J. Geophys. Res., 104 (D15), 18555–18581, 1999.

A coarse grid three-dimensional global inverse model of the atmospheric transport
2. Inversion of the transport of CO_2 in the 1980s

Thomas Kaminski and Martin Heimann

Max-Planck-Institut für Meteorologie, Hamburg, Germany

Ralf Giering[1]

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge

Abstract. Models of atmospheric transport can be used to interpret spatiotemporal differences in the observed concentrations of CO_2 in terms of its surface exchange fluxes. Inversion of the atmospheric transport is the systematic search for both a flux field that yields an optimal match between modeled and observed concentrations and, equally importantly, the uncertainties in this inferred flux field. The present inversion study combines observations of the CO_2 concentration at the global station network of the NOAA/CMDL in the 1980s with a representation of the atmospheric transport model TM2 by its Jacobian matrix, which has been previously computed by the adjoint model of TM2. This Jacobian matrix maps monthly fluxes on the approximately 8^o latitude by 10^o longitude horizontal model grid onto the resulting changes in the monthly CO_2 concentration at every station. Since the number of observational sites is much smaller than the number of grid cells, the inverse problem is highly underdetermined. A unique solution is determined by including a priori information on the surface exchange fluxes derived from output of high-resolution models of both the terrestrial biosphere and the ocean, combined with statistics of fossil fuel burning and land use change. Performing a Bayesian synthesis inversion, for a target period in the 1980s, the average seasonal cycle and the mean annual magnitude of CO_2 surface fluxes on the TM2 grid are inferred. The resulting simulated concentration compares well with independent observations. On a global scale, an oceanic sink of 1.5 ± 0.4 gigatons of carbon (GtC) is estimated. This sink is stronger in the northern than in the southern hemisphere. On a regional scale, however, the inferred exchange fluxes exhibit high uncertainty, indicating a low capacity of the global observational network to monitor regional trace gas emissions. These findings are relatively insensitive to the year of meteorological driving data, suggesting interannual changes in concentration should primarily result from source not transport changes.

1. Introduction

Owing to human activities such as fossil fuel burning and changes in land use, the atmospheric CO_2 concentration has

[1] Now at Jet Propulsion Laboratory, Pasadena, California.

rised by about 25% since preindustrial times. Observations of the atmospheric CO_2 concentration indicate that during the 1980s about 3 GtC/year of the estimated anthropogenic emissions of about 7 GtC/year remained in the atmosphere. On decadal timescales, the most important processes that can remove CO_2 from the atmosphere are uptake by the ocean and by the terrestrial biosphere. The net exchange flux with the ocean is driven by the difference between oceanic and atmospheric partial pressures of CO_2, while the net exchange flux with the terrestrial biosphere is the difference between net primary productivity (NPP) of the vegetation and heterotrophic respiration fluxes from plant litter and soil. Unbalances in the net exchange fluxes are caused by a number of terrestrial sink processes such as regrowth of forest following harvest, fertilization by the increasing atmospheric CO_2 concentration, or fertilization by nitrogen, but their magnitude is difficult to quantify.

For the 1980s, the Intergovernmental Panel on Climate Change (IPCC) estimates a yearly oceanic uptake of 2 ± 0.5 GtC based on a number of studies running models of the oceanic carbon cycle [*Houghton et al.*, 1995a] (unlike IPCC, who gives a 90 % confidence interval, throughout this paper, we quantify all uncertainties in terms of $\pm$ 1 standard deviation). A yearly uptake of about 2 GtC by the terrestrial biosphere is computed as the residuum in the budget equation of atmospheric CO_2.

Spatial differences in the atmospheric concentrations of CO_2, which are being measured at global networks of monitoring stations, reflect the spatial and temporal structure of the exchange flux fields. Prescribing surface flux fields, the atmospheric CO_2 concentration at observational sites can be simulated by atmospheric transport models. Consistency of simulated concentrations with observations has the potential to constrain flux fields. Investigating this consistency for a number of reasonable flux scenarios, *Keeling et al.* [1989] and *Tans et al.* [1990] inferred a sink in the midlatitudes of the northern hemisphere, for 1984 and 1981-1987, respectively. Both studies differ, however, in their interpretation of this sink: while Keeling et al. attributed much of this sink to an oceanic southward transport of carbon by the global thermohaline oceanic circulation (global ocean uptake of more than 2 GtC), Tans et al. found only a marginal oceanic uptake (less than 1 GtC) and concluded a major terrestrial sink. One reason for this disagreement is that in both studies the information from the atmospheric CO_2 observations is complemented by different pieces of additional information: while Keeling et al. use atmospheric measurements of the isotopic composition of CO_2, Tans et al. employed data of air–sea partial pressure differences of CO_2 .

Inversion of the atmospheric transport is an alternative to subjectively choose flux fields and compare the simulated atmospheric responses with observations. It consists in a systematic determination of a flux field that yields an optimal match between simulated and observed concentrations. A second, equally important goal of an inversion is to provide an estimate for the uncertainty in the inferred flux field.

A typical horizontal resolution for transport models employed in global inversion studies is 8^o latitude by 10^o longitude, i.e., in every time step the magnitudes of the flux into about 1000 grid cells have to be specified. On the other hand, the number of sampling sites at which global observational networks provide concentration measurements of atmospheric CO_2 is less than 100. Hence the inverse problem is highly underdetermined: i.e., there are many flux fields yielding the same modeled concentration at the observational sites. To obtain a unique solution, additional information has to be included in the inversion procedure. One way to do so consists in introducing further equations (hard constraints) for the flux field. The usual approach is to compose the flux field of spatiotemporal patterns with n_x unknown scaling coefficients. Spatial patterns such as latitude bands [*Brown*, 1993, 1995; *Ciais et al.*, 1995] or spherical harmonics in space [*Enting and Mansbridge*, 1989] can be derived from the geometry of the Earth and then be combined, e.g., with harmonics in time. Alternatively, a partitioning of the Earth's surface characterized by relevant processes can be constructed [*Hartley and Prinn*, 1993; *Hein and Heimann*, 1994; *Enting et al.*, 1995; *Hein et al.*, 1996; *Bousquet*, 1997; *Rayner et al.*, 1999a], or statistically motivated patterns like empirical orthogonal functions (EOFs) of the fluxes can be derived. Provided that n_x is small enough to yield an overdetermined inverse problem, the coefficients can be determined by a regression. (More precisely, n_x has to be smaller than the number of linear independent equations.) In another approach the inverse problem is made even determined, i.e., n_c is made equal to n_x, by interpolation or extrapolation of the observed concentrations [*Ciais et al.*, 1995; *Law*, 1999]. Here the actual inversion step is based on a statistical model rather than on a transport model. See *Enting* [1999] for a more detailed classification of inversion techniques.

An alternative to handle the underdetermined problem without reducing the number of unknowns is the so-called Bayesian approach, which allows to include a priori information on the fluxes in the inversion: Both atmospheric observations and a priori information are described in terms of probability densities. Using the transport as constraint, consistent probability densities are derived. The a priori information regularizes the underdetermined inverse problem (by acting as a soft constraint). Usually, the observations as well as the a priori information about the fluxes are described in terms of Gaussian probability densities. In combination with linearized transport, simple formulas for the posterior esti-

mate of the fluxes and its uncertainty can be derived [see, e.g., *Tarantola*, 1987].

Many of the abovementioned studies have applied Bayesian approaches [*Hartley and Prinn*, 1993; *Hein and Heimann*, 1994; *Enting et al.*, 1995; *Hein et al.*, 1996; *Bousquet*, 1997; *Rayner et al.*, 1999a]. Thus, by choice of their inversion method, they could have avoided to prescribe fixed patterns. Nevertheless, they could not get rid of these patterns for computational reasons. Formally, these inversions are based on a transport matrix mapping scaling coefficients of flux patterns onto concentrations. Computationally, this matrix has been derived by running the respective transport model separately with each of the prescribed flux patterns and recording the contributions to the corresponding concentration signal at each of the monitoring sites and times. Owing to the considerable computational cost of the necessary transport model runs the number of source components remains low; in the abovementioned studies the spatial resolution ranges from 5 to about 30 regions.

The two most recent of the abovementioned studies infer the surface exchange fluxes of CO_2 for the periods from 1985 to 1995 [*Bousquet*, 1997] and from 1980 to 1995 [*Rayner et al.*, 1999a], respectively. In addition to observations of atmospheric CO_2 at global networks and prior estimates for the surface exchange fluxes, they include further observations to better discriminate between land and ocean uptake: Both use measurements of the isotopic composition of CO_2. Rayner et al. also include measurements of the oxygen to nitrogen ratio. They estimate yearly ocean uptakes of 1.5 ± 0.5 GtC (Bousquet) and 2.1 GtC (Rayner et al.) respectively.

For the present study we also apply the Bayesian approach. Yet, as described in a companion paper [*Kaminski et al.*, this issue], for the atmospheric transport model TM2 in a cyclostationary setup, a matrix representation on the entire model grid is available. In contrast to the abovementioned studies, we are in a situation to perform an inversion for the model's spatial resolution of approximately 8^o latitude by 10^o longitude and monthly time scale in the flux space. This high resolution is greatly relaxing the hard constraint of prescribing a few fixed patterns and enables us to capture much more of the spatial variability of the fluxes, which allows a more realistic simulation of the concentration. Evidently, with increasing number of unknowns, the gain of information about the particular unknowns from the atmospheric observations becomes smaller: The higher the resolution in the space of fluxes, the lower the reduction of uncertainty for a particular flux component by the inversion. Yet, for sums of flux components representing large-scale quantities such as, e.g., the scaling coefficients of prescribed patterns, the gain of information is higher than for individual flux components. Another important advantage of a higher resolution has been pointed out by *Snieder* [1993] and *Trampert and Snieder* [1996]: In combination with inhomogeneous sampling of the observations, insufficient resolution in the space of unknowns causes the inversion to yield biased estimates. Our network (see Figure 2 of *Kaminski et al.* [this issue]) indeed seems to be characterized by an inhomogeneous spatial distribution of the observational sites. Hence, a high resolution appears favorable to reduce a possible bias, especially in sums of estimated fluxes. For our inversion of the transport of CO_2, we use a priori information on the surface fluxes derived from output of high-resolution models of both the terrestrial biosphere [*Knorr and Heimann*, 1995] and the ocean [*Six and Maier-Reimer*, 1996], combined with statistics of fossil fuel burning [*Andres et al.*, 1999] and land use change [*Houghton et al.*, 1987].

The layout of the remainder of this paper is as follows: Section 2 presents our inverse modeling approach. Section 3 describes the a priori information on the fluxes. Section 4 deals with the atmospheric observations, followed by a description of our inversion technique and a discussion of the singular value spectrum in section 5. The inferred flux field as well as its spatial and temporal means are presented in section 6 together with their uncertainties. A recipe to compute posterior uncertainties is given in the appendix. Section 7 discusses the resulting simulated concentration at observational sites. Section 8 presents the total fluxes for some oceanic regions and countries. An attempt to assess the reliability of the posterior flux field is made in section 9. Eventually, in section 10 we draw conclusions about the carbon cycle and discuss perspectives of the adjoint approach for inversion of the atmospheric transport.

2. Methodology

In a companion paper [*Kaminski et al.*, this issue], we have described our model for the simulation of the quasi-stationary seasonal cycle of atmospheric CO_2 at the sampling sites of an observational network, c_{qsc}:

$$c_{qsc} = Mf. \tag{1}$$

M essentially is the Jacobian matrix of the atmospheric transport model TM2 in a cyclostationary setup, and f is a prescribed surface flux field on the approximately 8^o latitude by 10^o longitude model grid. The Jacobian matrix has been computed by the adjoint model of TM2, which also has been introduced by *Kaminski et al.* [this issue]. The corresponding inverse problem consists in the algorithmic determination of a flux field f that for the observed quasi-stationary seasonal cycle satisfies (1).

We solve for monthly mean fluxes on the entire horizon-

tal (36×24) TM2 grid: i.e., we have $12 \times 36 \times 24 \approx 10000$ unknowns. As will be described in section 4, on the observational side we use monthly means at 25 stations plus a global trend: i.e., we have 301 equations. Hence the inverse problem defined by (1) is highly underdetermined: i.e., there are many flux fields yielding the same modeled concentration at the observational sites. Performing a "joint inversion" or "Bayesian inversion," a unique flux field is determined by including a priori information on the fluxes. Since the formalism for Bayesian inversions is derived and clearly explained in the book of *Tarantola* [1987], here we only give a brief description of the principle and the formulas we apply.

Rather than operating on the pure numbers c_{qsc} and f, the objects of Bayesian inversions are probability densities quantifying states of information on the respective quantities. The basic idea is illustrated in Figure 1: A prior state of information is characterized by independent probability densities of both atmospheric observations and a priori information on the fluxes. Using the atmospheric transport, i.e., (1), as constraint, a posterior state of information as characterized by a unique (if existing) probability density in a joint space of fluxes and concentrations is defined. Projections to the spaces of fluxes and concentrations yield a posteriori probability densities, which are consistent with the atmospheric transport.

As suggested by Figure 1, in our study, the observations as well as the a priori information about the fluxes are described in terms of Gaussian probability densities (see section 4 and section 3, respectively), which are determined by their respective means and covariance matrices f and C_f as well as c_{obs} and C_c, where C_c is the sum of observational uncertainty and uncertainty due to model error. Together with the linear model of (1), the inversion yields a Gaussian posterior probability density for the fluxes, whose mean $f^{'}$ and covariance matrix $C_f^{'}$ are given by

$$f^{'} = f + C_f^{'} M^{*} C_c^{-1} (c_{obs} - Mf) \qquad (2)$$

$$C_f^{'} = (M^{*} C_c^{-1} M + C_f^{-1})^{-1} \ . \qquad (3)$$

In the lower part of the figure, showing projections of the a posteriori density on the spaces of fluxes and concentrations, for comparison the respective a priori densities are depicted as dashed curves: The shift of the mean is quantified by (2), and the sharpening of the density, which reflects the gain of information through the inversion procedure, by (3).

Defining the (pseudo) inverse

$$M^{-1} := C_f^{'} M^{*} C_c^{-1} \qquad (4)$$

Equation (2) can be written in the form

$$f^{'} - f = M^{-1} (c_{obs} - Mf) \ : \qquad (5)$$

M^{-1} transforms the misfit between observations c_{obs} and the modeled concentration $c_{mod} = Mf$ resulting from the a priori fluxes to a correction of the a priori fluxes.

In our study, the off diagonal elements in the a priori covariance matrix of the fluxes C_f are unknown, and those in the covariance matrix of the observed concentrations C_c are small. For computational convenience, for both the observations and the a priori fluxes, we assume simple diagonal covariance matrices: i.e., we specify uncertainties in the form of standard deviations, but we avoid the definition of correlations among the uncertainties of different flux components or different observations. The a posteriori flux field can be shown to also minimize the cost function

$$J(\tilde{f}) = {}^{1}/_{2} [\sum_{i=1}^{n_f} (\frac{f - \tilde{f}}{\sigma_f})_i^2 + \sum_{i=1}^{n_c} (\frac{c_{qsc} - M\tilde{f}}{\sigma_c})^2 _i] \ : \qquad (6)$$

the a posteriori field, in a least squares sense, is close to the a priori field f, and the atmospheric response at the stations is close to the observations c_{qsc}. The weights of the individual terms are given by the reciprocals of the respective uncertainties, σ_f and σ_c. Since we are not only interested in the posterior flux fields, but also in an analysis of the posterior covariance matrix, we solve (2) instead of applying iterative minimization algorithms to (6). Technically, our inversion is based on a Singular Value Decomposition (SVD, see section 5).

The model resolution quantifies the ability of the observations to constrain the posterior estimate. If f_0 is a known flux field, then $c_{obs} := Mf_0$ would be the corresponding observation provided that the model was perfect. Inserting c_{obs} and an a priori estimate f together with their covariances into (5) yields

$$\begin{aligned} f^{'} - f &= M^{-1}(c_{obs} - Mf) \\ &= M^{-1} M (f_0 - f) \end{aligned}$$

for each component the correction suggested by the inversion procedure (5), is a weighted sum of the correction that would be necessary to recover f_0. These weights form the model resolution matrix, $R_m := M^{-1} M$. If R_m equals the identity matrix, the model resolution is perfect: by the inversion procedure the components of f_0 can be recovered independently of each other. Using the definitions of M^{-1} (equation (4)) and $C_f^{'}$ (equation (3)), links the model resolution matrix to the reduction of uncertainty $C_f^{'} C_f^{-1}$, which quantifies how much one learns from the observations:

$$\begin{aligned} R_m &= C_f^{'} M^{*} C_c^{-1} M \\ &= C_f^{'} (M^{*} C_c^{-1} M + C_f^{-1} - C_f^{-1}) \\ &= 1 - C_f^{'} C_f^{-1}. \end{aligned} \qquad (7)$$

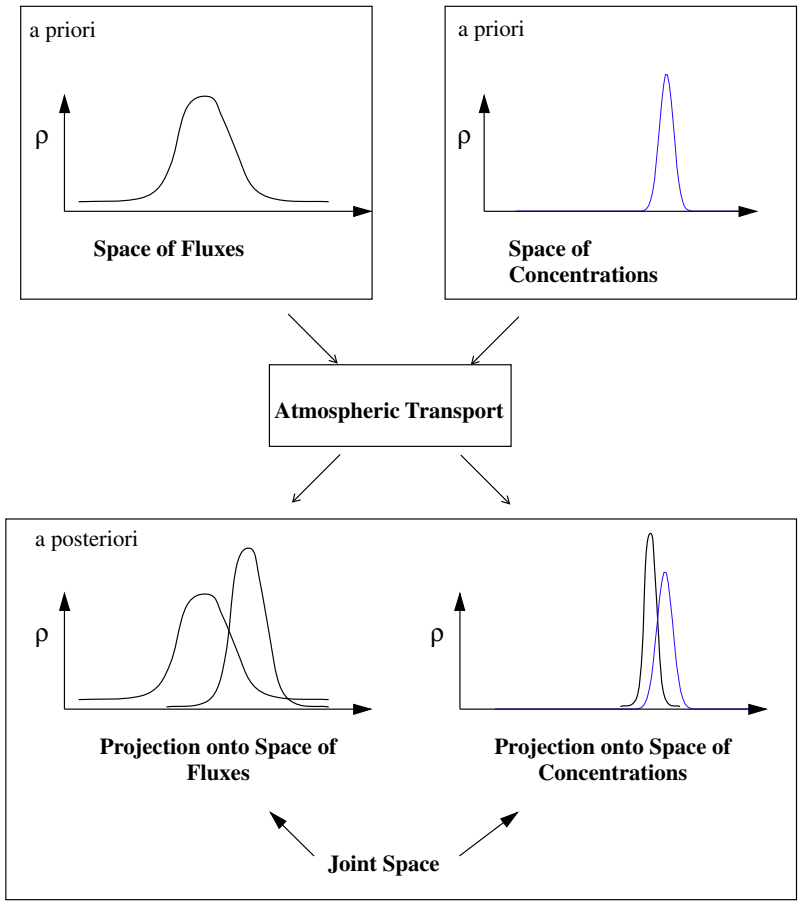


Figure 1. A schematic illustration of the Bayesian approach: The a priori state of information is represented by independent probability densities for fluxes and concentrations. Combining this information to the information about the atmospheric transport, represented by our numerical model, yields a consistent a posteriori state of information, represented by a probability density in the joint space of fluxes and concentrations; projections to the individual spaces, in general, are sharper than the a priori densities.

The higher the reduction of uncertainty, the closer the model resolution is to 1.

3. A Priori Fluxes and Uncertainties

3.1. A Priori Fluxes

In this section, we explain how we specify a prior flux into each grid cell and month, together with its uncertainty. This flux field is composed of the contributions from four components: a seasonal terrestrial biosphere in annual equilibrium, a correction for land use change, the ocean, and fossil fuel burning. Fluxes into the atmosphere are positive.

For the biospheric component the seasonal net exchange fluxes derived by the Simple Diagnostic Biosphere Model (SDBM, [*Knorr and Heimann*, 1995]) were interpolated from the models 0.5^o grid to the TM2 grid. The SDBM is driven by climate data, observed greenness from satellite derived global vegetation index data, and a drought stress indicator calculated with a one layer bucket model. NPP is the product of a globally constant photosynthetic light use efficiency, the observed greenness and the drought stress factor. Soil respiration is proportional to the drought stress factor and an exponential function of the soil temperature that is characterized by one global parameter Q_{10}; in each grid cell, the proportionality factor is chosen to achieve a locally

balanced yearly net flux. In the SDBM, two global model parameters, the light use efficiency and Q_{10}, are tuned by minimization of the misfit between observations of the seasonal cycle of atmospheric CO_2 and the seasonal cycle simulated by feeding the modeled fluxes into TM2. For this procedure the observations at the northern hemisphere stations BRW, CBA, AZR, and KUM from 1980 to 1990 and STM from 1982 to 1990 were used (see station map in Figure 2 of *Kaminski et al.* [this issue]). Through this parameter fit at least a part of the atmospheric observations that are used in our inversion have already influenced our a priori flux field. Strictly speaking, this constitutes a violation of the assumption of independent information about fluxes and atmospheric observations, upon which our inversion procedure is based. This violation, however, is slight for three reasons: first and most important, as described in section 3.2, the uncertainty in the a priori flux estimate is derived independently from the uncertainty in the atmospheric observations. Hence we do not risk to underestimate the uncertainty in our posterior flux estimate by double counting a fraction of the atmospheric information. Second, we would not expect significant flux changes, if instead the atmospheric observations from a period excluding our target period had been chosen in the construction of SDBM. Third, since only two global parameters have been tuned, most details of the flux field's structure are imposed by the climate and satellite data. Although, alternatively, fluxes computed from models that are not based on atmospheric observations can be used, we decided in favor of the SDBM, because it allows to include the satellite data and because the model performed well in intercomparison studies [*Heimann et al.*, 1998].

Since the fluxes computed by the SDBM represent a terrestrial biosphere in local equilibrium, i.e., the local annual mean flux is zero, perturbations of this equilibrium have to be quantified separately. The only perturbation for which we explicitly specify an a priori flux is land use change. On the basis of regional estimates by *Houghton et al.* [1987], an annual mean field of fluxes due to land use change has been compiled by *Heimann and Keeling* [1989]. This field yields a global annual mean source of 1.7 GtC.

For the oceanic component the seasonal net exchange fluxes computed by a simple plankton model [*Six and Maier-Reimer*, 1996] embedded in the Hamburg model of the oceanic carbon cycle [*Maier-Reimer*, 1993] were regridded from the models 3.5^o grid to the TM2 grid. The fluxes for the 1980s were taken from a transient run with prescribed observed atmospheric CO_2 concentrations, starting in 1756 from the models equilibrium for the preindustrial CO_2 concentration [*Enting et al.*, 1994]. Hence, in contrast to the biospheric component, the oceanic net exchange fluxes are not balanced; the global annual mean ocean uptake is 1.7 GtC.

From fossil fuel burning statistics of *Andres et al.* [1999] on a 1^o grid, annual mean fluxes on the TM2 grid have been interpolated. This yields a global annual net source of 5.3 GtC. Compared to the biospheric and oceanic components, the uncertainty is rather small, so that we exclude this component from the inversion by a procedure described in section 4.

The annual mean of the a priori flux field, i.e., the sum of the flux contributions from terrestrial biosphere and ocean, is shown in the middle panel of Figure 2.

3.2. Uncertainties

Compared to the uncertainties in the observed concentration, it is much more difficult to quantify the uncertainties in the fluxes. Yet these uncertainties are crucial parameters for the inversion as expressed by (6): A term with a large uncertainty has merely a small impact on the result of the inversion. Hence, by assigning large uncertainties to the fluxes we can perform a weighting in favor of the observed concentration. In the following we describe our standard choice, which is intended to emphasize the weight on the atmospheric observations. Among the individual flux components, the recipe gives a smaller weight to the a priori values of those components we consider more uncertain. By and large we consider large fluxes more uncertain than small ones, yielding heigher weights for the small fluxes. Since typical oceanic fluxes are much smaller than typical terrestrial fluxes, the first are more certain than the latter. In order not to fix any flux component that is erroneously small by assigning a too small uncertainty, we use a minimum uncertainty of 0.12 kg m^{-2} a^{-1}, which is about 100 % of the largest oceanic fluxes. Only those components for which nonzero fluxes do not make sense such as deserts or ice covered regions have the practically zero uncertainty of 10^{-12} kg m^{-2} a^{-1}. The remainder of this section describes our detailed recipe for assigning the uncertainties.

In every grid cell with a land fraction of more than 1%, the terrestrial flux is considered the sum of NPP, soil respiration, and land use change contributions. Each month the uncertainty for this sum of fluxes is determined by assuming independent uncertainties of 50% for the NPP and soil respiration components and an uncertainty of 100% for the land use change component. We use an annual mean NPP of 0.01 kg m^{-2} a^{-1} as limit to discriminate small fluxes from zero fluxes. Small flux components have the minimum uncertainty, while zero fluxes (deserts and ice covered regions) have the practically zero uncertainty.

The uncertainty for oceanic flux contributions is defined as follows: To every grid cell with an ocean fraction of more

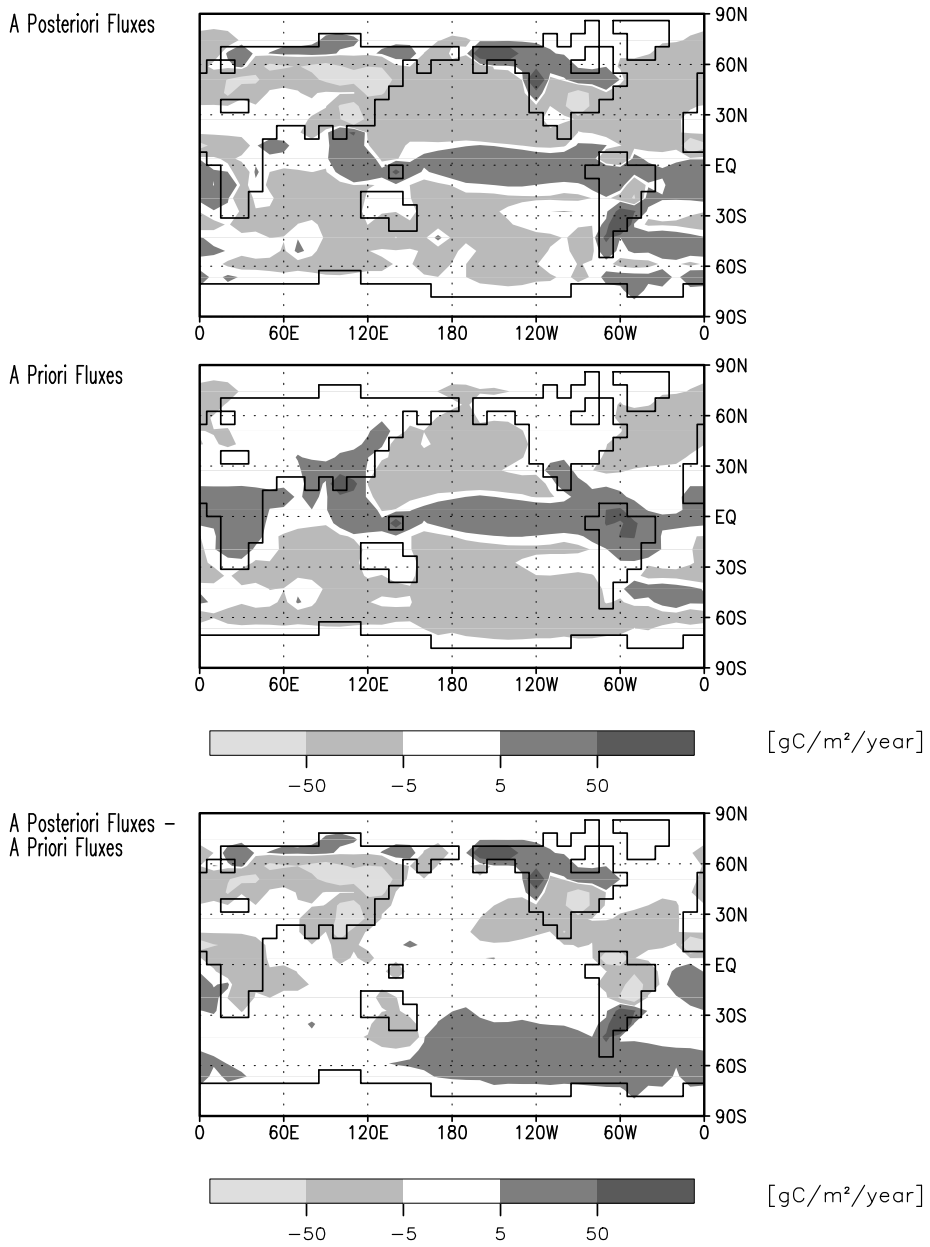


Figure 2. Annual mean of the sum of the flux contributions from terrestrial biosphere and ocean; (top) a posteriori, (middle) a priori, (bottom) and their difference; in the difference plot, positive values quantify an enhanced source or a reduced sink.

than 1% and with nonzero annual mean fluxes, we assign the minimum uncertainty, which, in any case, exceeds 100% of the respective flux. To discriminate fluxes that might erroneously be near zero from those that are definitely zero, an annual mean oceanic flux of 5×10^{-4} kg m^{-2} a^{-1} turned out a useful limit. Grid cells with smaller annual mean oceanic fluxes (ice covered areas) have the practically zero uncertainty.

According to this recipe, every grid cell has a terrestrial, or a ocean uncertainty, or both. In grid cells that have both, in proportion to the land fraction a mean uncertainty is assigned. The resulting uncertainties are displayed in Figure 3. The global and annual mean uncertainties are 0.46 GtC for the ocean and 1.33 GtC for the land.

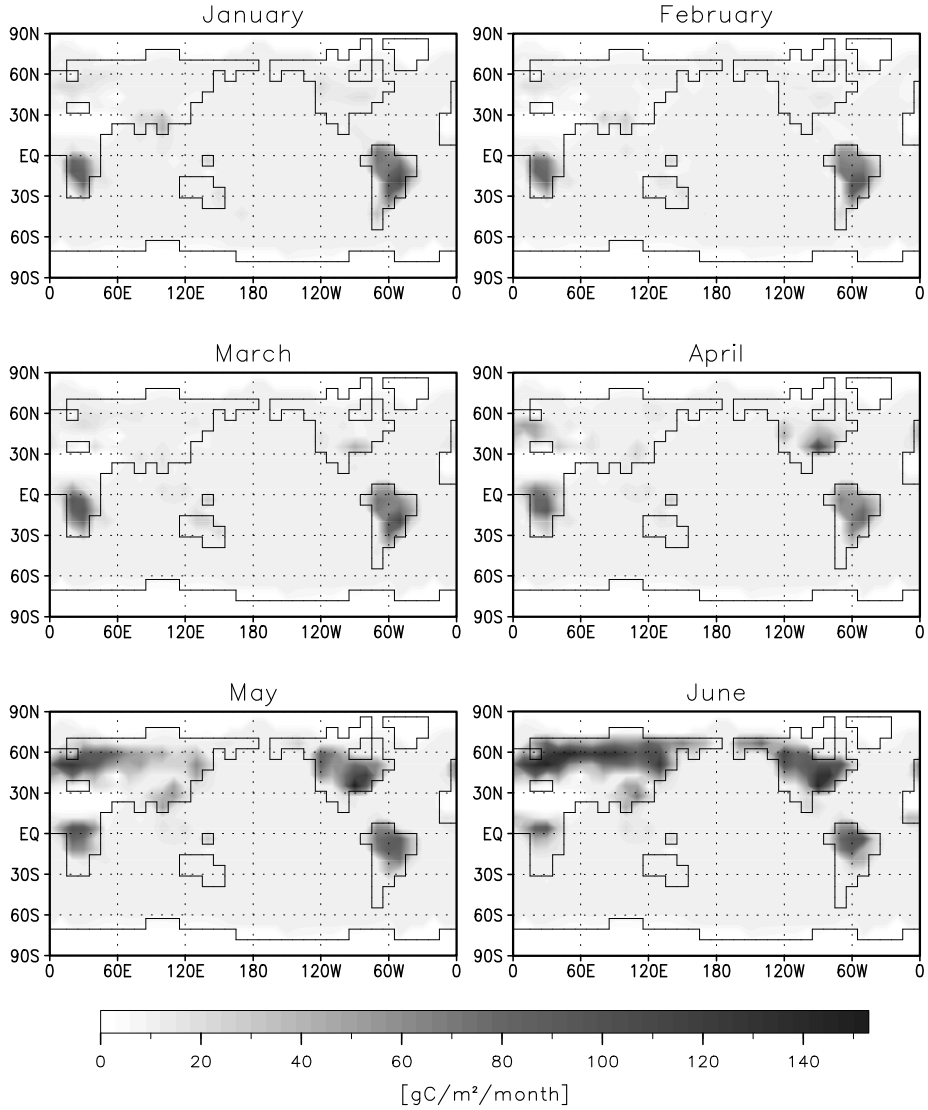


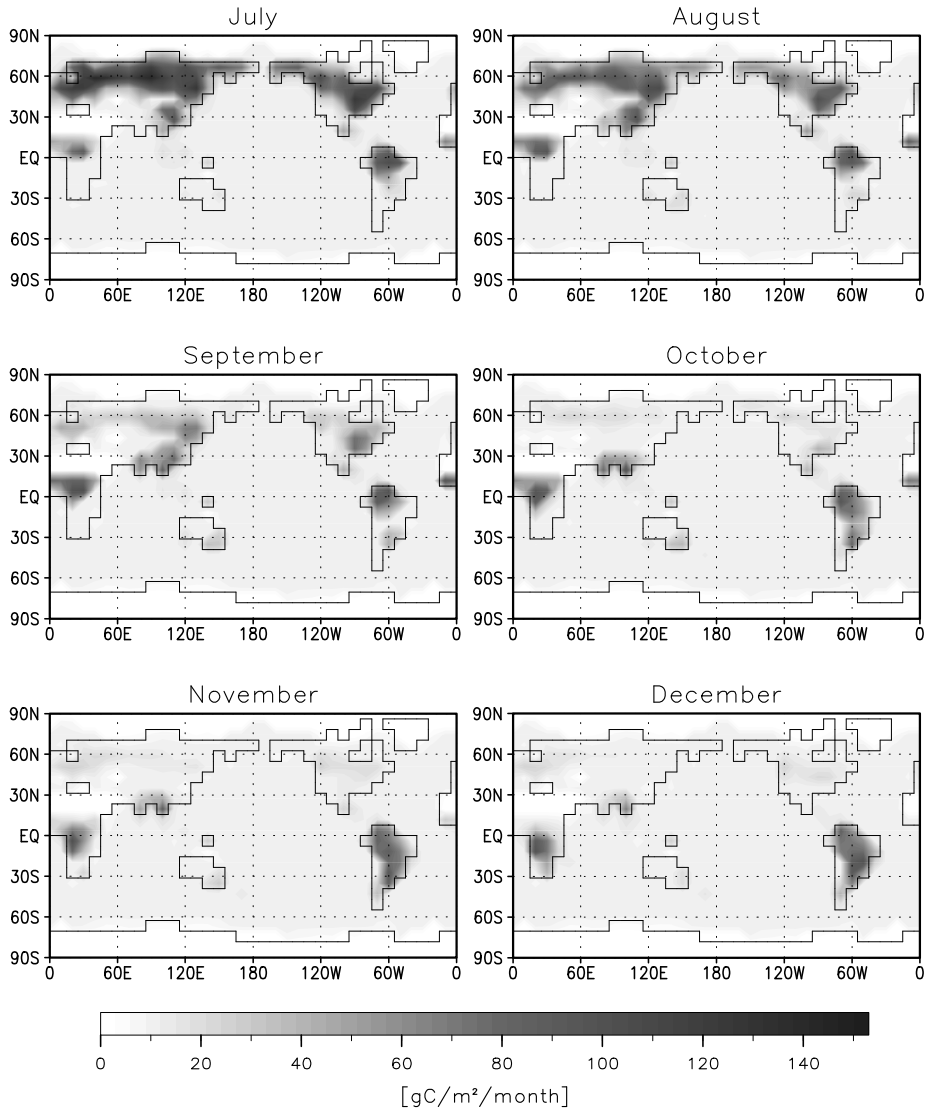
Figure 3. A priori uncertainties of the sum of the flux contributions from terrestrial biosphere and ocean.

Figure 3. (continued)

4. Observed Concentrations and Uncertainties

4.1. Observed Concentrations

Globalview - CO_2 is a database of high quality atmospheric measurements coordinated by the NOAA/CMDL. The observational net comprises more than 60 sites, for which smoothed weekly data together with an estimate of their uncertainties have been prepared [*Globalview–CO2*, 1996]. In order not to be affected by problems of intercalibration between different networks, we have restricted the data to those measured in the NOAA laboratory [*Conway et al.*, 1994]. As discussed in *Kaminski et al.* [this issue], our model is appropriate to simulate the mean quasi-stationary seasonal cycle for a target period of a few years. In their study *Tans et al.* [1990] interpreted the observations from 1981 to 1987. We choose a similar target period from January 1981 to January 1987 excluding the El Niño year 1987. For this target period data from the 25 NOAA sites displayed in Figure 2 of *Kaminski et al.* [this issue] were available. Unlike *Tans et al.* [1990] we have not excluded data from



any particular site of the network like the mountain stations MLO and NWR. We have not used the version of the data set in which temporal gaps in the records have been closed by statistical extension procedures [*Masarie and Tans*, 1995].

For comparison with our model, we extract from the observations the quasi-stationary seasonal cycle: At every station S and every month i in the target period, the mean concentration $c_{S,i}$ together with its uncertainty are computed. To quantify a periodic and a trend component, we use a statistical model similar to the one for the definition of the quasi-stationary seasonal cycle of *Kaminski et al.* [this issue]:

$$c_{S,i} = c_{p,S,i} + b \cdot t_i + N_{S,i} = \tilde{a}_S + S_{S,i} + b \cdot t_i + N_{S,i}. \quad (8)$$

Again, the periodic component c_p is decomposed into a periodic function S_S with zero mean and an offset $\tilde{a}_S$. For the observations, however, the offset in turn can be considered as the sum of two terms: The global mean concentration $\bar{c}_0$ at the beginning of the target period, and the spatial gradient a_S. The second term is the contribution of a global linear trend, where t_i is the length of the time interval from the beginning of the target period to the middle of the ith month. The noise term $N_{S,i}$ can be attributed to interannual variations in the fluxes and the transport as well as observational errors. The quantities c_p and b are estimated by a least squares fit together with their uncertainties.

Our model of the quasi-stationary seasonal cycle derived by *Kaminski et al.* [this issue] can be easily adapted by

adding the unknown global mean concentration $\bar{c}_0$ at the beginning of the target period,

$$b = \alpha \cdot \bar{f} \quad (9)$$

$$c_p = \bar{c}_0 + Tf - t \cdot \alpha \cdot \bar{f}, \quad (10)$$

where T denotes the Jacobian matrix. Composing the observables b and c_p to a single vector c_{cgt} and extending f by one additional "pseudoflux" component for $\bar{c}_0$, these equations define the matrix M of (1), which we actually invert.

As mentioned in section 3, in the inversion we want to consider the fossil fuel component in the fluxes f_f as known, because its uncertaixty is much smaller than the uncertainty of the oceanic and biospheric components. Owing to linearity of (1), from the observations we can subtract the modeled quasi-stationary seasonal cycle component at the station locations resulting from the fossil fuel source Mf_f. Eventually, we interpret f in (1) as the sum of the oceanic and biospheric components and c_{qsc} as the observed response of the concentration.

4.2. Uncertainties

For the data covariance matrix C_c in (3) and (2), we assume a diagonal structure: i.e., there are no correlations among uncertainties of different observations. Within the fitting procedure for (8), uncertainties for c_p and b are derived that are consistent with the uncertainties in the monthly mean observations. If our model was perfect, i.e., model error was zero, this would be the single source of uncertainty to be accounted for, when comparing model to observations. One obvious simplification, however, is the lack of interannually varying transport. The magnitude of interannual variations in transport, combined with that of the fluxes, is reflected in the degree to which the fit of a quasi-stationary seasonal cycle fails to match the monthly observations within the observational error bars. Hence we use these residue between the fitting curve determined by c_p and b and the monthly observations to derive an estimate of the uncertainty due the quasi stationary simplification in our model. For a given station and month of the year, say January, we compute a mean residuum as square root of the mean of the squares of all January residue in the target period. Finally, according to (3), we compose the variances of each component in c_p as the sum of both variances that determined within the fitting procedure (observational uncertainty) and that defined by the squares of these mean residue (uncertainty due to model error). This, however, might be over estimating the variances, because two overlapping estimates are added: Besides from interannual variations in transport the residue originate from two sources of error that we already account for elsewhere: accounting for observational error has been described above, and interannual variation of fluxes is accounted for trough the concept of a cyclostationary flux over the target period.

The fossil fuel component in the fluxes is uncertain as well. The IPCC [*Houghton et al.*, 1995b] estimates a 90% confidence interval of $\pm 10\%$ from the global annual mean flux. For convenience, we do not assume an uncertainty in the pattern of the emissions. Thus the resulting covariance of uncertainties in the simulated fossil fuel component in the quasi-stationary seasonal cycle can be easily computed and added to the estimates derived above, the off diagonal elements of this covariance are not taken into account.

In our inversion, we consider the global mean concentration $\bar{c}_0$ at the beginning of the target period unknown. In order to allow for a high flexibility, we assume the extremely high uncertainty of ± 1000 ppmv. The prior estimate is derived from the fit to the observations (see (8)) by computing the mean offset $\tilde{a}$ for all stations. The trend b is 1.41 ppmv a^{-1}. Its uncertainty determined within the fitting procedure is negligibly low, so that as prior uncertainty we assume the contribution from the uncertainty of the fossil fuel emissions which is 0.16 ppmv a^{-1}.

For a number of selected stations, Figure 4 shows the observed and modeled quasi-stationary seasonal cycles for 1981. The observations and their uncertainties have been composed of the fit as described above. Two versions of the modeled quasi-stationary seasonal cycle are displayed: The first results from the prior flux field, which has been described in section 3. The second results from the posterior flux field determined by the inversion and will be discussed in section 6.

5. Singular Value Decomposition

In our inversion, the information on the atmospheric transport is combined with atmospheric observations and the a priori information on the fluxes according to (2). Technically, the basis of our inversion consists of a singular value decomposition (SVD) of the model matrix M in (1). Since the SVD is derived, e.g., by *Menke* [1989] and well described, e.g., by *Press et al.* [1986], here we only give a brief summary. The aim of this section is to show how the SVD is applied to our problem and to discuss the singular values and vectors.

In the spaces of fluxes and concentrations, by the SVD two sets of n_c orthonormal vectors are derived, with respect to which the matrix M is diagonal: the right-hand singular vectors and the left-hand singular vectors. The diagonal elements are called singular values. The left-hand singular vectors span the complete space of concentrations, while the right-hand singular vectors only span a subspace of the space

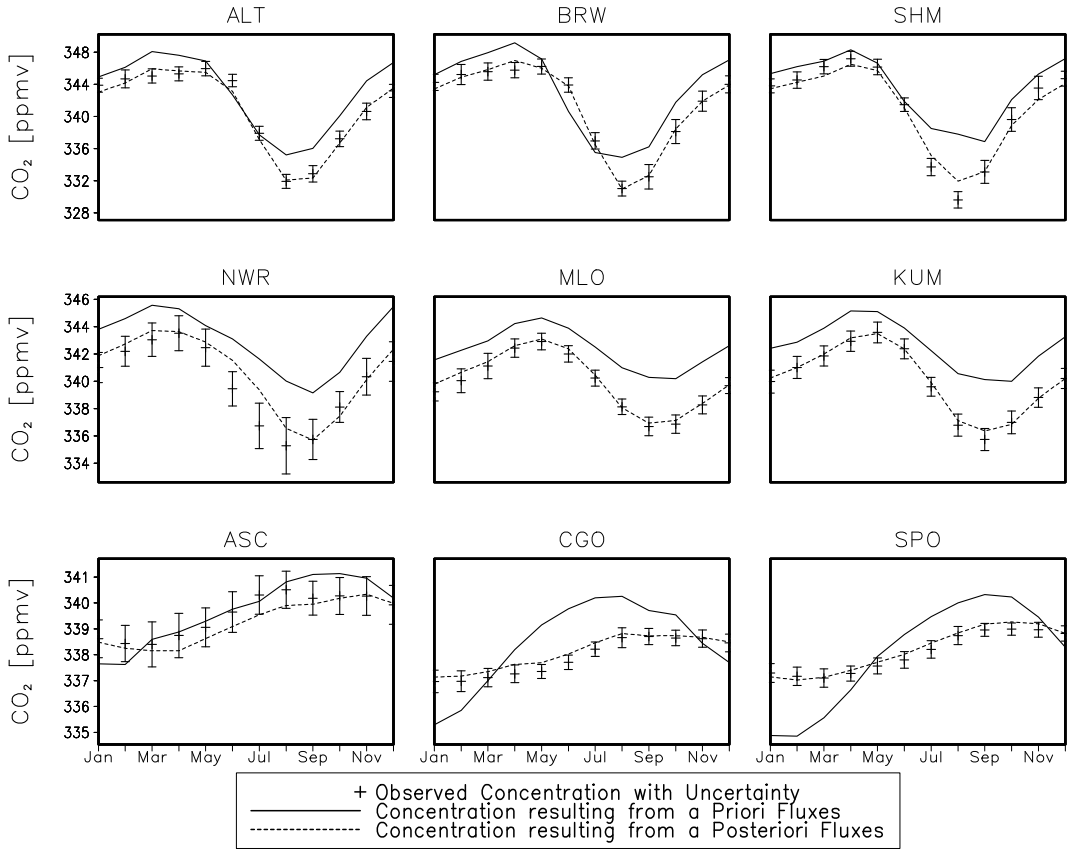


Figure 4. Observed concentration with mean quasi-stationary seasonal cycle for the first year of the target period, 1981. Error bars reflect observational uncertainties as well as uncertainties due to interannual variations. Modeled concentration resulting from a priori and a posteriori estimates of fluxes; see Figure 2 of *Kaminski et al.* [this issue] or Table 2 in *Kaminski et al.* [this issue] for station locations.

of fluxes. Arranging the associated singular vectors column by column in two matrices, U for the left-hand singular vectors and V for the right-hand singular vectors, and the singular values on the diagonal of a third matrix D, so that the associated singular vectors and values are in the same position within their respective matrices, our matrix M can be expressed as

$$M = UDV^T, \tag{11}$$

where V^T is the transposed of V.

Requiring orthonormality of the singular vectors and non negativeness of the singular values defines the singular values uniquely. U, D, and V are "almost" uniquely defined: (1) Permutations simultaneously changing the order in all three matrices are possible. (2) The singular values do not necessarily differ from each other (the identity matrix represents a pathological example). In such a case, simultaneous rotations of right-hand and left-hand singular vectors within the subspaces corresponding to the same singular value are possible, because these rotations do not disturb the orthonormality. For one-dimensional subspaces, i.e., for singular values appearing only once in D, such a rotation degenerates to a simultaneous flip of the signs of both the associated left-hand and right-hand singular vectors.

It is this uniqueness that makes any discussion of the singular vectors and values sensible. Through the orthonormality condition, however, the SVD depends on the units in the spaces of fluxes and concentrations. The only units being intrinsic to the problem are the prior uncertainties of the fluxes and the concentrations. Expressing the fluxes and concentrations in multiples of their uncertainties (we will refer to them as natural units) has the advantage that their prior covariances are represented by the respective identity matrices

(for convenience we denote both by 1), which also simplifies coding of the inversion algorithm. Although we transform to natural units to carry out and discuss the inversion, we keep on displaying fluxes and concentrations in original units.

Using natural units, for a flux component with a high prior uncertainty, a unit change is large (in original units) flux difference and thus tends to have a high impact on the simulated concentrations. Consequently, the right-hand singular vectors associated to the highest singular values, since they correspond to fluxes with high impact, tend to be dominated by flux components with high uncertainty. Apart from this, flux components that owing to atmospheric transport, have a high impact on the concentrations at one or several stations project well on the right-hand singular vectors corresponding to the highest singular values. Those flux components are well observable by the network.

The SVD of M is carried out by a library routine from NAGLIB [*NAGLIB*, 1987]. Figure 5 shows the spectrum of singular values of M in descending order. Except for the first singular value, which is by 3 orders of magnitude higher than the second one, the spectrum is concentrated on a relatively small interval on the positive axis: The difference between the second largest and the smallest singular values is less than three orders of magnitude. In particular, none of the singular values is zero, because the accuracy of our routine is higher than these 5–6 orders of magnitude. Hence the following three equivalent statements hold: (1) Our matrix M has full rank, i.e., the rows are linearly independent. (2) The space spanned by the right-hand singular vectors is perpendicular to the null space of M (denoted by $N(M)$) being defined as the subspace of the space of fluxes formed by all f with $Mf = 0$. (3) The range of M is the entire space of concentrations. Using the terminology of *Menke* [1989], our inverse problem is not overdetermined, since for any vector of observations at the stations we can find a flux vector that satisfies (1), i.e., that yields a consistent vector of simulated concentrations. In other words, the observations cannot contradict each other. Yet, of course, the problem is underdetermined: Together with one flux field satisfying (1) comes a whole $n_f - n_c$ dimensional subspace of flux vectors that satisfy (1) (All flux vectors satisfying (1) differ by a vector in $N(M)$).

In the SVD, all quantities of interest are expressed most naturally and most conveniently in natural units: Exploiting that U^TU, UU^T, V^TV, and the covariance matrices are identity matrices, the model resolution matrix of (7) can be represented by

$$\begin{aligned} R_m &= (1 + M^TM)^{-1}M^TM \\ &= (1 + VD^2V^T)^{-1}VD^2V^T \\ &= V\frac{D^2}{1+D^2}V^T, \end{aligned} \tag{12}$$

because $(1 + VD^2V^T)^{-1} = V\frac{1}{1+D^2}V^T$ holds in the range of V. According to Eq. (7), the posterior covariance matrix then takes the form

$$C_f^{'} = 1 - R_m = 1 - V\frac{D^2}{1+D^2}V^T \ . \tag{13}$$

In the directions of the right-hand singular vectors corresponding to the highest singular values, the uncertainty of the fluxes is most efficiently reduced. As explained above, those singular vectors are dominated by flux components with high uncertainty. This is consistent within the Bayesian framework, because it is easier to improve the degree of knowledge about those components that a priori are most uncertain. In $N(M)$, the uncertainty is not reduced at all. Of course, it is interesting to see which directions in the space of fluxes are constrained by the observations and how well they are constrained. The first singular value is about 28,400. The corresponding right-hand singular vector is dominated by a component of 1000 ppmv for the correction of the global mean concentration at the beginning of the target period, which we introduced as additional unknown in section 3; all flux components are close to zero with a global annual mean flux of -1.3 $\times 10^{-4}$ GtC. The corresponding left-hand singular vector consists of a uniform concentration of about 3.52 $\times 10^{-2}$ ppmv; the trend component is extremely small. These numbers are consistent: Multiplying M by this right-hand singular vector according to (11) yields 3.52 $\times 10^{-2}$ ppmv $\times$ $28,400 \approx 1000$ ppmv, which corresponds to the 1000 ppmv correction of the initial concentration component. According to (13), the uncertainty in the direction of the first right-hand singular vector is about 1/(1+28400^2) $\approx 10^{-9}$ natural units. The second singular value is 12. The component of the global mean concentration is –0.2 ppmv; the annual mean flux (see Figure 6) is positive in the southern hemisphere and negative in the northern hemisphere with a global mean of –0.6 GtC. The concentration is positive at the southern hemisphere stations and negative at the northern hemisphere stations; the trend component is only -0.02 ppmv a^{-1}. This singular value is mainly associated to the north-south gradient of the concentration. Its posterior uncertainty is of order 10^{-2} natural units. The third singular value is 4.9. The left-hand singular vector is dominated by a trend component of 0.12 ppmv a^{-1}, which is large given the low a priori uncertainty of 0.16 ppmv a^{-1}. On the flux side this corresponds to a sink that in natural units is rather

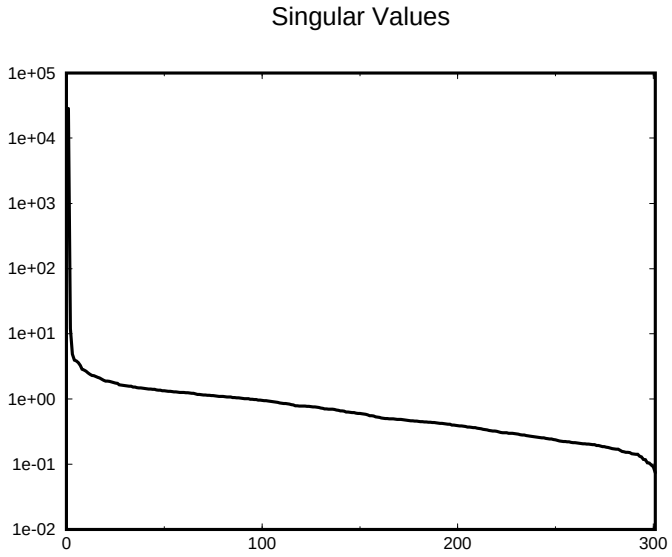


Figure 5. Spectrum of singular values.

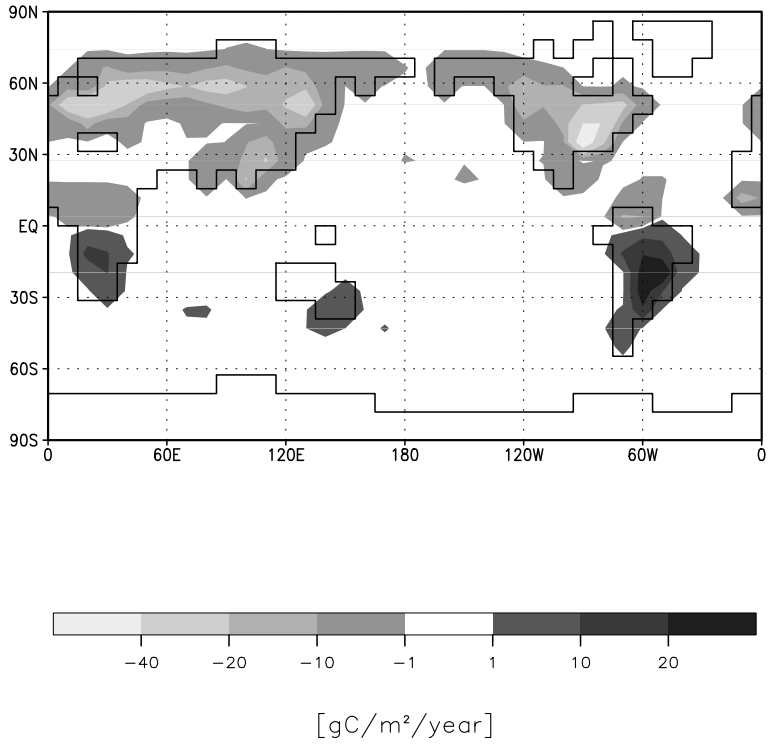


Figure 6. Annual mean fluxes computed from the second right-hand singular vector.

uniform and, hence, in original units exhibits the structure of the a priori uncertainties; the initial concentration component is small. Finally, the seasonal pattern on the concentration side shows the response to the seasonality of the prior

flux uncertainties. The next few singular values are primarily associated with the seasonal cycles in the fluxes and the concentrations. The remaining singular values range from 4.0 to 0.074, so that their posterior uncertainties range from 0.040 to 0.99 natural units. A detailed discussion of all singular values and vectors is far beyond the scope of this paper.

According to (3) and (4), the inverse is

$$M^{-1} = V \frac{D}{1+D^2} U^T \quad : \tag{14}$$

A misfit between observed and modeled concentrations yields a correction in the subspace of fluxes that is perpendicular to $N(M)$. The atmospheric data do not add any information to $N(M)$. A misfit in the direction of a left-hand singular vector that is associated to a low (high) singular value yields a large (small) correction of the fluxes in the direction of the associated right-hand singular vector. For the largest, the second largest, and the smallest singular value, these amplification factors $D/(1+D^2)$ take the values 3.52 $\times 10^{-5}$, 9.01 $\times 10^{-2}$, and 7.39 $\times 10^{-2}$, respectively.

Without the stabilizing effect of the a priori information, being reflected by the 1 in the denominator of (14), the amplification factor would be the pure reciprocal of the singular value. Systematic errors projecting well on the left-hand singular vectors associated to small singular values would be subject to tremendous amplifications. In this situation, the spectrum of singular values usually is truncated to get rid of these "nuisance" directions at the cost of reducing the subspace of the space of fluxes that can influence the concentrations [*Menke*, 1989; *Enting*, 1993; *Brown*, 1995].

In the direction of a particular right-hand singular vector, the reduction of variance and the adjustment of the fluxes by the inversion are coupled through the corresponding diagonal factors in the SVD. The adjustment, however, also depends on the misfit in the direction of the corresponding left-hand singular vector. If this misfit is small, the variance is reduced without any adjustment.

6. A Posteriori Fluxes and Reduction of Uncertainties

The a posteriori probability density in the space of fluxes is Gaussian (see section 2) and thus can be represented by its mean and its covariance matrix. The dimensions of our problem make a discussion of the full covariance matrix difficult, similar to the discussion of all singular vectors in section 5.

Single components of the flux vector, spatial or temporal means, however, are easier to discuss. Since they are derived from the flux vector by linear projections, their one-dimensional probability densities are Gaussian as well. The centers of these densities are our best estimates of the respective quantities; we will refer to these centers as posterior values. The variances of these densities quantify the uncertainties of the respective estimates. The square roots of the variances are the standard deviations; we will refer to them as posterior uncertainties. The appendix gives a recipe for efficient computation of the posterior uncertainties for these projections. One must keep in mind that a loss of information is the cost of this compression by projections.

6.1. A Posteriori Fluxes

Figure 7 shows the a posteriori sum of the terrestrial and oceanic flux components. The predominant feature is the seasonality of the land components: On the northern hemisphere, beginning in May at the midlatitudes and in June at the high latitudes, the terrestrial biosphere acts as a sink. From September in the high latitudes and October in the mid latitudes, in contrast, CO_2 is released by the biosphere. The fluxes over India exhibit a different seasonality: They are positive from April to August and negative from September to February. In the Tropics there is release in winter and uptake in summer and autumn: Between the northern and the southern hemisphere the phases are shifted by 6 months. In the South American mid latitudes the phase of the fluxes is opposite to the one in the northern mid latitudes. Australia has a peculiar seasonality: Its phase is similar to the northern hemispheric phase (see section 6.2). Over the ocean the seasonality is less pronounced. In the Southern Ocean, there is uptake from November to February and a slight release from April to July. From December to May the North Atlantic is a slight sink.

In the annual mean (Figure 2) there is a substantial terrestrial sink in the northern mid latitudes contrasted by a small source in the northern high latitudes. The African tropics are a sink, while the South American tropics exhibit a spatially alternating source-sink pattern. Australia is a small sink. The ocean takes up CO_2 in most regions. The Equatorial Pacific, however, is a strong source, and the South Atlantic and the Southern Ocean at high latitudes are a smaller source. In addition, localized sources are induced in the neighborhood of some of the stations (Cape Meares in Oregon, Point Barrow in Alaska), and localized sinks are induced around Cape Grim in Tasmania and Hawaii. Figure 8 shows the zonal means; the global annual mean sink is 2.3 ± 0.3 GtC.

6.2. Difference to a Priori Fluxes

To understand the behavior of the inversion procedure in detail, it is convenient to consider the cost function introduced in (6). The optimal flux field minimizes the sum of two contributions: the deviation of the posterior fluxes

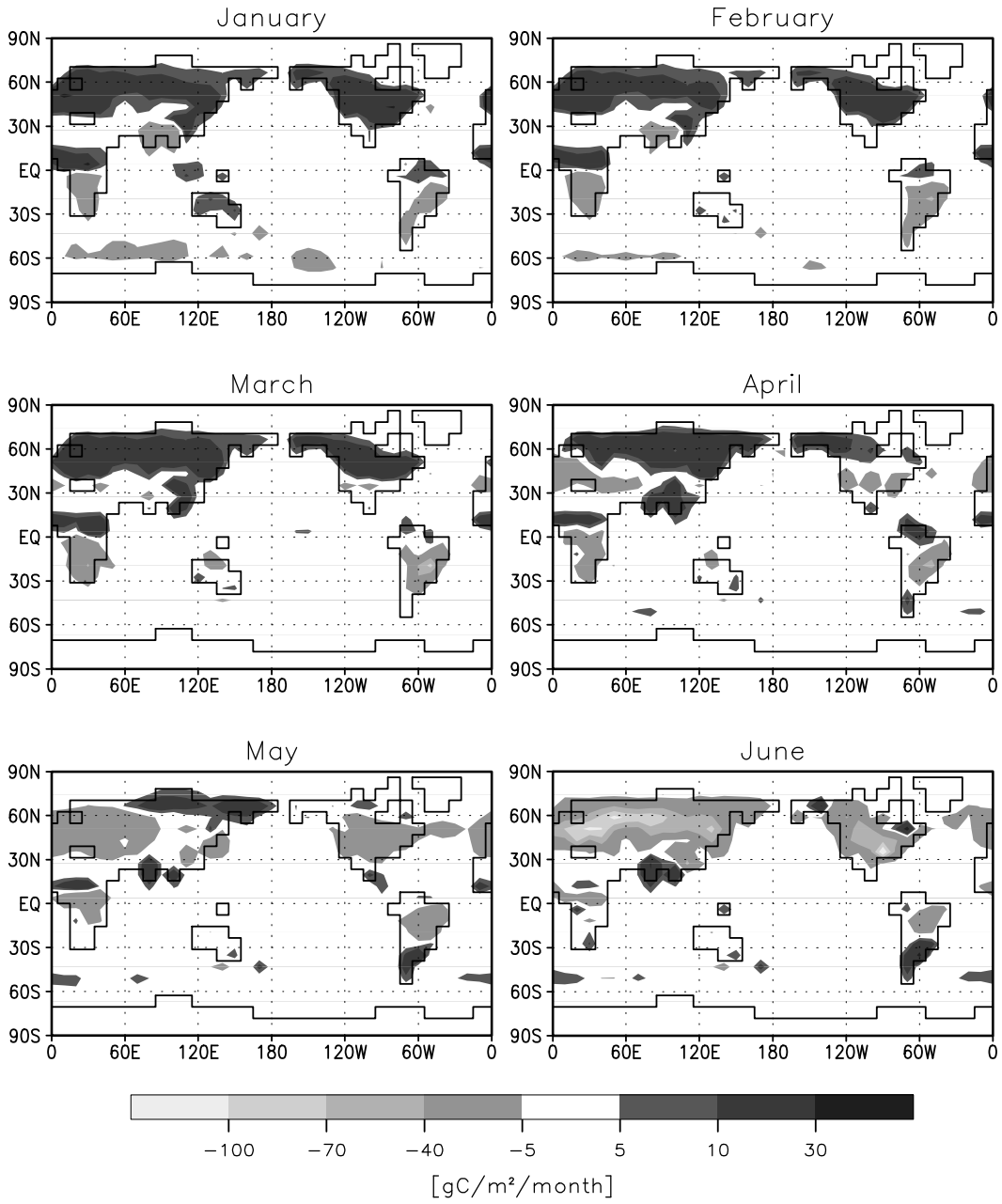
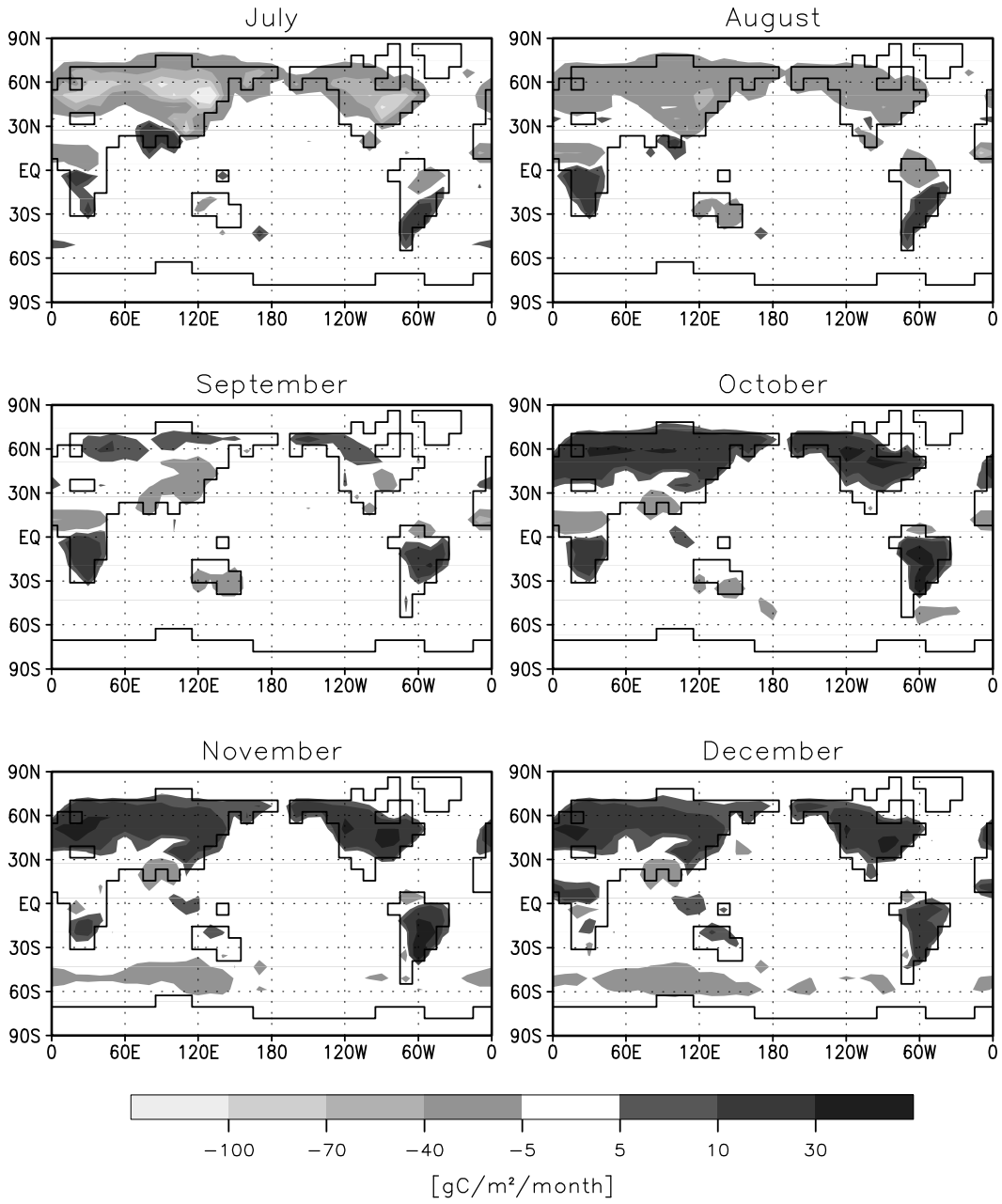


Figure 7. A posteriori sum of the terrestrial and oceanic flux components.

Figure 7. (continued)

from the prior fluxes and the misfit between modeled and observed concentrations, where the prior uncertainties are weighting factors. For a number of selected stations, Figure 4 shows the observed quasi-stationary seasonal cycle as well as the simulations with the prior and the posterior fluxes. In Figure 9, the difference of the posterior and prior fluxes is displayed. A priori, since the ocean uptake and the land use change flux cancel each other, the global annual mean flux is 0.0 ± 1.5 GtC. To match the global trend, however, a net sink of 2.3 GtC is needed. To achieve



this, the inversion procedure tends to reduce those flux components with relatively high uncertainty, because the corresponding deviation of the prior fluxes has a small weight in the cost function. According to the spatial distribution of the prior uncertainties (see Figure 3), this criterion favors primarily terrestrial locations for adjustments. At which locations the inversion actually performs adjustments also depends on the spatio temporal variations in the mismatch between observed and simulated concentrations.

In January 1981, the beginning of the target period, the inversion yields a global mean concentration of 338.9 ± 0.1 ppmv, which is by 0.4 ppmv smaller than the prior value. This difference can be clearly attributed to the higher weight of the northern hemisphere in the computation of the prior initial global mean concentration, which is caused by inhomogeneity of the network. Owing to an extremely high uncertainty of 1000 ppmv, the inversion procedure was essentially free to choose the initial concentration to match the

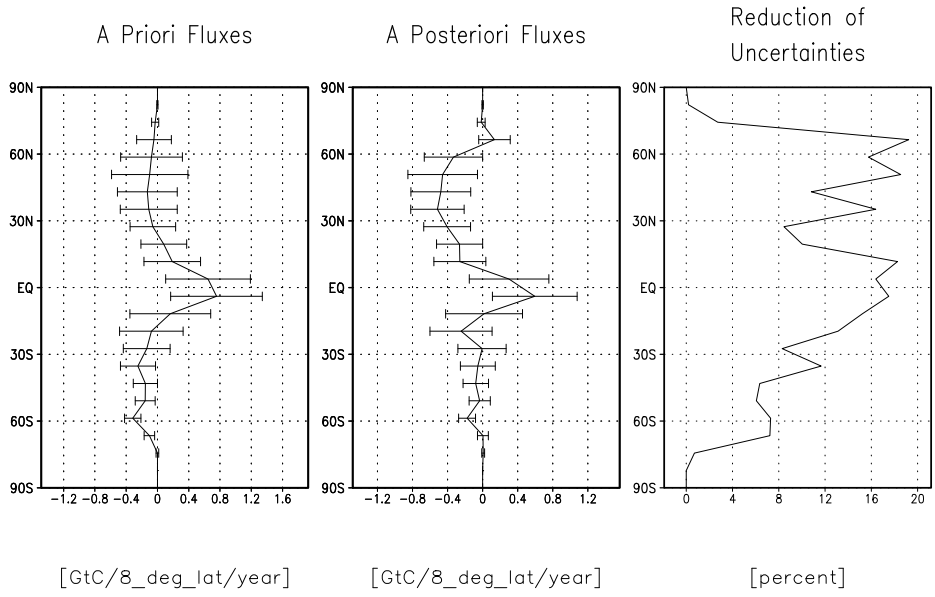


Figure 8. A priori and a posteriori uncertainties for the zonal and annual mean of the sum of the biospheric and oceanic flux components and the difference of their quotient from 1 in percent. Values close to 100 quantify small posterior uncertainty.

observations. Even after subtracting this small offset from the concentrations resulting from the prior fluxes (see Figure 4), in the northern hemisphere these concentrations are too high. In contrast, at the southernmost stations these concentrations are slightly too low. To flatten this north-south gradient, the inversion procedure enhances the net sink in the north and reduces the net sink in the south, which is obvious from a comparison of the posterior and prior zonal mean fluxes that are depicted in Figure 8. Because of the corresponding high a priori uncertainty, the largest adjustments (see Figure 2) are performed in terrestrial regions with high NPP, soil respiration, or land use change fluxes. A fraction of this adjustment also can be attributed to an enhanced uptake by the northern oceans and a reduced sink in the southern oceans, which will be discussed in section 8. These large scale adjustments have a seasonal structure (see Figure 9), which considerably improves the fit to the seasonal cycles in the concentration on hemispheric scale (see Figure 4): in the northern hemisphere midlatitudes, the terrestrial sink in May is reduced but enhanced in June and July, and the enhancement of the sink in the southern ocean takes place from October to January.

These adjustments that predominantly improve the fit on hemispheric scale tend to take place in large areas that are not well observed by the network. In contrast, localized adjustments in well observed areas tend to influence predominantly the concentration at particular observational sites. For instance, at a couple of stations, in the phase and the amplitude of the seasonal cycle there are mismatches between observed concentrations and those concentrations simulated with the prior fluxes. To improve the match, it is optimal, in terms of the cost function ((6)), to correct the fluxes locally at the grid cells and months in which the impact of the correction on the mismatching concentration is strongest. For a few stations and months, the impact has been depicted by *Kaminski et al.* [this issue]. For example, at point Barrow in Alaska (BRW) for the prior fluxes the resulting summer draw down in the concentration is early by about 1 month compared to observations. This yields strong mismatches in June and July, for which the inversion compensates by a correction of the fluxes. Inspection of the Jacobian matrix [*Kaminski et al.*, this issue] yields that, for the June mean concentration, a June flux correction in the few grid cells around BRW and slightly east of BRW has the largest impact. The northern grid cells are oceanic and have a smaller prior uncertainty, while the southern grid cells are terrestrial and have a much higher uncertainty. The difference in the uncertainties are so large, that adjusting primarily the southern grid cells is optimal. For equal uncertainties, however, equally distributed adjustments in all grid cells with high impact would yield a smaller sum of squared adjustments than unequally distributed adjustments and thus would be optimal: A least squares fit in general tends to smooth. For the July mean, the situation is similar. Possible reasons for

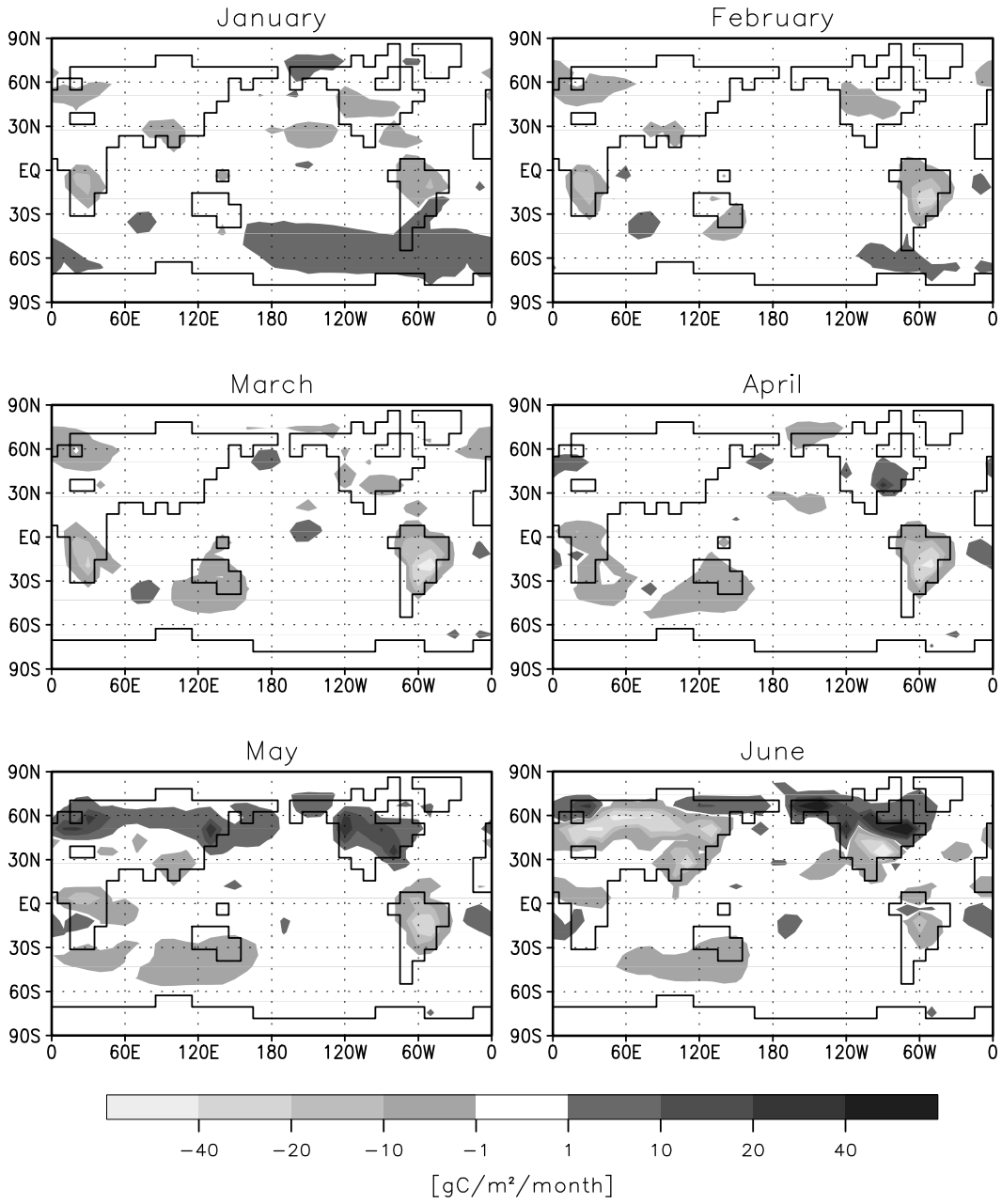
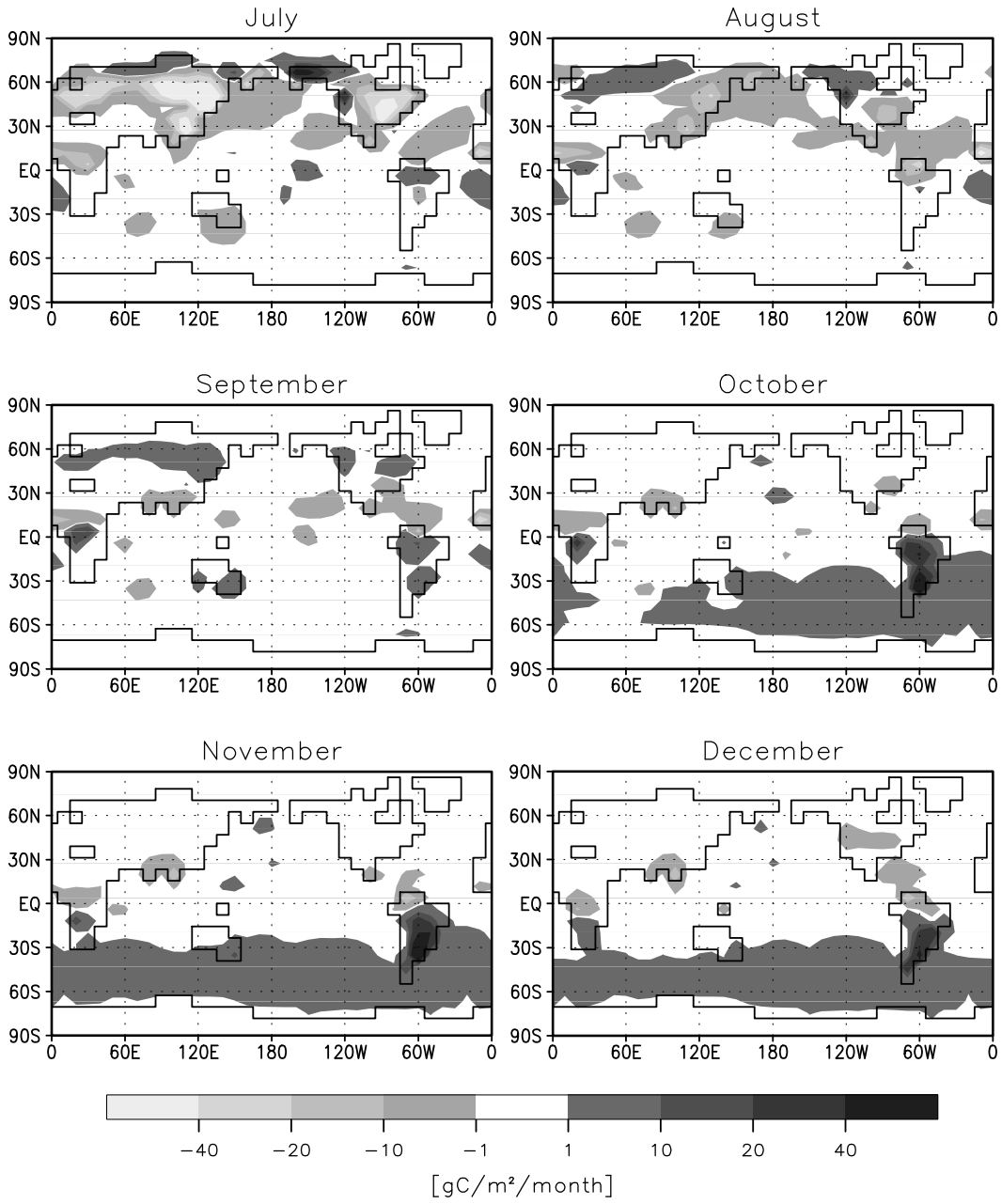


Figure 9. Seasonal cycle of the difference of the a posteriori and a priori flux estimates; positive values quantify an enhanced source or a reduced sink.

Figure 9. (continued)

the mismatch of concentrations are inaccuracies in the prior fluxes, in our model, or in the observations. In their publication, *Knorr and Heimann* [1995] name a number of possible reasons for an overestimation of the length of the summer draw down in their model. Furthermore, the satellite data used by the SDBM are less accurate in the high latitudes.

Concerning the model of the atmospheric transport, *Rehfeld* [1994] has performed a number of simulations for ra-



dioactive tracers with a 19-layer version of TM2. He found some disagreements between modeled and observed concentrations in the polar regions, which he traces back to deficiencies in the meteorological data or reduced numerical stability at very high latitudes due to smaller size of the grid cells.

Another possible reason for local mismatches is associated with the baseline selection procedure. This problem has been addressed by *Ramonet and Monfray* [1996], e.g., for the station at Cape Grim in Tasmania (CGO). To take samples being representative for large-scale air masses at the station observational data are rejected, whenever they are likely to be influenced by fluxes from southern Australia. Baseline conditions are defined according to criteria such as the weather regime or wind direction and speed at the station. *Ramonet and Monfray* [1996] successfully reduced the misfit by mimicking the baseline selection proce-

dure in a high resolution version (2.5^o by 2.5^o in the horizontal) of TM2. The modeled concentration became lower from March to August and higher from September to November. As depicted in Figure 9 around the south of Australia, our inversion procedure reduces the fluxes from February until August, while the fluxes are increased from September to November. Since in our model we did not mimic the selection procedure, at least a part of the correction of the fluxes compensates for this deficiency of our model. For Cape Meares in Oregon, they reported the same phenomenon. Unfortunately, they did not study the records of BRW.

Another problem of our model is its poor resolution of the planetary boundary layer and of the diurnal cycles of both turbulent vertical mixing and the biospheric fluxes. These diurnal cycles are covarying, because turbulent vertical mixing and photosynthetic activity are both driven by solar radiation. Using a model with a more sophisticated parameterization of turbulent vertical transport, *Denning* [1995] and *Denning et al.* [1995] found a significant contribution to the annual mean concentration from diurnally varying, annually balanced fluxes. Any quantification of the real magnitude of this contribution from observations is difficult. Since our model uses monthly mean fluxes, and the meteorological data driving the transport are only available every 12 hours, it can only simulate the fraction of this contribution that is due to covariances on the seasonal time scale and thus is likely to miss a 20% fraction as estimated by *Denning* [1995] with his model. Another advantage of a finer resolution of the vertical structure of the tracer concentration and its diurnal cycle is the opportunity to better represent the sampling process at the station locations, because the samples are usually taken at a particular time of the day.

6.3. Reduction of Uncertainty

According to (7) the diagonal of the model resolution can be expressed by subtracting the quotient of the variances of the respective components from 1. This reduction of uncertainty quantifies how the additional information from the atmospheric data has improved our knowledge of the fluxes. For the individual flux components (not shown), in general, the reduction of uncertainty by the atmospheric data is very small, reflecting the fact that on this small scale, a sparse network does not provide much information. For a particular flux component, the reduction of uncertainty is high, whenever this component projects well on one of the dominant right-hand singular vectors. As discussed in section 5, this gain of information is high for flux components having a high prior uncertainty or flux components being well observable by the network. For temporal and spatial means, the reduction of uncertainty, i.e., the difference between 1 and the quotient of uncertainties, is a more natural quantity than means of the diagonal of the model resolution matrix. Figure 10 shows the prior and posterior uncertainties for the annual mean as well as their quotient subtracted from 1, while Figure 8 shows the same quantities for the zonal means. Also, for these means the reduction of uncertainty is strongest in well observable areas and areas with high prior uncertainty.

Comparing the reductions of uncertainty for single flux components, annual means, and zonal and annual means, two points are worth noting: First, the average reduction of uncertainty increases significantly with increasing degree of accumulation of components: By the inversion we learn more on larger scales than on smaller scales. Second, although we have high prior uncertainties, even for the 24 zonal and annual mean fluxes, the reduction of uncertainty remains lower than 15%. In contrast, an alternative approach to the inverse problem, e.g., with 24 prescribed zonal and annual mean patterns and the same prior uncertainties, certainly would result in a much better reduction of uncertainty. This lower posterior uncertainty, however, would be to a certain degree artificial, because the additional information simply would be due to coupling the fluxes from many grid cells to flux patterns without allowing variations within the patterns. Simplification of the model by reducing its degrees of freedom would improve the reduction of variance. At the same time, however, the simulation of the concentrations at the stations would become less realistic. An alternative way to reduce the inverse problem's degrees of freedom without simplifying the model is to assume correlated prior uncertainties for the fluxes.

7. Simulated Concentrations

In this section we compare the posterior concentrations, i.e., the concentrations resulting from the posterior fluxes, to observations from two sets of stations: those whose data are used and those whose data are not used in our inversion.

For a number of the stations whose data are included in the inversion, Figure 4 shows the observations as well as the sets of simulated concentrations resulting from the prior and posterior fluxes. The contributions from the initial concentration, the global trend, and the seasonal cycle are used to compose the concentrations of 1981, the first year of our target period. In general, the posterior fluxes yield concentrations being consistent with the observations. On one hand, this can be interpreted as a consequence of the underdeterminancy of the inverse problem and the small weight of the prior flux estimates. On the other hand, the good fit indicates that there are no serious contradictions within and between both the observations and the a priori information on the fluxes. In detail, most of the remaining mismatch is due

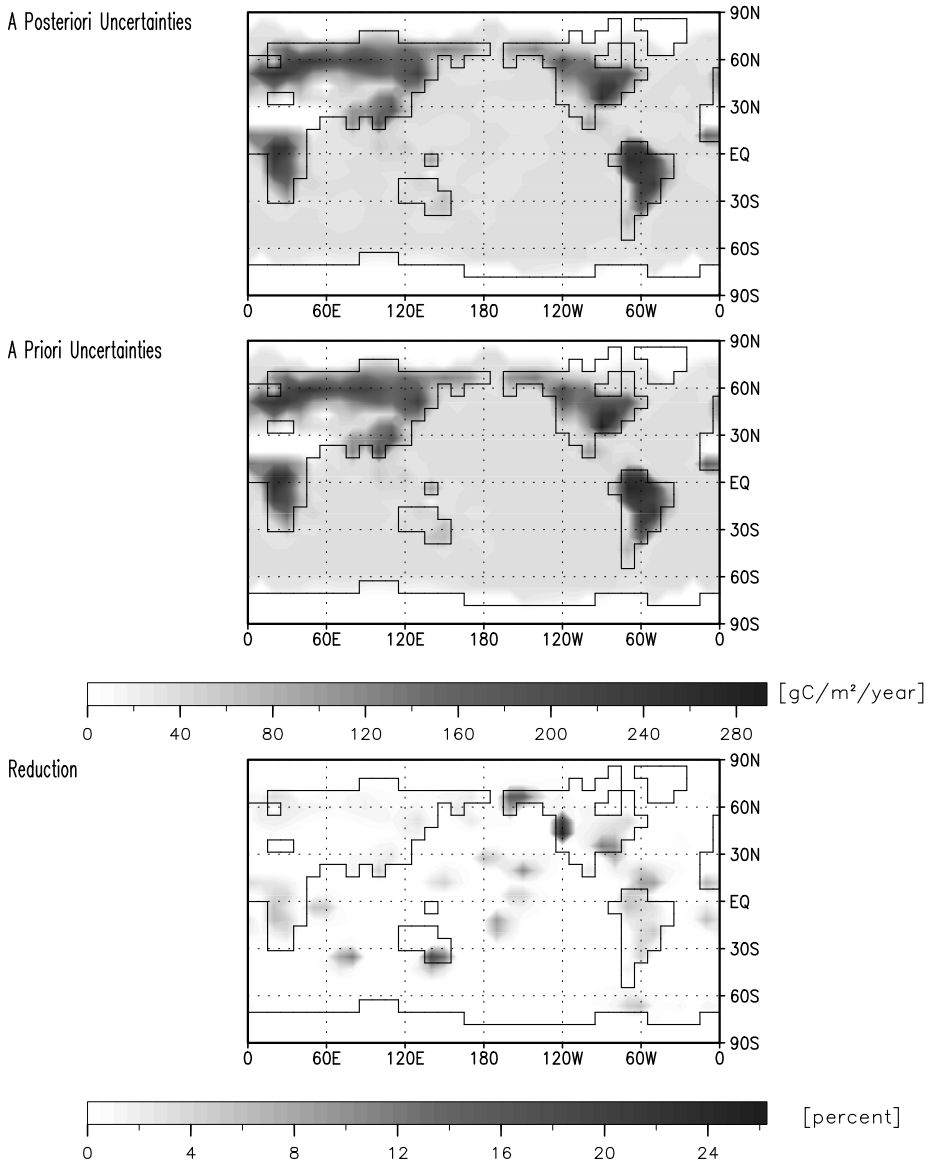


Figure 10. A priori and a posteriori uncertainties for the annual mean of the sum of the biospheric and oceanic flux components and the difference of their quotient from 1 in percent. Values close to 100 quantify small posterior uncertainty.

to a too weak simulated summer draw down at a number of sites on the northern hemisphere.

In Figure 11 we display a number of stations, whose observations we did not use for the inversion procedure. SYO has been left out by mistake, and the other stations did not make observations during the full target period. Yet, using a statistical model, *Masarie and Tans* [1995] managed to extend these records into our target period. Using data from other stations and other time periods, their data extension procedure constructed pseudodata. Clearly, these pseudodata are not independent from the data we already use, so that we could not include them in our inversion. Owing to the considerable amount of new information that nevertheless is contained, these data provide an opportunity to test our posterior flux fields.

Figure 12 shows the observed and modeled quasi-stationary seasonal cycle for 1981. The observations are composed as described in section 4 and above. The agree-

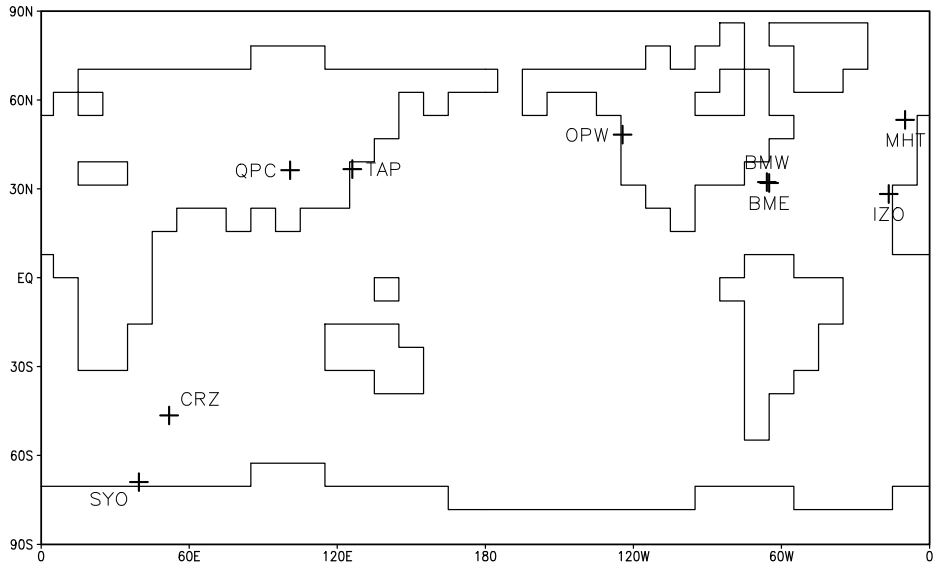


Figure 11. Monitoring stations whose observational data we use to test our a posteriori fluxes.

ment is improved by the inversion at all sites except for the stations QPC and TAP. These stations are mainly influenced by the southeastern part of the Asian continent, which is not well observed by our network (see Figure 2 of *Kaminski et al.* [this issue] and Figure 10), but which has a high prior uncertainty. Hence our inversion is not very sensitive even to larger changes of these fluxes (for wrong reasons). The inversion yields a flux field that over all is consistent with these additional observations although the weight on the observations is high. ’

8. Oceanic and Terrestrial Fluxes

For every grid cell and every month, as discussed in section 6, we can compute a posterior flux field and a posterior uncertainty. To infer information about the processes that control these fluxes, rather than a division into grid cells a division into regions associated to these processes is needed. For example, we would like to separate the oceanic and terrestrial contributions to the fluxes. In this section, first, we give a recipe to perform the necessary bookkeeping for accumulation of flux components. This recipe is then applied to infer the fluxes for a partitioning of the ocean into the regions used in the study of *Tans et al.* [1990], which allows to compare results. Furthermore, to explore the potential of the atmospheric observations to constrain mean fluxes over terrestrial regions we present these means and their uncertainties for some subjectively chosen key areas.

For grid cells crossed by a coastline, in general, there is no way to distinguish between land and ocean contributions. Splitting the flux in proportion to the grid cell’s land fraction provides at least a crude recipe. In many of these grid cells, however, a small oceanic flux is dominated by a much larger terrestrial flux. Hence this crude recipe is likely to yield an unrealistic estimate of the oceanic contribution, while the error for the land contribution in general is much lower. Therefore, to estimate regional oceanic fluxes, we slightly modify our recipe. As described in section 3 a quasi zero terrestrial uncertainty has been assigned to flux components with negligible terrestrial contribution to the flux. Except those flux components, all grid cells with a land fraction of more than 1% are regarded as land grid cells and their flux contribution is neglected in the computation of the regional mean. By this procedure we clearly miss a fraction of the oceanic fluxes, which we try to correct in a second step: Comparison of the posterior and prior flux fields indicates only small differences. To account for the missing fraction of the oceanic flux, we simply scale the posterior regional mean in the same proportion as the prior mean has to be scaled to recover its accurate oceanic fraction.

According to this scaling recipe, we compute the annual mean fluxes for a partitioning of the ocean into six regions defined by *Tans et al.* [1990]. For all regions the prior and posterior values as well as the scaling factor are listed in Table 1. The next column contains estimates derived by Tans et al. on the basis of observed air–sea differences in the partial pressure of CO_2. The following columns are discussed in section 9. The last but one line contains the sum,

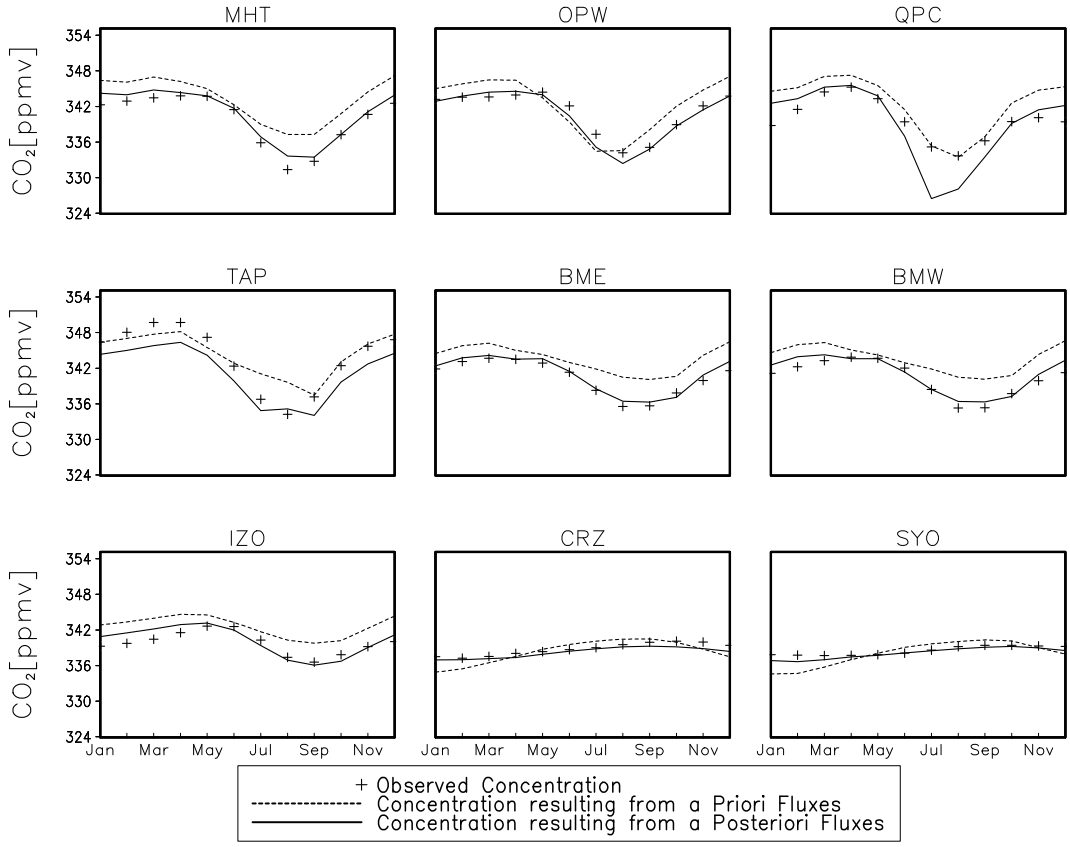


Figure 12. Observed concentration with mean quasi-stationary seasonal cycle composed to represent the first year of the target period, 1981. SYO does not belong to the NOAA/CMDL stations; data at the remaining stations have been extended to target period. These data are not included in our inversion. Modeled concentration resulting from a priori and a posteriori fluxes. Figure 11 shows station locations.

whose uncertainties are derived from the uncertainties of the regional estimates neglecting correlations. The numbers in the last line result from scaling the entire ocean with a single factor derived for the entire ocean and are thus less accurate. The sum of the posterior regional uptakes is by 0.1 GtC higher than the posterior uptake computed by scaling the entire ocean, which is caused by the lower regional scaling factor of the equatorial source as compared to the global scaling factor. The low reduction of uncertainty of about 10% indicates that even on global scale our data are insufficient to distinguish between oceanic and biospheric fluxes.

By means of our inversion we have constructed a global flux field that achieves a high degree of consistency with the atmospheric data. Its ocean uptake is about 1.5 GtC. In contrast, in their study Tans et al. concluded an ocean uptake of less than 1 GtC after comparing the atmospheric response to several flux scenarios. The oceanic flux field of Tans et al. is based on observations of the air–sea differences in the partial pressure of CO_2 during two periods: from January to April and from July to October. After closing spatial and temporal gaps by interpolation, employing an empirically derived expression for the monthly gas exchange coefficient Tans et al. transformed partial pressure differences into the regional CO_2 fluxes in Table 1. Yet combining these ocean flux fields with reasonable land flux fields, their simulated atmospheric response showed a significantly steeper north-south gradient in the atmospheric concentrations as observed, until they also varied the fluxes in the equatorial and southern oceans.

A difference to our study consists in the transport models: Tans et al. used the GISS model instead of TM2. Without performing our inversion with the GISS model, it is not possible to quantify the corresponding a posteriori fluxes. Results of a transport model intercomparison [*Rayner and Law*, 1995; *Law et al.*, 1996] suggest, however, that TM2

Table 1. Uptake/Release of Some Oceanic Regions and the Global Sum for 1981–1986

Location	Prior	Posterior	Scaling	Tans et al.	Mod. Met.	Mod. Net.	Mod. Prior
Atlantic subarctic[a]	–0.17 ± 0.04	–0.18 ± 0.03	1.75	–0.23	–0.19 ± 0.03	–0.18 ± 0.03	–0.18 ± 0.03
Atlantic gyre[b]	–0.17 ± 0.11	–0.33 ± 0.10	1.42	–0.30	–0.31 ± 0.10	–0.34 ± 0.10	–0.37 ± 0.10
North Pacific[c]	–0.55 ± 0.17	–0.72 ± 0.16	1.21	–0.06	–0.69 ± 0.16	–0.73 ± 0.16	–0.81 ± 0.15
Equatorial[d]	0.55 ± 0.28	0.55 ± 0.25	1.05	1.62	0.55 ± 0.25	0.53 ± 0.25	0.60 ± 0.22
Southern gyres[e]	–0.84 ± 0.26	–0.62 ± 0.23	1.19	–2.39	–0.63 ± 0.23	–0.66 ± 0.23	–0.42 ± 0.20
Antarctic[f]	–0.51 ± 0.14	–0.20 ± 0.12	1.03	–0.20	–0.22 ± 0.12	–0.21 ± 0.12	–0.15 ± 0.11
Sum	–1.69 ± 0.46	–1.50 ± 0.41		–1.6	–1.49 ± 0.41	–1.59 ± 0.41	–1.33 ± 0.37
Total ocean	–1.70 ± 0.45	–1.40 ± 0.40	1.26		–1.40 ± 0.40	–1.51 ± 0.40	–1.20 ± 0.32

Prior and posterior values, scaling factor, values derived from observed air–sea partial pressure differences by *Tans et al.* [1990], posterior values inferred in inversions with modified meteorology, network, and prior values. All values are in ppmv a^{-1}, except the dimensionless scaling factors.

[a] 50^oN–90^oN; 90^oW–20^oW

[b] 15^oN–50^oN; 90^oW–20^oW

[c] 15^oN–90^oN;110^oE–90^oW

[d] 15^oS–15^oN;180^oW–180^oE

[e] 50^oS–15^oS;180^oW–180^oE

[f] 90^oS–50^oS;180^oW–180^oE

would yield an even steeper north-south gradient for the scenarios of Tans et al., which even would have amplified the difference to our study. *Taylor et al.* [1991] and *Kurz* [1993] discussed problems in the procedure that Tans et al. used to derive their oceanic fields from measurements of the differences in partial pressure. In contrast, our prior flux field is consistent with the oceanic circulation and yields a partial pressure difference that is in the range of observations [*Kurz*, 1993]. Adjusting slightly this prior ocean flux field, our objective search algorithm succeeds in finding a flux scenario that is consistent with both the atmospheric observations and the partial pressure differences.

The ocean carbon cycle model providing the a priori flux fields does not simulate any significant north south transport of carbon by the thermohaline circulation [*Weber*, 1996]. By enhancing the oceanic sink in the northern hemisphere and reducing the oceanic sink in the southern hemisphere (see Table 1), the inversion suggests such a transport, confirming the conclusions of *Keeling et al.* [1989]. The magnitude of our global oceanic sink, however, is smaller than the 2.3 GtC they inferred for 1984. The location of our Atlantic sink is slightly farther south than expected. This, however, can be caused by our small uncertainty on the prior estimate for the "Atlantic subarctic" region. With 0.5 GtC our North Atlantic sink (15^oN–90^oN) is comparable to the results of *Rayner et al.* [1999a] (0.6 GtC) and *Bousquet* [1997] (0.7 GtC). Also the Antarctic sink (90^oS–50^oS) of 0.2 GtC is the same as that of Bousquet and only 0.1 GtC stronger than that of Rayner et al. Probably our posterior fluxes would become even more similar to theirs, if we increased the assumed uncertainties for the oceanic a priori fluxes to the values assumed in these studies (see section 3).

Assuming a global yearly fossil fuel emission of 5.3 GtC, the inversion reveals an oceanic sink of 1.5 GtC and a total sink of 2.3 GtC, so that the terrestrial biosphere has to account for the residual of 0.8 GtC. For some countries and continents Table 2 opposes industrial emissions to prior and posterior magnitudes of the biospheric sink as computed according to the simple recipe described above. Although the recipe tends to overestimate the biospheric uptake by including a fraction of the oceanic sink, none of the countries or continents can compensate its emissions. Maybe Australia is an exception. Part of the Australian sink, however, can be attributed to the failure of our model to mimic the baseline selection as discussed in section 6. The former Soviet Union and China appear to have taken up at least half of their respective fossil emissions. The inaccuracy of the simple recipe is illustrated by the last line: The global prior yearly biospheric flux is underestimated by 0.5 GtC = 1.7 GtC – 1.2 GtC (a priori value is 1.7 GtC from land use change), and the global posterior yearly biospheric flux is underestimated by 0.2 GtC = – 0.8 GtC – (– 1.0 GtC) (a posteriori biospheric

Table 2. Net Land Uptake and Release of Some Countries and Continents and in the Global Mean for 1981–1986

Country/Continent	Prior	Posterior	Fossil fuel	Mod. Met.	Mod. Net.	Mod. Prior
United States	0.01 ± 0.37	-0.16 ± 0.27	1.09	-0.16 ± 0.28	-0.14 ± 0.27	-0.01 ± 0.14
Australia	0.01 ± 0.11	-0.02 ± 0.10	0.06	-0.02 ± 0.10	-0.03 ± 0.10	-0.03 ± 0.08
China	0.07 ± 0.29	-0.38 ± 0.25	0.58	-0.30 ± 0.26	-0.35 ± 0.26	-0.11 ± 0.11
Europa	-0.02 ± 0.22	-0.08 ± 0.20	1.13	-0.10 ± 0.20	-0.06 ± 0.20	-0.08 ± 0.11
USSR	0.03 ± 0.47	-0.54 ± 0.32	0.99	-0.57 ± 0.32	-0.53 ± 0.32	-0.37 ± 0.18
India	0.02 ± 0.13	-0.02 ± 0.12	0.12	-0.02 ± 0.12	-0.01 ± 0.12	-0.03 ± 0.07
Total land	1.18 ± 1.33	-1.04 ± 0.53	5.27	-1.04 ± 0.53	-0.96 ± 0.53	-0.75 ± 0.31

Prior and posterior values, fossil emissions, posterior values inferred in inversions with modified meteorology, network, and prior values. All values are in ppmv a^{-1}.

flux via the budget is –3.0 GtC + 5.3 GtC – 1.5 GtC = – 0.8 GtC).

9. Sensitivity Experiments

For all posterior flux estimates the Bayesian approach enables us to compute uncertainties quantifying our posterior state of information. These posterior uncertainties are inferred from the prior uncertainties of fluxes and observations using our knowledge about the transport. The reliability of these ingredients and our ability to formalize our information in mathematical expressions determine to which degree our posterior state of information reflects reality, and in particular whether the true values corresponding to our posterior estimates are likely to be in the range specified by their posterior uncertainties. Clearly, these posterior uncertainties only reflect the fraction of uncertainty resulting from factors that we managed to incorporate in our inversion procedure.

While some of these sources of uncertainty not incorporated in our inversion procedure, such as the deviations from the Gaussian assumption, cannot be handled by our inversion procedure, others, such as errors in the transport model, can be investigated by feeding our inversion procedure with a different set of numbers. To explore at least the latter type of uncertainty we perform three inversions, in each of which we vary a particular subset of the numbers we provide to the inversion procedure: the transport matrix, the observational network, or the a priori information.

Our transport matrix represents a TM2 setup driven by meteorological data from 1987. To explore the sensitivity of the posterior fluxes to the transport matrix, we performed an inversion, for which we replace our matrix by a matrix derived with meteorological data from 1986. Comparing the transports of the El Niño year 1987 to that of the "ordinary" year 1986 can be expected to illustrate range for possible changes in the posterior fluxes that can be achieved by changing the year of meteorological data. In the annual mean there are differences on continents north of 40^oN as well as in the El Niño influenced regions. Compared to the posterior uncertainties (Figure 10), however, these differences are slight. This indicates the success of our attempt to include the uncertainty caused by the interannual changes in transport, which is described in section 4. On larger spatial scales, the differences remain low as well: comparing columns 4 and 7 of Table 1 or columns 3 and 5 of Table 2 reveals that the posterior uptake by the oceanic regions and by the selected countries or continents change by less than 10%. The single exception is China, whose yearly biospheric uptake is reduced from 0.38 to 0.30 GtC. Changes in the zonal mean are small, too. The total ocean uptake remains 1.5 GtC.

For their study *Tans et al.* [1990] excluded the data from the stations CMO, NWR, MLO, and RPB (see Figure 2 of *Kaminski et al.* [this issue]). To explore the sensitivity of our posterior fluxes to slight changes in the observational network we perform an inversion for their network. Comparing the annual means of the posterior fluxes, besides a strong local change around CMO, there are slight changes in North America, Asia, Africa, and even South America. Again, the changes in all oceanic regions defined in Table 1 (columns 4 and 8) and in all countries or continents named in Table 2 (columns 3 and 6) remain lower than 10%. Changes in the zonal mean are small, too. The total oceanic sink is slightly enhanced to 1.6 GtC, the fit of the observations is equally good.

Replacing the a priori information on the land fluxes by a more simple formulation, we explore the sensitivity of our posterior fluxes to changes in the a priori information. In

contrast to our standard case, the a priori estimate for the land flux is formed simply by the fields from the SDBM, i.e. we do not account for land use change. The prior uncertainties are based on the net exchange fluxes, rather than on the individual contributions of NPP and soil respiration as in our standard case: We use the absolute value of a flux component, whenever it is larger than 0.12 kg m^{-2} a^{-1}, which is about the value of a large oceanic flux. For most of the remaining flux components we assume an uncertainty equal to this value of 0.12 kg m^{-2} a^{-1}. For grid cells covered by deserts or ice, however, we do not want to assume large uncertainties. In order to exclude land grid cells without vegetation we used the net primary productivity computed by the SDBM. If its annual value is less than 0.01 kg m^{-2} a^{-1}, in the respective grid cell for all months we assume an uncertainty of 10^{-12} kg m^{-2} a^{-1}. Components with such a small prior uncertainty are essentially treated as constant by the inversion procedure, i.e. the a priori value is hardly changed. For ocean grid cells permanently covered by ice we assume as well the extremely low uncertainty of 10^{-12} kg m^{-2} a^{-1} in every month. Our criterion for identifying these grid cells is an annual oceanic flux of less than $5 \cdot 10^{-4}$ kg m^{-2} a^{-1}.

In a previous study [*Kaminski et al.*, 1998] we performed the inversion including the a priori information described above and discussed the a posteriori fields in detail. For the annual mean, Figure 13 shows the difference between the a posteriori fluxes for modified a priori information and the a posteriori fluxes for our standard case. In contrast to the two sensitivity experiments discussed above, Figure 13 reveals large differences between both posterior flux fields: By changing the a priori information, i.e., the spatial distribution of the prior uncertainty and the missing land use change contribution, terrestrial sources and sinks are shifted. This is also quantified by comparison of columns 4 and 9 in the ocean Table 1 or comparison of columns 3 and 7 in the land Table 2: The uptake by the Southern gyres is reduced by about one third, that by the USA has vanished, that of China is more than halved, and that of the former Soviet Union is considerably reduced as well. With a slightly reduced oceanic uptake of 1.3 GtC, the fit of the observations is as good as in the standard case.

10. Conclusions and Perspectives

From atmospheric observations at 25 stations, we inferred a cyclostationary flux field on the entire TM2 grid that is consistent with the observed quasi-stationary seasonal cycle during a target period at the beginning of the 1980s. In the inversion, we included a priori estimates of the fluxes to regularize the otherwise underdetermined inverse problem. This underdeterminancy is caused by the sparse network in conjunction with the diffusive nature of the atmospheric transport and is reflected by the poor reduction of the uncertainty for estimates of single flux components. This low reduction of uncertainty, however, is inherent to the problem and not an artifact of the high resolution; it reflects the classical trade off between resolution and variance of inverse problems. Reducing the number of unknowns by prescribing patterns only achieves an apparent reduction of uncertainty, because relations among unknowns are introduced, but the uncertainties of these relations are neglected. For larger-scale quantities such as spatial and temporal means, the reduction of uncertainty is higher.

To infer information about the processes controlling the fluxes, we had to use a few shortcomings: The fossil fuel contribution has been subtracted from the observations prior to inversion, and to untangle oceanic and biospheric fluxes, crude recipes have been applied. Problems of this type can be avoided by introducing, for every grid cell and month, as many unknowns as there are processes of interest. The inversion then distributes the correction of the flux onto the processes according to their respective prior uncertainties. This improved resolution of processes, however, requires the inversion of a matrix, whose size grows linearly with the number of processes per grid cell.

Our posterior estimate of 1.5 $\pm$ 0.4 GtC for the total ocean uptake contradicts the estimate of less than 1 GtC by *Tans et al.* [1990]. Replacing their simple interpretation of observed air–sea partial pressure differences by oceanic a priori information from a model that includes the population dynamics of phytoplankton in conjunction with an objective search algorithm are the main factors our higher estimate can be attributed to. On the other hand our estimate is lower than the 2.3 GtC inferred by *Keeling et al.* [1989] for 1984, although the structure of the sink supports their interpretation of the southward transport of carbon by the thermohaline circulation. On large scale, our posterior flux field is similar to those inferred by *Rayner et al.* [1999a] and *Bousquet* [1997]. Our estimate is not very sensitive to changes in the a priori information on the biospheric fluxes, changes of the network, and changes of the meteorological data that drive our transport model.

The reduction of uncertainty for the global net exchange fluxes with the ocean and the terrestrial biosphere is much lower as desirable. Measurements of additional tracers such as oxygen [*Keeling et al.*, 1996; *Stephens et al.*, 1998; *Rayner et al.*, 1999a], ratios of carbon isotopes in CO_2 [*Ciais et al.*, 1995; *Rayner et al.*, 1999a; *Bousquet*, 1997], or ratios of oxygen isotopes in CO_2 [*Ciais et al.*, 1997a, b; *Peylin et al.*, 1999] have been reported to impose more or less strong constraints on the partitioning between the processes. Since all these tracers are chemically inert, our matrix rep-

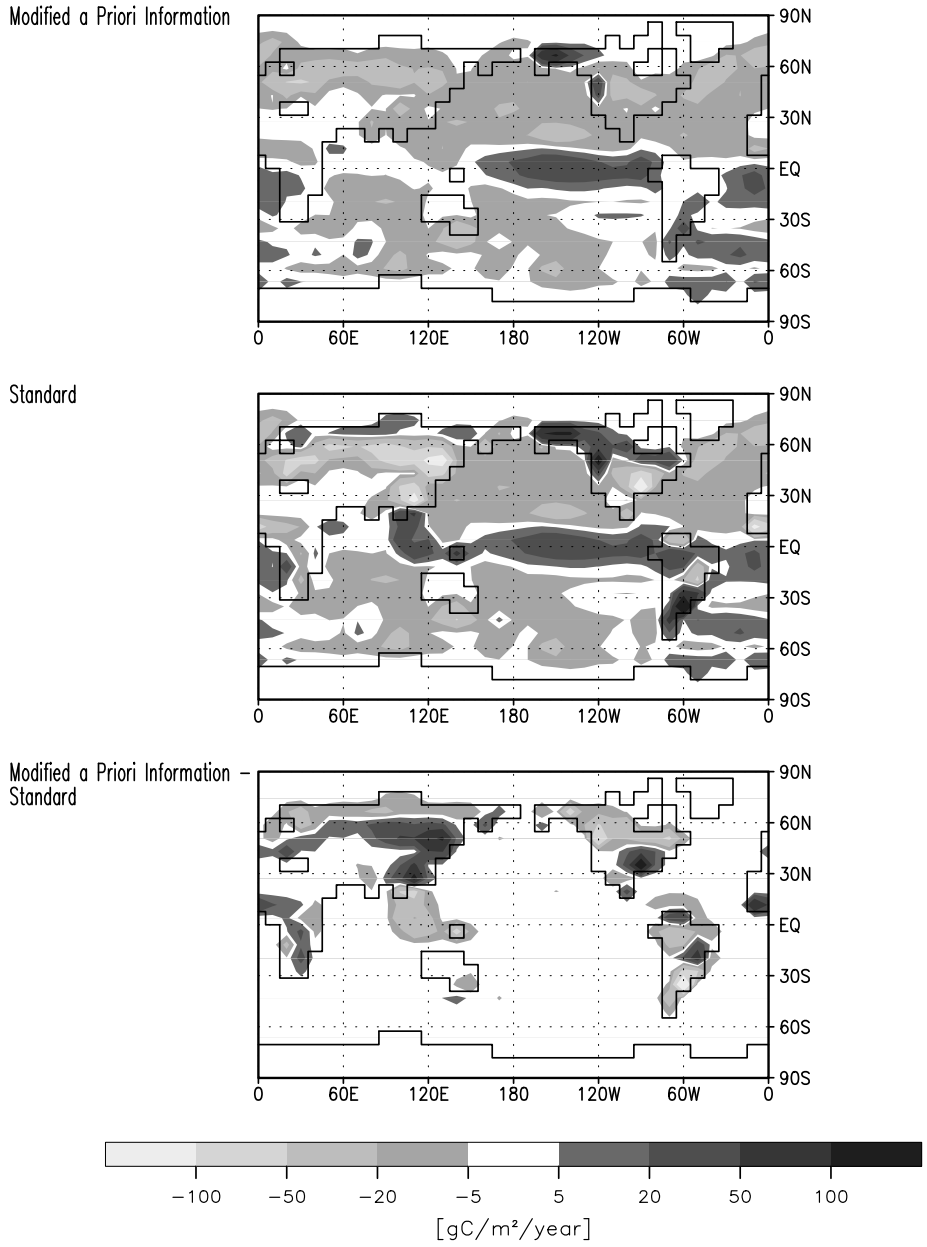


Figure 13. A posteriori estimate of annual mean of the sum of the flux contributions from terrestrial biosphere and ocean; (top) inversion with modified terrestrial a priori information, (middle) standard inversion, (bottom) and their difference; in the difference plot, positive values quantify an enhanced source or a reduced sink due to the modification.

resentation of the transport can be employed to include this additional information in the inversion procedure. In addition, a model of the processes that link these tracers' fluxes to the CO_2 fluxes is required, and the additional uncertainty introduced by the model has to be formalized, so that it can be transformed to an uncertainty for the flux estimates.

For a few countries and continents we estimated the magnitude of the biospheric sink to explore the capacity of the observations to detect the geographic origin of a tracer. In contrast to the resolution of processes, for problems of this type, in general, no further tracer can provide additional information. The only way to reduce the uncertainties is to in-

crease the number of observations as well as their precision. Methods for a systematic investigation of the optimal location of additional observational sites have been presented by *Rayner et al.* [1996]. In conjunction with search algorithms that need to try a high number of potential locations, for computing the atmospheric response at these locations the adjoint approach is clearly inferior to the forward approach, because the cost of the adjoint approach is proportional to the number of locations. The adjoint approach can be efficient, if the number of potential locations can be kept low, e.g., owing to logistic constraints, or if a search algorithm can get along with a small number of trials.

Our sensitivity experiments confirmed that for a sparse network the a priori information on the fluxes constitutes a crucial ingredient of the inversion. For convenience we assumed a Gaussian probability density, which is quantified by mean and covariance matrix. As mean we used output of process models. For the error covariance, however, no model results were available. Hence we invented an error covariance matrix. For simplicity of computation, we have not assumed correlations among the uncertainties of different components, although, especially on this small spatial scales, correlations are likely. Increasing correlations can be interpreted as a way to continuously reduce the degrees of freedom of the inverse problem. Hence an increased correlation would have two effects on our inversion: First, the prior uncertainty of large-scale mean fluxes such as the total ocean uptake would increase, because cancelling out of deviations from the mean with different sign becomes less likely. Second, spatial correlations would tend to couple groups of grid cells. Hence the localized source and sink spots would get less intense and more widespread. In contrast to prescribed spatiotemporal patterns, correlated prior uncertainties can be interpreted as a means to continuously reduce the number of degrees of freedom without neglecting uncertainties. By using output from different process models, the sensitivity of the inversion to the prior estimates can be investigated. Until information on correlated uncertainties from process models is available, the sensitivity of the inversion to this ingredient can only be explored by trying different assumptions on covariance matrices.

In the present study, we characterized the sources and sinks by their net exchange fluxes with the atmosphere, rather than the processes causing the fluxes. After coupling the transport model (or its Jacobian) to process models such as the SDBM [*Knorr and Heimann*, 1995], the corresponding adjoint can be applied to estimate the internal parameters of the process models. Coupling a model of the oceanic carbon cycle to the transport matrix would allow to simultaneously fit oceanic observations such as CO_2 partial pressure and atmospheric observations. Here again formalizing the prior uncertainties in the models is important. As a by product, by running (the linearization of) the optimized process model forward, the parameters in the process model and their uncertainties could be mapped onto the exchange fluxes and their uncertainties.

We have seen that the inversion tends to compensate for biases in our model by erroneous corrections of the fluxes. Hence improvement of the model is desirable. For the transport model, an improved version, TM3, driven by 6 hourly reanalyzed meteorological fields from the ECMWF is now available. It can be run in a finer horizontal and vertical resolution, so that a more realistic representation of the planetary boundary layer is possible. With this transport model, at least in part, the sampling procedure at the stations can be mimicked. Furthermore, covariances in the diurnal cycles of the surface fluxes and turbulent vertical transport can be better resolved. Compared to the forward approach, on a finer grid, the computation of a matrix representation by the adjoint is even more advantageous.

In our model, we assume cyclostationarity for the fluxes and the transport. An estimate of the corresponding model error was added to the observational uncertainty. Not only would it be interesting per se to study interannual variability of the fluxes [see, e.g., *Rayner et al.*, 1999a; *Law*, 1999; *Rayner et al.*, 1999b], also would a more flexible model allow to considerably reduce the uncertainty on the data side, which would improve the reduction of the fluxes' uncertainties. To consider interannual variations in the fluxes only a slight change in the setup is necessary: Instead of prescribing the same fluxes every year, interannual variation during the spin up is allowed. By the corresponding adjoint model, this more general matrix representation quantifying the impact of flux components up to 3 years ago can be derived at the same cost as the matrix for the cyclostationary case. In fact, it was not very smart to run the adjoint for the cyclostationary setup at all, because the general matrix can be easily transformed to the cyclostationary matrix. To include also the interannual variations in the transport, the setup has to be changed toward a simulation of the whole target period. For this case the cost of both the forward and the adjoint approach increases linearly with the length of the target period. In any case, resolving interannual variations in the transport imposes the challenge to computationally handle the inversion of a matrix whose size increases quadratically with the length of the target period. As a preliminary test of the impact of the interannual variations in the transport, the inversion has been performed in the same setup but with the meteorology of a different year: The resulting changes in the posterior fluxes were slight.

In our example, we employed the Jacobian to derive an estimate of the sources and sinks of CO_2. However, the

technique can be efficiently applied to other tracers in the same manner, as long as the number of observations is small compared to the number of source components of interest. Since at our observational sites also the concentrations of further tracers are measured, the same matrix can be used for modeling the quasi-stationary seasonal cycle resulting from those tracer's surface fluxes. (Of course a different conversion factor from mass to concentration has to be taken into account.) If the tracer, in addition, has sources or sinks above the ground, the transport matrix has to be complemented by further columns representing the sensitivity of the modeled concentration at the stations with respect to these additional sources or sinks. The cost for the computation of such an extended matrix is the same as for our matrix, so that compared to forward modeling the adjoint approach is even more advantageous.

The efficient computation of the transport matrix by the adjoint of TM2, which forms the basis of our approach, depends crucially on the sparsity of the network and on the linearity of the transport. For cases with as many observations as flux components or cases with important nonlinearities in the transport, the adjoint model allows an inversion without computing the full transport matrix: At the cost of 3–4 forward model runs, the adjoint can be employed to provide the gradient of the misfit between modeled and observed concentrations (equation (6)) with respect to all flux components. Exploiting this gradient information, most powerful algorithms [see, e.g., *Gill et al.*, 1981; *Press et al.*, 1986; *Tarantola*, 1987] can be applied to iteratively minimize the misfit by variation of the fluxes. If the inverse problem is well posed, these algorithms typically achieve a strong reduction of the misfit in a few iterations. However, although this approach is rather inexpensive, it does not yield reliable estimates of the uncertainties of the fluxes in an inexpensive way.

Adjoint models enable us to tackle efficiently the inversion of the atmospheric transport with an arbitrarily high resolution in the space of fluxes. Compared to many alternative methods adjoint models are a valuable tool for studying sources and sinks on smaller scales. They might have the potential to bridge the gap of scales between local process studies and global budgets. Of course, essential additional ingredients are high-quality atmospheric measurements of a dense network, a good model of the atmospheric transport, and accurate formalization of the available a priori information on the sources and sinks. Adjoint models especially provide a means of inferring anthropogenic trace gas emissions, which might be needed in the near future, e.g., to verify whether national emission targets are kept.

Appendix: Posterior Uncertainties of Spatial or Temporal Means

As described in section 2, our inversion results in a Gaussian posterior probability density in the space of fluxes, which is determined by a mean and a covariance matrix. The space of fluxes is of dimension $n_f \approx 10,000$. Hence the posterior covariance matrix has about $n_f^2 \approx 10^8$ entries, i.e., its size is about 100 MW. Since the matrix is symmetric, about 50 MW are redundant. The matrix could be computed and stored according to (13), but this is demanding in terms of both memory and CPU resources. Yet, since nobody wants to know all of these $^1/_2 \times 10^8$ entries, this computation is not necessary. Instead, what is needed is the posterior uncertainty of quantities defined by projections from the full space of fluxes to, e.g., single components, temporal or spatial means, or on the right hand singular vectors. In this appendix we give a recipe to compute the posterior uncertainties of these quantities without computing the full posterior covariance matrix.

Let P be a projection from the full space of fluxes to a quantity of interest $\tilde{y}$, such as the global annual mean ocean uptake:

$$\tilde{y} = P\tilde{f}. \tag{A1}$$

The uncertainty in the prior and posterior fluxes are

$$\sigma_y = PC_fP^T \qquad \sigma_y^{'} = PC_f^{'}P^T, \tag{A2}$$

where the superscript T denotes the transposed.

Exploiting the representation of $C_f^{'}$ in terms of the SVD of the model resolution matrix (see equation (13)) yields

$$\begin{aligned} \sigma_y^{'} &= PC_f^{^1/_2}(1 - V\frac{D^2}{1+D^2}V^T)C_f^{^1/_2}P^T \\ &= \sigma_y - PC_f^{^1/_2}V\frac{D^2}{1+D^2}V^TC_f^{^1/_2}P^T, \end{aligned} \tag{A3}$$

where $C_f^{^1/_2}$ transforms from natural to original units (definition of square roots of positive definite matrices is straitforward via a representation in terms of their eigenvalues and vectors, which can be found, e.g., in the work of *Tarantola* [1987]).

Since the matrix product is associative, instead of computing first the $n_f \times n_f$ matrix $C_f^{'}$, we can exploit the low dimension $n_c = 301$ of the square matrix $D^2/(1+D^2)$: Defining a $1 \times n_c$ matrix H by

$$H := PC_f^{^1/_2}V(\frac{D^2}{1+D^2})^{^1/_2}, \tag{A4}$$

the uncertainty of $y^{'}$ can be written in the form:

$$\sigma_y^{'} = \sigma_y - HH^T. \tag{A5}$$

Since $(D^2/(1+D^2))^{1/2}$ and $C_f^{1/2}$ are diagonal matrices, computation of H and, hence, computation of $\sigma_y^{'}$ via (A5) is very efficient.

Similarly, posterior covariances $\sigma_{1,2}^{'}$ of quantities $\tilde{y}_1$ and $\tilde{y}_2$ defined by projections P_1 and P_2 can be computed by

$$\sigma_{1,2}^{'} = \sigma_{1,2} - H_1 H_2^T, \tag{A6}$$

with

$$H_i := P_i C_f^{1/2} V \left(\frac{D^2}{1+D^2}\right)^{1/2} \qquad (i = 1,2). \tag{A7}$$

Acknowledgments. The authors thank Ian Enting and Peter Rayner for directing their attention to features that arise owing to the high spatial resolution of the fluxes, such as the impact of correlated prior uncertainties or the opportunity to reduce the bias in flux estimates. Thomas Kaminski thanks Rachel Law and Peter Rayner for many fruitful discussions about the atmospheric transport and its inversion. The authors thank two anonymous reviewers for many helpful suggestions and comments. Michael Voßbeck has produced all GrADS plots. This work was supported in part by the Commission of the European Communities under contract EV5V-CT92-0120. Computing support was provided by the Deutsches Klimarechenzentrum (DKRZ) in Hamburg.

References

Andres, R. J., G. Marland, T. Boden, and S. Bischoff, Carbon dioxide emissions from fossil fuel consumption and cement manufacture 1751 to 1991 and an estimate for their isotopic composition and latitudinal distribution, in *The Carbon Cycle*, edited by T. M. L. Wigley and D. Schimel, Cambridge Univ., New York, in press, 1999.

Bousquet, P., Optimisation des flux nets de CO_2: assimilation des mesures atmosphériques en CO_2 et en δ^{13}C dans un modele de transport tridimensionnel, Ph.D. thesis, Univ. Paris VI, 1997.

Brown, M., Deduction of emissions of source gases using an objective inversion algorithm and a chemical tranport model, *J. Geophys. Res.*, *98*, 12,639–12,660, 1993.

Brown, M., The singular value decomposition method applied to the deduction of the emissions and the isotopic composition of atmospheric methane, *J. Geophys. Res.*, *100*, 11,425–11,446, 1995.

Ciais, P., et al., Partitioning of ocean and land uptake of CO_2 as inferred by δ^{13}C measurements from the NOAA climate monitoring and diagnostics laboratory global air sampling network, *J. Geophys. Res.*, *100*, 5051–5070, 1995.

Ciais, P., et al., A three dimensional synthesis study of δ^{18}O in atmospheric CO_2 , I, surface fluxes, *J. Geophys. Res.*, *102*, 5857–5872, 1997a.

Ciais, P., et al., A three dimensional synthesis study of δ^{18}O in atmospheric CO_2 , II, simulations with the TM2 transport model, *J. Geophys. Res.*, *102*, 5873–5883, 1997b.

Conway, T. J., P. Tans, L. Waterman, K. Thoning, D. Buanerkitzis, K. Masarie, and N. Zhang, Evidence for interannual variability of the carbon cycle from the noaa–cmdl global air sampling network, *J. Geophys. Res.*, *99*, 831–855, 1994.

Denning, A. S., Investigations of the transport, sources, and sinks of atmospheric CO_2 using a general circulation model, Ph.D. thesis, Colo. State Univ., Fort Collins, 1995.

Denning, A. S., I. Y. Fung, and D. Randall, Latitudinal gradient of CO_2 due to seasonal exchange with biota, *Nature*, *376*, 240–243, 1995.

Enting, I. G., Inverse problems in atmospheric constituent studies, III, estimating errors in surface sources, *Inverse Prob.*, *9*, 649–665, 1993.

Enting, I. G., Green's function methods of tracer inversion, in *Geophysical Monograph Series*, edited by P. Kasibhatla, M. Heimann, D. Hartley, P. J. Rayner, N. Mahowald, and R. Prinn, AGU, Washington, D. C., in press, 1999.

Enting, I. G., and J. V. Mansbridge, Seasonal sources and sinks of atmospheric CO_2: Direct inversion of filtered data, *Tellus, Ser. B*, *41*, 111–126, 1989.

Enting, I. G., T. M. L. Wigley, and M. Heimann, Future emissions and concentrations of carbon dioxide: key ocean/atmosphere/land analyses, *Tech. Pap. 31*, CSIRO Div. of Atmos. Res., Aspendale, Victoria, Australia, 1994.

Enting, I. G., C. M. Trudinger, and R. J. Francey, A synthesis inversion of the concentration and δ^{13}C of atmospheric CO_2, *Tellus, Ser. B*, *47*, 35–52, 1995.

Gill, P. E., W. Murray, and M. H. Wright, *Practical Optimization*, Academic, San Diego, Calif., 1981.

Globalview–CO2, *Cooperative Atmospheric Data Integration Project - Carbon Dioxide* [CD-ROM], NOAA/CMDL, Boulder, Colo., 1996.

Hartley, D., and R. Prinn, Feasibility of determining surface emissions of trace gases using an inverse method in a three-dimensional chemical transport model, *J. Geophys. Res.*, *98*, 5183–5197, 1993.

Heimann, M., and C. D. Keeling, A three-dimensional model of atmospheric CO_2 transport based on observed winds, 2, Model description and simulated tracer experiments, in *Aspects of Climate Variability in the Pacific and the Western Americas, Geophys. Monogr. Ser.*, edited by D. H. Peterson, vol. 55, pp. 237–275, AGU, Washington, D. C., 1989.

Heimann, M., et al., Evaluation of terrestrial carbon cycle models through simulations of the seasonal cycle of atmospheric CO_2: First results of a model intercomparison study, *Global Biogeochem. Cycles*, *12*, 1–24, 1998.

Hein, R., and M. Heimann, Determination of global scale emissions of atmospheric methane using an inverse modelling method, in *Non-CO_2 Greenhouse Gases*, edited by J. van Ham et al., pp. 271–281, Kluwer, Norwell, Mass., 1994.

Hein, R., P. Crutzen, and M. Heimann, An inverse modeling approach to investigate the global atmospheric methane cycle, *Global Biogeochem. Cycles*, *11*, 43–76, 1996.

Houghton, J. T., L. M. Filho, J. Bruce, H. Lee, B. A. Callander, E. Haites, N. Harris, and K. Maskell (Eds.), *Climate Change 1995 - The Science of Climate Change: Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge Univ. Press, New York, 1995a.

Houghton, J. T., L. M. Filho, B. Callander, N. Harris, A. Dattenberg, and K. Maskell (Eds.), *Climate Change 1994 - Radiative Forcing of Climate Change*, Cambridge Univ., New York, 1995b.

Houghton, R. A., et al., 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, Ser. B, 39*, 122–139, 1987.

Kaminski, T., M. Heimann, and R. Giering, A matrix representation for an atmospheric transport model computed by its adjoint, in *Air Pollution Modelling and its Application XII*, edited by S.-E. Gryning and N. Chaumerliac, pp. 247–255, Plenum, New York, 1998.

Kaminski, T., M. Heimann, and R. Giering, A coarse grid three-dimensional global inverse model of the atmospheric transport, 1, Adjoint model and Jacobian matrix, *J. Geophys. Res.*, this issue.

Keeling, C. D., M. Heimann, and S. C. Piper, A three-dimensional model of atmospheric CO_2 transport based on observed winds, 4, Analysis of the mean annual gradients of CO_2, in *Aspects of Climate Variability in the Pacific and the Western Americas, Geophys. Monogr. Ser.*, edited by D. H. Peterson, vol. 55, pp. 305–363, AGU, Washington, D. C., 1989.

Keeling, R. F., S. C. Piper, and M. Heimann, Global and hemispheric CO_2 sinks deduced from changes in atmospheric O_2 concentration, *Nature, 381*, 218–221, 1996.

Knorr, W., and M. Heimann, Impact of drought stress and other factors on seasonal land biosphere CO_2 exchange studied through an atmospheric tracer transport model, *Tellus, Ser. B, 47*, 471–489, 1995.

Kurz, K. D., Zur Saisonalen Variation des ozeanischen Kohlendioxidpartialdrucks, Ph.D. thesis, Max-Planck-Institut f. Meteorol., Hamburg, Germany, 1993.

Law, R. M., CO_2 sources from a mass balance inversion: sensitivity to the surface constraint, *Tellus, Ser. B, 51*, 254–265, 1999.

Law, R. M., et al., Variations in modelled atmospheric transport of carbon dioxide and the consequences for CO_2 inversions, *Global Biogeochem. Cycles, 10*, 783–796, 1996.

Maier-Reimer, E., Geochemical cycles in an ocean general circulation model. Preindustrial Tracer Distributions, *Global Biogeochem. Cycles, 7*, 645–677, 1993.

Masarie, K. A., and P. P. Tans, Extension and integration of atmospheric carbon dioxide data into a globally consistent measurement record, *J. Geophys. Res., 100*, 11,593–11,610, 1995.

Menke, W., *Geophysical Data Analysis*, Academic, San Diego, Calif., 1989.

NAGLIB, *Fortran Library Manual - Mark 13*, Numerical Algorithms Group, Oxford, England, 1987.

Peylin, P., P. Ciais, A. Denning, and P. Tans, A three dimensional study of $\delta^{18}O$ in atmospheric CO_2: contribution of different land ecosystems, *Tellus, Ser. B*, in press, 1999.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge Univ. Press, New York, 1986.

Ramonet, M., and P. Monfray, CO_2 baseline concept in 3-d atmospheric transport models, *Tellus, Ser. B, 48*, 502–520, 1996.

Rayner, P. J., and R. M. Law, A comparison of modelled responses to prescribed CO_2 sources, *Tech. Pap. 36*, CSIRO Div. of Atmos. Res., Aspendale, Victoria, Australia, 1995.

Rayner, P. J., I. G. Enting, and C. M. Trudinger, Optimizing the CO_2 observing network for constraining sources and sinks, *Tellus, Ser. B, 48*, 433–444, 1996.

Rayner, P. J., I. G. Enting, R. J. Francey, and R. L. Langenfelds, Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, *Tellus, Ser. B, 51*, 213–232, 1999a.

Rayner, P. J., R. M. Law, and R. Dargaville, The relationship between tropical CO_2 fluxes and the southern oscillation, *Geophys. Res. Lett., 26*, 493–496, 1999b.

Rehfeld, S., Deposition Radioaktiver Tracer in einem Transportmodell der Atmosphäre, Ph.D. thesis, Max-Planck-Institut f. Meteorol., Hamburg, Germany, 1994.

Six, K. D., and E. Maier-Reimer, Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model, *Global Biogeochem. Cycles, 10*, 559–583, 1996.

Snieder, R., Global inversions using normal modes and long-period surface waves, in *Seismic Tomography: Theorie and Practice*, edited by H. M. Iyer and K. Hirahara, Chapman and Hall, New York, 1993.

Stephens, B., R. F. Keeling, M. Heimann, K. D. Six, R. Murane, and K. Caldera, Testing global ocean carbon models using measurements of atmospheric O_2 and CO_2 concentration, *Global Biogeochem. Cycles, 12*, 213–230, 1998.

Tans, P. P., I. Y. Fung, and T. Takahashi, Observational constraints on the global atmospheric CO_2 budged, *Science, 247*, 1431–1438, 1990.

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

Taylor, A. H., A. J. Watson, M. Ainsworth, J. E. Robertson, and D. R. Turner, A modelling investigation of the role of phytoplankton in the balance of carbon at the surface of the North Atlantic, *Global Biogeochem. Cycles, 5*, 151–171, 1991.

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

Weber, C., Zur Dynamik des interhemisphärischen CO_2-Transports im Ozean, Master's thesis, Max-Planck-Institut f. Meteorol., Hamburg, Germany, 1996.

R. Giering, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena CA 91109. (ralf@pacific.jpl.nasa.gov)

M. Heimann and T. Kaminski, Max-Planck-Institut für Meteorologie, Bundesstr. 55, D-20146 Hamburg, Germany. (heimann@dkrz.de; kaminski@dkrz.de)

September 2, 1998; revised February 26, 1999; accepted March 3, 1999.