Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS)

P. J. Rayner,[1,2] M. Scholze,[3,4] W. Knorr,[4,5] T. Kaminski,[6] R. Giering,[6] and H. Widmann[5]

[1]   This paper presents the space-time distribution of terrestrial carbon fluxes for the period 1979–1999 generated by a terrestrial carbon cycle data assimilation system (CCDAS). CCDAS is based around the Biosphere Energy Transfer Hydrology model. We assimilate satellite observations of photosynthetically active radiation and atmospheric CO_2 concentration observations in a two-step process. The control variables for the assimilation are the parameters of the carbon cycle model. The optimized model produces a moderate fit to the seasonal cycle of atmospheric CO_2 concentration and a good fit to its interannual variability. Long-term mean fluxes show large uptakes over the northern midlatitudes and uptakes over tropical continents partly offsetting the prescribed efflux due to land use change. Interannual variability is dominated by the tropics. On interannual timescales the controlling process is net primary productivity (NPP) while for decadal changes the main driver is changes in soil respiration. An adjoint sensitivity analysis reveals that uncertainty in long-term storage efficiency of soil carbon is the largest contributor to uncertainty in net flux.

1.   Introduction

[2]   The rising concentration of CO_2 in the atmosphere and the attendant potential for climate change has stimulated substantial scientific and policy scrutiny. Two scientific foci have been to quantify the space-time distribution of fluxes of the gas to the atmosphere or to predict the evolution of these fluxes into the future. Generally, the tools used to perform these two tasks have been different. The most commonly used tool for flux mapping on the global scale has been atmospheric inversion [e.g., *Law*, 1999; *Rayner et al.*, 1999; *Bousquet et al.*, 2000; *Enting*, 2002; *Rödenbeck et al.*, 2003a].

[3]   In this approach we deduce a space-time pattern of fluxes which, when subject to atmospheric transport, produces a set of concentrations close to those observed. The continuous nature of the flux field, the diffusive nature of atmospheric transport, and the sparsity of the observing network conspire to make this a poorly conditioned inverse problem so that many possible flux fields are compatible with the observations [e.g., *Kaminski and Heimann*, 2001]. Approaches to solving this problem usually involve some kind of regularization [e.g., *Fan et al.*, 1999] such as the use of prior information as a constraint on the available solution or solving for fluxes only in a highly restricted subspace. These approaches contain their own pitfalls as demonstrated by *Kaminski et al.* [2001]. A more fundamental problem is that the solution, no matter how good, contains no information on the processes responsible for the inferred flux pattern. We hence cannot use results from such a study to help predict the future behavior of the carbon cycle.

[4]   The other approach is the traditional one of forward modeling of the most important processes. This approach can allow us to make predictions and can include process understanding, but it also has a disadvantage. Most forward models will be checked ad hoc against various data sources or, in the worst case, an ensemble of other models, but there is usually no mechanism built in to the model to incorporate such testing data formally. That task is usually known as data assimilation and is, necessarily, the operational procedure when best guess predictions must be made on the basis of current knowledge, the paradigm being numerical weather prediction. In this paper we aim to bring such an approach to the modeling of the terrestrial carbon cycle, a carbon cycle data assimilation system (CCDAS).

[5]   The basic approach has been applied previously to a highly simplified model by *Kaminski et al.* [2002] (hereinafter referred to as K02). Briefly, they optimized the

_____

[1]CSIRO Atmospheric Research, Aspendale, Victoria, Australia.
[2]Now at Laboratoire des Sciences du Climat et de l'environnement–CEA de Saclay Orme des Merisiers, Gif sur Yvette, France.
[3]Max-Planck-Institut für Meteorologie, Hamburg, Germany.
[4]Now at QUEST–Department of Earth Sciences, University of Bristol, Bristol, UK.
[5]Max-Planck-Institut für Biogeochemie, Jena, Germany.
[6]FastOpt, Hamburg, Germany.

controlling parameters of the Simple Diagnostic Biosphere Model (SDBM) [*Knorr and Heimann*, 1995] with respect to a set of observations, in their case a seasonal cycle of atmospheric CO_2 concentrations. The model, driven by these optimized parameters, was then run to predict a range of quantities of interest. These quantities included net fluxes and net primary production. K02 also generated the sensitivity of concentration with respect to the input parameters as part of a process of calculating uncertainties on parameters. In K02, as here, model parameters are the various constants in the model that are not, themselves, influenced by the model state. Thus, while they may change in space to reflect different conditions and physiological mechanisms, they will not change in time. The study of *Vukićević et al.* [2001] can be considered somewhat complementary to K02 since it considered interannual variability of fluxes while neglecting spatial variations. Like K02, *Vukićević et al.* [2001] used a variational technique to optimize their model parameters and an adjoint of their model to calculate the required gradient. In one sense, this paper represents a fusion of the studies of K02 and *Vukićević et al.* [2001] since it considers both spatial and temporal variability of both fluxes and concentrations.

[6] Less formal versions of the same approach have been used previously for estimating parameters in other terrestrial biosphere models. *Wang et al.* [2001] used eddy-flux data from a field campaign to constrain parameters in a model combining photosynthesis and canopy turbulence. *Randerson et al.* [2002] used a direct mapping of the probability distribution to estimate the seasonal exchanges and isotopic discrimination of the terrestrial high latitudes. Finally, *Barrett* [2002] used a genetic algorithm, and a model of primary production and soil and litter decomposition to predict the turnover time of soil carbon within Australia. These studies are interesting since, while they apply the same basic approach, they do so to a range of different data and using quite different optimization techniques.

[7] In this study we replace SDBM with the Biosphere Energy Transfer HYdrology Scheme (BETHY) model described by *Knorr and Heimann* [2001a]. This model incorporates more mechanistic process understanding than SDBM and includes the possibility of net sources and sinks. Thus we will make more detailed and explicit comment on the behavior of the biosphere in this paper. BETHY is also capable of coupling to atmospheric general circulation models for climate change studies. This has the important consequence that it is fully prognostic and can, if model-generated driving fields are available, be run for future scenarios. We can, therefore, apply knowledge about the current terrestrial carbon cycle to predict its evolution into the future. Furthermore, the formal uncertainty estimates gained during the parameter optimization step can be propagated to calculate uncertainties of any of the predicted quantities, including the future evolution of the model.

[8] Recent projections of the interactions of climate change and the carbon cycle make it quite timely to investigate how well various parameters are constrained by available data and hence how well future trajectories can be predicted. Two studies [*Cox et al.*, 2000; *Dufresne et al.*, 2002] have modeled the response of the total carbon cycle (land and ocean) to climate change and, given the rather obvious influence of carbon cycle dynamics on the radiative balance, the strength of the feedback. They produced markedly different results with final CO_2 concentrations in 2100 differing by nearly 200 ppm. The processes underlying these differences were analyzed by *Friedlingstein et al.* [2003]. That study identified three key processes leading to the differences. The first was differences in ocean uptake of CO_2. The second was the differing responses of vegetation cover, the model of *Cox et al.* [2000] producing the much-cited Amazon die-back under climate change [*Cox et al.*, 2004]. The third process, responsible for most of the difference in terrestrial carbon storage, was the response of the soil carbon reservoir to climate change. The response is controlled not only by the sensitivity of heterotrophic respiration (usually embodied in an exponential function of temperature) but also by the allocation of biomass to the various carbon pools in the model. It is an important question to ask how well current data can constrain the parameters which describe this response and, alternately, what extra measurements may aid the constraint. *Vukićević et al.* [2001] generated estimates for some of these parameters in their globally averaged model but did not calculate the parameter uncertainties that would allow comment on the strength of the constraint.

[9] Finally, the data assimilation approach can be used to probe the behavior of the current carbon cycle. Again we aim to combine the forward and inverse modeling procedures which have usually remained separate. Several inverse studies [e.g., *Rayner et al.*, 1999; *Bousquet et al.*, 2000; *Rödenbeck et al.*, 2003a] have produced space-time distributions of fluxes and commented on the underlying processes. Conversely, forward modeling studies of the recent past [e.g., *Dargaville et al.*, 2002; *Cao et al.*, 2002; *Potter et al.*, 2003] have compared output from process models against observations. These models, coupled with an apparently reasonable atmospheric transport model, generally underestimate both the seasonal cycle and interannual variability in atmospheric CO_2 concentrations [*Dargaville et al.*, 2002]. Most studies of interannual variability of ocean fluxes using either observations [e.g., *Feely et al.*, 1999] or models [e.g., *Le Quéré et al.*, 2003] suggest that ocean flux variability is insufficient to bring simulated atmospheric CO_2 concentrations close to observed. Similarly, recent inverse studies [e.g., *Bousquet et al.*, 2000; *Rödenbeck et al.*, 2003a] suggest terrestrial fluxes dominate interannual variability. Thus we have reason to suspect either the model formulation or the controlling parameters in terrestrial models. It can be hard to separate which of the two is primarily responsible. By optimizing model parameters against an observational data set, we can at least define the best possible model performance for a given model structure. We may then choose to improve upon this structure based on the irreducible mismatches with the observations. We have already pursued this path several times in the preparation of this paper.

[10] We can also, if the optimized model simulation is acceptable, ask questions such as what processes are responsible for the apparent increase in terrestrial carbon uptake from the decade of the 1980s to the 1990s. This increase had been noted in earlier inverse studies using
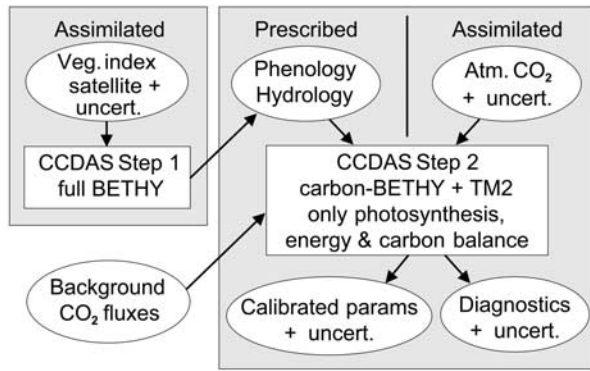
Figure 1. Sketch of the CCDAS setup. Ovals represent input and output data, and boxes represent calculation steps.

either atmospheric CO_2 alone [*Bousquet et al.*, 2000] or in combination with its isotopes [*Rayner et al.*, 1999] and was highlighted in the summary of *Prentice et al.* [2001]. The increase is directly inferred from the stability of the growth rate of atmospheric CO_2 in the face of substantial increases in fossil fuel emissions and only modest changes in other sources. The increase is rather rapid. *Prentice et al.* [2001, Table 3.1] give a land-atmosphere flux of -0.2 ± 0.7 GtC yr^{-1} for the 1980s and -1.4 ± 0.7 GtC yr^{-1} for the 1990s. This change in decadal mean flux suggests considerable sensitivity of the terrestrial carbon cycle to external forcing. Understanding this sensitivity may be one of the few observational clues we have of the multidecadal evolution of the carbon cycle during the 21st century.

2. Method

[11] The observed data used here are assimilated in two steps: First, the full BETHY model is used to assimilate remote sensing data for optimizing parameters controlling soil moisture and phenology. Second, a reduced version of BETHY is used to assimilate atmospheric CO_2 concentration observations. This simplified form of the model uses the optimized leaf area index (LAI) and plant available soil moisture fields from the first assimilation step as input data for the second step. The setup of the system is sketched in Figure 1.

2.1. Assimilating Remote Sensing Data

[12] The method of assimilating space-borne remote sensing data into the BETHY model has already been described in detail by *Knorr and Schulz* [2001b]. In brief, they defined a cost function by the squared deviation between global monthly fields of the fraction of Absorbed Photosynthetically Active Radiation (fAPAR) derived from satellite data and monthly fields of fAPAR as predicted by BETHY as well as initial and adjusted model parameters. Both observations and parameters are normalized by their assumed error variances. Model parameters are then adjusted until the cost function reaches a minimum using the downhill simplex method [*Press et al.*, 1992]. The chosen control parameters represent the phenology and hydrology part of the BETHY model and affect the leaf onset and shedding temperature and the maximum plant available soil moisture.

2.2. Assimilating Atmospheric CO_2 Data

[13] The method of assimilating atmospheric CO_2 data into BETHY follows closely the method as described by K02. Therefore it is only briefly sketched with emphasis put on differences between K02 and this study. CCDAS is operated first in a calibration mode and then in a prognostic mode. The calibration step of deriving an optimal parameter set **x** from atmospheric concentration observations **c** involves propagation of information in an inverse sense through a chain of models. In the forward direction, the flow of information is sketched in Figure 2. In the prognostic
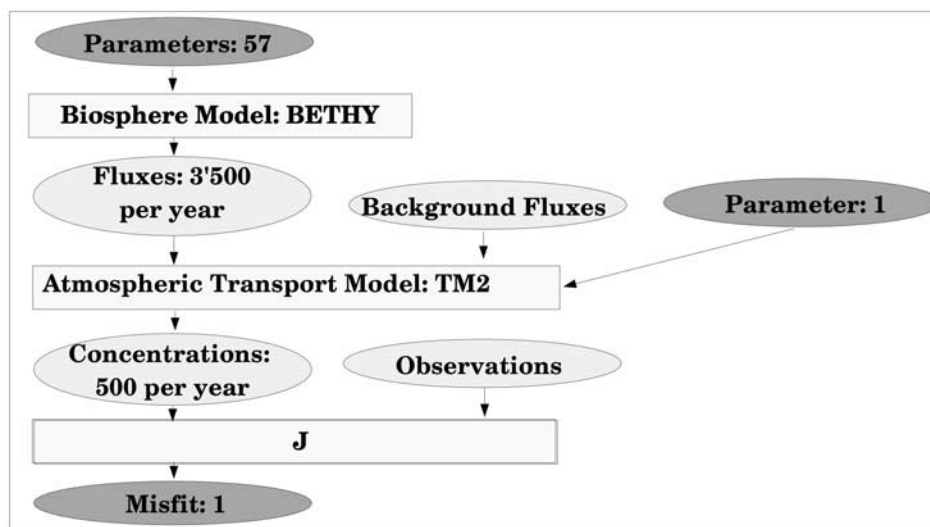

Figure 2. Model setup for the calibration mode. Oval boxes show the various quantities, dependent and independent variables are indicated by dark shading, and intermediate fields by light shading. Rectangular boxes denote the mappings between these fields.

mode, various quantities of interest (e.g., terrestrial carbon fluxes or atmospheric CO_2 concentrations) and their uncertainties can be calculated from the optimized parameter vector and its uncertainty as derived in the calibration mode.

[14] First, a terrestrial biosphere model **B** produces from a parameter vector **x** a space-time distribution of modeled fluxes $\mathbf{f}_M$ on a $2° \times 2°$ grid. These fluxes are mapped via an aggregation routine (**A**) onto the $7.8° \times 10°$ grid of an atmospheric transport model. Thus, small-scale forcing of the biosphere can propagate through to the atmospheric concentrations. The atmospheric transport model, denoted as **T**, finally maps fluxes onto atmospheric concentrations. The use of the adjoint approach described by *Kaminski et al.* [1999a] means that **T** can take fluxes at the full resolution of the underlying transport model and map them onto a predetermined set of observed concentrations. In fact, *Kaminski et al.* [1999a] demonstrated that if only modeled concentrations at selected sites and times are needed, the transport model can be included as a simple matrix multiplication without losing the required generality of flux fields.

[15] Combining the three steps yields a mapping, M, from biosphere model parameters to modeled concentrations $\mathbf{c}_M$,

$$\mathbf{c}_M = M(\mathbf{x}) = \mathbf{T} \times \mathbf{A} \times \mathbf{B}(\mathbf{x}). \quad (1)$$

The above equation represents a composition of functions. However, **T** and **A** are linear functions and therefore can be implemented as a matrix multiplication.

[16] The Bayesian approach is used here to formulate the optimization problem [*Tarantola*, 1987; *Enting*, 2002]. Thus a term embodying a priori knowledge **p** on the parameter vector is combined with the observational information to define an overall mismatch of an optimized parameter vector **x**. This mismatch is calculated by a cost function,

$$J(\mathbf{x}) = \frac{1}{2}\Big((M(\mathbf{x}) - \mathbf{c})^T \mathbf{C_c}^{-1}(M(\mathbf{x}) - \mathbf{c}) + (\mathbf{x} - \mathbf{p})^T \mathbf{C_p}^{-1}(\mathbf{x} - \mathbf{p}),\Big),$$
$$(2)$$

where $\mathbf{C_c}$ expresses the uncertainty for the observations **c** and $\mathbf{C_p}$ expresses the uncertainty for the prior **p** in the form of covariance matrices. This definition of the cost function reflects the assumption of Gaussian probability distributions for the observed concentrations and the a priori information about the parameters.

[17] Optimization of **x** in the nonlinear composite model M is done by minimizing J. Powerful minimization algorithms for functions like J rely on the availability of the gradient of J with respect to the parameters. Thus the optimization step uses the derivative of the cost function with respect to input parameters. The adjoint of the composite model, M, is a particularly efficient way of calculating this derivative in which M is run backward to compute the relationship between a small change in J and changes in **x**. All derivative code is generated by the tool Transformation of Algorithms in Fortran (TAF [*Giering and Kaminski*, 1998; *Giering et al.*, 2005]), which applies automatic differentiation [*Griewank*, 2000] to the source

code of the model. Hence TAF is applied in reverse mode to generate a subroutine to calculate $\nabla_\mathbf{x} J(\mathbf{x})$ for any **x**. A more detailed description of the use of TAF is given by *Kaminski et al.* [2003]. The adjoint of J evaluates the derivative of a scalar valued function. This derivative is a fairly compact expression, mapping a few dozen parameters onto their impact on the cost function. The cost function minimization allows the calculation of an optimal parameter set $\mathbf{x_{opt}}$.

[18] The dependency of the cost function on the inhomogeneous set of BETHY-parameters is highly nonlinear, which renders the minimization a complex task. Fortunately, in a few tens of iterations, the optimization can absorb the large-scale information from the observations, which is reflected by reduction in the value of the cost function by several orders of magnitude and a good fit to observations.

2.3. Calculation of Uncertainties

[19] The curvature of the cost function is the multivariate version of the second derivative. It is described by the Hessian, the matrix containing all the second and cross-derivatives at a point in parameter space. At a minimum in the cost function, the Hessian approximates the inverse covariance of the parameters [*Tarantola*, 1987]. The covariance matrix completely describes the formal uncertainty of the parameters provided their distribution is multivariate normal. We can calculate the Hessian by differentiating the gradient vector with respect to all parameters. We perform this by applying TAF a second time, to the function which calculates the gradient itself. Because the numbers of inputs and outputs for the gradient calculation are equal, this second differentiation is most efficiently calculated in forward mode, i.e., by the tangent linear model [*Kaminski et al.*, 2003].

[20] At the cost function minimum, the cost function must exhibit positive curvature in every direction. This will result in a positive definite Hessian, i.e., one for which all eigen-values are positive. The highly nonlinear nature of the model and the necessarily approximate numerical optimization routines mean we are unlikely to find the exact minimum in the cost function. This is clear from investigation of the gradient itself which should approach zero as we approach the minimum. In general, it does not. We also notice some negative eigen-values of the Hessian when we calculate it at $\mathbf{x_{opt}}$. This suggests the cost function contains small-scale noise in some directions. Meaningful uncertainty bounds in these directions should come from larger-scale curvature. A minimum value for this curvature is provided by the initial parameter uncertainty. We carry out our calculations in the dimensionless quantities formed by dividing each parameter by its prior uncertainty. In this modified parameter space every initial uncertainty is thus 1. We modify the Hessian to remove any information we suspect comes from this small-scale noise. We do this as follows:

[21] 1. Calculate the eigen-spectrum for the Hessian as

$$\mathbf{H} = \mathbf{V_H} \boldsymbol{\Lambda_H} \mathbf{V_H^T}, \quad (3)$$

where **H** is the Hessian, $\mathbf{V_H}$ is a matrix composed of the eigen-vectors of **H** (one per row) and $\boldsymbol{\Lambda_H}$ is the diagonal

Table 1. PFTs Defined in BETHY and Their Abbreviations

PFT Number	PFT Name	Abbreviation
1	tropical broadleaved evergreen tree	TrEv
2	tropical broadleaved deciduous tree	TrDec
3	temperate broadleaved evergreen tree	TmpEv
4	temperate broadleaved deciduous tree	TmpDec
5	evergreen coniferous tree	EvCn
6	deciduous coniferous tree	DecCn
7	evergreen shrub	EvShr
8	deciduous shrub	DecShr
9	C3 grass	C3Gr
10	C4 grass	C4Gr
11	tundra vegetation	Tund
12	swamp vegetation	Wetl
13	crops	Crop

matrix of eigen-values of $\mathbf{H}$ arranged in the same order as $\mathbf{V_H}$ (the usual diagonalization of $\mathbf{H}$).

[22] 2. In $\mathbf{\Lambda_H}$ replace any $\lambda < |\lambda_{min}|$ with 1. λ_{min} is the smallest (largest negative) eigen-value. We denote the modified matrix as $\mathbf{\Lambda'_H}$.

[23] 3. Reconstitute $\mathbf{H'}$ using $\mathbf{\Lambda'_H}$ in place of $\mathbf{\Lambda_H}$.

[24] If, as in our case, $\lambda_{min} < -1$, our reconstruction is conservative. Ideally, $\mathbf{\Lambda_H} > 1$ since the data should always add information to the prior estimate. The existence of negative eigen-values represents noise in the cost function, with an amplitude given by $|\lambda_{min}|$. Any direction less well constrained than this noise floor should be considered unconstrained, so we reset its uncertainty to the initial value of 1.

3. Models and Data

3.1. Terrestrial Carbon Cycle Model

[25] BETHY is a process-based model of the terrestrial biosphere [*Knorr*, 1997, 2000]. It simulates carbon assim-ilation and plant and soil respiration embedded within a full energy and water balance. Hence it is structured into four compartments: (1) energy and water balance, (2) photosynthesis, (3) phenology, and (4) carbon balance. BETHY is driven by observed climate data for the period 1979 to 2000 [*Nijssen et al.*, 2001] (which have been extended to the year 2000 (R. Schnur, personal communication, 2003)). It is run on a $2° \times 2°$ grid resolution. Global vegetation is mapped onto 13 plant functional types (PFT) based on work by *Wilson and Henderson-Sellers* [1985]. This is a condensation of the 23 PFTs in the original classification. A grid cell can contain up to three different PFTs, with the amount specified by their fractional coverage. Table 1 lists the PFTs, and the map in Figure 3 shows, in each grid cell, its dominant PFT.

[26] The model is used in two forms. In its full form it assimilates AVHRR data for the years 1989–1990 following the method of *Knorr and Schulz* [2001] as mentioned earlier. This procedure is used to optimize a set of time invariant parameters at each grid cell, related to water balance, temperature-limited phenology, and overall fractional vegetation cover. This first assimilation step thus provides monthly time series of LAI and plant available soil moisture ω (as a fraction of maximum soil water capacity) for later use in the simplified form of the model to assimilate atmospheric concentration observations. (Note that it is the parameters of the model that are fitted so that LAI and ω may respond to variations in climate.) In its simplified form (hereinafter known as Carbon-BETHY) the model is slightly reduced by leaving out the water balance from compartment 1 and the phenology compartment 3 completely. Therefore it can be run "off-line" with prescribed fields of LAI and ω from the full model. Time steps are 1 hour for the photosynthesis and autotrophic respiration part and 1 day for the heterotrophic respiration part. To
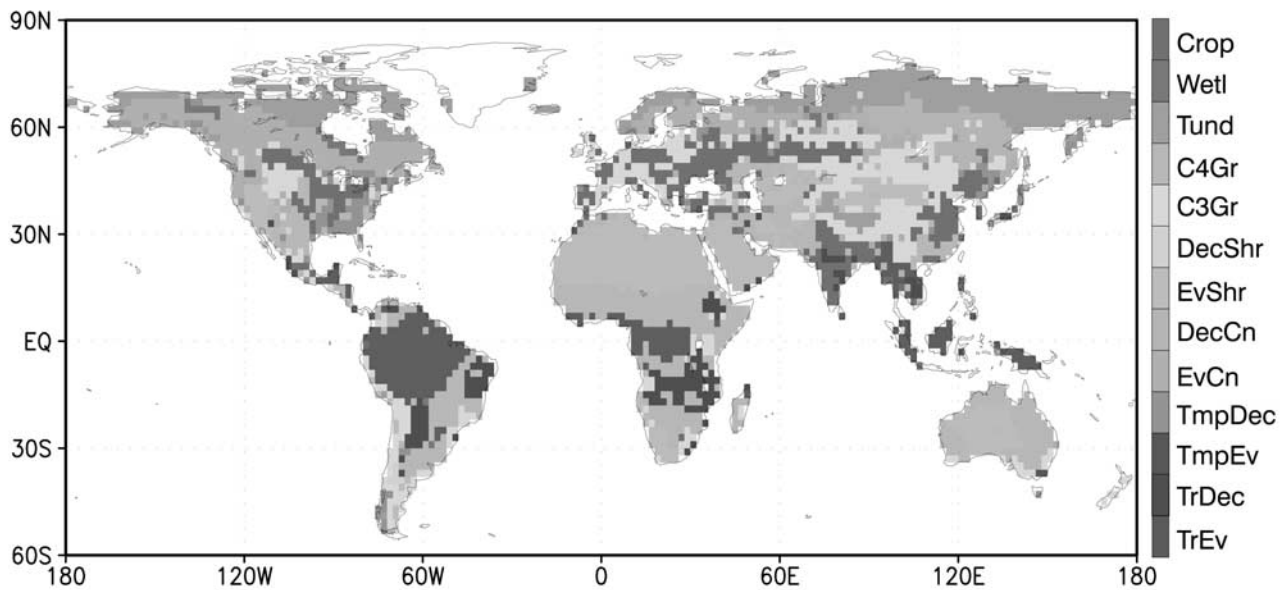

Figure 3. Distribution of the dominant PFT per grid cell. PFT labels are given in Table 1. See color version of this figure at back of this issue.

speed up computation, only one mean day is calculated per month preserving diurnal variations; the daily fluxes are then multiplied by the month length to generate monthly fluxes. Control parameters affect the photosynthesis scheme and both the autotrophic and heterotrophic respiration schemes. The parameters are described in the following paragraphs and listed with their a priori values and uncertainties in Table 2.

[27] At each model grid cell, photosynthesis is simulated following the formulations of *Farquhar et al.* [1980] or *Collatz et al.* [1992] for C_3 or C_4 metabolism, respectively. In the case of C_3 photosynthesis, gross primary productivity (GPP) is calculated as the minimum of an electron transport limited rate, J_E, and a rate, J_C, limited by the carboxylation enzyme Rubisco from which the leaf or dark respiration, R_d, is subtracted,

$$GPP = \min[J_C; J_E] - R_d, \qquad (4)$$

with

$$J_C = V_{max} \frac{C_i - \Gamma_*}{C_i + K_C(1 + O_x/K_O)} \qquad (5)$$

$$J_E = J \frac{C_i - \Gamma_*}{4\left(C_i + 2\Gamma_*\right)}, \qquad (6)$$

where

$$J = \frac{\alpha_q I J_{max}}{\sqrt{J_{max}^2 + \alpha_q^2 I^2}}, \qquad (7)$$

with the parameter α_q the quantum efficiency, C_i the leaf-internal CO_2 concentration, I the photosynthetically active radiation (PAR) absorption rate, and O_x the O_2 partial pressure. The maximum electron transport, J_{max}, varies linearly with the vegetation temperature, T_v in °C. The temperature sensitivity $a_{J,T}$ [*Farquhar*, 1988] is a control parameter,

$$J_{max}(T_v) = J_{max}^{25} \times a_{J,T} \times T_v, \qquad (8)$$

with

$$J_{max}^{25} = a_{J,V} \times V_{max}^{25}. \qquad (9)$$

Equation (9) is a simple recasting of the usual photosynthesis equations to use the parameters $a_{J,V}$ and $V_{max}(25°C)$ rather than the usual $J_{max}(25°C)$ and $V_{max}(25°C)$, and thus taking into account their correlation for the optimization. Initial values of $a_{J,V}$ are calculated from the a priori values of J_{max} and V_{max} at 25°C. The CO_2 compensation point, Γ_*, depends linearly on the vegetation temperature with the parameter $a_{\Gamma,T}$ being the dependency factor [*Farquhar*, 1988],

$$\Gamma_* = a_{\Gamma,T} \times T_v. \qquad (10)$$

Table 2. Controlling Parameters and Their Initial and Optimized Values and Uncertainties[a]

Parameter	Initial Value	Optimized Value	Initial Uncertainty	Optimized Uncertainty
V_{max}^{25} (TrEv)	60	60.2	2	1.98
V_{max}^{25} (TrDec)	90	89.9	2	1.88
V_{max}^{25} (TmpEv)	41	41.0	2	2.00
V_{max}^{25} (TmpDec)	35	34.9	2	2.00
V_{max}^{25} (EvCn)	29	28.5	2	2.00
V_{max}^{25} (DecCn)	53	52.8	2	2.00
V_{max}^{25} (EvShr)	52	51.9	2	2.00
V_{max}^{25} (DecShr)	160	160.1	2	2.00
V_{max}^{25} (C3Gr)	42	41.1	2	2.00
V_{max}^{25} (C4Gr)	8	7.9	2	1.99
V_{max}^{25} (Tund)	20	19.9	2	2.00
V_{max}^{25} (Wetl)	20	20.0	2	2.00
V_{max}^{25} (Crop)	117	117.7	2	1.92
$a_{J,V}$(TrEv)	1.96	1.96	0.5	0.50
$a_{J,V}$(TrDec)	1.99	1.98	0.5	0.50
$a_{J,V}$(TmpEv)	2.0	2.0	0.5	0.50
$a_{J,V}$(TmpDec)	2.0	2.0	0.5	0.50
$a_{J,V}$(EvCn)	1.79	1.79	0.5	0.50
$a_{J,V}$(DecCn)	1.79	1.79	0.5	0.50
$a_{J,V}$(EvShr)	1.96	1.96	0.5	0.50
$a_{J,V}$(DecShr)	1.66	1.66	0.5	0.50
$a_{J,V}$(C3Gr)	1.9	1.89	0.5	0.50
$a_{J,V}$(C4Gr)	17.5	17.5	0.5	0.50
$a_{J,V}$(Tund)	1.85	1.84	0.5	0.50
$a_{J,V}$(Wetl)	1.85	1.85	0.5	0.50
$a_{J,V}$(Crop)	1.88	1.88	0.5	0.50
α_q	0.28	0.27	0.5	0.50
α_i	0.04	0.04	0.5	0.50
K_C^{25}	460×10^{-6}	461×10^{-6}	0.5	0.50
K_O^{25}	0.33	0.34	0.5	0.50
$a_{J,T}$	0.04	0.04	0.5	0.50
$a_{\Gamma,T}$	1.7	1.71	0.5	0.50
E_{K_O}	35948	35971	0.5	0.50
E_{K_C}	59356	59229	0.5	0.50
$E_{V_{max}}$	58520	58725	0.5	0.50
E_k	50967	50967	0.5	0.50
E_{R_d}	45000	44884	0.5	0.50
$f_{R,leaf}$	0.4	0.24	2.5	1.26
$f_{R,growth}$	1.25	1.27	0.5	0.50
f_S	0.2	0.32	−0.013; +0.014	±0.013
κ	1.0	0.29	−0.09; +0.10	±0.006
$Q_{10,f}$	1.5	1.61	−0.07; +0.075	±0.068
$Q_{10,s}$	1.5	1.35	−0.07; +0.075	±0.006
τ_f	1.5	1.62	−0.16; +0.18	±0.017
β(TrEv)	1	1.44	25	4.37
β(TrDec)	1	0.35	25	1.24
β(TmpEv)	1	0.94	25	24.96
β(TmpDec)	1	2.48	25	7.37
β(EvCn)	1	0.92	25	0.93
β(DecCn)	1	0.73	25	2.01
β(EvShr)	1	1.39	25	4.79
β(DecShr)	1	0.39	25	2.27
β(C3Gr)	1	0.92	25	1.06
β(C4Gr)	1	1.56	25	2.20
β(Tund)	1	1.0	25	1.16
β(Wetl)	1	0.92	25	5.35
β(Crop)	1	3.36	25	24.81
offset	338	336.3	0.3	0.04

[a]Units are: V_{max}, μmol(CO_2)m^{-2}s^{-1}; $a_{J,T}$, (C)$^{-1}$; $a_{\Gamma,T}$, μmol(CO_2) mol(air)$^{-1}$(C)$^{-1}$; activation energies E, J/mol; τ_f, years; offset, ppm; all others are unitless. Uncertainties are in percentage except for lognormally distributed parameters for which a range is given. Uncertainties represent 1 standard deviation.

The temperature dependency of further enzyme kinetic rates such as the Michaelis-Menten constants K_O and K_C and the maximum carboxylation rate, V_{max}, are computed from the following equation with E the respective activation energy (with T_v in °C):

$$r(T_v) = r^{25} \exp\left\{ \frac{(T_v - 25°C)E}{298 \mathrm{K} R(T_v + 273°C)} \right\}. \tag{11}$$

R is the general gas constant in $\mathrm{J K^{-1} mol^{-1}}$, r stands for the rate in question, and r^{25} the respective rates at 25°C, i.e., K_O^{25}, K_C^{25} and V_{max}^{25} (the values of K_O, K_C, and V_{max} at 25°C). These rates and the respective activation energies (E_{K_O}, E_{K_C}, and $E_{V_{max}}$) are taken as parameters in this study.

[28] For C_4 photosynthesis, there are three limiting rates [*Collatz et al.*, 1992],

$$\mathrm{GPP} = \min[J_e; J_c; J_i] - R_d$$

$$J_e = V_{max}$$

$$J_c = k \times C_i \tag{12}$$

$$J_i = \alpha_i I,$$

where in the case of C_4 photosynthesis,

$$k^{25} = a_{J,V} \times V_{max}^{25}, \tag{13}$$

with k^{25} being the PEPcase (the initial CO_2 fixating enzyme in C_4 plants) CO_2 specificity at 25°C. The temperature dependency of k is then also calculated by equation (11) with the activation energy E_k as a parameter. Here α_i is the integrated C_4 quantum efficiency, and $a_{J,V}$ and α_i are both parameters. The dark respiration R_d is calculated as [*Farquhar et al.*, 1980; *Knorr*, 1997]

$$R_d(25°C) = \begin{cases} 0.011 V_{max}^{25} & (C_3) \\ 0.042 V_{max}^{25} & (C_4). \end{cases} \tag{14}$$

[29] The temperature dependency of R_d is also given by equation (11) with the activation energy E_R as a controlling parameter.

[30] Atmospheric CO_2 concentration as input to the photosynthesis model is kept constant, effectively excluding CO_2 fertilization. *Körner* [2000] noted a well-documented but highly variable response of terrestrial plants to increased CO_2 levels. Indirect effects, such as increased allocation to rapidly cycling forms of carbon [*Hungate et al.*, 1997], are even less well understood and can greatly reduce the effectiveness of CO_2 fertilization. We have, therefore, chosen to exclude it in the present study.

[31] The net primary productivity (NPP) is calculated as gross uptake of CO_2 by the leaves (GPP) minus total autotrophic respiration which includes plant maintenance respiration R_M and growth respiration R_G. Following *Knorr* [2000], R_M is calculated from the leaf respiration as

$$R_M = R_d / f_{R,leaf}, \tag{15}$$

with $f_{R,leaf}$ the leaf fraction of the maintenance respiration. Growth respiration is itself proportional to NPP and calculated as follows:

$$R_G = (f_{R,growth} - 1)\mathrm{NPP} = (f_{R,growth} - 1)(\mathrm{GPP} - R_M - R_G), \tag{16}$$

where $f_{R,growth}$ is the amount of carbon to be produced for a unit gain in vegetation biomass. Both $f_{R,leaf}$ and $f_{R,growth}$ are taken as parameters.

[32] The net ecosystem productivity (NEP) is now defined as

$$\mathrm{NEP} = \mathrm{NPP} - R_S = \mathrm{NPP} - R_{S,f} - R_{S,s}, \tag{17}$$

where the heterotrophic soil respiration (R_S) is composed of respiration from a short-lived litter pool with time-varying size and a long-lived soil carbon pool [*Knorr*, 2000]. Input to the litter pool is parameterized by the annual course of LAI in the case of deciduous PFTs or, in the case of evergreen PFTs, as a constant fraction of the leaf carbon pool. Soil respiration is assumed to be temperature and soil moisture dependent and calculated from the following equations:

$$R_S = (1 - f_S)k_f C_f + k_s C_s, \tag{18}$$

with the sizes of the fast or litter pool, C_f, and the slow pool, C_s. The rate constants are

$$k_f = \omega^\kappa Q_{10,f}^{T_a/10} / \tau_f \tag{19}$$

$$k_s = \omega^\kappa Q_{10,s}^{T_a/10} / \tau_s, \tag{20}$$

where f_S is the fraction of decomposition from the fast pool that goes to the long-lived soil carbon pool, T_a the air temperature, κ the soil moisture dependence parameter, $Q_{10,f}$ and $Q_{10,s}$ temperature dependence parameters, and τ_f and τ_s the pool turnover times at 25°C. Here f_S, κ, $Q_{10,f}$, $Q_{10,s}$, and τ_f are controlling parameters, whereas the turnover time of the slow carbon pool, τ_s, is not determined directly.

[33] In a normal forward run, the carbon pools must be spun up until respiration from these pools comes into equilibrium with material entering the pools. In fact, this is done for the fast litter pool by spinning up the model for 5 years. The slow carbon pool is treated differently as the spin-up can take centuries of simulation, thereby dominating the computational cost. This choice removes three influences that may affect the initial state of the model and hence the simulation: (1) the history of disturbance, (2) the history of CO_2 fertilization, and (3) the history of climate forcing. Of these, only the last can be captured in our simulation, and it has a relatively small effect on long timescales [*McGuire et al.*, 2001]. One cannot neglect slow respiration, however, since the relationship between this respiration flux and long-term mean NPP will determine the

overall carbon balance. In the absence of specific information regarding the effect of any of the three factors, the best estimate is to assume that NPP and SR at each grid cell are in balance [*Knorr*, 2000]. The size of the slow carbon pool is held constant through the simulation time of 21 years and is determined by the time averages of NPP and R_S (denoted by an overlying bar) at each grid cell and a scaling parameter β,

$$R_{S,s} = k_s C_s = \frac{\overline{NPP}/\beta - (1 - f_S)\overline{k_f C_f}}{\overline{k_s}/k_s}. \tag{21}$$

This formulation eliminates the constant factor $1/\tau_s$ present in equation (20). Substituting into equations (18) and (17) yields

$$\overline{NEP} = \overline{NPP}\left(1 - \frac{1}{\beta}\right). \tag{22}$$

From equation (22) it follows that $\overline{NEP}$ is zero for $\beta = 1$ (this is our first guess) and is positive (net uptake) for $\beta > 1$. This deliberate conflation of the turn-over time and pool size for the slow pool avoids the problem of determining these individually. As noted by *Trumbore* [2000], the turnover time is highly variable. Furthermore, the measurements we use in this study cannot distinguish pool size from turnover time since they appear together in the equations describing respiration. This is not such a serious problem with the fast pool, where the sensitivity of respiration to changes in pool size helps determine the turnover time. For the slow pool the initial size is so large that there will be little change over our study period. The disadvantage of our approach is that the response of the slow respiration to future climate changes and inputs of carbon is determined by both its size and its turnover time. We can only predict the ratio of the two.

3.2. Spatial Discretization

[34] A choice one always faces when setting up a parameter estimation problem such as this is the level of detail or the resolution of our parameter set. Carbon-BETHY has 21 controlling parameters for each PFT; there are 13 PFTs and 3462 grid points. Thus we have a range of choices from a single global description (21 controlling parameters) to an independent parameter set for every grid point (21 × 3462 controlling parameters). *Trampert and Snieder* [1996] and *Kaminski et al.* [2001] have pointed out the general problem of too coarse a parameter description resulting in biased estimates of parameters or, alternatively, an underestimation of the posterior uncertainty. In this study we apply 18 of the 21 parameters globally, mainly those concerning the soil model or general plant physiology. We vary two of the key photosynthetic parameters, V_{max} and the ratio J_{max}/V_{max}, for each PFT. Finally, the key carbon storage parameter β also varies with PFT. The full list of parameters with their initial values and uncertainties is given in Table 2. Many other choices are possible, and some parameters could be assigned by other criteria, for example, soil type. We will investigate the sensitivity of the results to the parameter

discretization in future studies. We note that while we do allow a spatially explicit treatment of some parameters and use a relatively sophisticated terrestrial model, many of our parameters are given quite low initial uncertainties. This was partly a result of early experience with the optimization scheme: the optimization would sometimes search unphysical regions in parameter space otherwise. Also, the choice represents an acceptance of the limited nature of the data set we are currently using. Clearly, there are a range of other observations that can be brought to bear on the terrestrial carbon cycle, and some of these will act to constrain processes not well observed by atmospheric concentrations and satellite greenness measures. The large excess of data over parameters does still allow the optimization to shift parameters a long way from their prior estimates (e.g., $f_{R,leaf}$) so we have probably not overly constrained the model with our choice of prior estimates.

3.3. Transport Model

[35] As with the K02 study, we use the adjoint form of the transport model TM2. The base model is described by *Heimann* [1995]. The model is an off-line tracer transport model with an about 8° latitude by 10° longitude grid and nine vertical levels. It is driven by analyzed winds from the European Centre for Medium-Range Weather Forecasts (ECMWF), in this case from the year 1987. The model features vertical transport by convection and turbulent eddy transport following the schemes of *Louis* [1979] and *Tiedtke* [1989], respectively.

[36] In K02 the critical observed data was seasonal cycles of atmospheric concentration. Consequently seasonal cycles of atmospheric transport were the most important simulated quantity for the transport model. In the current study, parameters will be constrained by long-term mean gradients, seasonal cycles, and interannual variability. TM2 has been a participant throughout the series of experiments in the Transport Model Comparison (TransCom) studies [*Law et al.*, 1996; *Gurney et al.*, 2003, 2004]. K02 already noted some of the idiosyncrasies of the seasonal transport of TM2, particularly unusually high simulated amplitudes for seasonal cycles over the tropics compared with high latitudes. We also note that the annual mean transport, diagnosed from spatial gradients of concentration arising from annually constant sources, is relatively fast with sources being mixed rapidly both zonally and meridionally. The model also exhibited relatively weak covariance between seasonal cycles of sources and transport. This combination of transport behavior meant that the model produced relatively low terrestrial uptake in the study of *Gurney et al.* [2003] (0.29 GtC yr^{-1} compared to 1.33 ± 1.12 GtC yr^{-1} for the average of all models). It also produced relatively low uptake over the Northern Hemisphere midlatitude continents (0.75 GtC yr^{-1} compared to 2.31 ± 0.61 GtC yr^{-1} for the average of all models). The overall terrestrial uptake in our calculation is fixed, but we expect these transport characteristics to produce smaller uptake over the Northern Hemisphere continents than many transport models and correspondingly larger uptake in the tropics.

[37] We use the adjoint form of TM2 as discussed by K02. The adjoint form of the model was constructed by

Kaminski et al. [1999a] and used in a previous flux inversion by *Kaminski et al.* [1999b]. As required by the shift to interannual rather than cyclostationary concentration fields, we use the pulse response form of the adjoint so that the decay of emissions back toward a uniform background can be captured. A limitation of this study is the use of meteorological driving data from a single year, in this case 1987. Thus interannual variations in measured concentration caused by interannual variations in transport will be misinterpreted as arising from interannual variations of sources. This may, in turn, corrupt estimates of parameters constrained by interannual source variations. *Kaminski et al.* [1999b] performed cyclostationary inversions using winds from 1986 and 1987 and the same adjoint form of TM2 and noticed only small differences in deduced fluxes. *Dargaville et al.* [2000], with a different model, noticed only a small difference in deduced fluxes when they used interannually varying transport and when they used the same year of meteorology throughout. *Rödenbeck et al.* [2003b] (with a different model again) reported larger differences. The results of *Kaminski et al.* [1999b], who used the same model as the current study, are probably the best guide. We will compare our inferred fluxes with a direct flux inversion of *Rödenbeck et al.* [2003a] using interannually varying transport as a cross check.

[38] The adjoint form of TM2 calculates the change in concentration at a given site in response to a source field. To compare model predictions with observations, we must specify an initial value for the concentration which we assume is well mixed. To avoid biasing the calculation with a wrong choice, we include this as an additional controlling parameter, offset.

3.4. Background Fluxes

[39] A problem in making any inferences about the terrestrial biosphere from atmospheric observations is the contribution of other CO_2 fluxes. In the language of inverse problems, these are known as "nuisance variables." They are not perfectly known, impact the data we wish to use, but are not the target of our study. We must use prior estimates of these fluxes. In this study we do not include any parameters concerning these fluxes in the controlling parameters for the optimization. This is a brave choice, indicating complete faith in the prior estimates. To ameliorate the position, we include an extra contribution to the data uncertainty arising from possibly incorrect specification of these fluxes. We consider three contributions to this background flux.

3.4.1. Fossil Fuel Emissions

[40] Combined emissions from fossil fuel burning and cement production are known to have increased over the study period and changed in spatial structure. We consider the first of these changes in our calculation but not the second. We use the flux magnitudes from *Marland et al.* [2001] for the years 1979 to 1995 and a constant magnitude of 6.5 GtC yr^{-1} for the years 1996–1999 that are not included in that study. The spatial pattern is that pertaining to 1990, taken from the data of *Andres et al.* [1996]. The most critical shortcoming in these data, as we use them, is probably the lack of seasonality in the flux. The differences

in growth rates of fossil fuel usage among different regions could also lead to misallocation of trends in the terrestrial flux.

3.4.2. Ocean Flux

[41] The specified ocean flux we use is taken from two sources. As in K02, we use the flux pattern and magnitude from *Takahashi et al.* [1999] to describe the flux climatology (both annual mean and seasonal cycle). We add to this an estimate of interannual variability in ocean flux taken from the study of *Le Quéré et al.* [2003] for the years 1980 to 1998. We first remove the seasonal climatology from this field and then add the anomalies for each month to the *Takahashi et al.* [1999] climatology. Interannual variability in ocean flux is generally believed to be small [*Prentice et al.*, 2001]. This is also true for the *Le Quéré et al.* [2003] study. Therefore, even if relative errors in that flux are large, the impact on our study will be small. More important is the long-term trend in flux with gradually increasing uptake in response to increasing atmospheric CO_2 concentrations. This trend is captured by the model of *Le Quéré et al.* [2003] with decadal average ocean uptake in the 1990s 0.3 GtC yr^{-1} greater than for the 1980s.

3.4.3. Land-Use Change

[42] This is undoubtedly the most problematic of the background fluxes for our study. Some care is even required in thinking about what quantity we need for such a study. We must include those processes which cannot be captured by our terrestrial model. These might well be different in different models. For example, a model using some diagnostic of leaf area can, in principle, capture the reduction in photosynthesis which may result from land clearing. Provided the model has been properly equilibrated with the previous vegetation cover, this will result in an imbalance in the biomass and soil carbon pools and, most likely, a net efflux. This effect may well be already included in a land-use change flux estimate which accounts for these changes in pool size. It is better to treat land-use change as part of a general model of disturbance within the terrestrial model. BETHY has no treatment of disturbance and so we must specify land-use change as an external flux. By assimilating only 2 years of satellite-derived fAPAR and adjusting time invariant parameters, we exclude observed interannual variations of LAI, some of which may be the result of land use change. Including these would bias the phenology model.

[43] As in K02, we use the estimate of *Houghton et al.* [1987] for the flux caused by land-use change. Importantly, this flux compilation contains no information about either seasonality or interannual variability. In K02 we noted the importance of errors in the seasonality of background fluxes. There the amplitude of the seasonal cycle was a measure of net primary productivity. BETHY is a more complex model so the effect of this error is harder to judge. The importance of neglecting interannual variability in the land-use change flux is also hard to assess. One focus of this study is differences in terrestrial fluxes between the decades of the 1980s and 1990s. The studies reported in section 3.4.2 [*Prentice et al.*, 2001] suggest smaller decade-to-decade differences in land-use flux with a fixed methodology than are produced by different methods of estimating

this flux for a given decade. The study of *Houghton* [2003] suggests a difference of only 0.2 GtC yr^{-1} for the average flux due to land-use change between the 1980s and 1990s. This estimate is close enough to the 0 value implicitly assumed in our study. More serious is the neglect of shorter-term variability. BETHY makes no attempt to account for emissions from fires. *Langenfelds* [2002] and *van der Werf et al.* [2004] suggest this flux is a large contributor to interannual variability, but this is a point of contention. The capacity of our optimized model to fit the events highlighted by *Langenfelds* [2002] might provide one clue.

3.5. Data

[44] As already mentioned, the model is optimized in two stages and against two different data sets. The first data set is of daily AVHRR fields from the NOAA Global Vegetation Index (GVI) satellite data archive for the years 1989 and 1990 with a resolution of $1/7°$ latitude by longitude [*Gutman et al.*, 1995]. From this field the Global Environment Monitoring Index (GEMI) [*Pinty and Verstraete*, 1992] is computed and then averaged to $1°$ latitude by longitude over space and to monthly values over time. The annual average data coverage is almost 75% of the global land area. Values of GEMI are then translated into fAPAR following a regression derived by *Knorr and Schulz* [2001]. The error in fAPAR is estimated to lie between 0.05 and 0.1.

[45] The second data set is of monthly mean atmospheric concentration data from *GLOBALVIEW-CO$_2$* [2001]. These data consist of pseudo-weekly interpolation of spatial and temporal fits to flask and continuous in situ CO$_2$ concentration data measured at approximately 100 sites globally. We limit ourselves to stations from the NOAA/CMDL network [*Conway et al.*, 1994] to avoid possible calibration problems among different laboratories. We use the time series of residual standard deviations (RSD) from the compilation to assign a data uncertainty to the observations. We only use data from years when sufficient measurements are made to assign values without the gap-filling procedures in the *GLOBALVIEW-CO$_2$* [2001] compilation. We effectively remove gap-filled values from the cost function by assigning a very high value to the data uncertainty. We also impose a floor of 0.5 ppm to the data uncertainty via the equation

$$\text{unc} = \sqrt{0.25 + \text{RSD}^2}, \qquad (23)$$

where unc is the data uncertainty used in the cost function. Equation (23) also accounts for model error [*Tarantola*, 1987]. Our data uncertainties range from 0.51 ppmv to 4.7 ppmv.

[46] The choice of CO$_2$ monitoring stations to use in a study such as this is also a trade-off among competing requirements. The more data one can include the lower the predicted uncertainty on parameters will be and, correspondingly, the better the constraint on predicted quantities. However, changes in station density throughout the study period may be misinterpreted by the optimization procedure as interannual variability in concentration, whereas it may be only the observation of previously undetected phenom-
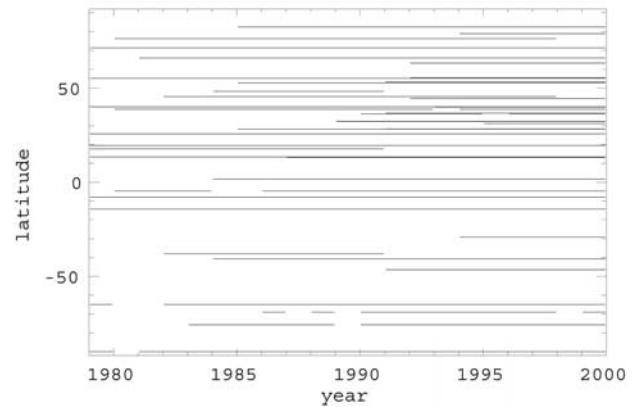

Figure 4. Temporal coverage of the data network as a function of latitude. Each station is plotted as a horizontal line at its latitude. The line is broken when data are not present.

ena. Further, stations with unrealistically large responses to the background fluxes will bias estimates of the mean concentration climatology. This climatology also helps constrain parameters. In summary, then, we need an observing network sufficiently dense for a good constraint, sufficiently invariant to allow realistic estimates of interannual variability, and containing monitoring stations for which we can expect a reasonable simulation by the transport model. Most of these criteria also held for K02. We therefore used the same observing network as that study. Figure 4 shows a representation of the temporal coverage of the network as a function of latitude and the actual location of stations is shown in Figure 10 in section 4.3.2.

[47] The geographical coverage shown in Figure 10 in section 4.3.2 demonstrates a common problem in determining terrestrial carbon fluxes from atmospheric observations: the predominantly marine focus of the observing network. There is an almost complete absence of data over the tropical continents, and even Europe has few sites that meet the criteria. Most of the European sites have too strong a response to the fossil fuel source, a result of the coarse resolution of TM2. Figure 4 shows that the major gap in latitude is south of $15°S$, a less serious concern for a study focusing on terrestrial fluxes. Also, while coverage clearly increases throughout the study period, there is a reasonable sampling of latitudes even at the beginning. This suggests that a PFT that is observed at all within the study period is likely to be observed throughout the whole period.

4. Results

[48] As might be expected, there is a rich suite of possible results to explore from such a calculation. In this paper we summarize and show examples of the fit to the concentration data we used in the calibration step. Next we show some of the parameters and their uncertainties, along with interpretation. Finally we show a range of estimated quantities and associated uncertainty. We focus on the difference in net uptake between the 1980s and 1990s. We present results from a single control optimization with initial

Table 3. Long-Term Mean Error (Model - Observed) (ppmv) and Normalized Misfit Value for the Worst Fitted Stations in the Control Optimization

Station	Location	Mean Mismatch, ppm	χ^2
Tae-Ahn Korea	36.7°N 126.8°E	−2.30	0.8
Baltic Sea	55.5°N 16.7°E	−1.59	0.29
Easter Island	29.2°S 109.4°W	1.16	1.06
Alert Canada	82.5°N 62.5°W	−0.40	2.78
Mould Bay Canada	76.3°N 119.4°W	−0.15	3.10
Ascension Island	7.9°S 14.4°W	0.97	3.39
Cold Bay Alaska	55.2°N 162.7°W	0.00	3.63
Barrow Alaska	71.3°N, 156.6°W	0.34	3.92
Shemya Island	52.7°N 174.1°E	0.42	4.56

parameter values and uncertainties taken from Table 2. We carried out a second optimization with initial parameter values perturbed by ±10% of their prior uncertainties. It converged to almost the same point; only the f_S parameter was noticeably different. This suggests the optimization is moderately robust.

4.1. Fit to Data

[49] The overall quality of fit to the data is embodied in the value of the cost function. This is dominated by the mismatch between predicted and observed concentrations, with only a small contribution from the mismatch between prior and posterior parameters. The final value of the cost function is 9588. More usefully, the average squared mismatch or reduced χ^2 statistic is the most common measure of the quality of fit. It is obtained by dividing twice the cost function by the number of observations. In our case the denominator is ambiguous. Should we use the gap-filled values described in section 3.5 or not? As a conservative choice, we do not include these months in our reduced χ^2 statistic, leaving 6936 observations. Our reduced χ^2 value is hence 2.76. In general, we would like a χ^2 of ≈1. The larger value here suggests that our model is incapable of fitting the data as well as the data and model uncertainties demand. We

should, therefore, regard the uncertainty estimates on parameters and predicted quantities as slightly optimistic.

[50] Table 3 lists some of the worst mismatches between predicted and observed concentration values. The table lists both the long-term mean mismatch and the normalized misfit value for some of the worst stations. The normalized misfit is the reduced χ^2 statistic considering only data from one station. We first notice considerable separation between those sites with bad long-term mean mismatches and those with high χ^2 values. For some of the bad mean mismatches, it is most likely not the terrestrial model that is responsible. For example, two of the worst mean mismatches are at Easter Island and Ascension Island. The concentrations here are under oceanic control. Similarly, the sites at Tae-Ahn Peninsula and Baltic Sea have large responses to the background fossil fuel fluxes. Small errors in these fluxes or their atmospheric transport to these observing sites could easily explain these errors. The large χ^2 values, on the other hand, are generally associated with sites of high seasonality. Note that four of the highest six mismatches occur for high-latitude North American sites. We have previously noted concerns with the seasonality of the transport in the TM2 model, but we must suspect that the terrestrial model has difficulty simulating the seasonal cycle of fluxes at high latitudes. Ascension Island shows a high χ^2 as well as a bad long-term mean mismatch. The high χ^2 value is harder to attribute, but errors in the seasonality of the ITCZ would cause large concentration errors here. This station was problematic in the simulation of *Knorr and Heimann* [1995] using the same transport model. They also noticed a significant impact from biomass burning.

[51] We can decompose the time series at a station into a climatological seasonal cycle and interannual variability, the latter often filtered to emphasize signals of key scientific interest. Here we use the procedures of *Thoning et al.* [1989] for this decomposition. Figures 5 and 6 show seasonal cycles fitted to the simulated and observed concentration time series at Barrow, Alaska, and Niwot Ridge,
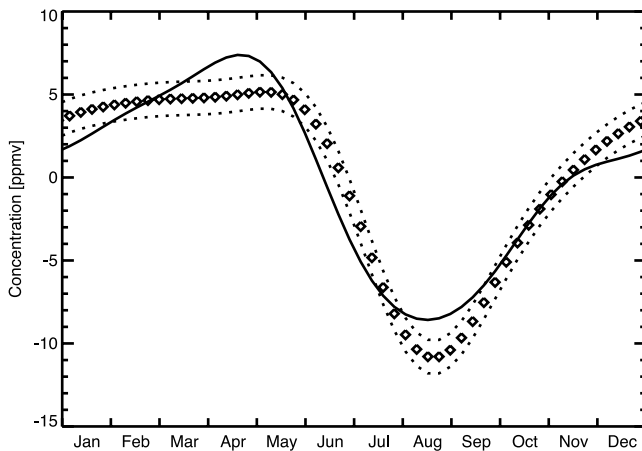
Figure 5. Modeled (solid line) and observed (diamonds) climatological seasonal cycle of CO_2 concentration from Barrow, Alaska. The dotted lines show the 1σ data uncertainty averaged over the whole record.
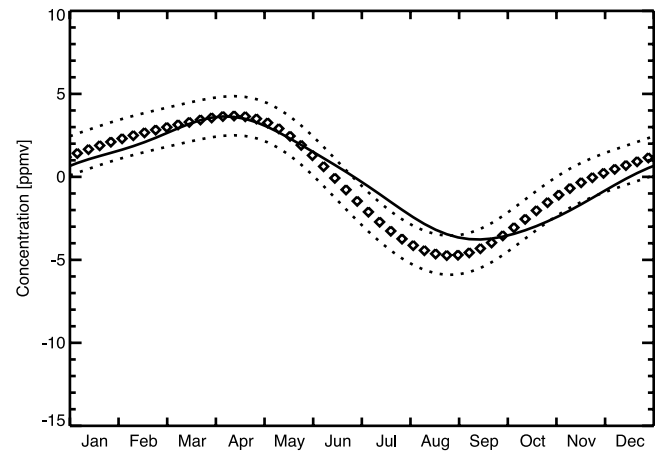
Figure 6. Modeled (solid line) and observed (diamonds) climatological seasonal cycle of CO_2 concentration from Niwot Ridge, Colorado. The dotted lines show the 1σ data uncertainty averaged over the whole record.
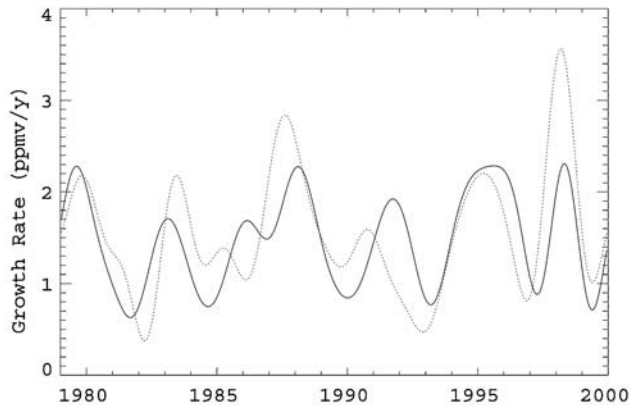
Figure 7. Observed (dotted line) and modeled (solid line) growth rates smoothed at approximately 1.5 year time-scales. Growth rates are calculated from equation (24). See color version of this figure at back of this issue.

Colorado. The uncertainties indicated by the dotted lines are the 1 standard deviation uncertainties averaged over the whole record.

[52] At Barrow, the most serious errors in the seasonal cycle occur in spring and summer. There is an overestimate of concentration in spring and a failure to match the depth of the summer drawdown. The underestimate of the seasonal cycle was noted by *Nemry et al.* [1999] in their comparison of many terrestrial carbon cycle models. This intercomparison relied on the same transport model we use (TM2). *Law et al.* [1996] found that TM2 simulated a relatively weak summer minimum compared to the other models in that study. We stress, again, the potential dependence of our results on the transport model we use. Likewise, the failure to match the slight increase in spring was common to many models in the study of *Law et al.* [1996], leading them to suspect the flux field they used as input. Finally, there is an underestimate of the concentrations in winter. This is most commonly ascribed [e.g., *Dargaville et al.*, 2002] to a failure to predict any soil respiration in winter, usually because the model disregards the insulating behavior of the snowpack. The behavior was also noted by *Knorr and Heimann* [2001a, 2001b] in the unoptimized version of BETHY, and by *Knorr and Heimann* [1995] for the simple model used in K02.

[53] We see the same underestimate of winter concentrations at Niwot Ridge. The spring peak is well matched here, but the failure to match the summer drawdown is compounded by a delay of a month in the minimum value.

[54] A good deal of the rest of the analysis in this paper focuses on the interannual variability in terrestrial fluxes. We can assess the verisimilitude of this variability by comparing interannual changes in the global CO_2 growth rate with those observed. We must use a consistent measure for the model and observed growth rates. While we can calculate the modeled growth rate exactly from the surface integral of fluxes, we must use some proxy for the observed. We can test the quality of the proxy by comparison with the

global average of the Globalview Marine boundary layer (MBL) curve. We find that, on average, the expression

$$C_{GLOB} = 0.25 C_{SPO} + 0.75 C_{MLO} \qquad (24)$$

(where SPO and MLO refer to South Pole and Mauna Loa, respectively) gives a good match to the globally integrated MBL value. Figure 7 shows C_{GLOB} for both the model and observations. Here we do use the extended Globalview record even though it is not used in the actual fit. The growth rate is calculated as the derivative of the trend curve fitted by the procedure of *Thoning et al.* [1989]. A third-order polynomial for the long-term trend and four harmonics of the seasonal cycle are removed and the residuals fitted with a smoothing spline with cutoff at 80 and 650 days. The curve represents variations in growth rate at interannual frequencies, but decadal variability is suppressed. The fit is generally good. In particular the timings of most major changes in growth rate are matched, although with some suggestion of a delay. Two significant divergences are the period around 1986 and, more strikingly, around 1991. In both years, modeled growth rates continue to increase at times of decrease in the observed growth rate. More significant, perhaps, is the failure to match the major growth rate maxima in the record, around 1983, 1989, and 1998. In each case the model underpredicts the peak in the observed growth rate. We interpret this failure as a lack of a disturbance parameterization in BETHY. There is strong evidence that fire was a significant contributor to the high growth rate observed around 1997–1998, either from trace gas measurements [*Langenfelds et al.*, 2002] or from satellite evidence [*Page et al.*, 2002; *van der Werf et al.*, 2004]. We can calculate the anomalous growth rates for the period August 1997 to September 1998, the period of the *Langenfelds et al.* [2002] and *van der Werf et al.* [2004] studies. We subtract the growth rates for this period from the 1980–2000 mean for both the modeled and observed time series shown in Figure 7. The difference between them is equivalent to a source of 2.5 GtC yr^{-1} for this period. This is within the range of 1.9 ± 0.8 GtC yr^{-1} attributed to biomass burning by *van der Werf et al.* [2004] for the same period. This is indirect support for the role of biomass burning in this event. We can say that the model cannot match the event even when its behavior has been tuned as much as possible. This suggests that some process not included in the model is responsible, but we cannot say which process. We note that the minima in the growth rate are better modeled, suggesting that the model does contain the necessary processes to explain these events.

[55] We can summarize the comparison between the model performance in fitting the seasonal cycle and the interannual anomalies in concentration. For the seasonal cycle, we construct time series like those shown in Figures 5 and 6. We calculate the RMS error across all 41 stations, at 1.15 ppmv. For the interannual anomalies, we calculate time series as the integral of the curves in Figure 7. We adjust these integrated curves so that they have the same mean. The RMS difference between the two integrated time series is 0.64 ppmv. We do not account for the data uncertainty

here since the calculation of such an error for the interannual anomalies is not well defined. When we consider that the ensemble of stations includes many with very small seasonal cycles, it is clear that the average performance is better for interannual than seasonal data.

[56] As a final test of the model performance, Figure 8 shows the variation in terrestrial flux computed by our optimization and two inversion calculations from *Rödenbeck et al.* [2003a]. Our calculations and those of *Rödenbeck et al.* [2003a] are not completely independent since both use some of the same atmospheric data. The inversion of *Rödenbeck et al.* [2003a] has the advantage that it is not subject to potential bias from assumptions about underlying processes. *Rödenbeck et al.* [2003b] performed several inversion calculations with different station networks; they faced the same trade-off of length of run versus station density as we do here. We compare with their 16 and 19 station cases since these cover the period of most of the interesting events. Their 42 station case, more comparable in some ways, is too short for useful comparison. We apply the same smoothing to their results as to our fluxes and the growth rates in Figure 7. Again we see a good fit to most of the major events in the record. This is encouraging, especially since *Rödenbeck et al.* [2003a] use a different transport model and, most importantly, use time-varying meteorology to drive it. The good match suggests that some of the weaknesses in our study do not fatally compromise the results. Curiously, the departures occur at the same time as the departures from the observed growth rate curve in Figure 7. This supports the idea that these departures really are failures in the terrestrial model and not, for example, some problem with the background fluxes. In summary, the optimized model does a reasonable job of fitting both observed concentration variations and fluxes inferred from these concentrations by relatively independent methods. The match is better on interannual than seasonal timescales, suggesting we can make clearer inferences at longer timescales.

4.2. Optimized Parameters

[57] Table 2 shows prior and predicted values and uncertainties for the parameters. We concentrate on parameters showing either large uncertainty reduction or large shifts from their prior values. Recall from section 3.2 that there is a division between globally uniform and spatially explicit parameters in our set-up. In general, we expect greater possibilities for uncertainty reduction for global parameters since these are observed by a larger data set. We see first that the spatially explicit parameters pertaining to photosynthesis show very little reduction in uncertainty. There are two reasons for this. The first is the size of the prior uncertainty on these parameters. Generally, the smaller the prior uncertainty the more difficult it is for the data to add any new information about this parameter. The second is the indirectness of observing photosynthesis itself. Recall that the quantity directly related to atmospheric concentration is the net flux in which the photosynthesis appears only as part of a sum. We can contrast this with the β parameters. Equation (22) shows that the observed NEP is related to β by multiplication with the NPP, making it easier to observe.
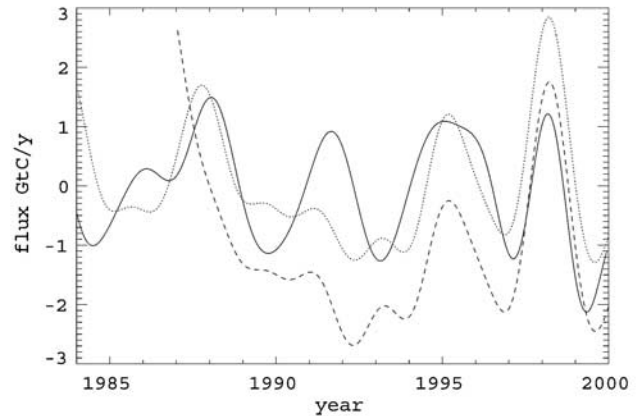

Figure 8. Comparison of our optimized terrestrial flux (solid line) with inversions using 16 stations (dotted line) and 19 stations (dashed line) of *Rödenbeck et al.* [2003a]. See color version of this figure at back of this issue.

[58] Two parameters controlling autotrophic respiration, $f_{R,leaf}$ and $f_{R,growth}$, show substantial shifts from their prior values, at least relative to their prior uncertainties. In both cases, they are associated with tight prior estimates (2.5% and 0.5%, respectively) and only slight reductions in their uncertainty. We see from equation (15) that $f_{R,leaf}$ appears as the denominator in transforming dark respiration into maintenance respiration. The large fractional change in its value hence nearly doubles maintenance respiration and so greatly changes the relationship between GPP and NPP. The change in $f_{R,growth}$ is, fractionally, much smaller and only appears significant because of the small prior uncertainty applied to this number.

[59] The model of soil respiration in BETHY has two pools and respiration depending on pool sizes, temperature, and moisture. The temperature and moisture dependence are controlled by global parameters while the mean value of respiration from the slow pool is controlled by the β parameters which are spatially explicit. The optimization adds substantial information for many of these parameters. The τ_f, κ, and $Q_{10,s}$ all show large reductions in uncertainty. $Q_{10,s}$ and τ_f show relatively small shifts from their prior values suggesting these values are compatible with atmospheric concentration observations. The κ is greatly reduced, weakening the dependence of soil respiration on soil moisture, except when soil moisture is very low. This contrasts somewhat with both K02 and *Saleska et al.* [2003].

[60] The β parameter, as we noted in section 3.1, controls the storage efficiency of a given ecosystem. It, along with the NPP, defines the net carbon flux (equation (22)). The optimization will adjust the net carbon flux to match the long-term mean spatial gradients and growth rates of concentration. These are both strong constraints, so it is not surprising that unlike the spatially explicit photosynthesis parameters, the β parameters for many PFTs are well constrained. In particular, most PFTs with large NPP show large reductions in uncertainty for β. This is gratifying but predictable. The NEP, the quantity actually observed by the

Table 4. Various Global Fluxes From the Prior and Optimized Model

Name[a]	Value, Gt C/yr			
	1980–2000 (Prior)	1980–2000	1980–1990	1990–2000
GPP	135.70	134.80	134.30	135.30
Growth respiration	23.50	22.35	22.31	22.39
Maintenance respiration	44.04	72.70	72.13	73.28
NPP	68.18	40.55	40.63	40.46
Fast respiration	53.83	27.40	27.60	27.21
Slow respiration	14.46	10.69	10.71	10.67
NEP	−0.11	2.453	2.318	2.587

[a]GPP, gross primary productivity; NPP, net primary productivity; NEP, net ecosystem productivity.

atmosphere, contains the product of β and NPP. Hence large NPP leads to a larger sensitivity of the cost function to the β parameter and hence better observability of the parameter. We notice some very large shifts from the prior (balanced) value. The form of equation (22) means that variations below 1 are more dramatic than variations above. For example, a value of 2 means the ecosystem is storing half its NPP, while a value of $\frac{1}{2}$ means the ecosystem is losing carbon at a rate equal to its NPP; that is, heterotrophic respiration is twice NPP. Departures to both these extremes are unlikely and cast some doubt on some of the large deviations seen in Table 2. The general form is interesting, however. The large values of β(TrEv) and β(C4Gr) suggest substantial uptake in the tropics, although this is partly balanced by the small value of β(TrDec). We also see high values for β(TmpDec) and β(Crop), characteristic of the large midlatitude uptakes commonly seen in atmospheric inversions. The β(Crop) is extraordinarily large. We should note, though, that β(Crop) exhibits the substantial shift and

small reduction in uncertainty about which we have already warned.

4.3. Derived Fluxes

4.3.1. Global Fluxes

[61] Table 4 lists various components of the biospheric flux for the prior and optimized model. It also splits the optimized fluxes by decade for analysis in the next section. There is little difference between the GPP for the prior and optimized parameter set, reflecting the small shifts in most of the photosynthesis parameters. For the maintenance respiration the shift is dramatic, following directly from the change in $f_{R,leaf}$ seen in Table 2. This change drives a large reduction in NPP. The reduction in NPP, in turn, implies reduced input to the fast soil pool which reduces the size of this pool and consequently the fast respiration. This must be the cause of this large reduction since the parameters which control the timescales of fast respiration do not change much. It is tempting to go further and suggest that the change in slow respiration follows from the same logic, but the simplification we make, using the β parameter, removes the slow pool size from the governing equation for slow respiration altogether. Instead, the slow respiration is reduced by the generally positive shift in the β parameters coupled with the reduction in NPP (see equation (22)). Finally, the NEP, the quantity actually seen by the atmosphere, moves from near zero (forced by the initial value of β) to the value required to match the atmospheric growth rate. The NEP and NPP imply a global β value of 1.06. Roughly, then, terrestrial ecosystems are storing 6% of their NPP long-term.

4.3.2. Spatial Distributions of Flux

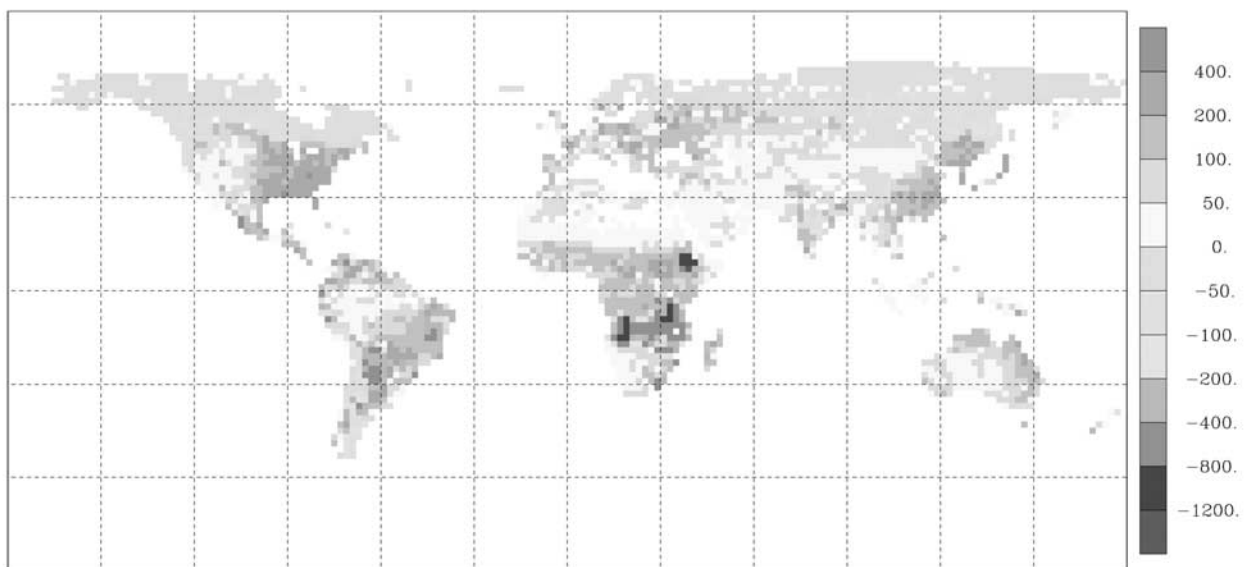[62] Figures 9 and 10 show the long-term mean net flux to the atmosphere and its uncertainty as computed from

Figure 9. Mean net flux to the atmosphere for the period 1980–2000 (gC m^{-2} yr^{-1}). See color version of this figure at back of this issue.
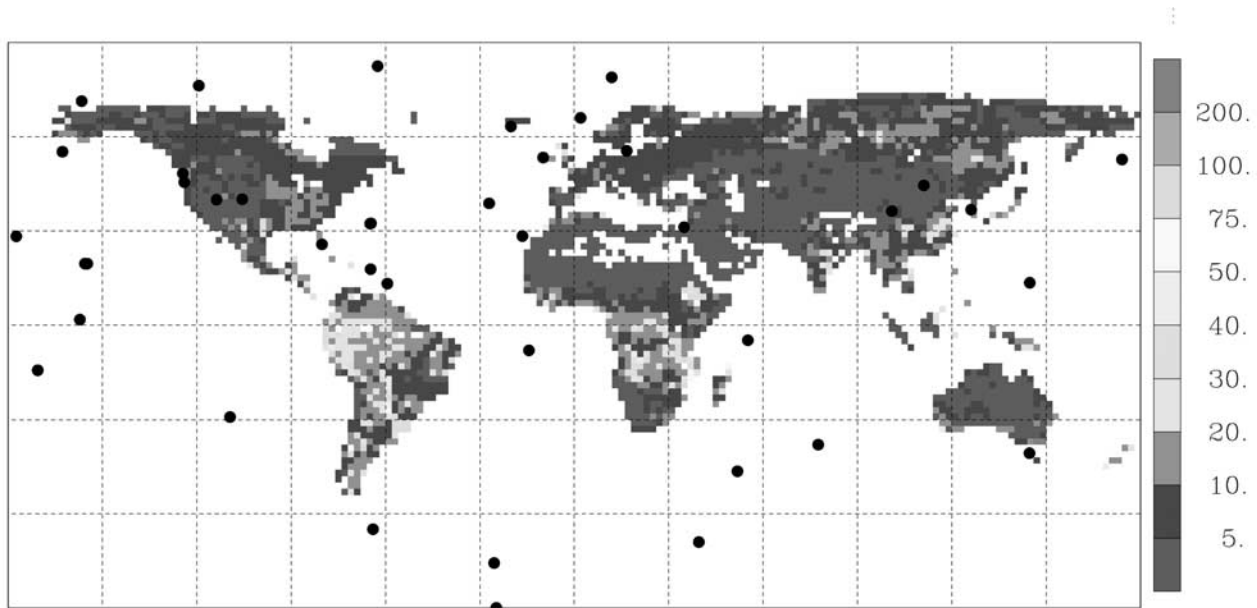
Figure 10. Uncertainty of mean net flux to the atmosphere for the period 1980–2000 (gC m^{-2} yr^{-1}) expressed as the standard deviation. Dots indicate locations of observational sites. See color version of this figure at back of this issue.

the optimized model. As in K02, the flux uncertainty is calculated by propagating the covariance in the uncertainty of derived parameters through the Jacobian of fluxes with respect to parameters calculated at the optimum.

[63] We can consider these fluxes from two viewpoints, either as an estimate of the fluxes which is guaranteed consistent with the behavior of a credible terrestrial model or in terms of the processes which give rise to the fluxes. Qualitatively, the picture of the net flux in Figure 9 is similar to that from direct flux inversions [e.g., *Tans et al.*, 1990; *Rayner et al.*, 1999; *Gurney et al.*, 2002]. We see relatively large uptake over the Northern Hemisphere continents, and uptake over the tropical continents, which partly balances the large background source from land-use change. The first result is an extremely common outcome in direct flux inversion studies. While one would expect a smaller sink if the optimization had the freedom to reduce the northern ocean sink, we note that the total continental sink north of 30°N for 1992–1996 is 1.65 GtC yr^{-1}, compared with the 2.31 ± 0.61 GtC yr^{-1} for northern land regions found by *Gurney et al.* [2002] for the comparable time period. The model distributes the continental uptake roughly equally between Eurasia and North America, implying more intense sinks in the smaller North American region. This is also clear from Figure 9. For the region 30°S–30°N the optimized model generates an NEP of 0.64 GtC yr^{-1}. When subtracted from the specified land-use flux of 1.6 GtC yr^{-1}, this represents a net source of 0.96 GtC yr^{-1} to the atmosphere. This is, again, close to the 1.14 GtC yr^{-1} for the *Gurney et al.* [2002] study, and, once again, the model estimate for 1992–1996 is little different from the long-term mean.

[64] Equation (22) shows that the net flux is determined by a combination of the productivity and the storage

efficiency. We can hence understand Figure 9 by referring to the relevant entries in Table 2. The large uptakes in northern midlatitudes are caused by the large values of β(TmpDec) and β(Crop). The tropical uptakes are associated with the large values of β(TrEv) and β(C4Gr). The occasional sources seen in the tropics are associated with the dominance of the tropical deciduous PFT which has a small β value. The value implies that $R_S > \overline{NPP}$, which seems unlikely and hints at some missing process in the model. The sources in high northern latitudes are associated with the two coniferous PFTs, each of which have small β values.

[65] Comparison of uncertainties generated from the optimized model and those from a direct flux inversion is difficult because they have quite different spatial characteristics. In particular, the behavior of the two classes of inversion is different as we aggregate estimates to larger and larger scales. While a direct flux inversion returns an estimate and uncertainty for each region, the parameter inversion generates parameter estimates with global scale or with influence scattered across the globe. In direct flux inversions, the limited resolving power of the current observing network often gives rise to negative correlations among neighboring regions so that in flux density terms at least, uncertainties decrease as we integrate to successively larger regions. For the parameter inversion, many neighboring points in space may depend on a single parameter in a similar way, giving rise to positive uncertainty correlations among neighboring points. However, the relationship is reversed at small scales. Within any of the large regions traditionally used in direct flux inversions, all fluxes are completely correlated since only one magnitude is used to scale the whole pattern. The different sensitivities of different points to multiple parameters can break down this complete correlation in the parameter inversion, meaning

that not even neighboring points have perfectly correlated uncertainties.

[66] However we consider them, the pointwise, formal uncertainties shown in Figure 10 are low compared to equivalent uncertainties from a direct flux inversion such as that of *Gurney et al.* [2002]. For example, the average uncertainty (expressed as a standard deviation) over tropical America is 16 gC m^{-2} yr^{-1} while the average of the models of *Gurney et al.* [2002] is equivalent to 100 gC m^{-2} yr^{-1}. This is true despite the use of a tighter prior estimate and larger observing network in the *Gurney et al.* [2002] study (76 stations compared to 41 here) and the estimate of a smaller number of parameters in that study (22 source magnitudes compared to our 57 parameters). This is true even when we account for the underprediction of uncertainty mentioned in section 4.1. There are several reasons for this lower uncertainty of predicted fluxes. First, some parameters have quite low prior uncertainty, as mentioned in section 3.2. Another reason is the number of observations in our study, 6936 compared to the 76 of *Gurney et al.* [2002]. This does not really explain the performance in tropical America which is unobserved in either network. The more fundamental reason is the ability of the terrestrial model to propagate information gained from one location to wide areas, a result of having single parameters apply to entire PFTs or the entire globe. Whether this propagation is reasonable depends on whether our regional distribution of parameters provides a good description of the terrestrial biosphere. Attractive as they may seem, the low formal uncertainties of the assimilation approach are not an especially strong reason for preferring it. First, they may be obtained at the cost of higher biases in the estimates since we do not, for example, match the atmospheric concentration records as well as a flux inversion. Second, they are contingent on the spatial detail in the parameter description. The ability, however, to comment on which processes can be constrained and to what degree is an important capability.

[67] Consideration of the area-integrated uncertainty gives us one check on the approximations we made in calculating the parameter covariance. The atmospheric growth rate over the 2 decades of the study is known quite precisely. With the statistical model we use for the concentrations, the uncertainty on the average growth rate is about 0.007 ppmv. We assume zero uncertainty for the other contributing fluxes, and so the uncertainty in the globally integrated long-term mean terrestrial flux should also be very small. Our uncertainty propagation yields 0.006 ppmv. This near cancellation of larger pointwise errors is caused by an interplay of flux sensitivities, uncertainties, and uncertainty correlation. The close agreement suggests that the chain of calculations in the uncertainty analysis is behaving well.

[68] Analysis of the sensitivities embodied in the flux Jacobian can yield great insight into the contributions to the uncertainty in fluxes and even suggest methods to reduce it. One remarkable feature of Figure 10 is some notable hot spots in Japan and southeastern Australia. The approximate variance of the flux at a point can be written as

$$\mathbf{C}(\vec{F})_{i,i} = \frac{\partial F_i}{\partial \vec{P}} \ \mathbf{C}(\vec{P}) \ \frac{\partial F_i}{\partial \vec{P}}^T \ = \sum_{j,k} \mathbf{C}(\vec{P})_{j,k} \frac{\partial F_i}{\partial P_j} \frac{\partial F_i}{\partial P_k}. \quad (25)$$

We can hence quantify the contribution of each term in the parameter covariance and the sensitivity to the overall uncertainty of the long-term net flux. For the southeast Australian hot spot, we find this comes almost entirely from the variance for β(TmpEv). This PFT does, indeed, dominate this grid cell. The β(TmpEv) has posterior uncertainty almost unchanged from the initial value of 0.25. Differentiating equation (22) we see that a small change in net flux is given by

$$\delta\overline{\text{NEP}} = \delta\overline{\text{NPP}}\left(1 - \frac{1}{\beta}\right) - \frac{\delta\beta}{\beta^2}\overline{\text{NPP}}. \quad (26)$$

We can use equation (26) to propagate uncertainties in $\overline{\text{NPP}}$ and β to uncertainties in $\overline{\text{NEP}}$. This is a productive grid cell with $\overline{\text{NPP}}$ = 978 gC m^{-2} yr^{-1}. Neglecting any uncertainty on $\overline{\text{NPP}}$ (a good approximation) Equation (26) yields an uncertainty in $\overline{\text{NEP}}$ of 245 gC m^{-2} yr^{-1}, close to the 216 gC m^{-2} yr^{-1} produced by the numerical calculation. The sensitivity analysis has revealed what processes are responsible for the uncertainty in the long-term net flux estimates. In this case it is the storage efficiency of the temperate evergreen PFT. This suggests that this study would benefit greatly by the introduction of data such as soil carbon measurements or direct NEP measurements for this PFT. The other uncertainty hot spot over Japan is dominated by the same PFT. A consequence of using these measurements taken in one of these hot spots would be a reduction in uncertainty over both of them. This reflects the particular regionalization of parameters we have chosen. The above calculation points out the utility of having both the sensitivity and uncertainty results available. Future studies using these techniques will introduce existing data sets of point measurements into these calculations as has been demonstrated by K02.

4.3.3. Decadal Differences

[69] One of the more striking inferences drawn from measurements of atmospheric CO_2 concentration is the stability of its growth rate on decadal timescales. This occurs despite the large variability on shorter timescales evident in Figure 7 and despite a large change in some terms in the atmospheric carbon balance. In our calculation, the decadal mean input from fossil fuels rises from 5.50 GtC yr^{-1} in the 1980s to 6.32 GtC yr^{-1} in the 1990s. The decadal mean ocean uptake, as reported by *Le Quéré et al.* [2003] and used in this study, increases by only 0.3 GtC yr^{-1}. Yet the atmospheric growth rate reported by *Prentice et al.* [2001] is 3.3 ± 0.1 GtC yr^{-1} for the 1980s and 3.2 ± 0.1 GtC yr^{-1} for the 1990s. In our calculation, there is no change in the flux from land-use change between the 1980s and 1990s. Our calculation should hence show an increase of 0.6 GtC yr^{-1} in the decadal average NEP from the 1980s to the 1990s. It shows an increase of only 0.3 GtC yr^{-1} in the optimized model (Table 4). One clear reason for this is the treatment of atmospheric CO_2 concentrations. We held this fixed at a mean value for the 2 decades in both the assimilation of the AVHRR and the concentration data. There is no possibility, then, of a CO_2 fertilization effect. This was the effect reported by *Cao et al.* [2002], who noted a large increase in NEP resulting from an increase in NPP (arising from increases in CO_2 concentration) and smaller increase
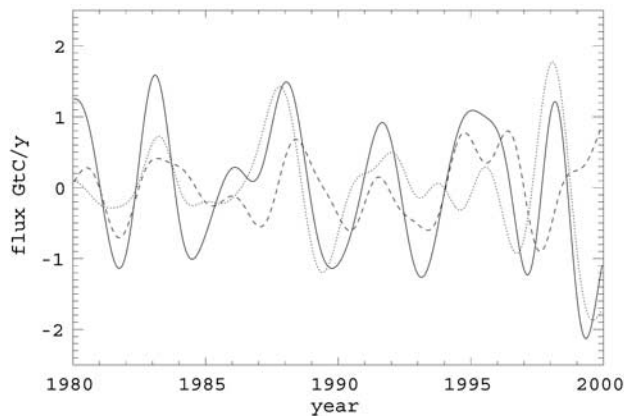
Figure 11. Global (solid line), tropical ($20°S−20°N$) (dotted line), and northern extratropical ($20°N−90°N$) (dashed line) anomalies in flux to the atmosphere from the optimized model. See color version of this figure at back of this issue.

in heterotrophic respiration forced by increased temperature. Interestingly, our smaller increase in NEP arises from the opposite combination, a slight decrease in NPP but a larger decrease in heterotrophic respiration. The decrease in heterotrophic respiration occurs despite an increase in temperature of $0.3°C$ between the 1980s and 1990s in the driving data for the model. Average soil moisture decreases by 0.5% between the 2 decades, which would explain a slight reduction although it is so seasonally variable that long-term average changes are not necessarily a good guide.

[70] We can make an estimate of the impact of CO_2 fertilization in BETHY using a previous calculation of W. Knorr (personal communication). They estimated a sensitivity of GPP to CO_2 concentration of 0.24 GtC yr^{-1}/ppmv. Using our proxy for the global concentration, we obtain a mean concentration difference of 14.7 ppmv between the 2 decades, yielding an increase of 3.53 GtC yr^{-1} in GPP. We can calculate the ratio of NPP to GPP from Table 4 and estimate an increase in NPP of 1.06 GtC yr^{-1}. We then have two options for calculating the impact on NEP, corresponding to a lower and upper bound. As a lower bound, we can use the effective β of 1.06 (section 4.3.1). This reflects an assumption that the disequilibrium in the slow pool from our optimization is the correct value for the impact of CO_2 fertilization. Using equation (22), this yields an increase in NEP of 0.06 GtC yr^{-1}. The upper bound is to assume that none of the carbon added to the slow pool by CO_2 fertilization is returned to the atmosphere. Thus we can use the allocation rate from the fast to the slow pool (f_S) from Table 2. This yields a decadal increase in NEP of 0.34 GtC yr^{-1}. This is smaller than the increases noted by either *Cao et al.* [2002] or *Potter et al.* [2003]. Such an increase would, however, nearly correct the prediction of decadal changes in atmospheric growth rate from our model. An obvious future direction for this work is to repeat the optimization in the presence of CO_2 fertilization.

4.4. Interannual Variability

[71] The spatial and temporal distributions of the components of NEP from the optimized model allow us to ask what regions and fluxes are responsible for driving global interannual variability. Figure 11 shows the interannual variability in global, tropical, and northern extratropical flux to the atmosphere. Tropical variability is generally larger than extratropical and explains more of the interannual variability in the global flux. The northern extratropics play the dominant role in both the negative flux anomaly after the eruption of Mount Pinatubo in 1991 and the positive flux anomalies around 1995. In contrast, the large positive anomaly in 1997–1998 and negative anomaly in 1999–2000 occur in the tropics with the northern extratropics slightly counteracting them. Recall from Figures 7 and 8 that we believe the model is underpredicting the strength of both anomalies. Unfortunately, the tropical concentration records for both events are too sparse to locate the error in the tropics or extratropics, but the model does agree with the conclusion of *Langenfelds* [2002] who located the 1997–1998 anomaly in the tropics.

[72] Figure 12 decomposes the NEP by process rather than by region. Note that we have reversed the signs of several of the fluxes so that anomalies to the atmosphere are positive for all terms. We have neglected the slow respiration for clarity; it is nearly constant throughout. It is clear from the figure that unlike the decadal differences, the interannual variations in NPP drive those of NEP. To quantify this, we calculate the linear regression of the NEP against its components. The coefficients are of no interest since we know the relationship among the various flux components, but we see a correlation coefficient of 0.88 between NPP and NEP and only $−0.33$ between fast respiration and NEP. The picture, generally, is of flux anomalies driven by production anomalies slightly modulated by changes in respiration. For example, the high growth events of 1987–1988 and 1991–1992 show a production anomaly (in fact negative, but shown here as a
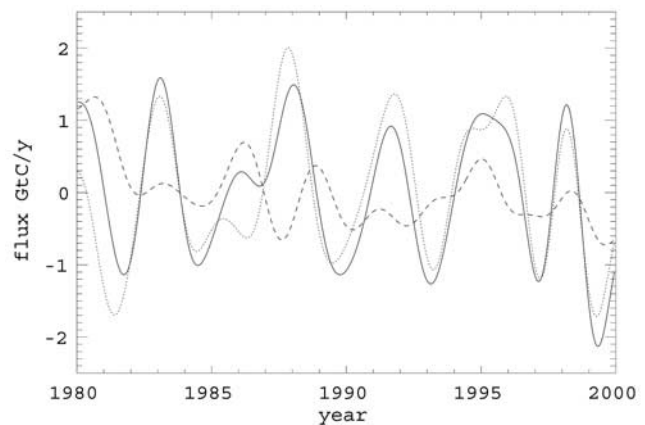

Figure 12. Global anomalies in negative NEP (solid line), negative NPP (dotted line), and fast respiration (dashed line) on interannual timescales. See text for details of time filtering. See color version of this figure at back of this issue.

peak in flux to the atmosphere) offset by a decrease in respiration, while the low growth rate events of 1993 and 1999, as well as the high growth rate event of 1998, show production anomalies exacerbated by respiration anomalies. There are mechanisms for the model to exhibit either of these behaviors. For example, an increase in production will lead to an increase in litterfall and hence the size of the fast soil pool. Without any changes in the driving data, this will lead to an increase in soil respiration offsetting the production anomaly. With an approximate 1-year turnover time and a smoothing time for the plots of about 1 year this will appear as a synchronous response. In contrast, high temperatures and low soil moisture often occur together. High temperature increases respiration (equation (19)) while low soil moisture decreases production. Equation (19) suggests that low soil moisture will reduce respiration, but the low optimized value of κ in Table 2 almost removes this effect. The net impact of these soil moisture and temperature events is of reinforcing anomalies in production and respiration. The global integral in Figure 12 also hides a mass of probably competing processes occurring in different regions. Separating these effects, especially in the responses to major climate anomalies such as ENSO, is the topic of a future paper using this optimized model.

[73] Finally, we can compare the interannual variability from our optimized model with a recent study of *Potter et al.* [2003]. Both curves, using the atmospheric sign convention and the same smoothing as for Figure 12, are shown in Figure 13. We also show our prior model for comparison. The agreement is good through the middle of the record (1986–1992) but relatively poor outside this period. The positive anomalies in 1994–1995 and 1997–1998, evident both in our model and in the atmospheric growth rate, are not seen in the *Potter et al.* [2003] record. Recall from Figure 7 that we underestimate the strength of the latter event. We also produce a stronger minimum around 1993. As a generalization, interannual variability in the *Potter et al.* [2003] record is smaller than our optimized model which is, in turn, smaller than our prior model. It is noteworthy that optimization against atmospheric data reduces the interannual variability in our model while it would, presumably, increase that in the model of [*Potter et al.*, 2003] to bring its predicted concentrations closer to observations. We can also see, by the comparison of the prior to the optimized fluxes, that the good fit we obtain to the interannual variability comes from the formulation of the model itself, with the optimization tuning some of the magnitudes. The ability of the BETHY model to match interannual variability was previously noted by *Knorr* [2000].

5. Discussion

[74] Although we have discussed most of the major results as we showed them, there are a few general points we should make. The first is that our validation of the model is rather superficial. Although we have compared its behavior with a number of other calculations, we have not compared with many data sources completely independent of the data we used to optimize the model. The model produces a rich suite of observables such as diurnal cycles
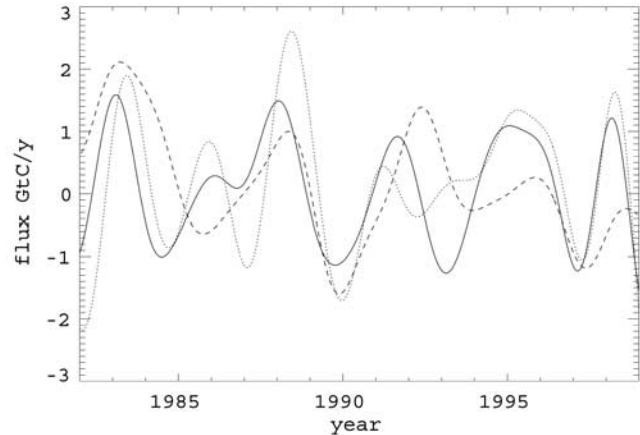

Figure 13. Global negative NEP anomalies for our prior (dotted line) and optimized (solid line) models and that of *Potter et al.* [2003] (dashed line). See color version of this figure at back of this issue.

of flux. We can compare our model predictions with observations for such quantities only if we can match the conditions at the observing sites. These will inevitably be different from the large-scale averages we use to force the global model. However, there is no reason, in principle, why we cannot run the model with the required forcing data at a point but with relevant optimized parameters taken from the global model. We can also optimize a pointwise version of the model across the range of PFTs used in the global study and then cross validate against the concentration data used in this study. We are pursuing both these approaches.

[75] We can also contrast the results of this study with those of K02. They noticed a general tendency to increase NPP in order to increase the seasonal cycle of flux (and concentration) at high latitudes. We see the opposite effect here with NPP being greatly reduced. Our final estimate is about 20% lower than K02 obtained fitting only the seasonal cycle. Of course that was a much simpler model. They also saw relatively weak constraints on the Q_{10} values. They speculated that there might be more information about Q_{10} available from multiannual records. This does seem to be the case with Q_{10} at least for the slow pool, being well constrained, even when we allow the increase to account for our high χ^2 value. Again we are fitting two global Q_{10} values in this study while K02 were fitting 12 spatially explicit parameters. Finally, we note that the values we obtain for Q_{10} are quite close to those from K02 and generally lower than most other estimates. Again this difference is an artifact of model formulation; we use atmospheric temperature as the input to the respiration parameterization rather than the more usual soil temperature. *Jones and Cox* [2001], using ad hoc tuning of Q_{10} in a coupled climate biosphere model, also obtained a good constraint for Q_{10} around the more conventional value of 2.0 (but again using soil temperature).

6. Conclusions

[76] We have presented the carbon cycle data assimilation system (CCDAS) and have employed CCDAS to generate a

record of terrestrial carbon fluxes for the 1980s and 1990s by optimizing the parameters of a biosphere model to obtain the best match to records of atmospheric CO_2 concentration. The obtained fit is moderately good for seasonal cycles and very good for interannual variability. Some parameters are shifted greatly from their prior values and some are well constrained by the atmospheric observations although, most notably, there is little constraint on parameters controlling photosynthesis. The long-term mean net flux shows the usual pattern of substantial uptake over the Northern Hemisphere continents, in this case divided evenly between North America and Eurasia. There is a moderate uptake over the tropics, partly balancing the imposed flux due to land-use change. Uncertainties of net fluxes are much lower than for a comparable synthesis inversion. Interannual variability is dominated by the tropics and by changes in NPP, with respiration sometimes reinforcing and sometimes offsetting NPP anomalies. Respiration is more important on decadal timescales, although this may reflect the neglect of CO_2 fertilization in the experiment.

[77] This paper documents a rapidly evolving modeling system in which the assimilation capacity we describe is closely coupled with the development of the BETHY model. New parameterizations and inputs for BETHY are now routinely tested in assimilation mode. We are also removing some of the constraints of this study, for example, the assumption of a perfectly known ocean flux or small prior uncertainties on some parameters. We are adding important missing processes to the terrestrial model, particularly fire. Finally, we will test the sensitivity to the transport model by replacing TM2 with the TM3 model used by *Rödenbeck et al.* [2003a].

References

Andres, R. J., G. Marland, I. Fung, and E. Matthews (1996), A $1° \times 1°$ distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990, *Global Biogeochem. Cycles*, *10*, 419–429.

Barrett, D. J. (2002), Steady state turnover time of carbon in the Australian terrestrial biosphere, *Global Biogeochem. Cycles*, *16*(4), 1108, doi:10.1029/2002GB001860.

Bousquet, P., P. Peylin, P. Ciais, C. L. Quéré, P. Friedlingstein, and P. P. Tans (2000), Regional changes in carbon dioxide fluxes of land and oceans since 1980, *Science*, *290*, 1342–1346.

Cao, M., S. D. Prince, and H. H. Shugart (2002), Increasing terrestrial carbon uptake from the 1980s to the 1990s with changes in climate and atmospheric CO_2, *Global Biogeochem. Cycles*, *16*(4), 1069, doi:10.1029/2001GB001553.

Collatz, G. J., M. Ribas-Carbo, and J. A. Berry (1992), Coupled photosynthesis-stomatal conductance model of leaves for C_4 plants, *J. Aust. Plant Physiol.*, *19*, 519–538.

Conway, T. J., P. P. Tans, L. S. Waterman, K. W. Thoning, D. R. Kitzis, K. A. Masarie, and N. Zhang (1994), Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, *J. Geophys. Res.*, *99*, 22,831–22,855.

Cox, P. M., R. A. Betts, C. D. Jones, S. A. Spall, and I. J. Totterdell (2000), Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, *Nature*, *408*, 184–187.

Cox, P. M., R. A. Betts, M. Collins, P. P. Harris, C. Huntingford, and C. D. Jones (2004), Amazonian forest diaback under climate-carbon projections for the 21st century, *Theor. Appl. Climatol.*, *78*, 137–156.

Dargaville, R. J., R. M. Law, and F. Pribac (2000), Implications of interannual variability in atmospheric circulation on modeled CO_2 concentrations and source estimates, *Global Biogeochem. Cycles*, *14*, 931–943.

Dargaville, R. J., et al. (2002), Evaluation of terrestrial carbon cycle models with atmospheric CO_2 measurements: Results from transient simulations considering increasing CO_2, climate, and land-use effects, *Global Biogeochem. Cycles*, *16*(4), 1092, doi:10.1029/2001GB001426.

Dufresne, J.-L., P. Friedlingstein, M. Berthelot, L. Bopp, P. Ciais, L. Fairhead, H. Le Treut, and P. Monfray (2002), On the magnitude of positive feedback between future climate change and the carbon cycle, *Geophys. Res. Lett.*, *29*(10), 1405, doi:10.1029/2001GL013777.

Enting, I. G. (2002), *Inverse Problems in Atmospheric Constituent Transport*, Cambridge Univ. Press, New York.

Fan, S., J. L. Sarmiento, M. Gloor, and S. W. Pacala (1999), On the use of regularization techniques in the inverse modelling of atmospheric carbon dioxide, *J. Geophys. Res.*, *104*, 21,503–21,512.

Farquhar, G. D. (1988), Models relating subcellular effects of temperature to whole plant response, *Symp. Soc. Exper. Biol.*, *42*, 395–409.

Farquhar, G. D., S. V. Cammerer, and J. A. Berry (1980), A biochemical model of photosynthesis in leaves of C_3 species, *Planta*, *149*, 78–90.

Feely, R. A., R. Wanninkhof, T. Takahashi, and P. Tans (1999), Influence of El Niño on the equatorial Pacific contribution to atmospheric CO_2 accumulation, *Nature*, *398*, 597–601.

Friedlingstein, P., J.-L. Dufresne, P. M. Cox, and P. Rayner (2003), How positive is the feedback between climate change and the carbon cycle?, *Tellus, Ser. B*, *55*, 692–700, doi:10.1034/j.1600-0560.2003.01461.x.

Giering, R., and T. Kaminski (1998), Recipes for adjoint code construction, *ACM Trans. Math. Software*, *24*, 437–474.

Giering, R., T. Kaminski, and T. Slawig (2005), Applying TAF to a Navier-Stokes solver that simulates an Euler flow around an airfoil, *Future Gen. Comput. Syst.*, in press.

GLOBALVIEW-CO_2 (2001), *Cooperative Atmospheric Data Integration Project - Carbon Dioxide* [CD-ROM], NOAA CMDL, Boulder, Colo. (Available via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/ GLOBALVIEW)

Griewank, A. (2000), *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, Frontiers Appl. Math.*, vol. 19, Soc. of Ind. and Appl. Math., Philadelphia, Pa.

Gurney, K. R., et al. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, *415*, 626–630.

Gurney, K. R., et al. (2003), TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus, Ser. B*, *55*, 555–579, doi:10.1034/j.1600-0560.2003.00049.x.

Gurney, K. R., et al. (2004), Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks, *Global Biogeochem. Cycles*, *18*, GB1010, doi:10.1029/2003GB002111.

Gutman, G., D. Tarpley, A. Ignatov, and S. Olson (1995), The enhanced NOAA global land dataset from the Advanced Very High-Resolution Radiometer, *Bull. Am. Meteorol. Soc.*, *76*, 1141–1156.

Heimann, M. (1995), The global atmospheric tracer model TM2, *Tech. Rep. 10*, Dtsch. Klimarechenzentrum, Hamburg, Germany.

Houghton, R. A. (2003), Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850–2000, *Tellus, Ser. B*, *55*, 378–390.

Houghton, R. A., et al. (1987), The flux of carbon from terrestrial ecosystems to the atmosphere in 1980 due to changes in land use: Geographic distribution of the global flux, *Tellus*, *39*, 122–139.

Hungate, B. A., E. A. Holland, R. B. Jackson, F. S. Chapin III, H. A. Mooney, and C. B. Field (1997), The fate of carbon in grasslands under carbon dioxide enrichment, *Nature*, *388*, 577–579.

Jones, C. D., and P. M. Cox (2001), Constraints on the temperature sensitivity of global soil respiration from the observed interannual variability in atmospheric CO_2, *Atmos. Sci. Lett.*, *2*, 166–172.

Kaminski, T., and M. Heimann (2001), Inverse modeling of atmospheric carbon dioxide fluxes, *Science*, *294*, 5541.

Kaminski, T., M. Heimann, and R. Giering (1999a), A coarse grid three-dimensional global inverse model of the atmospheric transport: 1. Adjoint model and Jacobian matrix, *J. Geophys. Res.*, *104*, 18,535–18,553.

Kaminski, T., M. Heimann, and R. Giering (1999b), A coarse grid three-dimensional global inverse model of the atmospheric transport: 2. Inversion of the transport of CO_2 in the 1980s, *J. Geophys. Res.*, *104*, 18,555–18,581.

Kaminski, T., P. J. Rayner, M. Heimann, and I. G. Enting (2001), On aggregation errors in atmospheric transport inversions, *J. Geophys. Res.*, *106*, 4703–4715.

Kaminski, T., W. Knorr, P. Rayner, and M. Heimann (2002), Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, *Global Biogeochem. Cycles*, *16*(4), 1066, doi:10.1029/2001GB001463.

Kaminski, T., R. Giering, M. Scholze, P. Rayner, and W. Knorr (2003), An example of an automatic differentiation-based modelling system, in *Computational Science–ICCSA 2003, International Conference Montreal, Canada, May 2003, Proceedings, Part II, Lecture Notes Comput. Sci.*, vol. 2668, edited by V. Kumar et al., pp. 95–104, Springer, New York.

Knorr, W. (1997), Satellitengestützte Fernerkundung und Modellierung des globalen CO₂-Austauschs der Landvegetation: Eine Synthese, Ph.D. thesis, Max-Planck-Inst. für Meteorol., Hamburg, Germany.

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

Knorr, W., and M. Heimann (1995), Impact of drought stress and other factors on seasonal land biosphere CO₂ exchange studied through an atmospheric tracer transport model, *Tellus, Ser. B*, *47*, 471–489.

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

Knorr, W., and M. Heimann (2001b), Uncertainties in global terrestrial biosphere modeling: 2. Global constraints for a process-based vegetation model, *Global Biogeochem. Cycles*, *15*, 227–246.

Knorr, W., and J.-P. Schulz (2001), Using satellite data assimilation to infer global soil moisture and vegetation feedback to climate, in *Remote Sensing and Climate Modeling: Synergies and Limitations*, edited by M. Beniston and M. Verstraete, pp. 273–306, Springer, New York.

Körner, C. (2000), Biosphere responses to CO₂ enrichment, *Ecol. Appl.*, *10*, 1590–1619.

Langenfelds, R. L. (2002), Studies of the global carbon cycle using atmospheric oxygen and associated tracers, Ph.D. thesis, Univ. of Tasmania, Hobart, Tasmania, Australia.

Langenfelds, R. L., R. J. Francey, B. C. Pak, L. P. Steele, J. Lloyd, C. M. Trudinger, and C. E. Allison (2002), Interannual growth rate variations of atmospheric CO₂ and its isotope δ¹³C, H₂, CH₄ and CO between 1992 and 1999 linked to biomass burning, *Global Biogeochem. Cycles*, *16*(3), 1048, doi:10.1029/2001GB001466.

Law, R. M. (1999), CO₂ sources from a mass-balance inversion: Sensitivity to the surface constraint, *Tellus, Ser. B*, *51*, 254–265.

Law, R. M., et al. (1996), Variations in modelled atmospheric transport of carbon dioxide and the consequences for CO₂ inversions, *Global Biogeochem. Cycles*, *10*, 783–796.

Le Quéré, C., et al. (2003), Two decades of ocean CO₂ sink and variability, *Tellus, Ser. B*, *55*, 649–656, doi:10.1034/j.1600-0560.2003.00043.x.

Louis, J. F. (1979), A parametric model of vertical eddy fluxes in the atmosphere, *Boundary Layer Meteorol.*, *17*, 187–202.

Marland, G., T. A. Boden, and R. J. Andres (2001), Global, regional, and national CO₂ emissions, in *Trends: A Compendium of Data on Global Change*, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., U.S. Dep. of Energy, Oak Ridge, Tenn.

McGuire, A. D., et al. (2001), Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO₂, climate and land use effects with four process-based ecosystem models, *Global Biogeochem. Cycles*, *15*, 183–206.

Nemry, B., L. François, J. C. Gêrard, M. Heimann, and the participants of the Potsdam NPP Model Intercomparison (1999), Comparing global models of terrestrial net primary productivity (NPP): Analysis of the seasonal atmospheric CO₂ signal, *Global Change Biol.*, *5*, 65–76.

Nijssen, B., R. Schnur, and D. Lettenmaier (2001), Retrospective estimation of soil moisture using the VIC land surface model, 1980–1993, *J. Clim.*, *14*, 1790–1808.

Page, S. E., F. Siegert, J. O. Rieley, H. D. Boehm, A. Jaya, and S. Limin (2002), The amount of carbon released from peat and forest fires in Indonesia during 1997, *Nature*, *420*, 61–65.

Pinty, B., and M. M. Verstraete (1992), GEMI—A nonlinear index to monitor global vegetation from satellites, *Vegetatio*, *101*, 15–20.

Potter, C., S. Klooster, R. Myneni, V. Genovese, P.-N. Tan, and V. Kumar (2003), Continental-scale comparisons of terrestrial carbon sinks estimated from satellite data and ecosystem modeling 1982–1998, *Global Planet. Change*, *39*, 201–213.

Prentice, I. C., G. D. Farquhar, M. J. R. Fasham, M. L. Goulden, M. Heimann, V. J. Jaramillo, H. S. Kheshgi, and C. Le Quéré (2001), The carbon cycle and atmospheric carbon dioxide, in *Climate Change 2001: The Scientific Basis*, edited by J. T. Houghton et al., chap. 3, pp. 183–237, Cambridge Univ. Press, New York.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1992), *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge Univ. Press, New York.

Randerson, J. T., et al. (2002), Carbon isotope discrimination of arctic and boreal biomes inferred from remote atmospheric measurements and a biosphere-atmosphere model, *Global Biogeochem. Cycles*, *16*(3), 1028, doi:10.1029/2001GB001435.

Rayner, P. J., I. G. Enting, R. J. Francey, and R. L. Langenfelds (1999), Reconstructing the recent carbon cycle from atmospheric CO₂, δ¹³C and O₂/N₂ observations, *Tellus, Ser. B*, *51*, 213–232.

Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003a), CO₂ flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem. Phys.*, *3*, 2575–2659.

Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003b), Time-dependent atmospheric CO₂ inversions based on interannually varying tracer transport, *Tellus, Ser. B*, *55*, 488–497.

Saleska, S. R., et al. (2003), Carbon in Amazon forests: Unexpected seasonal fluxes and disturbance-induced losses, *Science*, *302*, 1554–1557.

Takahashi, T., R. H. Wanninkhof, R. A. Feely, R. F. Weiss, D. W. Chipman, N. Bates, J. Olafsson, C. Sabine, and S. C. Sutherland (1999), Net sea-air CO₂ flux over the global oceans: An improved estimate based on the sea-air pCO₂ difference, paper presented at 2nd International CO₂ in the Oceans Symposium, Cent. for Global and Environ. Res., Natl. Inst. for Environ. Stud., Tsukuba, Japan.

Tans, P. P., I. Y. Fung, and T. Takahashi (1990), Observational constraints on the global atmospheric CO₂ budget, *Science*, *247*, 1431–1438.

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

Thoning, K. W., P. P. Tans, and W. D. Komhyr (1989), Atmospheric carbon dioxide at Mauna Loa observatory: 2. Analysis of the NOAA/GMCC data, 1974–1985, *J. Geophys. Res.*, *94*, 8549–8565.

Tiedtke, M. (1989), A comprehensive mass flux scheme for cumulus parameterization in large-scale models, *Mon. Weather Rev.*, *117*, 1779–1800.

Trampert, J., and R. Snieder (1996), Model estimations biased by truncated expansions: Possible artifacts in seismic tomography, *Science*, *271*, 1257–1260.

Trumbore, S. (2000), Age of soil organic matter and soil respiration: Radiocarbon constraints on belowground C dynamics, *Ecol. Appl.*, *10*, 399–411.

van der Werf, G. R., J. T. Randerson, G. J. Collatz, L. Giglio, P. S. Kasibhatla, J. Arellano, F. Avelino, S. C. Olsen, and E. S. Kasischke (2004), Continental-scale partitioning of fire emissions during the 1997 to 2001 El Niño/La Niña period, *Science*, *303*, 73–76.

Vukićević, T., B. H. Braswell, and D. Schimel (2001), A diagnostic study of temperature controls on global terrestrial carbon exchange, *Tellus, Ser. B*, *53*, 150–170.

Wang, Y. P., R. Leuning, H. Cleugh, and P. A. Coppin (2001), Parameter estimation in surface exchange models using non-linear inversion: How many parameters can we estimate and which measurements are most useful?, *Global Change Biol.*, *7*, 495–510.

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

————————————

R. Giering and T. Kaminski, FastOpt, Schanzenstrasse 36, 20357 Hamburg, Germany. (ralf.giering@fastopt.com; thomas.kaminski@fastopt.com)

W. Knorr and M. Scholze, QUEST–Department of Earth Sciences, University of Bristol, Wills Memorial Building, Bristol BS8 1RJ, UK. (wolfgang.knorr@bristol.ac.uk; marko.scholze@bristol.ac.uk)

P. Rayner, Laboratoire des Sciences du Climat et de l'environnement–CEA de Saclay Orme des Merisiers, 91191 Gif sur Yvette, France. (peter.rayner@cea.fr)

H. Widmann, Max-Planck-Institut für Biogeochemie, Postage 100164, D-07701 Jena, Germany. (heiner.widmann@dkrz.de)
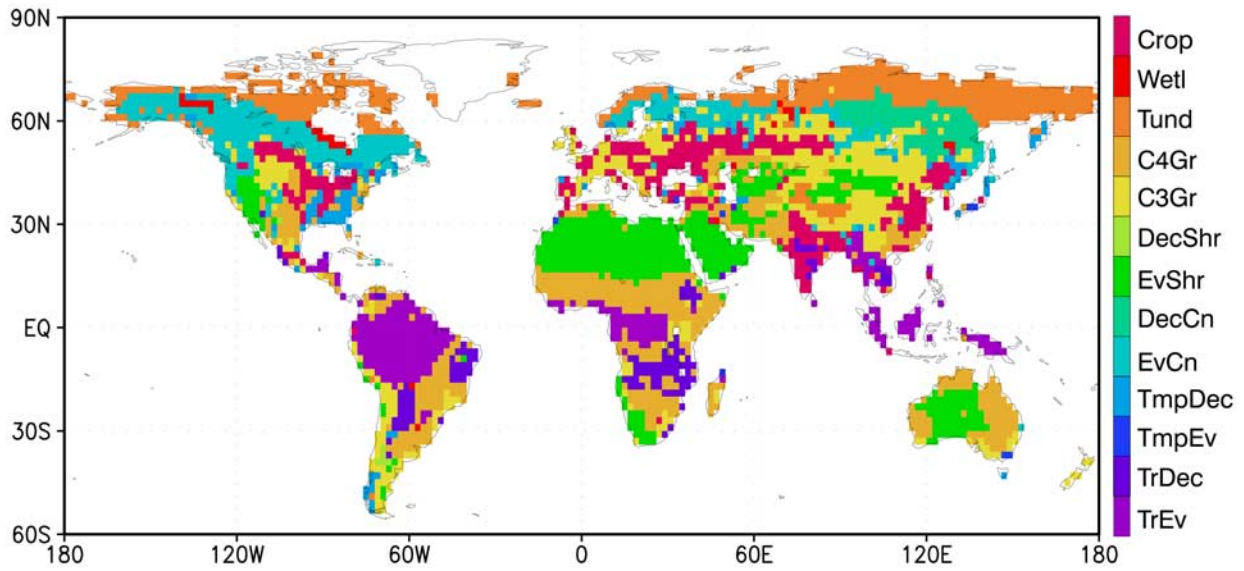
Figure 3. Distribution of the dominant PFT per grid cell. PFT labels are given in Table 1.
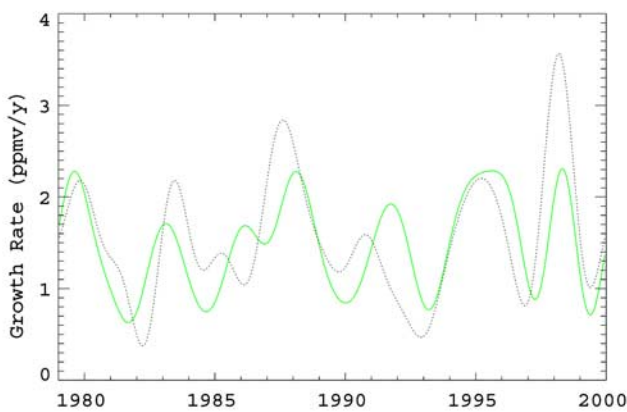

Figure 7. Observed (dotted line) and modeled (solid line) growth rates smoothed at approximately 1.5 year timescales. Growth rates are calculated from equation (24).
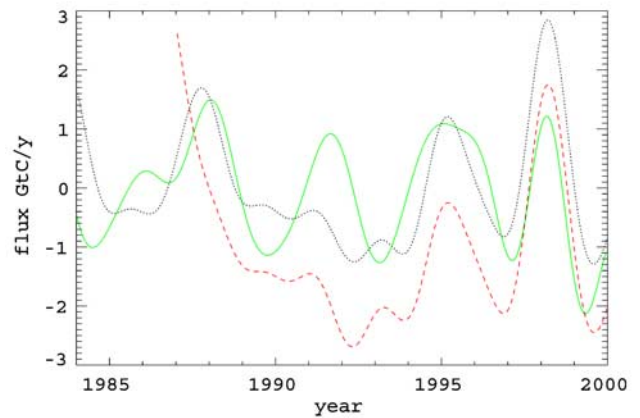
Figure 8. Comparison of our optimized terrestrial flux (solid line) with inversions using 16 stations (dotted line) and 19 stations (dashed line) of *Rödenbeck et al.* [2003a].
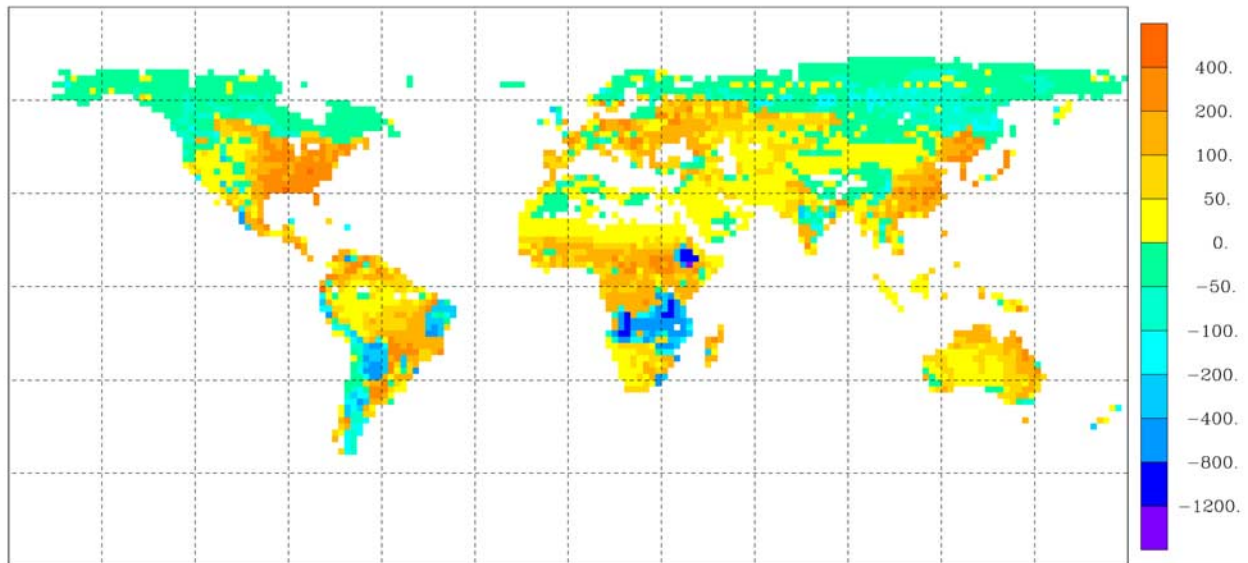
Figure 9. Mean net flux to the atmosphere for the period 1980–2000 (gC m^{-2} yr^{-1}).
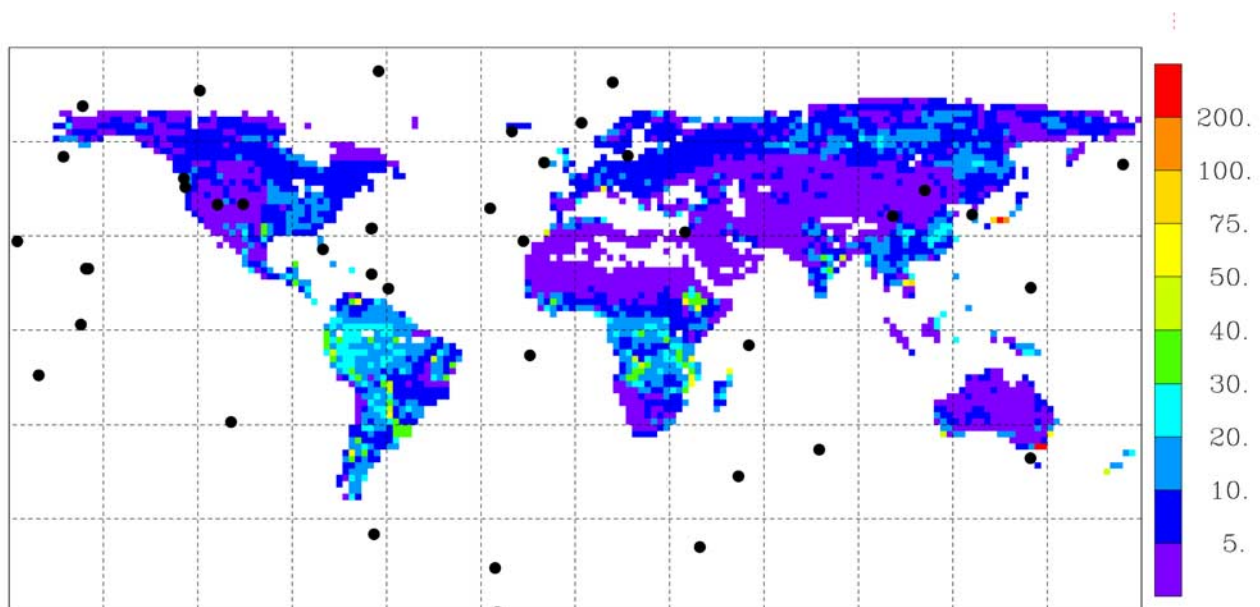
Figure 10. Uncertainty of mean net flux to the atmosphere for the period 1980–2000 (gC m^{-2} yr^{-1}) expressed as the standard deviation. Dots indicate locations of observational sites.
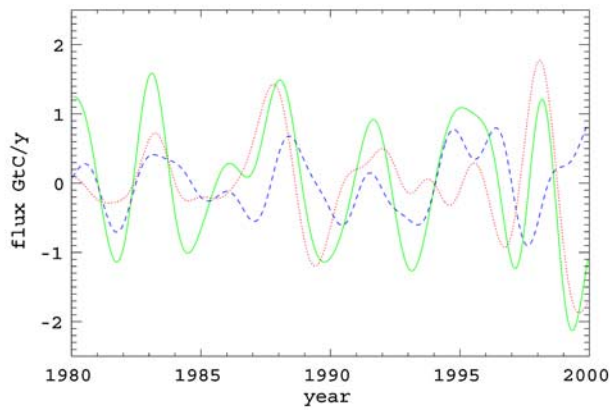
Figure 11. Global (solid line), tropical (20°S–20°N) (dotted line), and northern extratropical (20°N–90°N) (dashed line) anomalies in flux to the atmosphere from the optimized model.
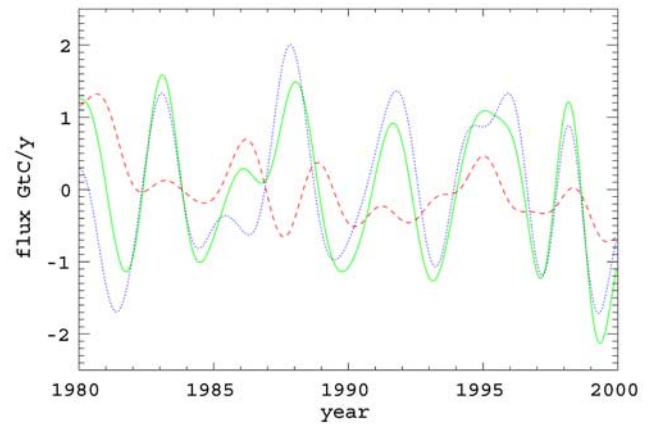

Figure 12. Global anomalies in negative NEP (solid line), negative NPP (dotted line), and fast respiration (dashed line) on interannual timescales. See text for details of time filtering.
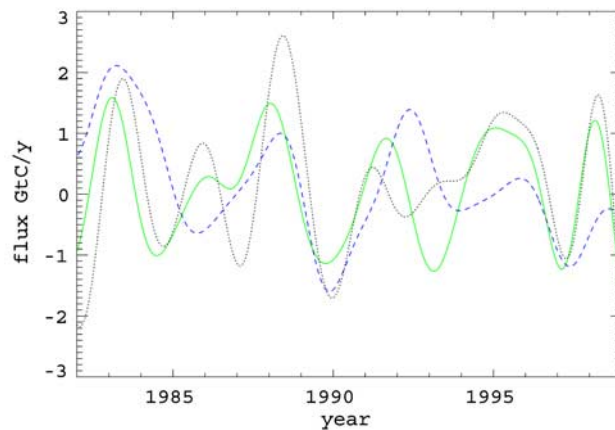

Figure 13. Global negative NEP anomalies for our prior (dotted line) and optimized (solid line) models and that of *Potter et al.* [2003] (dashed line).