Tellus (1986), *38B*, 11–26

A study of linear inversion schemes for an ocean tracer model

By J. V. MANSBRIDGE and I. G. ENTING, *CSIRO, Division of Atmospheric Research, Private Bag # 1, Mordialloc, Victoria, 3195, Australia*

(Manuscript received July 3; in final form December 9, 1985)

ABSTRACT

Several linear biased estimation techniques are applied to a 12-reservoir ocean tracer model. The procedures reduce the sensitivity of the calibration to noise in the data, but the results still indicate a persistent problem with the phosphorus distribution in the Antarctic region. The airborne fraction of CO_2 is found to be insensitive to the changes in circulation associated with different calibration procedures.

1. Introduction

Traditionally, the most important part of models of the global carbon cycle has been the ocean model. Early modelling studies of the carbon cycle tended to either ignore the possibility of biospheric changes or else invoke biospheric changes to account for discrepancies between observations and model predictions for the amount of fossil fuel CO_2 remaining in the atmosphere. In recent years, a more critical appraisal of the role of the biosphere has suggested that agricultural clearing is currently releasing large amounts of CO_2 into the atmosphere (Houghton et al., 1983). Such large releases would exacerbate the discrepancies between the ocean models and the CO_2 observations and so the results of the ecosystem modellers have prompted a re-examination of the ocean models used in carbon cycle studies.

The earliest ocean models used in carbon cycle studies were essentially one-dimensional, resolving only the vertical direction. They consisted of a number of levels ranging from as few as two in the earliest models (e.g., Craig, 1957), to 5 (Keeling, 1973), to the limit of a continuum with diffusive transport (Oeschger et al., 1975). There have been many refinements to this basic pattern. Differentiation between warm and cold regions and advective injection of cold surface waters into deep layers (with a consequent upwelling) have been used in various combinations by Björkström (1979), Pearman (1980), Bacastow and Björkström (1981), Crane (1982) and Peng et al. (1983). These various refinements have not proved adequate to allow the uptake of the amount of CO_2 implied by the ecosystem modelling.

A number of new models have been introduced in order to determine whether a more detailed resolution of ocean processes might lead to significant changes in the CO_2 uptake. Three such models are the "outcrop" model of Siegenthaler (1983), the regional model of Broecker et al. (1980) and the 12-reservoir model of Bolin et al. (1983) henceforth denoted BBHM. The 12 reservoir model is regarded as a fore-runner of more detailed models. A preliminary description of an 84 reservoir model of the Atlantic has been given by Moore and Björkström (1986). These studies employ a suite of tracers to calibrate the models by a constrained inversion technique. The use of constrained inversion techniques is necessary because, as discussed by Broecker (1981), as the resolution of an ocean model increases, the number of model parameters grows faster than the number of distinct items of data. The study by BBHM uses 5 tracers: carbon, carbon-14, alkalinity, oxygen and phosphorus. It is notable for its detailed treatment of detrital fluxes. The detrital flux is important in carbon cycle modelling since, although it does not provide a mechanism for sinking significant amounts of excess CO_2, the total carbon gradient that it establishes must be taken into account when natural ^{14}C distributions are used to calibrate ocean uptake. This has been discussed by Broecker (1981, p. 440). One way of allowing for this effect without explicitly modelling detritus is to artificially deepen the model oceans to hold the excess carbon (Bacastow and Björkström, 1981). Enting (1985) has described how the failure to allow for detrital fluxes shows up as a distinct anomaly in the ^{14}C balance in a box-diffusion model calibrated by constrained inversion.

The main aim of this paper is to explore the calibration technique used by BBHM. We have related their approach to more general linear inversion techniques which we discuss within the general framework of biased estimation (see for example Goldstein and Smith, 1974). The techniques that we present have a number of advantages compared to those used by BBHM. Firstly, the behaviour of the calibration procedure is amenable to a general analysis so that we can assess its stability. Secondly we appear to have improved the stability of the calibration. Finally we are able to unite the two cases, "incompatible" and "indeterminate", considered by BBHM into one formalism. The layout of the remainder of this paper is as follows: Section 2 reviews some techniques of biased estimation and discusses them in relation to the linear estimation problem involved in the BBHM calibration. Section 3 describes the way in which we have implemented the model described by BBHM. We have made a number of small changes, partly for simplicity and partly because of incomplete specification by BBHM. Section 4 describes the results of a range of estimation procedures applied to the model. It is seen that some aspects of the solution are sensitive to the details of the calibration. A comparison of least-square solutions with BBHM also indicates that some aspects are sensitive to the small changes in model formulation. The analysis in Section 5 shows, however, that this sensitivity is not reflected in the value of the airborne fraction of CO_2 predicted by the model so that the model is inconsistent with the large CO_2 releases estimated from ecosystem modelling.

2. Linear estimation techniques

The problem considered by BBHM was to obtain estimates of a parameter vector $\boldsymbol{x}$, of n parameters x_i, given a relation of the form:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{2.1}$$

where $\boldsymbol{b}$ is of dimension m. They considered two cases. Their method I (the incompatible case) was used when eq. (2.1) did not have an exact solution. In this case they took the least squares solution, i.e., the solution that minimised $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|$ where $\|\cdot\|$ denotes the Euclidean 2-norm (i.e. $\|\boldsymbol{a}\|^2 = \sum a_i^2$).

Their method II (the indeterminate case) was used when eq. (2.1) had multiple solutions. In this case they took the solution of (2.1) that had the minimum $\|\boldsymbol{x}\|^2$. They stated that they "have not found a logically comparable methodology that applies to all systems". In each case BBHM applied additional constraints. The two cases can be combined by use of pseudo-inverse matrices. An additional third case not considered by BBHM, namely that eq. (2.1) is simultaneously incompatible and indeterminate, is also encompassed by the use of pseudo-inverses. The pseudo-inverse, $\boldsymbol{A}^+$, of a matrix $\boldsymbol{A}$ is defined (Deutsche, 1965) by the relations

$$\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A} = \boldsymbol{A} \tag{2.2a}$$

$$\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^+. \tag{2.2b}$$

If $\boldsymbol{A}$ is written in terms of its singular value decomposition as

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{W}^{\mathrm{T}}, \tag{2.3}$$

where $\boldsymbol{U}$ and $\boldsymbol{W}$ are orthogonal matrices and $\boldsymbol{\Lambda}$ is a diagonal matrix whose non-zero elements are $\lambda_1 \lambda_2 \dots \lambda_N$ then

$$\boldsymbol{A}^+ = \boldsymbol{W}\boldsymbol{\Lambda}^+\boldsymbol{U}^{\mathrm{T}} \tag{2.4}$$

where $\boldsymbol{\Lambda}^+$ is diagonal with non-zero elements

$$\Lambda_{ii}^+ = \lambda_i^{-1} \quad \text{if } \lambda_i > 0 \tag{2.5a}$$

$$\Lambda_{ii} = 0 \quad \text{if } \lambda_i = 0. \tag{2.5b}$$

The pseudo-inverse solution of (2.1) is

$$\hat{\boldsymbol{x}} = \boldsymbol{A}^+\boldsymbol{b} \tag{2.6}$$

and has the property that $\hat{\boldsymbol{x}}$ minimizes $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|$ and, subject to this requirement, also minimizes $\|\hat{\boldsymbol{x}}\|$.

The purpose of the present study is to analyse a more general class of linear estimation procedures. The statistical model that is used to analyse the estimation problem (2.1) is to put

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_{\text{true}} + \boldsymbol{\varepsilon}, \tag{2.7}$$

where $\boldsymbol{\varepsilon}$ represents a set of independent error components with zero mean. If the variances of the ε_i are known then they can be taken as 1 without loss of generality. The least squares estimator is defined by

$$\hat{\boldsymbol{x}}_{\text{LSQ}} = (\boldsymbol{A}^{\text{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\text{T}}\boldsymbol{b} \tag{2.8}$$

and has the property that if $\text{E}[\cdot]$ denotes expectations taken over the error distribution then

$$\text{E}[\hat{\boldsymbol{x}}_{\text{LSQ}}] = \boldsymbol{x}_{\text{true}}, \tag{2.9}$$

i.e., the estimator $\hat{\boldsymbol{x}}_{\text{LSQ}}$ is unbiased. It is the best linear unbiased estimator in that if the ε_i are normally distributed then

$$\text{var}[\hat{\boldsymbol{x}}_{\text{LSQ}}] = \text{E}[\|\hat{\boldsymbol{x}}_{\text{LSQ}} - \text{E}[\hat{\boldsymbol{x}}_{\text{LSQ}}]\|^2] \tag{2.10}$$

is smaller than for any other linear unbiased estimator. The use of biased estimators arises from the consideration that such estimates can reduce the variance, and more importantly they can give smaller values of mean-square error, (MSE), which is

$$\text{MSE}[\hat{\boldsymbol{x}}_{\text{LSQ}}] = \text{E}[\|\hat{\boldsymbol{x}} - \boldsymbol{x}_{\text{true}}\|^2] \tag{2.11}$$

In general the least-squares estimator will not be best linear estimator for minimizing the MSE. As stated by Golub et al. (1979), "Allowing a bias may reduce the variance tremendously". Goldstein and Smith (1974) have considered a wide class of linear estimators that give smaller MSE than least-squares. They define these estimators in terms of the singular value decomposition of the matrix $\boldsymbol{A}$. They constructed a class of regularized inverses denoted $\boldsymbol{A}_{\text{R}}^{-1}$ of the form

$$\boldsymbol{A}_{\text{R}}^{-1} = \boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{U}^{\text{T}} \tag{2.12}$$

where $\boldsymbol{\Psi}$ is a diagonal matrix whose non-zero elements ψ_i are related to those of $\boldsymbol{\Lambda}$.

Goldstein and Smith considered a class of linear estimators

$$\hat{\boldsymbol{x}} = \boldsymbol{A}_{\text{R}}^{-1}\,\boldsymbol{b}, \tag{2.13}$$

with the ψ_i (and thus $\boldsymbol{A}_{\text{R}}^{-1}$) defined in terms of a function with a single parameter Γ. Specifically

$$\psi_i = c(\lambda_i, \Gamma) \tag{2.14}$$

where

$$c(\lambda_i, 0) = \lambda_i^{-1}, \tag{2.15a}$$

$$\frac{\partial}{\partial \Gamma} c(\lambda_i, \Gamma)/\lambda_i < 0 \qquad \text{for all } \Gamma \geqslant 0. \tag{2.15b}$$

They showed that for any $\boldsymbol{x}_{\text{true}}$, there existed a $\Gamma > 0$ such that the biased estimator $\boldsymbol{A}_{\text{R}}^{-1}(\Gamma)\boldsymbol{b}$ had a smaller MSE than the least-squares estimator which corresponds to $\Gamma = 0$. The practical problem involved in using this result is that the best Γ will depend on the unknown $\boldsymbol{x}_{\text{true}}$ and on the variance of the ε_i.

The class of functions specified by Goldstein and Smith includes the case

$$\psi_i = \frac{\lambda_i^N}{\lambda_i^{N+1} + \Gamma^N}. \tag{2.16}$$

For $N = 1$ this gives the ridge estimator introduced by Hoerl and Kennard (1970). The $N \to \infty$ limit corresponds to a truncation of the singular value spectrum, putting

$$\psi_i = \lambda_i^{-1} \qquad \text{for } \lambda_i > \Gamma, \tag{2.17a}$$

$$= 0 \qquad \text{for } \lambda_i < \Gamma. \tag{2.17b}$$

This truncation was discussed by Jackson (1972).

For each of the specific regularization schemes, and particularly for ridge regression, there has been extensive discussion of the best way to determine Γ, the regularization parameter. (See Golub et al., 1979 for references.) We have considered a number of the simpler approaches.

Our first approach was to follow Twomey (1977) and note that since Γ is, from a practical point of view, an arbitrary parameter, solutions that are highly sensitive to Γ should not be accepted and so we sought solutions in the region where they depended only slightly on Γ. The second approach was to use the value suggested by Hoerl et al. (1975),

$$\Gamma = n\|\boldsymbol{A}\hat{\boldsymbol{x}}_{\text{LSQ}} - \boldsymbol{b}\|^2/(m\|\hat{\boldsymbol{x}}_{\text{LSQ}}\|^2). \tag{2.18}$$

Another aspect that we considered was the property of ridge estimators mentioned by Goldstein and Smith, namely that in cases where least-squares estimators are poorly conditioned and some components have the wrong sign, then ridge estimators tend to correct these errors in sign. This is of considerable interest in connection with the ocean model of BBHM because they applied a number of positivity constraints on various parameters and, as can be seen from their solution, a number of these constraints were active, indicating that an unconstrained least-squares solution would have a number of parameters with unphysical negative values.

In our studies this expected behaviour of the ridge estimators did not in general occur and problems with signs of parameters persisted in the regularized solutions.

3. The ocean model

We use a 12-box ocean model which is essentially the same as that used by BBHM. In each box the continuity equation is applied to each of six tracers—carbon, carbon-14, alkalinity, phosphorus, oxygen and water. Fig. 1 shows the 12 boxes, some of which are connected by flux arrows. The double-headed arrows represent diffusive fluxes whose directions depend on the tracer gradients. The single-headed arrows show the nominal direction of each advective flux. (The figure is based on Fig. 8 of BBHM.)

The reader is referred to BBHM pages 208–213 for a complete description of the 12-box model. We will only note here the three ways in which our formulation differs from theirs.

(1) The air-sea exchanges of C and ^{14}C have been taken from BBHM Fig. 8 or Fig. 9 or deduced from their specifications on pages 211 and 212.

(2) The air-sea exchange of oxygen is not modelled. This causes no difficulties as we never attempt to solve the oxygen conservation equation in the four surface boxes.

(3) One of the water conservation equations is superfluous and so we only use eleven of these equations.

A series of experiments has shown that these differences produce only slight quantitative effects.

We always use version (A) of the model structure given by BBHM on page 213 of their paper. The size of each box and the values of q, the tracer concentration in each box, are taken from page 217 of the same paper.

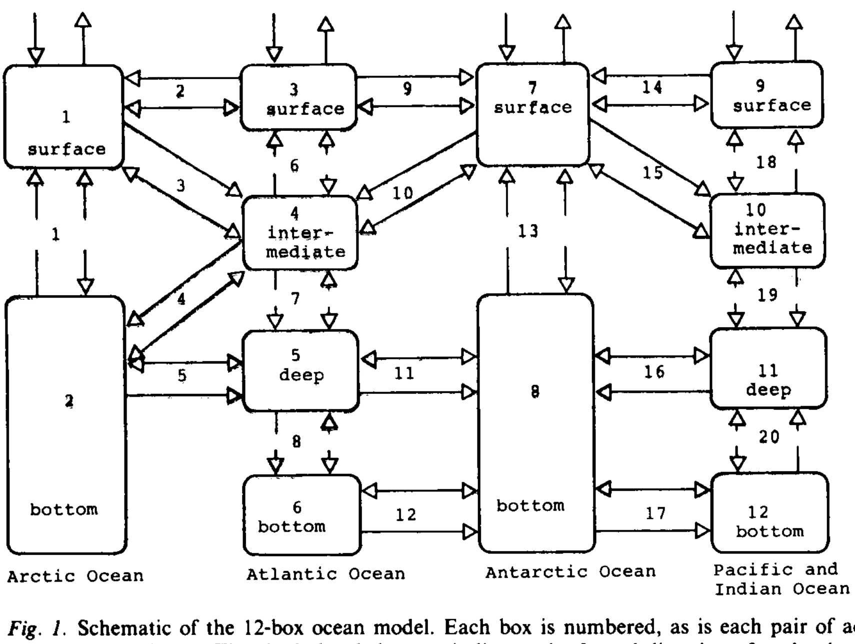

Fig. 1. Schematic of the 12-box ocean model. Each box is numbered, as is each pair of advective and diffusive fluxes between boxes. The single headed arrow indicates the formal direction of each advective flux. The air-sea fluxes are given by a flux into and a flux out of each box.

For the steady state model there are 64 unknowns—20 advective fluxes, 20 diffusive fluxes, and 24 detritus production terms. There are 75 conservation equations—12 each for C, ^{14}C, A and P, 11 for H_2O, 8 for O, and 8 for detritus.

The conservation equations can be expressed in the general matrix form as

$$A(q)\,x = b(q). \tag{3.1}$$

The vector q represents the observed concentrations of the tracers in each box and the vector x contains the 64 unknowns.

Due to modelling and observational errors this set of equations will be inconsistent to some degree. The final solution will therefore depend partly on how we chose to scale the equations. We choose the same units for the q_i^y as BBHM; this causes each of the q_i^y to be of order unity. Following BBHM the C, ^{14}C, A, H_2O and detritus conservation equations are multiplied by 25, the O-equations are left at order unity, and the P-equations are multiplied by 0.1. The above scaling was determined semi-empirically by BBHM. The small scaling factors of O and P are the result of the large errors involved in these quantities. These errors are partly observational, but may be substantially due to the model formulation. In particular, the Redfield ratios may be numerically wrong or even an inadequate description of the chemistry involved.

Even with the O and P equations scaled down we find, as did BBHM, that they dominate the mean square errors. In general about 80% of this error is due to poor conservation of phosphorus, 15% is due to poor conservation of oxygen, and the other 5% is spread fairly evenly through the other equations.

4. Comparison of estimators

Having set up the ocean model in the form $A(q)x = b(q)$, we applied some of the biased estimation techniques described in Section 2. Since the various procedures are described in terms of the singular values of $A(q)$ we show these 64 singular values in Fig. 2. They range from 524.1 to 0.0099. Hence, the condition number (the largest singular value divided by the smallest one) is approximately 5.3×10^4. This indicates that there will be no significant rounding errors in the numerical solution. Fig. 2 also shows an indication of some of the errors in these eigenvalues. These error estimates were obtained from a series of 10 numerical experiments which added independent normal "errors" δq_i^y to the q_i^y and constructed a perturbed matrix $A + \delta A = A(q + \delta q)$. The standard deviations of the δq_i^y were:

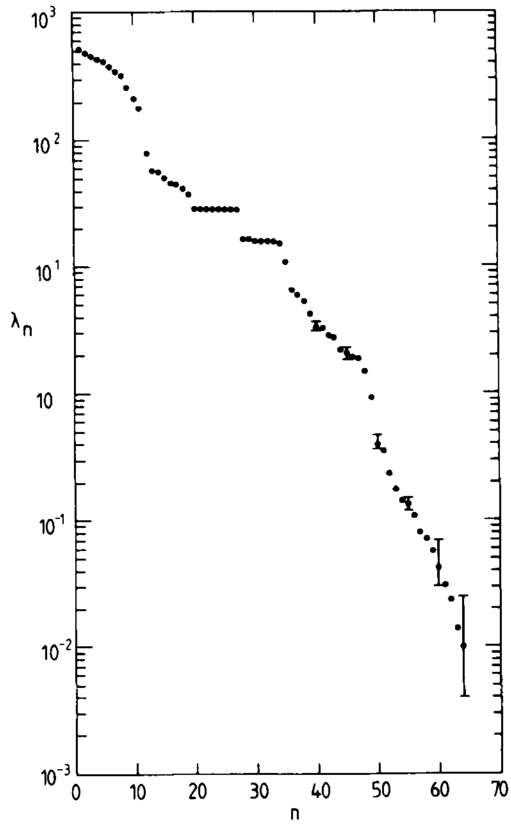

Fig. 2. The magnitudes of the 64 singular values for the basic set of equations. The error bars on selected singular values show the extreme values found from 10 perturbation experiments.

0.01 mol m^{-3} for C
0 for ^{14}C
0.004 eq m^{-3} for alkalinity
0.1 mmol m^{-3} for P
0.01 mol m^{-3} for O.

These same values were used by BBHM in their error analysis. The error bounds in Fig. 2 are the extreme values from the set of numerical experiments. We see that the relative error in λ_1 is about 0.2% while the relative error in λ_{64} is over 100%. Recalling that it is the inverses of the singular values that influence $\hat{x}$, it is clear that the errors in the small singular values are very important.

It is the possibility of examining the structure of A as in Fig. 2 that leads us to favour estimation techniques that can be related to the singular value decomposition. Since the 5 smallest eigenvalues are uncertain by a factor of 2 or more it follows that the problem is effectively underdetermined as well as being inconsistent. Computationally, the solution of (3.1) is well-determined for the given vector q since the inverse of the condition number greatly exceeds the size of the errors arising from the computer arithmetic. However, the numerical uncertainties in the q values used to construct A mean that a subspace of possible x vectors will be poorly defined by the solution of (3.1). A study of eigenvalues indicates the dimensionality of the underdetermined subspace while the occurrence of a condition number whose inverse is less than the proportional errors in the q merely warns that such a subspace exists.

The first of the regularized solutions that we consider uses the truncation scheme defined by (2.17a, b). We specify the truncation in terms of l, the number of singular values used rather than giving a cutoff level. We considered l from 40 to 64. Fig. 3 shows $\|\hat{x}\|^2$, the squared magnitude of the solution and $\|A\hat{x} - b\|^2/\|b\|^2$, the relative error in the defining equations. As l decreases, $\|\hat{x}\|^2$ decreases and $\|A\hat{x} - b\|^2/\|b\|^2$ increases. This behaviour is quite general for the type of regularized solution defined by (2.12). It is shown in the Appendix that if any of the ψ_i are decreased while none are increased then $\|\hat{x}\|^2$ will decrease and $\|A\hat{x} - b\|^2$ will increase.

We regard the solution with $l = 64$ (i.e., the least-squares solution) as unacceptable because of its strong dependence on the poorly determined λ_{64}. The solution was also unacceptable on physical grounds because it shows many large negative K's and incorrect circulation patterns. With $l = 64$, the solution is corrupted by the attempt to accurately fit errors in the original matrix equation.

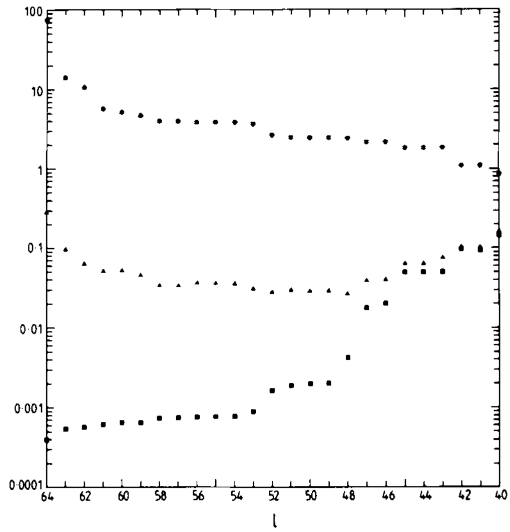

Fig. 3. $\|\hat{x}\|^2$ (indicated by ●), $\|A\hat{x} - b\|^2/\|b\|^2$ (indicated by ■), and $\|(A + \delta A)\hat{x} - b - \delta b\|^2/\|b + \delta b\|^2$ (indicated by ▲), for truncations $l = 64$ to 40.

In order to choose an appropriate truncation we have investigated the manner in which errors in the observations affect the solution for various l. For each truncation we took the estimate $\hat{x}^{(l)}$ and substituted it into a perturbed matrix equation. This was repeated for each of 10 sets of perturbed equations obtained by generating δq vectors as described above. On Fig. 3 is plotted the mean of the perturbed relative error

$$\|A(q + \delta q)\, \hat{x}^{(l)} - b(q + \delta q)\|^2/\|b(q)\|^2. \tag{4.1}$$

For $l = 64$ this is large because the solution $\hat{x}^{64}$ is quite specific to $A(q)$. When l is small the perturbed relative error is large and similar to the relative error because in each case $\hat{x}$, and therefore $A\hat{x}$ is becoming small.

The best fits to the perturbed matrix equation are in the range $l = 48$ to $l = 58$. Qualitatively the solutions change little over this range and in Fig. 4 we present the result for $l = 58$. For comparison we have shown the solution from p. 219 of BBHM in Fig. 5. Since we have used their solution to define the sign of our v_i any negative value that we find indicates a solution opposite to theirs. This occurs where we find downwelling from the surface to deep water in the Arctic Ocean.

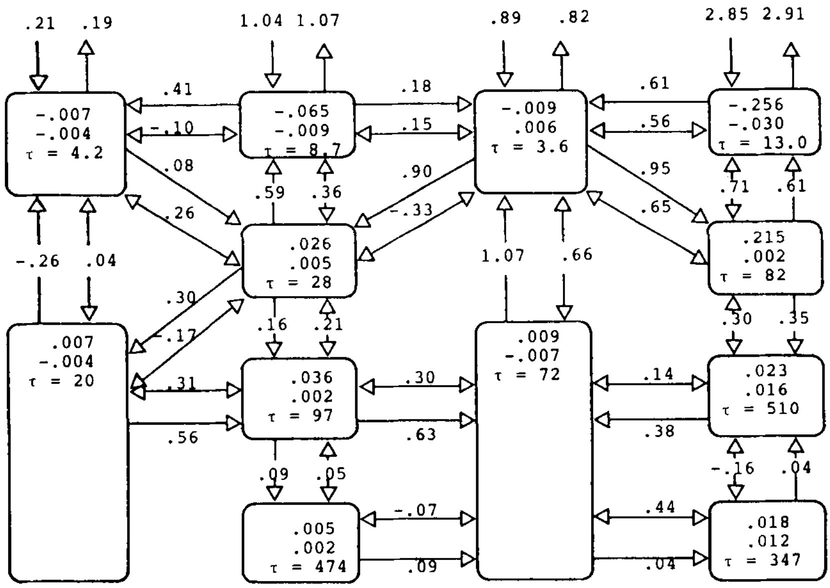

Fig. 4. Deduced fluxes for basic set of equations with truncation at $l = 58$. Between boxes: advective and turbulent fluxes of water (10^{15} m^3 yr^{-1}). Flow into or out of surface boxes 10^{15} mol C yr^{-1}. In boxes: net organic detritus formation and loss of carbon (−) or decomposition and gain of carbon (+) in 10^{15} mol C yr^{-1} (upper line), net carbonate formation and loss of carbon (−) or dissolution and gain of carbon in 10^{15} mol C yr^{-1} (middle line), and turnover time (τ) for water in box (lower line).

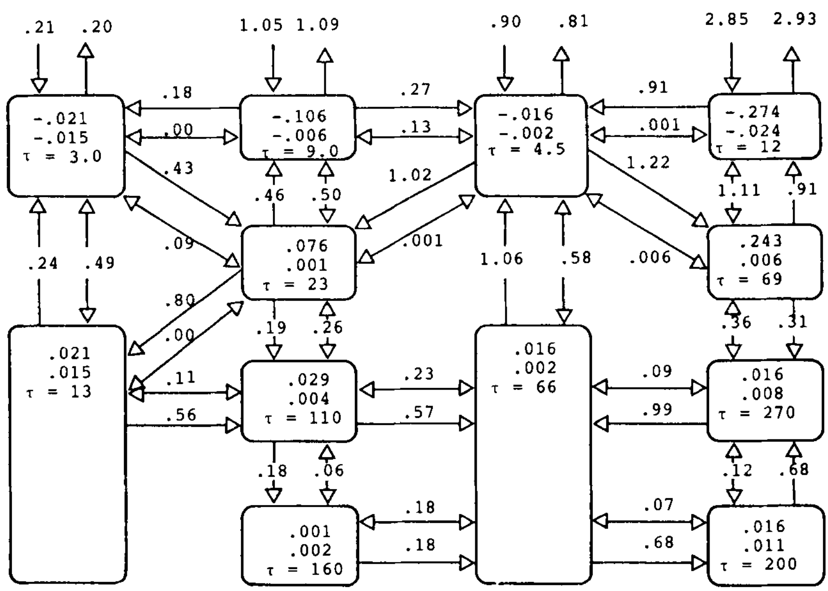

Fig. 5. Deduced fluxes in Fig. 8 of BBHM. Notation as in our Fig. 4.

Five of our diffusion coefficients are negative. The same difficulty was experienced by BBHM. They solved it by imposing positivity constraints on the coefficients. We will describe our use of constraints later. Interestingly, the most negative diffusion is associated with the Antarctic Ocean where the inorganic detritus formation also has the wrong sign and the organic detritus formation is too weak. The Antarctic Ocean is also associated with the largest mean square errors. Phosphorus conservation errors in boxes 7 and 10 (the Antarctic surface and the Pacific/Indian intermediate) account for 67.5% of the mean square error. If we include phosphorus conservation errors from the other 4 boxes that directly connect with the Antarctic surface (i.e., boxes 3, 4, 8 and 9) then these account for 77% of the total mean square error.

Antarctic detritus formation is also a problem in BBHM. They find $-0.030 \leqslant F_{o7} \leqslant -0.02$ and $-0.006 \leqslant F_{c7} \leqslant 0$. They note that the magnitudes are too small, and their error bounds clearly indicate that in some experiments F_{c7} is determined by their lower bound.

In order to consider a different class of solutions, we have also investigated the ridge-estimators defined by

$$\psi_i = \lambda_i/(\lambda_i^2 + \Gamma). \tag{4.2}$$

This approach is equivalent to finding the $\hat{\boldsymbol{x}}$ which minimizes

$$\theta = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2 + \Gamma\|\boldsymbol{x}\|^2. \tag{4.3}$$

Fig. 6 describes the solutions as Γ varies showing $\|\hat{\boldsymbol{x}}\|^2$, $\|\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{b}\|^2/\|\boldsymbol{b}\|^2$ and the perturbed error $\|\boldsymbol{A}(\boldsymbol{q} + \delta\boldsymbol{q})\hat{\boldsymbol{x}} - \boldsymbol{b}(\boldsymbol{q} + \delta\boldsymbol{q})\|^2/\|\boldsymbol{b}(\boldsymbol{q} + \delta\boldsymbol{q})\|^2$ as in Fig. 3, as well as θ, the function that was minimized. The curves in Fig. 6 are smoother versions of those in Fig. 3. For $\Gamma = 0.004$, we have $\|\boldsymbol{x}\|^2 = 3.95$, $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2/\|\boldsymbol{b}\|^2 = 0.663 \times 10^{-3}$ and $\|\boldsymbol{A}(\boldsymbol{q} + \delta\boldsymbol{q})\hat{\boldsymbol{x}} - \boldsymbol{b}(\boldsymbol{q} + \delta\boldsymbol{q})\|^2/\|\boldsymbol{b}(\boldsymbol{q} + \delta\boldsymbol{q})\|^2 = 0.371 \times 10^{-1}$. For the $l = 58$ truncation, the same quantities have values 4.07, 0.728×10^{-3} and 0.347×10^{-1}, respectively, as would be expected since $\lambda_{59}^2 < \Gamma < \lambda_{58}^2$. The solution for $\hat{\boldsymbol{x}}$ for $\Gamma = 0.004$ is shown in Fig. 7 and is similar to that from the $l = 58$ truncation. The same 5 diffusion coefficients are negative, the Antarctic carbonate transfer is still the wrong sign and there is still a strong downwelling from the Arctic surface to deep water.

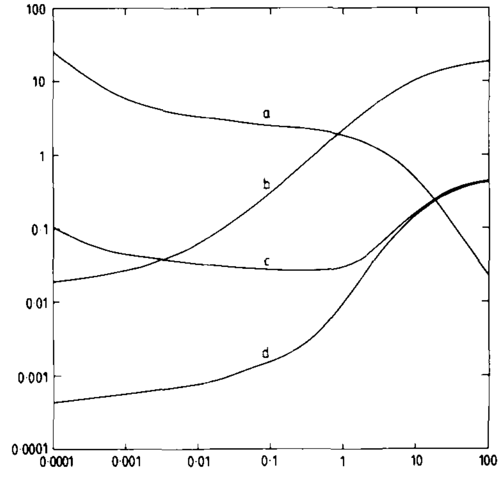

Fig. 6. Solution diagnostics as Γ, the smoothing parameter, varies: (a) $\|\hat{\boldsymbol{x}}\|^2$; (b) $\|\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{b}\|^2 + \Gamma\|\hat{\boldsymbol{x}}\|^2$; (c) $\|(\boldsymbol{A} + \delta\boldsymbol{A})\hat{\boldsymbol{x}} - \boldsymbol{b} - \delta\boldsymbol{b}\|^2/\|\boldsymbol{b} + \delta\boldsymbol{b}\|^2$; (d) $\|\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{b}\|^2/\|\boldsymbol{b}\|^2$.

Although there are some substantial quantitative differences between the two solutions it is clear that the biased estimation procedures are not proving adequate for avoiding incorrect signs in the manner mentioned by Goldstein and Smith (1974). In the work of BBHM particular coefficients were constrained by strict inequalities to be of the required sign. If a particular inequality constraint is active then the corresponding x_i will be zero. This approach presents two difficulties. Firstly, the results cannot be readily interpreted in terms of a singular value decomposition, the advantages of which have been described earlier. Secondly, because of the large value of the condition number the problem is effectively underdetermined and could display an instability analogous to that described by Björkström (1985) for an underdetermined problem. Accordingly, we have introduced constraints as extra equations that, like all our other equations, are only required to be satisfied approximately. We began by requiring that the five negative diffusion coefficients be (approximately) equal to zero. The equations specifying this condition were given a weighting of 25.

In Fig. 8, $\|\boldsymbol{x}\|^2$, $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2/\|\boldsymbol{b}\|^2$ and $\|\boldsymbol{A}(\boldsymbol{q} + \delta\boldsymbol{q})\boldsymbol{x} - \boldsymbol{b}(\boldsymbol{q} + \delta\boldsymbol{q})\|^2/\|\boldsymbol{b}(\boldsymbol{q} + \delta\boldsymbol{q})\|^2$ are plotted as functions of the truncation l. Setting the 5 coefficients to zero has severely reduced $\|\boldsymbol{x}\|^2$ for

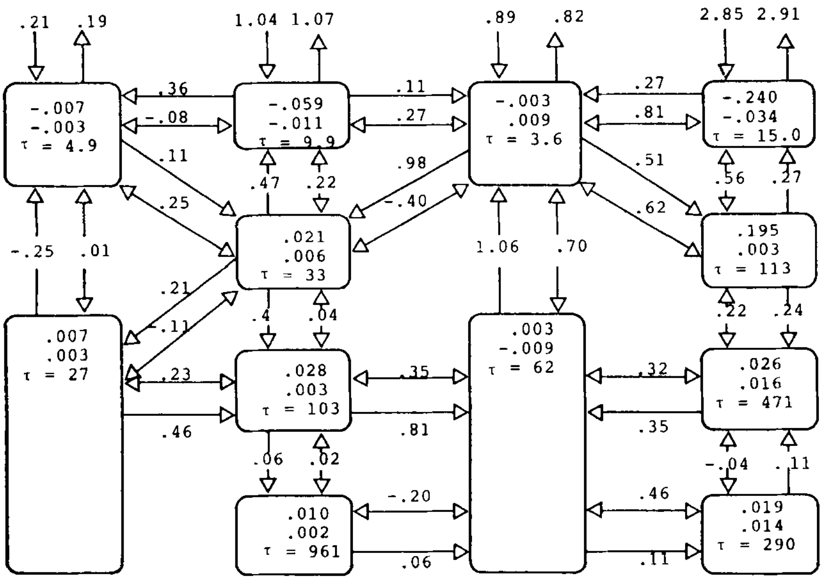

Fig. 7. Deduced fluxes for basic set of equations with smoothing $\Gamma = 0.004$. Notation as in Fig. 4.

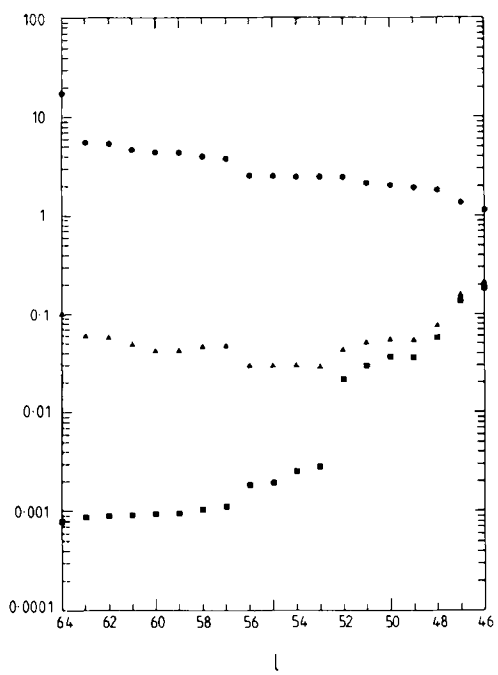

Fig. 8. $\|\hat{x}\|^2$ (indicated by ●), $\|A\hat{x} - b\|^2/\|b\|^2$ (indicated by ■), and $\|(A + \delta A)\hat{x} - b - \delta b\|^2/\|b + \delta b\|^2$ (indicated by ▲), for truncations $l = 64$ to 46.

$l = 64$, but examination of the solution still shows it to be unrealistic.

The greater stability of this constrained solution is seen by examining the "error" estimates for the smallest eigenvalue. Compared to the unconstrained problem λ_{64} is larger (0.018 *cf.* 0.0099) and the "error bounds" are also closer.

The variation of the $\|A(q + \delta q)\hat{x} - b(q + \delta q)\|^2/\|b(q + \delta q)\|^2$ with l suggests that solutions for l from 56 to 60 might be appropriate. These are illustrated in Figs. 9 and 10. Fig. 11 shows the $l = 58$ solution. Taken together, these figures illustrate the uncertainties in the solution method. Overall, the $l = 60$ truncation appears in closest agreement with BBHM, but both Figs. 9 and 10 still show Arctic downwelling. Although Fig. 11 shows a weak Arctic upwelling, it is the only truncation to do so. This behaviour is an indication that the Arctic circulation has a high degree of sensitivity to the solution technique.

The transfer of detrital carbonate is incorrect for all the 3 cases. For $l = 60$ and $l = 58$ this does not seem to adversely influence other aspects of the solution but for $l = 56$ the circulation associated with the Antarctic Ocean is unrealistic.

Since $(\lambda_{60})^2 \approx 0.004$ the solution from the ridge estimator with $\Gamma = 0.004$ is very similar to the

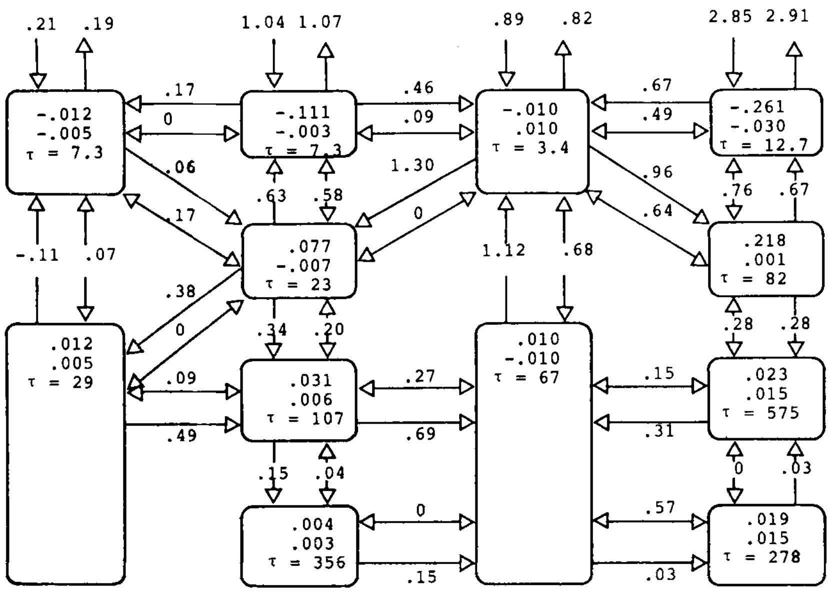

Fig. 9. Deduced fluxes for basic set of equations and diffusion fluxes 2, 4, 10, 12 and 20 constrained to be near zero. The truncation parameter $l = 60$. Notation as in Fig. 4.

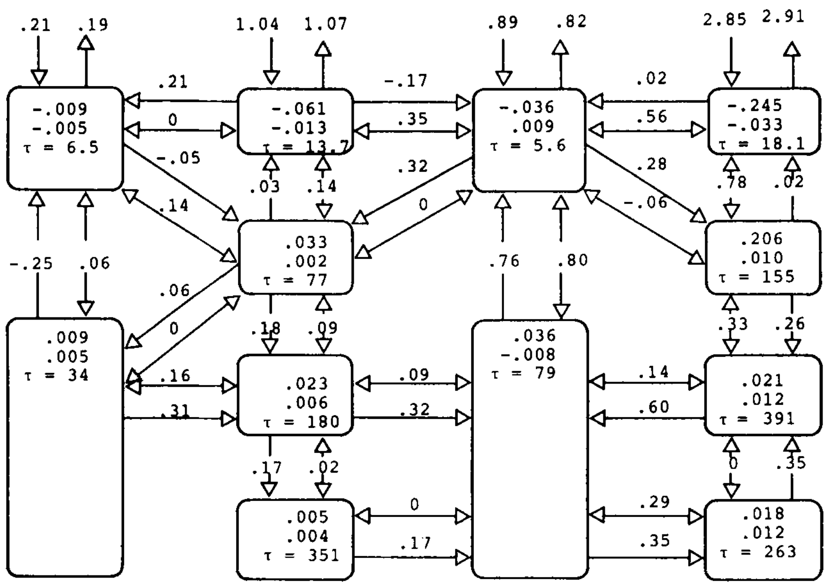

Fig. 10. Deduced fluxes for the equations as in Fig. 9 except that $l = 56$.

circulation in Fig. 9. One interesting difference however is that the weak circulation from the Atlantic deep water to the bottom box and then to the Antarctic has almost totally vanished. The circulation found earlier was opposite to the common view and our result suggests that this feature is not a stably determined part of the solution. We also found the ridge estimator solution with $\Gamma = 0.0014$, the value suggested by the expression of Hoerl et al. (1975) (see eq. 2.18

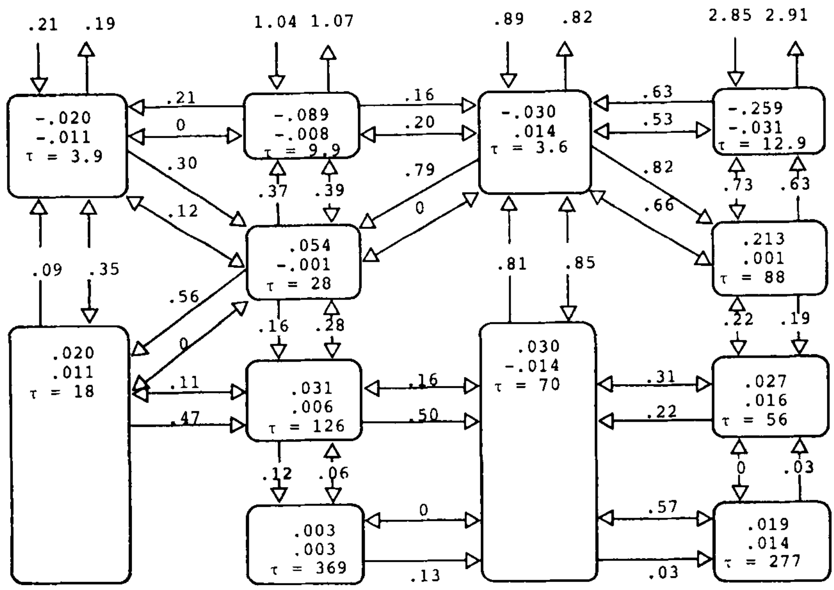

Fig. 11. Deduced fluxes for the equations as in Fig. 9 except that $l = 58$.

above). The solution is little different from the $\Gamma = 0.004$ case except that the circulation from the Atlantic deep water to bottom water to the Antarctic has changed sign and is now 0.19 units in the expected direction.

In further experiments we have applied additional constraints to the detrital fluxes. We have found that fixing one of F_{o7} and F_{c7} simply causes the other to be worse in the solution. When we set $F_{o7} = -0.040$ and $F_{c7} = -0.005$ with a scaling factor of 100, the rsulting solution has additional negative diffusion coefficients. The solution when all the diffusion coefficients involving the Antarctic surface water and other oceans are set to zero (in addition to all the other constraints described above) is shown in Fig. 12 (for the $l = 58$ truncation).

Although the circulation is quantitatively different to the others it still shows many of the same features: the Arctic downwelling and the Atlantic bottom water formation are as usually found in the model. However, among the differences, we observe upwelling of Atlantic deep water for the first time. This also occurs in one case from BBHM (their Fig. 11) when they use different Redfield ratios. Also there is a strong flow of bottom Antarctic water to the bottom Pacific/Indian water—something not seen before in our experiments.

5. The airborne fraction

As indicated above, one of the main reasons for the development of more refined ocean models such as that of BBHM is the question of whether discrepancies in the atmospheric CO_2 budget are due to inadequacies in existing ocean models.

In order to examine the oceanic uptake of CO_2 arising from the various solutions found in the previous section we refer back to eq. (3.1). Up to this point we have taken $\partial q_i^y/\partial t = 0$ but in this section we consider time varying perturbations about a steady state. For simplicity we consider an exponential growth in atmospheric partial pressure of CO_2 so we put

$$p_a = p_a(-\infty) + \alpha\, e^{\eta t}. \tag{5.1}$$

We assume a uniform buffer factor of $\xi = 10$ and so take the partial pressure of CO_2 in box i as

$$p_i = p_i(-\infty)[1 + \xi\, \Delta q_i^C\, e^{\eta t}/q_i^C(-\infty)]. \tag{5.2}$$

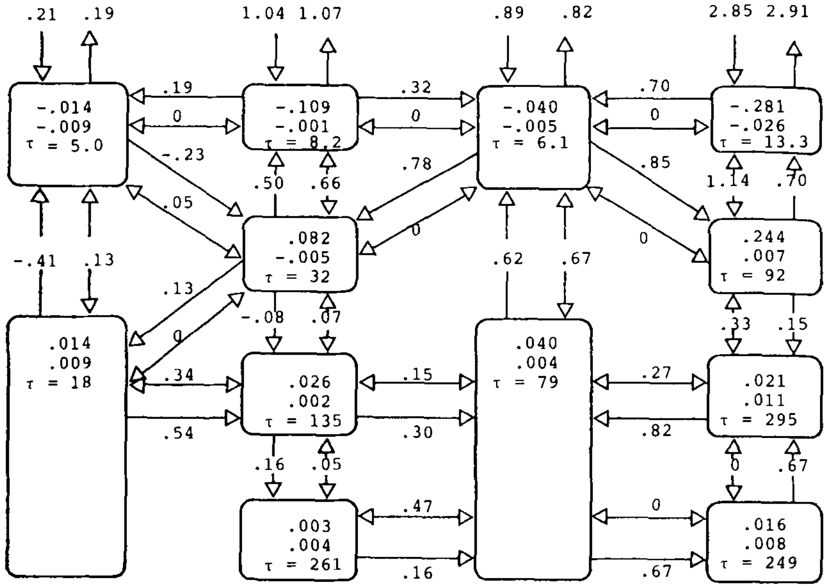

Fig. 12. Deduced fluxes for basic set of equations with diffusion fluxes 2, 4, 9, 10, 14, 15, 17, and 20 set to zero and Atlantic organic and carbonate surface detrital terms set to -0.040 and -0.005.

The quantities $p_a(-\infty)$, $p_i(-\infty)$, $q_i^C(-\infty)$ were defined by the steady state solution. Writing each ocean carbon concentration in the form

$$q_i^C(t) = q_i^C(-\infty) + \Delta q_i^C\, e^{\eta t} \tag{5.3}$$

gives

$$\frac{\partial}{\partial t} q_i^C = \eta\, \Delta q_i^C\, e^{\eta t}. \tag{5.4}$$

Substituting these expressions into the carbon conservation equations and removing a common factor of $e^{\eta t}$ gives a set of equations

$$\boldsymbol{D}\boldsymbol{\Delta q} = \alpha \boldsymbol{e}. \tag{5.5}$$

The $\alpha\boldsymbol{e}$ term arises from substituting (5.1) into the term for the surface fluxes of carbon

$$F_{ai}^C = kA_i(p_a - p_i). \tag{5.6}$$

Here k is a piston velocity and A_i is the surface area of a given ocean.

The matrix $\boldsymbol{D}$ is square and non-singular and so eq. (5.5) is readily solved for $\boldsymbol{\Delta q}$. The airborne fraction is easily calculated by converting the atmospheric CO_2 increase in (5.1) into a mass and expressing this as a proportion of the total of this atmospheric perturbation and the mass perturbation of ocean carbon. Fig. 13 shows the airborne fraction as a function of η, the growth rate, for the transports shown in Figs. 5 and 9. Although there are substantial differences between these circulations the computed airborne fractions differ by less than 1%.

Bacastow and Björkström (1981) have carried out a detailed comparison of the CO_2 uptake by different ocean models and have prepared a figure directly comparable to our Fig. 13 (their Fig. 11 p. 46). This shows that the present model barely differs from other ocean models in its CO_2 uptake.

As none of these ocean models appears able to explain the missing carbon problem it is interesting to examine what changes would enable the present model to do so. We have calculated the airborne fraction using the advective and diffusive transport coefficients of BBHM scaled by a factor β. As Fig. 13 shows, it is not very important whether we use the coefficients in BBHM or any of those found earlier in this section.

The airborne fraction a_f is a monotonic decreasing function of β. It is almost linear in $\ln \beta$ until $\beta \approx 10$, and charges slowly. For $\eta = 0.03478$ (an e-folding time of 28.75 years) a_f does not

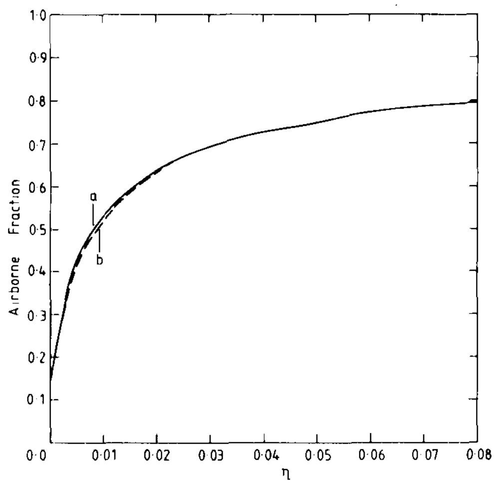

Fig. 13. Airborne fraction as a function of η for the transports shown in Fig. 5 (curve a) and in Fig. 9 (curve b).

reach 0.5 until $\beta = 9$; it has an asymptotic value of 0.3553 as $\beta \to \infty$. For a faster CO_2 growth rate where $\eta = 0.04546$ (a 22 year *e*-folding time) a_f is not 0.5 until $\beta = 18$, with an asymptotic value of 0.3985. Such large values of advective and diffusive fluxes are totally unrealistic. It therefore appears that no reasonable "tuning" of the fluxes will be able to substantially alter the calculated airborne fraction.

6. Conclusion

In the preceding sections we have investigated the application of a number of biased estimation procedures to the ocean model described by Bolin et al. (1983), (BBHM). These procedures were considered as alternatives to the least squares fits used by BBHM in their "incompatible case". We found, as expected, that the least-squares estimator was unduly sensitive to errors in the data and it will also be sensitive to errors in the model formulation. The biased estimators of the "singular value truncation" or "ridge" type were found to be suitable for reducing this sensitivity.

Nevertheless, in spite of the theoretical expectations of improved performance from ridge estimators and our numerical experiments on the perturbed relative error, our use of biased estimators did not succeed in avoiding many of the anomalous features found in BBHM. In particular serious problems were encountered with the detrital flux in the Antarctic region. We would conclude that this anomaly is due to either gross errors in the data or a severe "mismatch" between the data that is fitted and the way in which the Antarctic region is being modelled. As noted by BBHM, the treatment of phosphorus seems to be a particular problem.

While the various calibration procedures gave a variety of different ocean circulations there was very little variation in the corresponding airborne fractions of CO_2. Our study (and that of Bacastow and Björkström, 1981) indicates that this type of ocean model cannot take up the large amounts of carbon suggested by the results of ecosystem modelling. To absorb such amounts would require some mechanism such as an outcrop of deep water (as considered by Peng et al., 1983 or Seigenthaler, 1983) that serves to bring undersaturated water to the surface in polar regions.

7. Appendix

A.1. Behaviour of $\|\boldsymbol{x}\|^2$ *under regularization*

We choose

$$\hat{\boldsymbol{x}} = \boldsymbol{W\Psi U}^{\mathrm{T}}\boldsymbol{b}. \tag{A1}$$

Let us write

$$\mu_i = (\boldsymbol{U}^{\mathrm{T}}\boldsymbol{b})_i = \sum_{j=1}^{m} U_{ji} b_j. \tag{A2}$$

Now as Ψ is a diagonal $n \times m$ matrix

$$(\boldsymbol{W\Psi})_{ij} = \begin{cases} \psi_j W_{ij}, & j \leqslant p', \\ 0, \text{ otherwise}, \end{cases} \tag{A3}$$

where $p' = \min(m, n)$. Substituting eqs. (A2) and (A3) into eq. (A1) gives

$$\hat{x}_i = \sum_{k=1}^{p'} \psi_k W_{ik} \mu_k, \tag{A4}$$

and so, because $\boldsymbol{W}$ is orthogonal, we get

$$\|\hat{\boldsymbol{x}}\|^2 = \sum_{i=1}^{p'} \psi_i^2 \mu_i^2. \tag{A5}$$

Any decrease in the magnitude of ψ_{ik} therefore causes $\|\hat{\boldsymbol{x}}\|^2$ to decrease.

A.2. $\|A\hat{x} - b\|^2$ *in terms of a singular value decomposition*

Combining eqs. (2.3), (2.12) and (2.13) gives

$$A\hat{x} = U\Lambda W^{T} W\Psi U^{T} b$$

whence

$$A\hat{x} = U\Lambda\Psi\mu, \tag{A6}$$

using eq. (A2) and the orthogonality of W. Now

$$[\Lambda\Psi]_{ij} = \begin{cases} \lambda_i \psi_i \delta_{ij}, & i, j \leqslant p', \\ 0 \text{ otherwise}, \end{cases} \tag{A7}$$

where δ_{ij} is the Kronecker delta, and so

$$[U\Lambda\Psi]_{ij} = \begin{cases} \lambda_j \psi_j U_{ij}, & j \leqslant p' \\ 0 \text{ otherwise}. \end{cases} \tag{A8}$$

Substituting eq. (A8) into (A6) gives

$$(A\hat{x})_i = \sum_{k=1}^{p'} \lambda_k \psi_k \mu_k U_{ik}, \tag{A9}$$

and $A\hat{x}$ is expressed as a sum of the column vectors of U. These column vectors, being orthonormal, span the space of m-vectors and so, using eq. (A2), we find that

$$b_i = \sum_{k=1}^{m} \mu_k U_{ik}. \tag{A10}$$

Combining eqs. (A9) and (A10), and again using the orthonormality of U, we obtain

$$\|A\hat{x} - b\|^2 = \begin{cases} \sum_{k=1}^{n} (1 - \lambda_k \psi_k)^2 \mu_k^2 + \sum_{k=n+1}^{m} \mu_k^2, & m > n, \\ \sum_{k=1}^{m} (1 - \lambda_k \psi_k)^2 \mu_k^2, & m \leqslant n. \end{cases} \tag{A11}$$

Therefore if $m > n$ we obtain a least squares best fit by taking $\psi_k = \lambda_k^{-1}$ when $\lambda_k \neq 0$. The same definition of ψ_k will give an exact solution if A has rank $m \leqslant n$.

Additionally, as ψ_k decreases from λ_k^{-1}, then $|A\hat{x} - b|^2$ will increase.

A.3. $\|A\hat{x} - b\|^2 + \Gamma\|x\|^2$ *in terms of a singular value decomposition*

Combining eqs. (A5) and (A11) gives

$$\|A\hat{x} - b\|^2 + \Gamma\|\hat{x}\|^2 = \begin{cases} \sum_{k=1}^{n} (1 - \lambda_k \psi_k)^2 \mu_k^2 + \sum_{k=n+1}^{m} \mu_k^2 \\ \quad + \Gamma \sum_{k=1}^{n} \psi_k^2 \mu_k^2, & m > n \\ \sum_{k=1}^{m} (1 - \lambda_k \psi_k)^2 \mu_k^2 \\ \quad + \Gamma \sum_{k=1}^{m} \psi_k^2 \mu_k^2, & m \leqslant n. \end{cases} \tag{A12}$$

On substituting

$$\psi_k = \lambda_k/(\lambda_k^2 + \Gamma), \tag{A13}$$

we find, after some straightforward algebra, that

$$\|A\hat{x} - b\|^2 + \Gamma\|\hat{x}\|^2 = \begin{cases} \sum_{k=1}^{n} \Gamma\mu_k^2/(\lambda_k^2 + \Gamma) + \sum_{k=n+1}^{m} \mu_k^2, & m > n \\ \sum_{k=1}^{n} \Gamma\mu_k^2/(\lambda_k^2 + \Gamma), & m \leqslant n. \end{cases} \tag{A13}$$

Clearly $\|A\hat{x} - b\|^2 + \Gamma\|\hat{x}\|^2$ is a monotonic increasing function of Γ.

Notation

a	used as a subscript to denote atmosphere.
a_f	airborne fraction of CO_2.
A	Matrix defining estimation problem (dimension $m \times n$).
A_i	Surface area of ocean region i.
b	"Data" vector in linear estimation problem.
$c(\lambda_i, \Gamma)$	Function defining a class of regularizations.
C	Superscript denoting carbon.
D	Matrix used to express differential equation for exponential growth as an algebraic equation.
e	Forcing vector for exponential growth of CO_2.
$E[\cdot]$	Statistical expectation with respect to errors ε.
i, j, k	Reservoir indices or general numerical indices.

k	Piston velocity for air-sea gas exchange.
l	Cut-off number for singular value truncation method.
m	dimension of $\boldsymbol{b}$.
n	Dimension of $\boldsymbol{x}$.
p	CO_2 partial pressures.
q_i^y	Concentration of tracer y in reservoir i.
$\boldsymbol{U}$	Orthogonal matrix in singular value decomposition of $\boldsymbol{A}$.
$\boldsymbol{W}$	Orthogonal matrix in singular value decomposition of $\boldsymbol{A}$.
$\boldsymbol{x}$	Parameter vector.
$\boldsymbol{x}_{\text{true}}$	Idealized unknown "true" parameter vector.
$\hat{\boldsymbol{x}}$	Any statistical estimator for $\boldsymbol{x}_{\text{true}}$.
y	Tracer index (one of C, ^{14}C, A, P, O or W).
α	Scale factor for exponential CO_2 increase.
β	Scale factor for transport terms.
Γ	Ridge parameter.
ε	Random errors in $\boldsymbol{b}$.
η	growth rate of atmospheric CO_2.
λ_n	nth (in decreasing order) of singular values of $\boldsymbol{A}$, i.e., Λ_{nn}.
λ_y	Decay rate for tracer y. Zero except for carbon-14.
$\boldsymbol{\Lambda}$	Diagonal matrix in singular value decomposition of $\boldsymbol{A}$.
ξ	Buffer factor for CO_2.
ψ_n	nth diagonal element in regularized inverse of $\boldsymbol{\Lambda}$, i.e., Ψ_{nn}.
$\boldsymbol{\Psi}$	Regularized inverse of $\boldsymbol{\Lambda}$.

REFERENCES

Bacastow, R. and Björkström, A. 1981. Comparison of ocean models for the carbon cycle. *Carbon cycle modelling* SCOPE 16, ed. B. Bolin. John Wiley and Sons, Chichester, pp. 29–79.

Björkström, A. 1979. A model of CO_2 interaction between atmosphere, oceans, and land biota. *The global carbon cycle*. SCOPE 13, eds. B. Bolin, E. T. Degens, S. Kempe and P. Ketner. John Wiley and Sons, Chichester, pp. 403–457.

Björkström, A. 1985. A numerical instability that occurs when optimizing an underdetermined set of linear equations. In *Seminars on geophysical instabilities*, ed. E. Källén. Report DM-47, Department of Meteorology, University of Stockholm.

Bolin, B., Björkström, A., Holmén, K. and Moore, B. 1983. The simultaneous use of tracers for ocean circulation studies. *Tellus 35B*, 206–236.

Broecker, W. S. 1981. Geochemical tracers and ocean circulation. *Evolution of physical oceanography*, ed. B. A. Warren and C. Wunsch. MIT Press, Cambridge, Mass., pp. 434–460.

Broecker, W. S., Peng, T.-H. and Engh, R. 1980. Modelling the carbon system. *Radiocarbon 22*, 565–598.

Craig, H. 1957. The natural distribution of radiocarbon and the exchange time of carbon dioxide between the atmosphere and sea. *Tellus 9*, 1–17.

Crane, A. J. 1982. The partitioning of excess CO_2 in a five-reservoir atmosphere-ocean model. *Tellus 34*, 398–405.

Deutsche, R. 1965. *Estimation theory*. Prentice Hall, Englewood Cliffs, N.J.

Enting, I. G. 1985. Principles of constrained inversion in the calibration of carbon cycle models. *Tellus 37B*, 7–27.

Goldstein, M. and Smith, A. F. M. 1974. Ridge-type estimators for regression analysis. *J. R. Stat. Soc. B36*, 284–291.

Golub, G. H., Heath, M. and Wahba, G. 1979. Generalised cross-validation as a method for choosing a good ridge parameter. *Technometrics 21*, 215–223.

Hoerl, A. E. and Kennard, R. W. 1970. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics 12*, 55–67.

Hoerl, A. E. Kennard, R. W. and Baldwin, K. F. 1975. Ridge regression: some simulations. *Communications in statistics, A—Theory and methods*, pp. 105–123.

Houghton, R. A., Hobbie, J. E., Melillo, J. M., Moore, B., Peterson, B. J., Shaver, G. R. and Woodwell, G. M. 1983. Changes in the carbon content of terrestrial biota and soils between 1860 and 1980: a net release of CO_2 to the atmosphere. *Ecol. Monogr. 53*, 235–262.

Jackson, D. D. 1972. Interpretation of inaccurate, insufficient and inconsistent data. *Geophys. J. R. Astron. Soc. 28*, 97–109.

Keeling, C. D. 1973. The carbon dioxide cycle: Reservoir models to depict the exchange of atmospheric carbon dioxide with the oceans and land plants. *Chemistry of the lower atmosphere*, ed. S. I. Rasool. Plenum, New York, pp. 251–329.

Moore, B. and Björkström, A. 1986. Developing parameters for ocean carbon models by the inverse method. In *The changing carbon cycle: a global analysis*, eds. J. Trabalka and D. Reichle. Springer-Verlag, in press.

Oeschger, H., Siegenthaler, U., Schotterer, U. and Gugelmann, A. 1975. A box diffusion model to study the carbon dioxide exchange in nature. *Tellus 27*, 168–192.

Pearman, G. I. 1980. Preliminary studies with a new global carbon cycle model. *Carbon dioxide and climate: Australian research*, ed. G. I. Pearman. Australian Academy of Science, Canberra, pp. 79–91.

Peng, T.-H., Broecker, W. S., Freyer, H. D. and Trumbore, S. 1983. A deconvolution of the tree-ring based $\delta^{13}C$ record. *J. Geophys. Res. 86C*, 3609–3620.

Redfield, A. C. 1958. The biological control of chemical factors in the environment. *Am. Sci. 46*, 206–226.

Siegenthaler, U. 1983. Uptake of excess CO_2 by an outcrop diffusion model of the ocean. *J. Geophys. Res. 88C*, 3599–3608.

Twomey, S. 1977. *Introduction to the mathematics of inversion in remote sensing and indirect measurements.* Elsevier Scientific Publishing Co., Amsterdam.