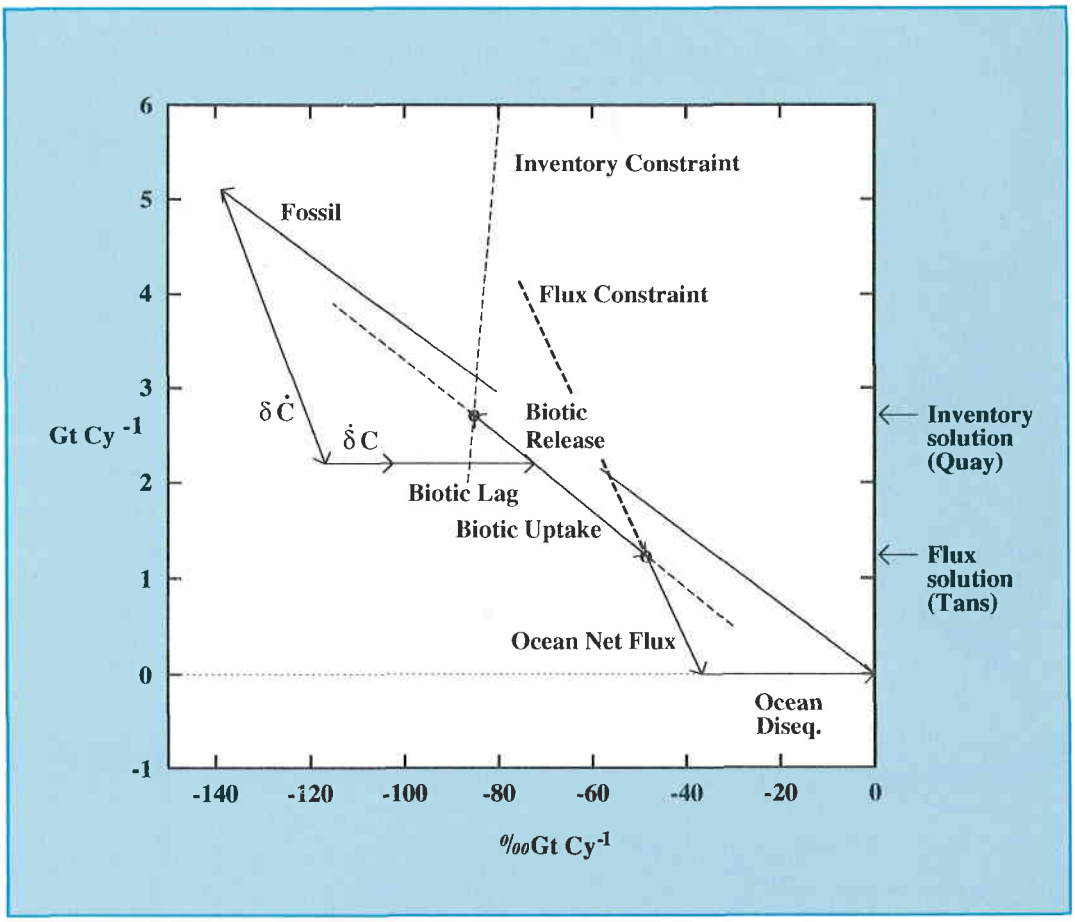
CSIRO
AUSTRALIA

Synthesis Inversion of Atmospheric CO_2 Using the GISS Tracer Transport Model

I.G. Enting, C.M. Trudinger, R.J. Francey and H. Granek



CSIRO DIVISION OF ATMOSPHERIC RESEARCH TECHNICAL PAPER No. 29

ERRATUM

Equations (3), (4a) and (4b) should read

$$w_k = \sum_{\mu=1}^{N'} A_{k\mu}\sigma_\mu \qquad (3)$$

$$w_k = c_k \quad \text{for } k \leq M \qquad (4a)$$

$$w_k = s_{k-M} \quad \text{for } k > M \qquad (4b)$$

Synthesis Inversion of Atmospheric CO_2 Using the GISS Tracer Transport Model

by I.G. Enting, C.M. Trudinger, R.J. Francey and H. Granek

Division of Atmospheric Research Technical Paper No. 29

National Library of Australia Cataloguing-in-Publication Entry

Synthesis Inversion of Atmospheric CO_2 Using the GISS Tracer Transport Model Bibliography.
ISBN 0 643 05251 8
1. Atmospheric carbon dioxide – Measurement. 2. Carbon cycle (Biogeochemistry) – Mathematical models. 3. Atmospheric circulation – Mathematical models.
I. Enting, I.G.. II. CSIRO. Division of Atmospheric Research. (Series: Division of Atmospheric Research Technical Paper; no. 29).
551.51120287

Synthesis Inversion of Atmospheric CO_2 Using the GISS Tracer Transport Model

I.G. Enting, C.M. Trudinger, R.J. Francey and H. Granek
CSIRO, Division of Atmospheric Research
Private Bag 1, Mordialloc, Vic 3195, Australia

Contents

Synthesis Inversion of Atmospheric CO_2 Using the GISS Tracer Transport Model

Abstract

The atmospheric CO_2 budget for the period around 1986–7 is estimated using a Bayesian synthesis analysis. The source strengths of a set of industrial, biotic and ocean processes are estimated by matching observed concentrations to modelled concentrations calculated using the GISS three-dimensional tracer transport model. The Bayesian synthesis uses prior estimates of these process strengths to stabilise what is otherwise an ill-conditioned inversion. The analysis gives preliminary estimates of the ranges of uncertainty arising from the inversion.

The inversion uses observations of the concentration and $\delta^{13}C$ of atmospheric CO_2. The ability to estimate uncertainties makes it possible to compare the relative importance of various data items in reducing these uncertainties. The main result is that global totals of oceanic versus biotic exchange are determined primarily by the combined global budgets for CO_2 and $^{13}CO_2$. The atmospheric transport model constrains the regional sources and sinks but these constraints make only a small contribution to reducing the uncertainty in the global budget. Our best estimate of the global net air-sea carbon flux is 1.4 ± 1.0 Gt C y^{-1}. The net carbon storage by the oceans will be some 0.6 Gt C y^{-1} greater than this CO_2 flux due to carbon transported into the oceans by rivers. The corresponding estimate of air-to-biota CO_2 flux is -2.5 ± 0.9 Gt C y^{-1}, implying a net storage of -1.3 Gt C y^{-1} after accounting for carbon lost from the biota via rivers or as CO.

We find that within the range of ≈ 1 Gt C y^{-1} our estimate of air-sea flux is very sensitive to the choice of data that are fitted. This confirms the appropriateness of our formal error estimates that are based on *a priori* statistics. These large uncertainties mean that our budget estimates must be regarded as preliminary and indicate the need for a refined statistical analysis of the observational data to assess their representativeness.

We present a simple preliminary study of oxygen distributions. The results indicate that, once the CO_2 distribution is fitted, the spatial gradient of oxygen does not give significant additional discrimination between alternative carbon budgets and that the long-term trend of atmospheric oxygen will make the greatest contribution to resolving carbon budgets.

1. Introduction

The atmospheric budgets of the major greenhouse gases are subject to very considerable uncertainty (see for example IPCC, 1990). This uncertainty in the budget leads to corresponding uncertainties in projections of the future concentrations to be expected from various possible release scenarios and thus complicates any choices regarding emission-reduction strategies. Clarification of the atmospheric carbon budget is therefore highly desirable.

For CO_2 and CH_4 most recent estimates of the atmospheric budget have been based on the use of atmospheric transport modelling to interpret the space-time distributions of concentrations in terms of the space-time distributions of sources and sinks (Keeling et al., 1989b; Tans et al., 1990a; Sarmiento and Sundquist, 1992; Fung et al., 1991). The principle is that, to the extent that the spatial distribution of sources and sinks can be deduced from observations, the location will act as a constraint on the possible processes.

The determination of surface sources and sinks from observations of surface concentrations is a poorly-determined inverse problem (Newsam and Enting, 1988; Enting and Newsam, 1990). As such, direct estimates of sources are subject to arbitrarily large errors arising from errors in any or all of the observations, the transport model or the inversion technique. In general, meaningful estimates can only be obtained in the context of a set of prior constraints.

A range of inversion calculations estimating the spatial distribution of sources of trace gases have been performed: for CO_2 by Enting and Mansbridge (1989, 1991), Tans et al. (1989, 1990a), Keeling et al. (1989b), Taylor (1989), Law et al. (1992); for CH_4 by Fung et al. (1991), Brown (1992, 1993) and for CFCs by Hartley (1992), Hartley and Prinn (1992, 1993).

Most of the two-dimensional inversions have used some form of surface mass balance to deduce sources from concentrations (Enting and Mansbridge, 1989, 1991; Tans et al., 1989; Brown, 1992, 1993). This method does not readily generalise to three-dimensional models because the observational data do not adequately specify the full surface concentration field — see however some preliminary studies by Law et al. (1992). The solutions of such 'mass-balance' inversions are generally stabilised by assuming some smoothness properties of the concentrations and sources. Most of the inversions using three-dimensional transport models have been based on some sort of synthesis approach (Keeling et al., 1989b; Tans et al., 1990a; Fung et al., 1991). The method involves seeking a linear combination of source/sink processes such that the corresponding linear combination of their calculated responses matches the observational data. In synthesis calculations the solutions are stabilised by the assumptions about the spatial distributions of the source components.

We present a new synthesis estimate of the atmospheric CO_2 budget, applying to 1986–1987. Particular features of our calculation are: the use of a Bayesian estimation procedure, and an attempt to obtain systematic (and realistic) estimates of the uncertainties in the estimated budget. Our error analysis is preliminary — a full error analysis would include the effects of uncertainties in spatial distributions (we discuss some possibilities in Section 6 but do not present any results) and the effects of transport model error (which we have been unable to assess because a suitable methodology has not been developed).

The outline of the remainder of this report is as follows: Section 2 describes the inversion technique that we adopted, firstly in general terms and then in terms of the specific way in which the model calculations are combined to give estimates of the source/sink strengths. Section 3 describes the way in which the source processes are divided into distinct components for use in the synthesis analysis. It gives the values of the scale factors involved in the various atmospheric responses to the processes and in addition gives the prior estimates required by the Bayesian formalism. Section 4 describes the observational data used in the studies. The results of the inversions are presented in Section 5. Section 6 discusses several possible extensions to the calculations, particularly extensions to the error analysis. Section 7 concludes the body of the report, discussing the implications of our results for understanding the global carbon cycle and considering the areas in which a refinement of the present studies seem possible and desirable. The details of the transport model are given in Appendix A. Appendix B describes the use of a simple biotic model to estimate the amount of isotopic disequilibrium between the atmosphere and the terrestrial biota. Appendix C defines our notation.

2. Inversion techniques

2a Generic formalism

The synthesis process seeks to estimate the strengths, σ_μ, of N source/sink processes by comparing M observed concentrations, c_j, to responses calculated using an atmospheric transport model. If the model response for observation j to a source μ of unit strength is $T_{j\mu}$ then the fit is made on the assumption that

$$c_j = \sum T_{j\mu}\sigma_\mu + \text{observational noise} \tag{1}$$

In the Bayesian formalism, the fit is constrained to take account of prior estimates, s_μ, of the source strengths. In this report we consider only least squares fits. This corresponds to assuming independent normal distributions for the prior estimates and the data. This Bayesian inversion is obtained by minimising

$$\Theta = \sum_{j=1}^{M}(c_j - \sum_{\mu=1}^{N'} T_{j\mu}\sigma_\mu)^2/u_j^2 + \sum_{\mu=1}^{N}(\sigma_\mu - s_\mu)^2/v_\mu^2 \tag{2}$$

where u_j is the standard deviation of the observational noise of c_j and v_μ is the standard deviation of the prior estimate s_μ. Taking $N' > N$ allows for some additional 'pseudo-sources' which we exclude from the fit to the prior estimates (see below).

Our implementation solves the minimisation problem as a weighted least-squares solution of the equations

$$v_k = \sum_{\mu=1}^{N'} A_{k\mu}\sigma_\mu \qquad \text{for } k = 1 \text{ to } M+N \tag{3}$$

with the 'data' defined as:

$$v_k = c_k \qquad \text{for } k \leq M \tag{4a}$$

$$v_k = s_{k-M} \qquad \text{for } k > M \tag{4b}$$

with weights:

$$d_k = u_k \qquad \text{for } k \leq M \tag{5a}$$

$$d_k = v_{k-M} \qquad \text{for } k > M \tag{5b}$$

and responses:

$$A_{k\mu} = T_{k\mu} \qquad \text{for } k \leq M \tag{6a}$$

$$A_{k\mu} = \delta_{k-M,\mu} \qquad \text{for } k > M \tag{6b}$$

The equations are solved using a modified version of routine SVDFIT (and the routines which it calls) from Press et al. (1986). This produces estimates $\hat{\sigma}_\mu$ of the source strengths. The routine SVDVAR is used to calculate the covariance matrix $V_{\mu\eta}$ for these estimates.

2b Implementation

In order to apply the formalism from the previous section in a synthesis study, we need to be more specific about the types of data that are fitted. For each j in the range 1 to M we define 3 indicator functions:

$\nu(j)$ — which takes values 1, 2 or 3 to specify whether the data item refers to CO_2, $^{13}CO_2$ or O_2 respectively;

$\rho(j)$ — which specifies the location of the observation;

$\omega(j)$ — which specifies the frequency of the time variation.

The frequency index takes values 0 for the annual mean, -1 and $+1$ for the sine and cosine of $2\pi t$ and -2 and $+2$ for the sine and cosine of $4\pi t$. The special value $\omega(j) = 10000$ refers to the global trend. In this case $\rho(j)$ is undefined.

The trend data have:

$$T_{j\mu} = G_{\nu(j)}(\mu) \qquad \text{if } \omega(j) = 10000 \tag{7a}$$

where the factors $G_\nu(\mu)$ specify the contribution of source μ to the trend in constituent ν. The $G_\nu(\mu)$ are listed in Table 1, with μ as a row index.

Another 'special case' involves the need to estimate a global mean offset for the concentration of each constituent. These offsets are included in the formalism as 'pseudo-sources' contributing to the data with a factor:

$$T_{j\mu} = 1 \qquad \text{if } \omega(j) = 0 \text{ and } \mu \text{ is the 'pseudo-source' for constituent } \nu(j) \tag{7b}$$

The pseudo-sources for CO_2, $^{13}CO_2$ and O_2 are given indices $\mu = N+1, N+2, N+3$. Thus the sum in (3) runs from 1 to $N' = N + 3$.

Each of the 'real' sources has a characteristic space-time distribution which in turn gives a characteristic space-time distribution in the atmospheric concentration. We define an index $\phi(\mu)$ to refer to the space-time distributions for process μ since, in a few cases, several processes have the same distribution (which implies that they can *only* be distinguished on the basis of distinctions between the constituents). The distributions of responses are denoted $R_{\rho\omega\phi}$ and so

$$T_{j\mu} = R_{\rho(j)\omega(j)\phi(\mu)} F_{\nu(j)}(\mu) \qquad \text{for } \omega(j) \in [-2, 2] \text{ and } \mu \leq N \tag{7c}$$

The factors $F_\nu(\mu)$ specify the relative responses for each of the constituents and are listed in Table 1.

In the present study, we have fitted each observational record to the model concentration for the cell containing the observing site. We have not attempted any refined estimation based on interpolation between grid-point values, although such an extension could be readily incorporated within the formalism described here. Similarly, we have not adopted the approach used by Fung et al. (1991) of fitting data to neighbouring grid cells for coastal sites where the model will have a source that acts on the whole grid cell (with reduced strength) while the data will generally be selected for onshore transport. (We have tested such 'corrections' and found that the changes in the solution are small compared to the range of uncertainty.)

Figure 1 gives a schematic representation of the information flow in our analysis.

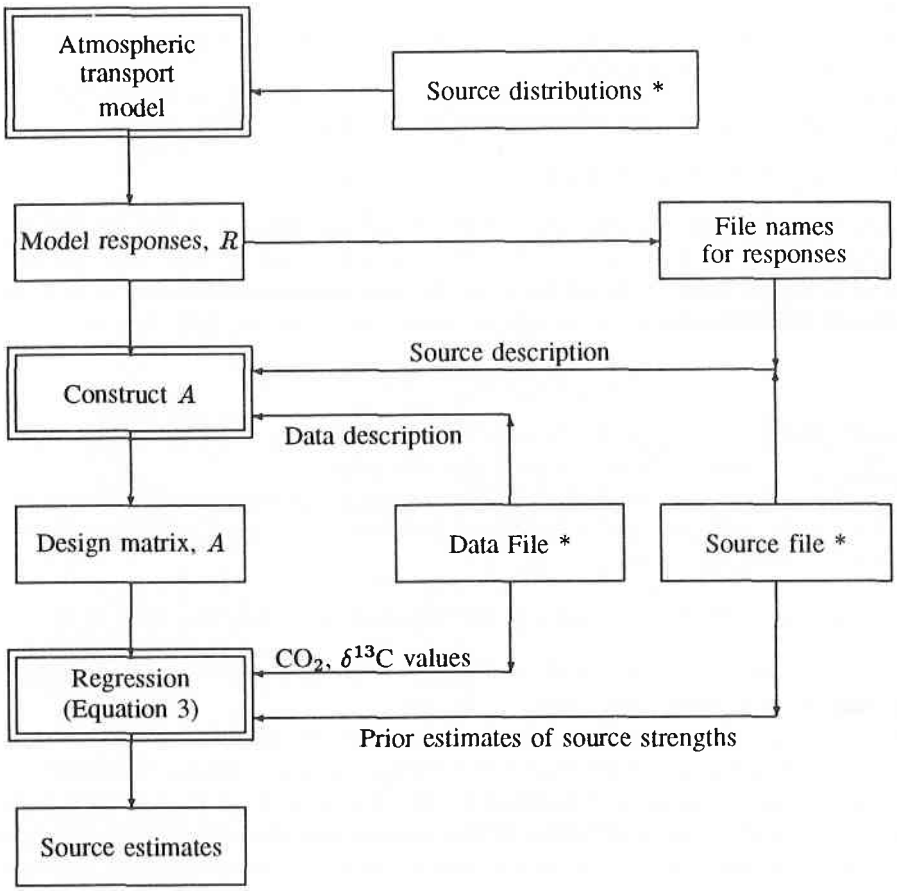


Figure 1: Information flow in the synthesis analysis. Double boxes denote processing, single boxes denote information. The source distributions (described in Section 3b) are used in the GISS transport model to produce sets of model responses, $\boldsymbol{R}$. These are combined with source specifications ($\boldsymbol{F}(\mu)$, $\boldsymbol{G}(\mu)$ and $\boldsymbol{\phi}(\mu)$ from Table 1) and data specifications ($\nu(j)$, $\rho(j)$ and $\omega(j)$) (Section 4) to produce the design matrix $\boldsymbol{A}$. The final stage of processing takes the prior estimates of the sources ($\boldsymbol{s}_\mu$ and $\boldsymbol{v}_\mu$ from Section 3e) and the data values (c_j and u_j from Section 4) to produce source estimates $\hat{\sigma}_\mu$ and the covariance matrix V. The *'s denote input files.

2c Units and normalisations

- The observational data are expressed in ppmv (parts per million by volume) for CO_2 and O_2, except for the trend terms which are in ppmv per year;
- The $^{13}CO_2$ data are expressed as a ^{13}C anomaly, X, as defined below with units of ‰ppmv.
- The pseudo-sources defining global mean concentrations are expressed in the units of the corresponding concentrations, for convenience.
- Real sources, σ_μ, are in Gt C y^{-1} (gigatonnes of carbon per year), specifying the annual mean source strength, except for the seasonal component. We choose the sign convention that sources to the atmosphere are positive.
- The calculated responses are in units of ppmv of CO_2 per Gt C y^{-1} except for the seasonal biotic flux. This means that the CO_2 scale factors $F_1(\mu)$ will generally be equal to 1. The other $F_\nu(\mu)$ convert the CO_2 responses to responses for ^{13}C anomalies and O_2 concentrations.
- The seasonal biotic source is expressed as a numerical multiple of the prior source distribution which is based on the work of Fung et al. (1983) and which has a NPP of 45.6 Gt C y^{-1}.
- The gross fluxes between atmosphere and biota and atmosphere and oceans are expressed as Gt C y^{-1} for consistency, even though the gross flux terms do not appear in the carbon budget, being required only in calculating the effect of isotopic disequilibrium.

Isotopic compositions are frequently expressed in terms of the 'δ' notation that expresses the amount by which an isotopic ratio differs from a standard. For $^{13}CO_2$ the definition is

$$\delta^{13}C = [^{13}C{:}^{12}C_{\text{sample}}/^{13}C{:}^{12}C_{\text{standard}} - 1] \times 1000 \qquad (8a)$$

in units of per mil (denoted ‰). We denote the standard $^{13}C{:}^{12}C$ ratio by r_{standard}. (Its measured value of 0.0112372 is not explicitly needed.)

The ^{13}C anomaly used in the calculation is

$$X = ([^{13}C]/R_{\text{ref}} - [C]) \times 1000 \qquad (8b)$$

where R_{ref} is a reference $^{13}C{:}C$ ratio. X acts as a conservative tracer. If we used $R_{\text{ref}} = R_{\text{standard}}$ (where $R_{\text{standard}} = r_{\text{standard}}/(1 + r_{\text{standard}})$ is the standard $^{13}C{:}C$ ratio for a $\delta^{13}C$ of zero), then X would be equal (to a very good approximation) to $[C] \times \delta^{13}C$. This choice of anomaly definition is used in Section 4d when discussing global budgets. For the purposes of the inversion calculation however, rather than choose R_{ref} to correspond to a $\delta^{13}C$ of zero, we choose R_{ref} to correspond to a $\delta^{13}C$ of -8, close to the observed atmospheric mean. With this definition, we have, to a good approximation

$$X = (\delta^{13}C + 8)[C] \qquad (8c)$$

The approximation in our calculations arises from the fact that the $\delta^{13}C$ quantification of isotopic variations is expressed in terms of $^{13}C{:}^{12}C$ ratios while the present calculations

need ^{13}C:C ratios. We assume that all the ratios are sufficiently close to the standard ratio 0.0112372 that we can convert from ^{13}C:^{12}C ratios (denoted by lower case 'r') to ^{13}C:C ratios (denoted by upper case 'R') by multiplying by $1/(1 + r_{\mathrm{standard}})$. We shall use this approximation without further comment.

The reason for the choice of offset in defining X can be seen when we consider the trend in X. We have:

$$\frac{d}{dt}X = (\delta^{13}\mathrm{C} + 8)\frac{d}{dt}[\mathrm{C}] + [\mathrm{C}]\frac{d}{dt}\delta^{13}\mathrm{C} \tag{9}$$

By choosing R_{ref} so that the factor $(\delta^{13}\mathrm{C} + 8)$ is close to zero, the uncertainty in the trend in X is almost entirely due to the uncertainty in $\frac{d}{dt}\delta^{13}\mathrm{C}$. Similarly, for seasonal and spatial variations about the mean atmospheric isotopic composition, the uncertainties in X can be treated as due solely to the uncertainties in the δ^{13}C measurements. This simplifies the calculations of the effect of these uncertainties on the estimated sources.

The CO_2 conversion factor 0.471 ppmv per Gt C is the ratio of the number of moles ($\frac{1}{12.01} \times 10^{15}$) in a gigatonne of carbon to the number of moles of dry air in the atmosphere (a mass of 512.4×10^{19} g, Trenberth, 1981, divided by a mean molecular weight of 28.9644) with the ratio multiplied by 10^6 to convert the mixing ratio to ppmv.

2d Supplementary analysis

Beyond the regression analysis that forms the basis of our inversion technique there are three additional calculations that assist in the interpretation of the results and which we have implemented as routine parts of the inversion calculation.

The first is the ability to look at the covariances of linear combinations of the data. Perhaps the most obvious example is the relation between the estimates of summed biotic exchange and summed ocean exchange. The formalism is straightforward. For a set of K quantities defined by the linear combinations $x_\beta = \sum_\mu q_{\beta\mu}\sigma_\mu$, the covariance of x_β and x_γ is given by $\sum_{\mu\eta} V_{\mu\eta} q_{\beta\mu} q_{\gamma\eta}$.

Secondly, we include the capability of calculating the residuals $D_j = c_j - \sum_\mu A_{j\mu}\hat{\sigma}_\mu$ and the normalised residuals D_j/u_j.

The final related feature is the ability to look at the concentration distributions implied by the inversion. In particular we consider the zonal means. For each response calculated using the GISS model, we have included the zonal mean response. These zonal mean responses are then combined using the solution coefficients, $\hat{\sigma}_\mu$, and the weights, F_ν, to give zonal mean distributions of concentration, ^{13}C anomaly and oxygen. The anomaly distribution is combined with the concentration to produce the δ^{13}C distribution.

3. Source components

Process	Distrib.	F_1	F_2	F_3	G_1	G_2	G_3	Prior	sd
Mean CO_2	Uniform	1	0	0	0	0	0	n.a.	n.a.
Mean $^{13}CO_2$	Uniform	0	1	0	0	0	0	n.a.	n.a.
Mean O_2	Uniform	0	0	1	0	0	0	n.a.	n.a.
Fossil	Fossil	1	−19	−1.41	0.471	−8.95	−0.664	5.3	0.3
Land-use	Land-use	1	−17	−1.05	0.471	−8.01	−0.495	2.2	1.0
Bio. uptake	Bio-uptake	1	−17	−1.05	0.471	−8.01	−0.495	−2.8	1.0
CO oxid.	CO	1	−17	−0.5*	0.471	−8.01	−0.471	0.9	0.2
Seasnl. bio.	Season	1	−17	−1.05	0	0	0	1.0	0.5
Bio. diseq	Bio-release	0	0.30	0	0	0.141	0	100	50
n:Atlant(B)	Atlant(B)	1	−1.70	0	0.471	−0.80	0	−0.2	0.5
n:Pacif(B)	Pacif(B)	1	−1.70	0	0.471	−0.80	0	0.2	0.5
n:Atlant(N)	Atlant(N)	1	−1.45	0	0.471	−0.68	0	−0.3	1.0
n:Pacif(N)	Pacif(N)	1	−1.45	0	0.471	−0.68	0	−0.3	1.0
n:Equat.	Equatorial	1	−1.45	0	0.471	−0.68	0	1.7	1.5
n:Atlant(S)	Atlant(S)	1	−1.45	0	0.471	−0.68	0	−0.5	1.0
n:Pacif(S)	Pacif(S)	1	−1.45	0	0.471	−0.68	0	−1.0	1.5
n:Indian(S)	Indian(S)	1	−1.45	0	0.471	−0.68	0	−0.7	1.0
n:South.	Southern	1	−1.45	0	0.471	−0.68	0	−0.2	1.0
g:Atlant(B)	Atlant(B)	0	−0.94	0	0	−0.44	0	2.7	1.4
g:Pacif(B)	Pacif(B)	0	−1.08	0	0	−0.51	0	2.5	1.3
g:Atlant(N)	Atlant(N)	0	0.67	0	0	0.32	0	5.1	2.6
g:Pacif(N)	Pacif(N)	0	0.32	0	0	0.15	0	8.7	4.4
g:Equat.	Equatorial	0	1.53	0	0	0.72	0	17.2	8.6
g:Atlant(S)	Atlant(S)	0	0.43	0	0	0.20	0	4.0	2.0
g:Pacif(S)	Pacif(S)	0	0.74	0	0	0.35	0	8.4	4.2
g:Indian(S)	Indian(S)	0	0.46	0	0	0.21	0	5.7	2.9
g:South.	Southern	0	−1.50	0	0	−0.71	0	22.3	11.0

Table 1: Summary of source components used in the synthesis studies. Regional suffices for oceans refer to Far north (B for boreal), North (N) and South (S). Prefices 'n:' and 'g:' refer to the contributions of net air-sea flux and gross air-sea flux (through isotopic disequilibrium) respectively. The effect of CO on the spatial distributions of oxygen (marked with '*') is discussed in the text.

3a Summary

Table 1 summarises the source components that we use in the synthesis studies. Each line represents a distinct process with a source strength to be estimated. (The first three lines are 'pseudo sources' — we need to fit these quantities as part of the estimation process even though they are not of interest.) Each process, μ, is characterised by a specified space-time distribution of surface flux. Column 2 specifies the distribution for each process. The details of these distributions are given in Section 3b below.

Each process will have a characteristic response factor, $F_\nu(\mu)$, for CO_2, the ^{13}C anomaly and for O_2 (denoted by $\nu = 1, 2, 3$). These factors are listed in columns 3 to 5 and are