 2011

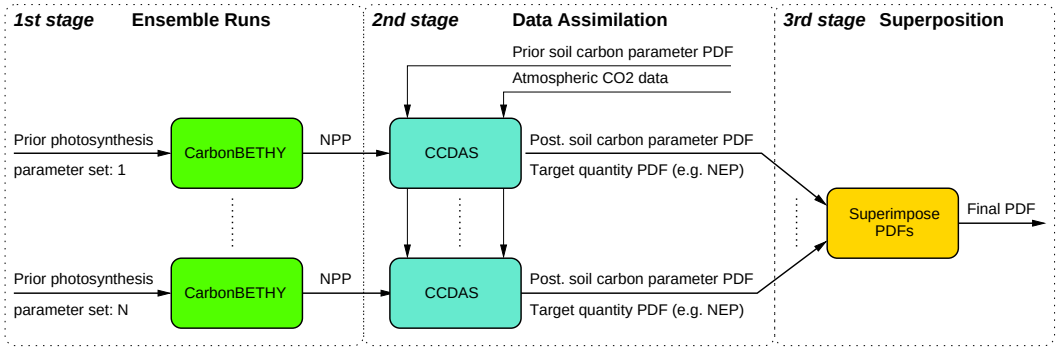

Fig. 1. Flow chart of the ensemble-adjoint optimization approach.

GMDD
4, 1513–1531, 2011

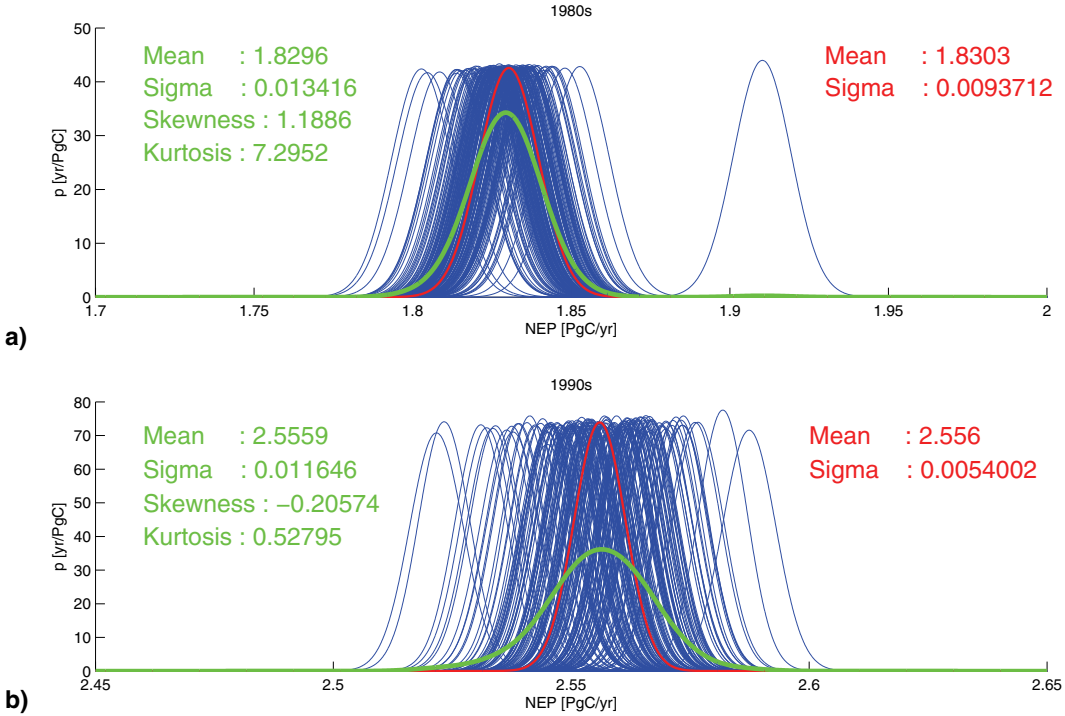

Fig. 2. PDFs for global NEP for **(a)** the 1980s and **(b)** the 1990s. Blue: 170 individual PDFs, red: base case PDF (prior photosynthesis parameters, not part of the ensemble), green: superimposed PDF from the 170 individual PDFs.

GMDD
4, 1513–1531, 2011

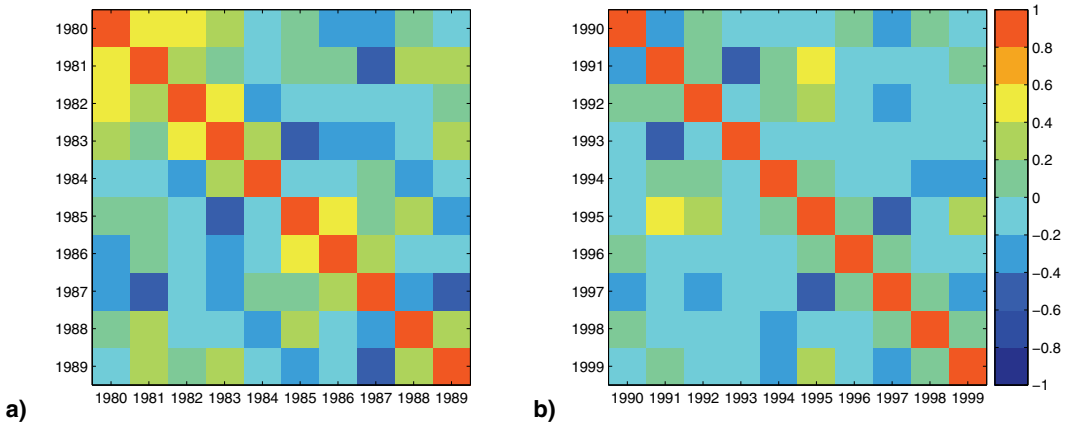

Fig. 3. Uncertainty correlation matrix of global mean NEP for **(a)** the 1980s and **(b)** the 1990s.

GMDD
4, 1513–1531, 2011