

# UC Irvine

## Faculty Publications

### Title

Optimization and sensitivity of a global biogeochemistry ocean model using combined in situ DIC, alkalinity, and phosphate data

### Permalink

<https://escholarship.org/uc/item/1tf9c12w>

### Journal

Journal of Geophysical Research, 113(C8)

### ISSN

0148-0227

### Authors

Kwon, Eun Young  
Primeau, Francois

### Publication Date

2008-08-07

### DOI

10.1029/2007JC004520

### Supplemental Material

<https://escholarship.org/uc/item/1tf9c12w#supplemental>

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Optimization and sensitivity of a global biogeochemistry ocean model using combined in situ DIC, alkalinity, and phosphate data

Eun Young Kwon<sup>1,2</sup> and François Primeau<sup>1</sup>

Received 23 August 2007; revised 29 November 2007; accepted 21 April 2008; published 7 August 2008.

[1] We present a systematic parameter optimization and sensitivity analysis of a three-dimensional global ocean biogeochemistry model. We use the global data sets of dissolved inorganic carbon (DIC), alkalinity, and phosphate to constrain the parameters of a biogeochemistry model which include the stoichiometric ratios rC:P and rN:P, the fraction  $\sigma$  of organic material production allocated to dissolved organic matter (DOM), the lifetime  $1/\kappa$  of DOM, the exponent  $\alpha$  in the power law for the depth profile of the remineralization of particulate organic carbon (POC), the rain ratio  $R$  of  $\text{CaCO}_3$ , and the e-folding length scale  $d$  for the depth profile of  $\text{CaCO}_3$  dissolution. The data-constrained parameter values are rC:P =  $137 \pm 11$ ,  $\sigma = 0.74 \pm 0.04$ ,  $1/\kappa = 1.7 \pm 0.5$  years,  $\alpha = -0.97 \pm 0.07$ ,  $R = 0.081 \pm 0.008$ , and  $d = 2100 \pm 300$  m. The postoptimization carbon export from POC is  $15 \pm 1$  Gt/a and from  $\text{CaCO}_3$  is  $1.2 \pm 0.1$  Gt/a of which  $67 \pm 4\%$  dissolves above 2000 m. The  $\pm$  ranges indicate an average 1% decrease in the fraction of the spatial variance in the observed tracer data captured by the model. The sensitivity of the model to its parameters is presented in terms of sensitivity patterns defined as the derivative of the model's equilibrium tracer distribution with respect to the parameters ( $S$  patterns). The soft-tissue, carbonate, and gas exchange pump mechanisms responsible for the sensitivities are presented. The pump decomposition of the  $S$  patterns illustrates quantitatively how changes in organic and inorganic carbon fluxes are coupled with the large-scale ocean circulation and how the gas exchange pump couples to the global ventilation patterns through changes in surface chemistry.

**Citation:** Kwon, E. Y., and F. Primeau (2008), Optimization and sensitivity of a global biogeochemistry ocean model using combined in situ DIC, alkalinity, and phosphate data, *J. Geophys. Res.*, 113, C08011, doi:10.1029/2007JC004520.

## 1. Introduction

[2] A difficult problem in climate modeling is selecting the proper level of complexity for ocean biogeochemistry modules in order to minimize the uncertainty of atmospheric carbon dioxide predictions. While complex models can in principle better represent potential climate feedbacks and regional ecosystem variations, they have the disadvantage of having a larger number of parameters that are difficult to constrain from limited observational and experimental data. For example, the study of Friedrichs *et al.* [2006] has demonstrated in the context of a 1-D ecosystem model, that complex models can have worse predictive skill than their simpler counterparts when there is insufficient data to properly constrain a large number of parameters. The study of Friedrichs *et al.* [2006] also suggests that a proper

comparison of different models requires that each model's parameter set be first optimized objectively, and that the sensitivity of the model parameters should be taken into account. For global models there have been few systematic parameter sensitivity and optimization studies. Models often use parameter values that have been tuned in different circulation models, and parameter correlations are rarely taken into account.

[3] It is our opinion that one of the goals of global ocean biogeochemistry modeling should be to perform systematic parameter optimization and sensitivity analysis of a hierarchy of models of varying complexity. The usual approach to analyzing the sensitivity of a model consists of changing a parameter, rerunning the model long enough for the transients to die off and recording the change in the output. Because the time for a global model to reach a new equilibrium after a parameter is changed is on the order of several thousand years the usual "hand-tuning" approach to optimizing parameters can yield only limited results. We need an alternative that can overcome the excessive computational costs associated with the slow model spin-up. For this we have recently developed a fully implicit solver for obtaining steady state solutions of a global ocean biogeochemistry

<sup>1</sup>Department of Earth System Science, University of California, Irvine, California, USA.

<sup>2</sup>Now at Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey, USA.

**Table 1.** Description of Parameters

Parameter	Description
rC:P	Stoichiometric ratio of carbon to phosphorus
rN:P	Stoichiometric ratio of nitrogen to phosphorus
$\sigma$	fraction of organic carbon production allocated into DOC
$R$	ratio of CaCO <sub>3</sub> to POC production
$\kappa$	first-order decay rate constant for DOC
$\alpha$	exponent in the power law for the depth profile of the remineralization of POC
$d$	length scale for the depth profile of CaCO <sub>3</sub> dissolution

model [Kwon and Primeau, 2006, hereinafter KP06]. The new solver eliminates the need to explicitly time step the governing equations and thus makes it possible, for the first time, to conduct a systematic parameter sensitivity analysis of a global ocean biogeochemistry model.

[4] Here we continue the line of study begun in KP06 and present a systematic parameter optimization and sensitivity analysis for a biogeochemistry model based on the formulation used for phase 2 of the Ocean Carbon-Cycle Model Intercomparison Project (OCMIP-2), [Najjar *et al.*, 2007]. The results related to the phosphorous-cycle component of the model were reported previously in KP06. Here we focus on the model's carbon and alkalinity cycles.

[5] In OCMIP-2 type models, the carbon and nitrogen cycles are keyed to phosphorus using C:N:P stoichiometry ratios. Net community production is obtained semidiagnostically by restoring PO<sub>4</sub> toward the observed PO<sub>4</sub> field whenever the simulated surface PO<sub>4</sub> concentration exceeds the observed value. The production and cycling of particulate organic matter (POM), semilabile dissolved organic matter (DOM) and CaCO<sub>3</sub> are then related to the new production through the following parameters: (1) rC:P, the stoichiometric ratio of carbon to phosphorus, (2) rN:P, the stoichiometric ratio of nitrogen to phosphorus, (3)  $\sigma$ , the fraction of organic material production allocated to DOM, (4)  $R$ , the ratio of CaCO<sub>3</sub> export production to particulate organic carbon export production, (5)  $\kappa$ , the first-order decay rate constant for DOM, (6)  $\alpha$ , the exponent in the power law for the depth profile of the remineralization of POM, and (7)  $d$ , the length scale for the depth profile of CaCO<sub>3</sub> dissolution (see Table 1 and section 2.2). The reference OCMIP-2 model [Najjar *et al.*, 2007] uses parameter values culled from multiple studies, in which subsets of the parameters were tuned separately [Anderson and Sarmiento, 1994; Yamanaka and Tajika, 1996, 1997].

[6] A novel aspect of our study is that we combine global observations of PO<sub>4</sub>, dissolved inorganic carbon (DIC) and total alkalinity (TA) to optimize the entire parameter set simultaneously. Parameter interactions are thus taken into account, and by computing the shape of the cost function we are also able to obtain important information about the degree to which various parameters are constrained independently from the rest. We also perform a sensitivity analysis of the global equilibrium distributions of DIC and TA to changes in model parameters by computing sensitivity patterns defined as partial derivatives of equilibrium solutions with respect to the parameters. The sensitivity patterns help us explain the parameter interactions and the degree to which different parameters are constrained in our multiparameter optimization results.

[7] The sensitivity patterns also provide a unique opportunity to elucidate the inner workings of the ocean's carbon cycle. By decomposing the sensitivity patterns using the pump separation method of Volk and Hoffert [1985] as modified by Gruber and Sarmiento [2002] we can learn how different parameterized biogeochemical processes couple with the global circulation and the air-sea carbon exchange to redistribute carbon within the ocean. The result is a decomposition of the DIC sensitivities into components associated with the soft-tissue, carbonate, and gas exchange pumps that are particularly useful for obtaining a mechanistic understanding of the model's behavior.

[8] This paper is organized as follows. We present the implicit ocean biogeochemistry model in section 2, the parameter sensitivity analysis in section 3, followed by the parameter optimization study in section 4. Discussion and conclusions are presented in section 5.

## 2. Implicit Ocean Biogeochemistry Model

### 2.1. Transport Model

[9] As in KP06, we use an offline transport model coupled to an OCMIP-2 biogeochemistry model. The transport operator is derived from the time-averaged velocity and eddy-diffusivity tensor fields taken from the solution of a fully spun-up dynamical ocean general circulation model (OGCM). The offline transport model like the parent OGCM has a spatial horizontal resolution of approximately 3.75° × 3.75° and 29 vertical levels ranging in thickness from 50 m at the surface to 300 m near the bottom. The parent OGCM uses the KPP vertical mixing scheme [Large *et al.*, 1994] and the GM isopycnal eddy-mixing scheme [Gent and McWilliams, 1990]. Extensive details about the model's tracer transport characteristics can be found in the works of Primeau [2005], Primeau and Holzer [2006], and Holzer and Primeau [2006, 2008].

### 2.2. Biogeochemistry Model

[10] The biogeochemistry model uses the OCMIP-2 formulation described in Najjar *et al.* [2007]. Biological production is driven by the supply of phosphate according to the parameterization

$$J_{prod} = \frac{rC:P}{\tau} ([PO_4] - [PO_4]_{obs}) \Theta([PO_4]), \quad (1)$$

where rC:P is the C:P ratio and  $\tau$  is the restoring timescale.  $\Theta([PO_4]) \equiv \frac{1}{2}[1 + \tanh\{([PO_4] - [PO_4]_{obs})/\lambda\}]$  is a smoothed step function used to switch off production when the surface [PO<sub>4</sub>] drops below the observed value. The smoothing is needed for our implicit solver based on Newton's method to converge (KP06) but gives essentially the same solution as the original discontinuous step function formulation.

[11] For each mole of organic carbon produced, a fraction  $\sigma$  is allocated to dissolved organic carbon (DOC). The remainder is exported as particulate organic carbon (POC). Through rC:P the DOC field is exactly proportional to the dissolved organic phosphorous (DOP) field and does not require its own prognostic equation. The export flux of POC at the base of the euphotic zone ( $z = z_c$ ) is given by

$$F_o = (1 - \sigma) \int_{z_c}^0 J_{prod} dz. \quad (2)$$

[12] Since the phosphorus cycle is independent of the DIC and TA cycles, we first obtain equilibrium solutions for  $\text{PO}_4$  and DOP separately as was done in KP06, and then use the resulting  $J_{prod}$  and DOP fields to drive the TA and DIC models.

[13] To compute the air-sea fluxes of carbon the solution of the seawater  $\text{CO}_2$  system is needed. For this the surface water [TA] is needed. The TA field is obtained (independently of the DIC solution) by solving

$$\frac{\partial[\text{TA}]}{\partial t} + \mathbf{u} \cdot \nabla[\text{TA}] - \nabla(K\nabla[\text{TA}]) = J_{v\text{TA}} + J_{b\text{TA}}. \quad (3)$$

[14]  $J_{v\text{TA}}$  is the virtual flux of TA due to the concentrating and diluting effects of freshwater evaporation and precipitation.  $J_{b\text{TA}}$  is the source-sink term of TA due to the biological production and remineralization of organic nitrogen and  $\text{CaCO}_3$ ,

$$J_{b\text{TA}} = -\frac{r\text{N:P}}{r\text{C:P}} J_{b\text{C}_{org}} + 2J_{b\text{CaCO}_3}, \quad (4)$$

where rN:P is the N:P ratio. The organic carbon sink-source term for DIC above and below the base of the euphotic zone is given by

$$J_{b\text{C}_{org}} = r\text{C:P} \times \kappa[\text{DOP}] - J_{prod} \text{ for } z > z_c \quad (5)$$

and

$$J_{b\text{C}_{org}} = r\text{C:P} \times \kappa[\text{DOP}] + \frac{\partial}{\partial z} F_{\text{POC}}(z) \text{ for } z < z_c, \quad (6)$$

in which  $F_{\text{POC}}(z)$  is the downward flux of POC parameterized using a power law as in the work of *Martin et al.* [1987] and  $z_c = -75$  m,

$$F_{\text{POC}}(z) = \left(\frac{z}{z_c}\right)^\alpha F_o. \quad (7)$$

[15] In addition to exporting organic carbon, marine organisms can also export  $\text{CaCO}_3$  hard shells. This export flux is taken to be proportional to the export of POC. The rain ratio  $R$  is the proportionality constant at the base of the euphotic zone. The exported  $\text{CaCO}_3$  is assumed to dissolve in the water column according to the vertical flux profile represented by an exponential curve. Above and below the base of the euphotic zone, the  $\text{CaCO}_3$  sink-source term for DIC takes the form:

$$J_{b\text{CaCO}_3} = -(1 - \sigma)R \times J_{prod} \text{ for } z > z_c \quad (8)$$

and

$$J_{b\text{CaCO}_3} = \frac{\partial}{\partial z} F_{\text{CaCO}_3}(z) \text{ for } z < z_c \quad (9)$$

where

$$F_{\text{CaCO}_3}(z) = R \times F_o \exp\left(\frac{z - z_c}{d}\right) \quad (10)$$

is the downward flux of  $\text{CaCO}_3$ . The parameter  $d$  is the length scale for the remineralization of  $\text{CaCO}_3$ . In this model, there is no sedimentation of  $\text{C}_{org}$  and  $\text{CaCO}_3$ . All the flux that enters through the top of the bottom grid box is assumed to be remineralized within the bottom box.

[16] Once the solutions for the  $\text{PO}_4$  and TA cycles are known, the DIC field can be obtained by solving

$$\frac{\partial[\text{DIC}]}{\partial t} + \mathbf{u} \cdot \nabla[\text{DIC}] - \nabla(K\nabla[\text{DIC}]) = J_{v\text{DIC}} + J_{g\text{DIC}} + J_{b\text{DIC}}. \quad (11)$$

[17]  $J_{v\text{DIC}}$  is the virtual flux of DIC due to the concentrating and diluting effects of freshwater evaporation and precipitation.  $J_{g\text{DIC}}$  denotes the surface flux of carbon due to air-sea gas exchange and is a function of both DIC and TA. The formulation and parameterization for air-sea carbon fluxes follow the OCMIP-2 protocol [*Najjar et al.*, 2007] of which the computation of the gas transfer velocity is adapted from *Wanninkhof* [1992].  $J_{b\text{DIC}}$  is the source-sink term of DIC due to the biological production and remineralization of carbon,

$$J_{b\text{DIC}} = J_{b\text{C}_{org}} + J_{b\text{CaCO}_3}. \quad (12)$$

### 2.3. Implicit Equilibrium Solver

[18] The model's governing equations can be expressed symbolically as follows

$$\frac{d}{dt} \begin{bmatrix} [\text{PO}_4] \\ [\text{DOP}] \end{bmatrix} = G_1 \left( \begin{bmatrix} [\text{PO}_4] \\ [\text{DOP}] \end{bmatrix}; \mu \right), \quad (13)$$

$$\frac{d}{dt} [\text{TA}] = G_2 \left( [\text{TA}]; \mu, \begin{bmatrix} [\text{PO}_4] \\ [\text{DOP}] \end{bmatrix} \right), \quad (14)$$

$$\frac{d}{dt} [\text{DIC}] = G_3 \left( [\text{DIC}]; \mu, [\text{TA}], \begin{bmatrix} [\text{PO}_4] \\ [\text{DOP}] \end{bmatrix} \right). \quad (15)$$

where  $\mu = [d, \alpha, \kappa, \sigma, R, r\text{C:P}, r\text{N:P}]$  is the set of model parameters. Equation (13) governs the phosphorous cycle. Equation (14) governs the total alkalinity cycle and equation (15) governs the carbon cycle. Note that  $[\text{PO}_4]$  and  $[\text{DOP}]$ , the dependent state variables in (13), appear only parametrically in  $G_2$  and  $G_3$  and that  $[\text{TA}]$ , the dependent state variable in (14), appears only parametrically in  $G_3$ . Consequently, equations (13), (14), and (15) can be solved sequentially without having to solve the fully coupled system simultaneously.

[19] Our method of obtaining steady state solutions is thus to set the time derivatives in equations (13), (14), and (15) to zero and use Newton's method [e.g., *Kelly*, 2003] to solve

$$G_1 \left( \begin{bmatrix} [\text{PO}_4] \\ [\text{DOP}] \end{bmatrix}; \mu \right) = 0, \quad (16)$$

**Table 2.** Reference (OCMIP-2) and Optimized (OPT) Parameter Values

Parameter	rC:P	rN:P	$\sigma$	$1/\kappa$ (years)	$\alpha$	$R$	$d$ (m)
OCMIP-2 <sup>a</sup>	117	16	0.67	0.5	-0.9	0.07	3500
OPT	137	24	0.74	1.7	-0.97	0.081	2100
1% Range <sup>b</sup>	126–149	13–36	0.69–0.78	1.2–2.3	-1.04–-0.90	0.072–0.089	1800–2400
OPT(rN:P = 16) <sup>c</sup>	142	16	0.73	1.8	-0.96	0.073	2100

<sup>a</sup>OCMIP-2 reference parameter values as given by *Najjar et al.* [2007].

<sup>b</sup>The ranges indicate the parameter sensitivities and are computed on the basis of an average 1% decrease in the fraction of the spatial variance in the tracer data captured by the model while the other parameters are held constant at their optimal values.

<sup>c</sup>Optimized parameter values when rN:P is held constant at its reference value rN:P = 16.

for  $[\text{PO}_4]$  and  $[\text{DOP}]$ , and then to solve

$$G_2\left([\text{TA}]; \mu, \begin{bmatrix} [\text{PO}_4] \\ [\text{DOP}] \end{bmatrix}\right) = 0, \quad (17)$$

for  $[\text{TA}]$ , and finally to solve

$$G_3\left([\text{DIC}]; \mu, [\text{TA}], \begin{bmatrix} [\text{PO}_4] \\ [\text{DOP}] \end{bmatrix}\right) = 0, \quad (18)$$

for  $[\text{DIC}]$ .

[20] The Newton's solver consists of iterating

$$X \leftarrow X - \left(\frac{\partial G}{\partial X}\right)^{-1} G(X) \quad (19)$$

until  $\max|G(X)| < \text{tol}$  where  $\text{tol}$  is chosen to be as small as possible given the finite machine precision. Newton's method requires that the Jacobian matrix,  $(\partial G/\partial X)$ , be inverted at every iteration. For this we use a factor-and-solve approach as implemented in the MUMPS sparse matrix solver [*Amestoy et al.*, 2001]. The factored matrix is then available for the sensitivity analysis (see section 3.1). For our model, steady solutions to the full model can be obtained in approximately 1 h on a single processor workstation. Obtaining the same solution using a time stepping approach would take on the order of a week.

#### 2.4. Equilibrium Solutions Based on OCMIP-2 Reference Parameters

[21] In this section, we present the equilibrium solutions of TA and DIC obtained using the OCMIP-2 reference parameters and compare with the observed fields. The reference parameters used for the OCMIP-2 simulation are listed in Table 2. Figure 1 shows depth-averaged concentrations of TA in the left column and DIC in the right column with the simulated fields in the top row and observed fields in the middle. Figures 2 and 3 show respectively the zonal averages of  $[\text{TA}]$  and  $[\text{DIC}]$ . The bottom rows in Figures 1 to 3 show the equilibrium solutions based on our optimized parameter set. A discussion of the optimized solutions is deferred until section 4.5.

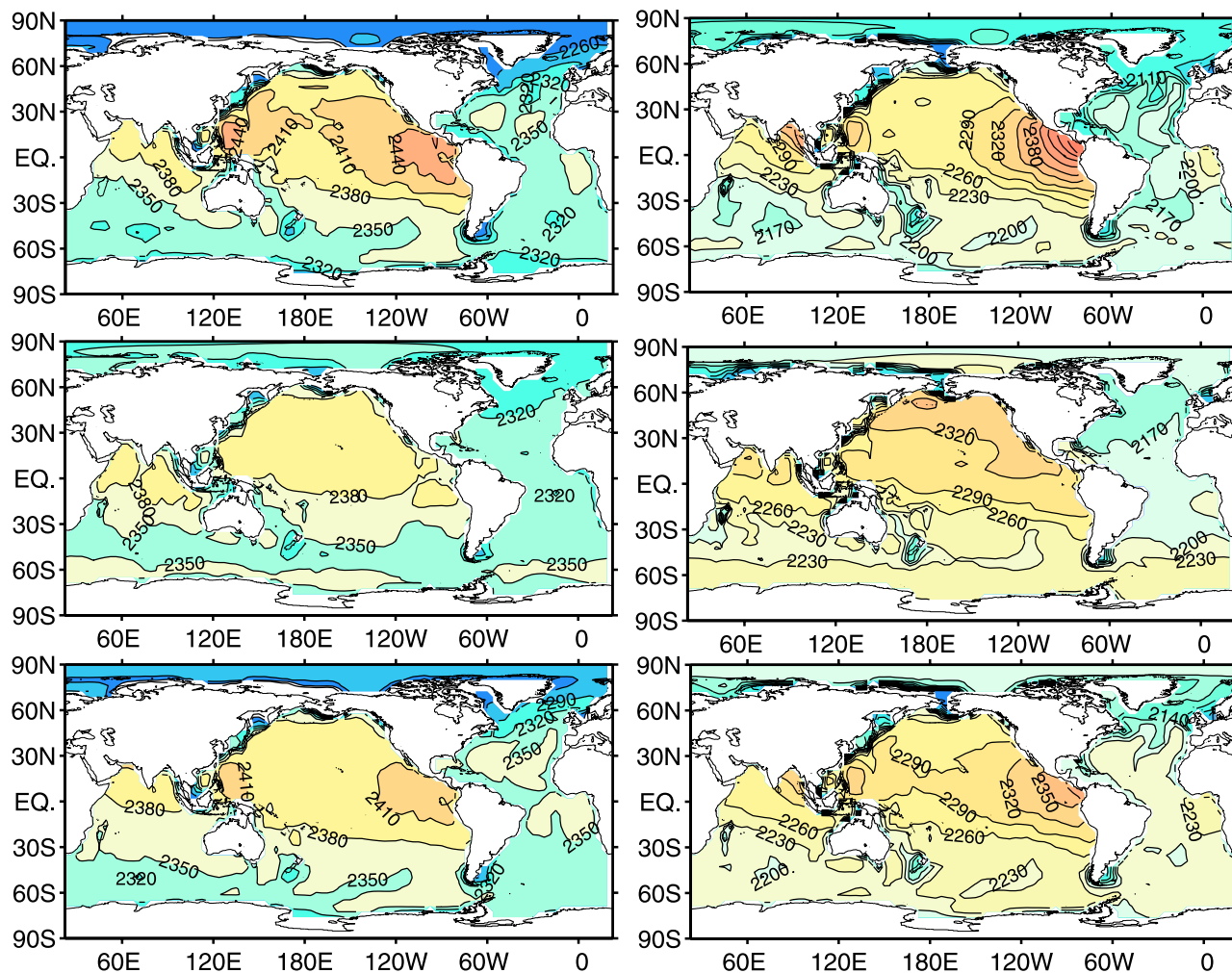
[22] There is a clear contrast in both  $[\text{DIC}]$  and  $[\text{TA}]$  between the Atlantic and Pacific that is well captured by the

model (Figure 1). The model however has a pronounced maximum in the equatorial Pacific that is absent in the observations. The  $\text{PO}_4$  field, not shown here, but plotted in KP06 shows a similar pattern. In the study of KP06, optimizing the parameters that control the production and consumption of DOP reduced the excess  $[\text{PO}_4]$  in the equatorial Pacific Ocean by approximately 70%. The remaining discrepancy suggests that the errors in the transport model are responsible for at least part of the excess  $[\text{PO}_4]$  in the eastern equatorial Pacific in accord with previous studies by *Aumont et al.* [1999] and *Matear and Holloway* [1995]. The same errors in the circulation and in the parameters that control DOP production are likely to be also responsible for part of the deficiencies in the simulated TA and DIC fields. Our results presented in section 4.5 show that this is indeed the case.

[23] The zonally averaged TA distribution is presented in Figure 2 for the Atlantic, Indian, and Pacific oceans. Our model captures the gross features of the observed TA distribution with depleted  $[\text{TA}]$  near the surface and a gradual increase in  $[\text{TA}]$  with depth. As in the observations, the highest  $[\text{TA}]$  values are located at depth in the northern parts of the Indian and Pacific oceans. The model also captures the elevated  $[\text{TA}]$  in the subtropical surface waters of each basin. In particular, the higher  $[\text{TA}]$  in the surface subtropical Atlantic caused by excess evaporation over precipitation is well captured by the model.

[24] Our simulation with the OCMIP-2 parameters overestimates the basin-to-basin contrast of the TA concentration.  $[\text{TA}]$  is underestimated in the upper North Atlantic waters and overestimated in the deep North Pacific. Another discrepancy can be seen in the deep Atlantic Ocean. Simulated  $[\text{TA}]$  is higher than observed by approximately 30  $\mu\text{eq/kg}$  below 3000 m. As mentioned in KP06 for the case of  $\text{PO}_4$ , the discrepancy in the deep Atlantic ocean is likely due to the shortcomings of the transport model because Antarctic Bottom Water (AABW) intrudes too far north, pushing the North Atlantic Deep Water (NADW) to a depth that is shallower than observed. Because the AABW has higher levels of preformed TA than does NADW, the transport error results in elevated  $[\text{TA}]$  near the bottom of the Atlantic Ocean, where the contribution of AABW is higher than that of NADW. In addition, the downward  $\text{CaCO}_3$  fluxes and dissolution at depth further increase  $[\text{TA}]$  along the pathway of bottom water as it reaches farther northward than in the real ocean.

[25] The overall patterns of zonally averaged  $[\text{DIC}]$  are well reproduced by the model (Figure 3). In both the



**Figure 1.** Depth-averaged [TA] in  $\mu\text{eq kg}^{-1}$  (left column) and [DIC] in  $\mu\text{mol kg}^{-1}$  (right column). Top panels: equilibrium solutions obtained using the OCMIP-2 reference parameters, Middle panels: GLODAP observed fields, Bottom panels: equilibrium solutions obtained using the optimized parameters.

observation and the simulation, [DIC] is lowest above the thermocline and increases gradually with depth. The model also captures the observed basin-to-basin contrast within the deep ocean. However there is a discrepancy in the location of the [DIC] maximum. The model produces the highest [DIC] levels in the area centered near  $10^{\circ}\text{N}$  at a depth of approximately 2500 m in the Pacific whereas the observed [DIC] maximum is located near 1500 m northward of  $30^{\circ}\text{N}$ . The discrepancy in the highest [DIC] zone is also found in the  $\text{PO}_4$  simulation (KP06) and is likely due to the combined effect of biogeochemistry and circulation errors in the model. Too weak upwelling of North Pacific waters compared to the real ocean produces a nutricline that is too deep in our model. As a result, primary production in the Northern Pacific surface ocean is nearly shut off because of the lack of nutrient supply from below.

[26] In the next section we present sensitivity patterns that show how changes in the biogeochemistry parameters affect the spatial distribution of TA and DIC. The sensitivity analysis then paves the way for understanding the results

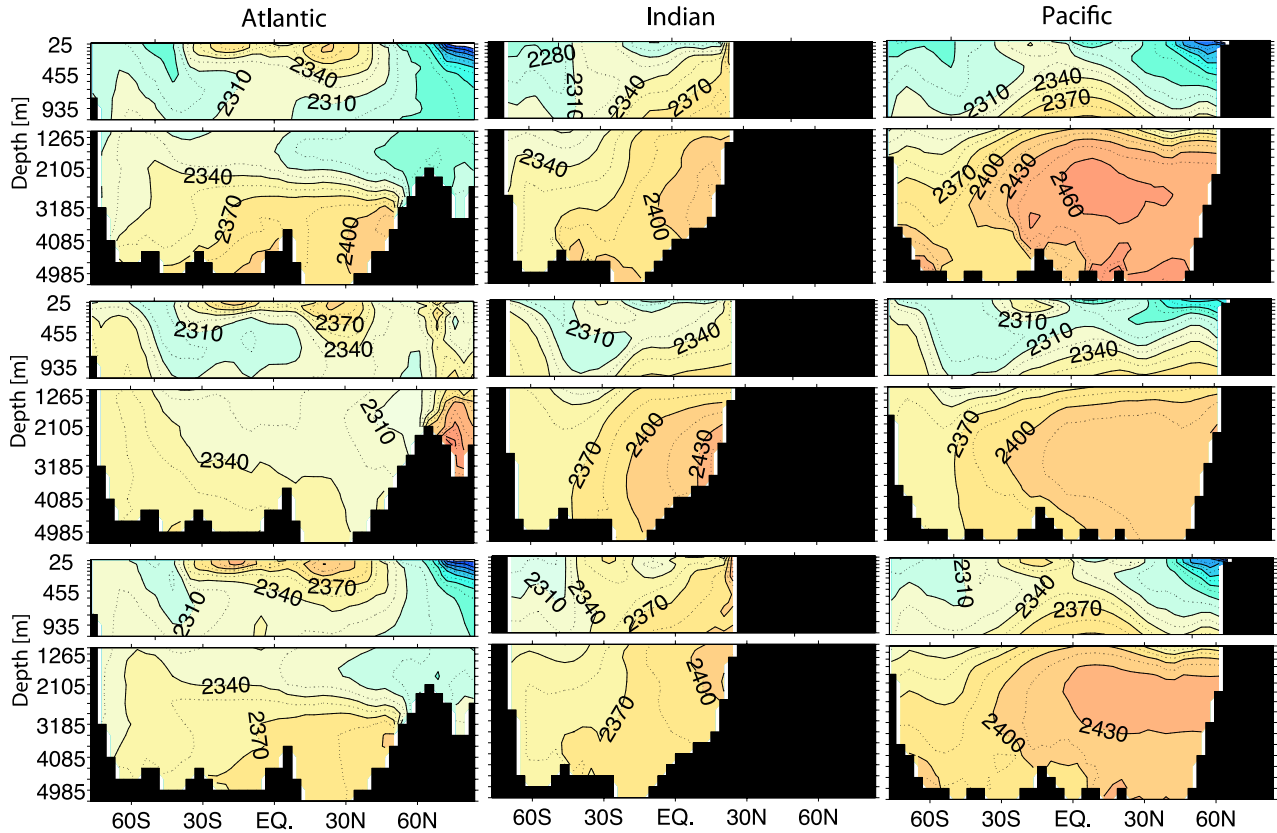
of the parameter optimization study which we present in section 4.

### 3. Model Sensitivity Analysis

#### 3.1. Definition of Sensitivity Patterns and Computational Method

[27] The standard approach for obtaining the sensitivity of a model to its parameters consists of running the model to obtain a reference solution, varying an input parameter, rerunning the model and recording the change in the model's output. Albeit being conceptually straightforward the method is computationally expensive and greatly limits the scope of parameter sensitivity analysis for global ocean biogeochemistry models.

[28] We obtain the sensitivity of the equilibrium tracer fields using an alternative approach that can take advantage of the computations already done as part of the Newton solver to greatly reduce the overall computational costs. Our approach is to compute the partial derivatives of the



**Figure 2.** Zonally averaged [TA] distribution in  $\mu\text{eq kg}^{-1}$  in the Atlantic (left column), Indian (middle column), and Pacific (right column) oceans for the simulation using the OCMIP-2 reference parameters (top row), the GLODAP observed field (middle row), and for the simulation using the optimized parameter set (bottom row).

equilibrium tracer fields directly. In the limit of a small parameter perturbation the standard approach to sensitivity analysis yields precisely a finite difference approximation to these partial derivatives.

[29] To obtain an equation for the sensitivity of the equilibrium TA field to one of the model parameters, we take the partial derivative of equation (17) with respect to the parameter. Denoting the particular parameter by  $\mu$ , we have

$$\frac{\partial[\text{TA}]_{eq}}{\partial\mu} = - \left[ \frac{\partial G_2}{\partial[\text{TA}]_{eq}} \right]^{-1} \cdot \left( \frac{\partial G_2}{\partial\mu} + \left[ \frac{\partial G_2}{\partial[\text{PO}_4; \text{DOP}]_{eq}} \right] \frac{\partial[\text{PO}_4; \text{DOP}]_{eq}}{\partial\mu} \right). \quad (20)$$

Similarly, for the sensitivity of the equilibrium DIC field to a parameter  $\mu$ , we take the partial derivative of equation (18) with respect to  $\mu$  to obtain

$$\frac{\partial[\text{DIC}]_{eq}}{\partial\mu} = - \left[ \frac{\partial G_3}{\partial[\text{DIC}]_{eq}} \right]^{-1} \left( \frac{\partial G_3}{\partial\mu} + \left[ \frac{\partial G_3}{\partial[\text{TA}]_{eq}} \right] \frac{\partial[\text{TA}]_{eq}}{\partial\mu} + \left[ \frac{\partial G_3}{\partial[\text{PO}_4; \text{DOP}]_{eq}} \right] \frac{\partial[\text{PO}_4; \text{DOP}]_{eq}}{\partial\mu} \right). \quad (21)$$

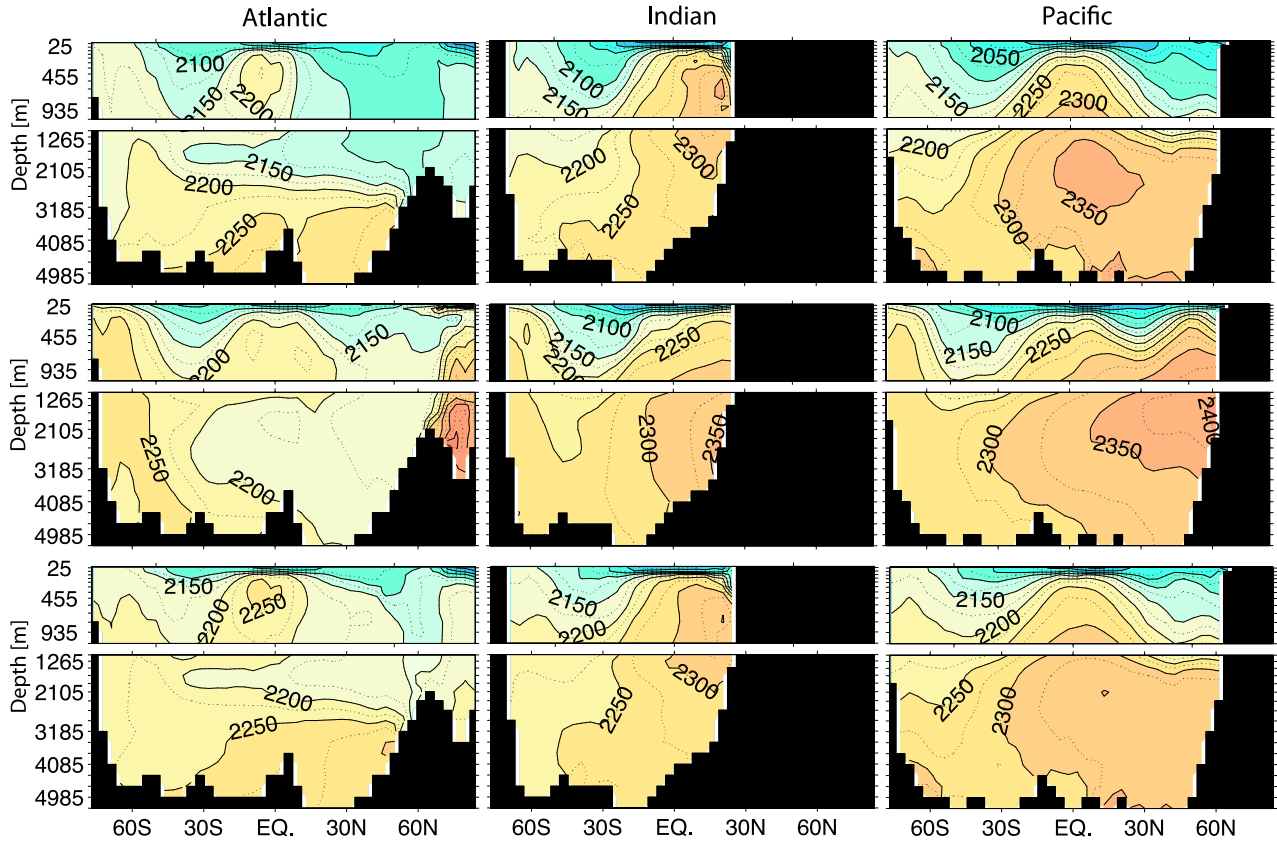
[30] In equations (20) and (21) all terms are vectors except those in square brackets which are matrices. The two matrices requiring inversion in equations (20) and (21) are exactly the Jacobian matrices that are already factored as part of the Newton solver applied to equations (17) and (18). The sensitivity of the equilibrium solutions can therefore be obtained with a negligible increase in the overall computational cost.

[31] In the remainder of this paper we will refer to the three-dimensional sensitivity patterns given by the partial derivatives of the equilibrium [TA] and [DIC] with respect to a model parameter as  $S$  patterns and denote them using the following shorthand notation

$$S_{\mu}^{TA} \equiv \frac{\partial[\text{TA}]}{\partial\mu} \text{ and } S_{\mu}^{DIC} \equiv \frac{\partial[\text{DIC}]}{\partial\mu}, \quad (22)$$

where  $\mu$  can be any one of the model parameters [rC:P, rN:P,  $\sigma$ ,  $\kappa$ ,  $\alpha$ ,  $R$ ,  $d$ ].

[32] In Figures 4 and 6, we present the resulting  $S^{TA}$  and  $S^{DIC}$  patterns, respectively. The reference state for which we compute the sensitivity patterns is the equilibrium solution obtained with the standard OCMIP-2 parameter values listed in Table 2. It turns out that the structure of the  $S$  patterns is not too sensitive to the reference parameter state. Thus our discussion in this paper is valid regardless we used



**Figure 3.** Zonally averaged [DIC] distribution in  $\mu\text{mol kg}^{-1}$  in the Atlantic (left column), Indian (middle column) and Pacific (right column) oceans for the simulation using the OCMIP-2 reference parameters (top row), the GLODAP observed field (middle row), and for the simulation using the optimized parameter set (bottom row).

the OCMIP-2 parameters or the optimized parameters that will be discussed in section 4.

### 3.2. Decomposition Method of $S^{DIC}$ Patterns

[33] The mechanisms that are responsible for the oceanic carbon cycle can be better understood with the concept of the carbon pumps [Volk and Hoffert, 1985]. Here we adopt the pump separation method developed by Gruber and Sarmiento [2002] to separate the  $S^{DIC}$  patterns into three sensitivity components: sensitivities due to the soft-tissue pump, the carbonate pump and the gas exchange pump. Analogous to the three pumps in the work of Gruber and Sarmiento [2002], the soft-tissue pump produces the sensitivity component that would be obtained if there were only organic material fluxes in the absence of gas exchange. The carbonate pump produces the sensitivity component that would be obtained if only  $\text{CaCO}_3$  sources-sinks took place in the absence of air-sea gas exchange. The remaining  $S^{DIC}$  pattern explains the net effect of the evasion or invasion of carbon to and from the atmosphere due to a parameter change. The  $S^{DIC}$  pattern is thus separated into the three components,

$$S_{\mu}^{DIC} = S_{\mu}^{DIC_{\text{soft}}} + S_{\mu}^{DIC_{\text{carb}}} + S_{\mu}^{DIC_{\text{gasx}}}. \quad (23)$$

[34] To derive the decomposition of the  $S^{DIC}$  patterns, we begin by decomposing the deviation of the preindustrial

DIC distribution from its global mean value into the three components as follows:

$$\Delta[\text{DIC}] = \Delta[\text{DIC}]_{\text{soft}} + \Delta[\text{DIC}]_{\text{carb}} + \Delta[\text{DIC}]_{\text{gasx}}, \quad (24)$$

where the  $\Delta$  operator yields the anomaly field with respect to the global mean, i.e.,

$$\Delta[\text{DIC}] = [\text{DIC}] - \langle \text{DIC} \rangle_{\text{ref}}, \quad (25)$$

where  $\langle \text{DIC} \rangle_{\text{ref}}$  represents the global mean of [DIC] at the reference parameter set,

$$\langle \text{DIC} \rangle_{\text{ref}} = \frac{1}{V} \int_V [\text{DIC}]_{\text{ref}} dV \quad (26)$$

where  $V$  is the total volume of the global ocean. Likewise the anomaly fields for  $[\text{PO}_4]$  and  $[\text{TA}]$  can be obtained from

$$\Delta[\text{PO}_4] = [\text{PO}_4] - \langle \text{PO}_4 \rangle_{\text{ref}}, \quad (27)$$

and

$$\Delta[\text{TA}] = [\text{TA}] - \langle \text{TA} \rangle_{\text{ref}}, \quad (28)$$



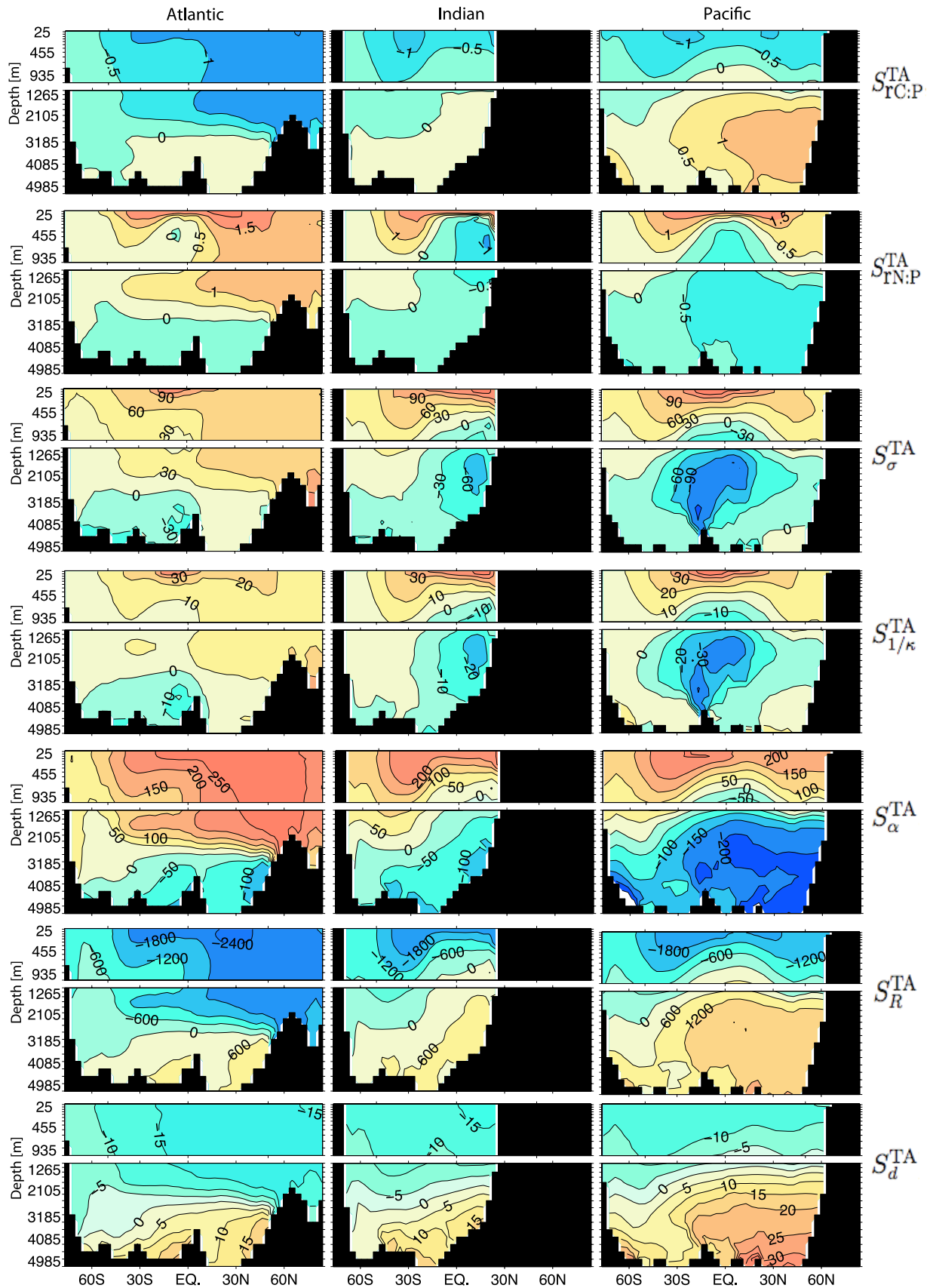


Figure 4

respectively. Because we seek to explain the net change in the DIC distribution induced by a parameter change, the anomaly field should retain all changes that have occurred including a net increase or decrease of carbon storage. Therefore the global mean value  $\langle \text{DIC} \rangle_{ref}$  is referenced to a fixed parameter set whereas the tracer distribution  $[\text{DIC}]$  is subject to change with varying parameter values. The soft-tissue component is obtained by multiplying the C:P ratio to the deviation of phosphate and the carbonate component is computed using the nitrogen-corrected TA deviation [Brewer *et al.*, 1975],

$$\Delta[\text{DIC}]_{soft} = \text{rC:P} \times \Delta[\text{PO}_4] \quad (29)$$

and

$$\Delta[\text{DIC}]_{carb} = \frac{\Delta[\text{TA}] + \text{rN:P} \times \Delta[\text{PO}_4]}{2}. \quad (30)$$

The remaining DIC deviation is the gas exchange component, thus equation (24) becomes

$$\Delta[\text{DIC}] = \text{rC:P} \times \Delta[\text{PO}_4] + \frac{\Delta[\text{TA}] + \text{rN:P} \times \Delta[\text{PO}_4]}{2} + \Delta[\text{DIC}]_{gasx}. \quad (31)$$

The partial derivative of equation (31) with respect to a parameter  $\mu$  yields

$$\begin{aligned} \frac{\partial \Delta[\text{DIC}]}{\partial \mu} &= \frac{\partial(\text{rC:P} \times \Delta[\text{PO}_4])}{\partial \mu} \\ &+ \frac{1}{2} \left( \frac{\partial(\Delta[\text{TA}] + \text{rN:P} \times \Delta[\text{PO}_4])}{\partial \mu} \right) \\ &+ \frac{\partial \Delta[\text{DIC}]_{gasx}}{\partial \mu}. \end{aligned} \quad (32)$$

Note that the term on the left hand side is equivalent to  $S_{\mu}^{DIC}$ , and the first, second and third terms on the right hand side are respectively contributions from the soft-tissue pump, carbonate pump and gas exchange pump to the  $S_{\mu}^{DIC}$  pattern. Equation (32) can be expressed using the  $S$  pattern notation as

$$S_{\mu}^{DIC} = \text{rC:P} \times S_{\mu}^{PO4} + \frac{1}{2} \left( S_{\mu}^{TA} + \text{rN:P} \times S_{\mu}^{PO4} \right) + S_{\mu}^{DIC_{gasx}}, \quad (33)$$

if  $\mu \in \{\sigma, 1/\kappa, \alpha\}$ , as

$$S_{\mu}^{DIC} = \frac{1}{2} S_{\mu}^{TA} + S_{\mu}^{DIC_{gasx}}, \quad (34)$$

if  $\mu \in \{R, d\}$ , as

$$S_{\mu}^{DIC} = \Delta\text{PO}_4 + \frac{1}{2} S_{\mu}^{TA} + S_{\mu}^{DIC_{gasx}}, \quad (35)$$

if  $\mu = \text{rC:P}$  and as

$$S_{\mu}^{DIC} = S_{\mu}^{DIC_{gasx}}, \quad (36)$$

if  $\mu = \text{rN:P}$ . The missing terms in equations (34) through (36) are due to the fact that  $S_R^{PO4} = S_d^{PO4} = S_{\text{rC:P}}^{PO4} = S_{\text{rN:P}}^{PO4} = 0$  and  $S_{\text{rN:P}}^{TA} + \Delta\text{PO}_4 = 0$  (see equation (43) in the next section). As a result, the soft-tissue pump components for  $\mu = R, d, \text{rN:P}$  and the carbonate pump component for  $\mu = \text{rN:P}$  are absent. The resulting decomposed  $S^{DIC}$  patterns are presented in Figure 7 and discussed in section 3.5.

### 3.3. Decomposition Method of $S^{TA}$ Patterns

[35] We also decompose the  $S^{TA}$  patterns to the contributions from the soft-tissue pump and the carbonate pump,

$$S_{\mu}^{TA} = S_{\mu}^{TA_{soft}} + S_{\mu}^{TA_{carb}}. \quad (37)$$

[36] To derive the decomposition of the  $S^{TA}$  patterns, we begin by decomposing the deviation of TA distribution from its global mean value into the two components as follows:

$$\Delta[\text{TA}] = \Delta[\text{TA}]_{soft} + \Delta[\text{TA}]_{carb}. \quad (38)$$

The soft-tissue component is given by

$$\Delta[\text{TA}]_{soft} = -\text{rN:P} \times \Delta[\text{PO}_4], \quad (39)$$

and the remaining  $[\text{TA}]$  deviation becomes the carbonate component. Taking partial derivatives of equation (38) with respect to a parameter  $\mu$ , we obtain

$$\frac{\partial \Delta[\text{TA}]}{\partial \mu} = \frac{\partial(-\text{rN:P} \times \Delta[\text{PO}_4])}{\partial \mu} + \frac{\partial \Delta[\text{TA}]_{carb}}{\partial \mu}. \quad (40)$$

Using the  $S$  pattern notation, equation (40) can be rewritten as

$$S_{\mu}^{TA} = -\text{rN:P} \times S_{\mu}^{PO4} + S_{\mu}^{TA_{carb}}, \quad (41)$$

if  $\mu \in \{\sigma, 1/\kappa, \alpha\}$ , as

$$S_{\mu}^{TA} = S_{\mu}^{TA_{carb}}, \quad (42)$$

if  $\mu \in \{R, d, \text{rC:P}\}$  and as

$$S_{\mu}^{TA} = -\Delta[\text{PO}_4], \quad (43)$$

**Figure 4.**  $S^{TA}$  patterns showing the sensitivity of the equilibrium solution of TA to changes in the parameter values. From top to bottom,  $S_{\text{rC:P}}^{TA} [\mu\text{eq kg}^{-1} (\text{molC/molP})^{-1}]$ ,  $S_{\text{rN:P}}^{TA} [\mu\text{eq kg}^{-1} (\text{molN/molP})^{-1}]$ ,  $S_{\sigma}^{TA} [\mu\text{eq kg}^{-1}]$ ,  $S_{1/\kappa}^{TA} [\mu\text{eq kg}^{-1} \text{ year}^{-1}]$ ,  $S_{\alpha}^{TA} [\mu\text{eq kg}^{-1}]$ ,  $S_R^{TA} [\mu\text{eq kg}^{-1}]$  and  $S_d^{TA} [\mu\text{eq kg}^{-1} \text{ km}^{-1}]$ . The zonally averaged TA sensitivities in the Atlantic (left column), Indian (middle column), and Pacific (right column) are shown. Warm colors correspond to positive sensitivities and cold colors correspond to negative sensitivities.

if  $\mu = \text{rN:P}$ . The soft-tissue components for  $\mu = R, d, \text{rC:P}$  are absent in equation (42) because  $S_R^{PO_4} = S_d^{PO_4} = S_{\text{rC:P}}^{PO_4} = 0$ . Since  $S_{\text{rN:P}}^{TA}$  itself is equal to the minus anomaly field of  $\text{PO}_4$ , the carbonate component for  $\mu = \text{rN:P}$  is absent in equation (43). The resulting decomposed  $S^{TA}$  patterns are presented in Figure 5 and discussed in the following section.

### 3.4. S Patterns for TA: $S^{TA}$ Patterns

[37] The  $S^{TA}$  patterns (Figure 4) show how parameter perturbations that result in a decrease of [TA] in surface waters are accompanied by a commensurate increase in deeper waters, and vice versa. The pattern reflects changes in uptake in sunlit surface waters tied to a necessary change in export flux and subsurface remineralization. The sensitivity patterns further show how the biologically driven fluxes couple to the circulation to produce a large-scale redistribution of the tracer within and across ocean basins. For example, the zero sensitivity contour above which [TA] increases and below which [TA] decreases (or vice versa) shows a clear contrast across the basins. The zero sensitivity line lies deeper in the Atlantic relative to the Indian and Pacific oceans. The formation and transport of NADW are responsible for the penetration of water properties acquired at the surface into the interior of the Atlantic Ocean. At the same time, changes in the downward flux of biogenic particles result in changes in [TA] at depths that are then carried into the Indian and Pacific oceans by the lower branch of the global overturning circulation. The accumulated changes then upwell in the tropics and northern hemisphere causing the zero sensitivity line to bow upward in those basins.

[38] The  $S^{TA}$  patterns show that an increase in  $\text{rC:P}$ ,  $R$  or  $d$  tends to enhance the vertical and basin-to-basin contrasts of [TA] while an increase in  $\text{rN:P}$ ,  $\sigma$ ,  $1/\kappa$  or  $\alpha$  tends to reduce the contrasts. The tendency is the result of the parameter change influencing either organic nitrogen or  $\text{CaCO}_3$  cycles. However, as illustrated in Figure 5, changes in  $\text{CaCO}_3$  fluxes dominate the  $S^{TA}$  patterns over changes in organic nitrogen fluxes for all parameters except  $\text{rN:P}$ . Parameter changes that lead to increased export production of  $\text{CaCO}_3$  enhance the vertical and basin-to-basin contrasts of [TA] while parameter changes that lead to decreased export of  $\text{CaCO}_3$  tend to reduce the contrasts. For example, increases in  $\text{rC:P}$  and  $R$  both raise the  $\text{CaCO}_3$  export production from the surface waters and subsequent dissolution in the abyssal waters. An increase in the dissolution length scale  $d$  also increases the downward  $\text{CaCO}_3$  transport. Therefore increasing  $\text{rC:P}$ ,  $R$  and  $d$  causes a reduction of [TA] in the upper oceans and newly formed NADW and, at the same time, causes an accumulation of [TA] in the older water masses of the North Pacific. On the other hand, increasing  $\sigma$ ,  $1/\kappa$  or  $\alpha$  results in decreased  $\text{CaCO}_3$  production. Increases in  $\sigma$  and  $1/\kappa$  both decrease the export production of POC (see Figure 8 in KP06) to which the production of  $\text{CaCO}_3$  is proportional. A deeper remineralization of organic

matter resulting from an increased  $\alpha$  also reduces new production of organic carbon and  $\text{CaCO}_3$  by decreasing the supply of nutrients into the euphotic layer. Therefore, increasing  $\sigma$ ,  $1/\kappa$  or  $\alpha$  erodes the vertical and basin-to-basin contrasts of [TA]. While all the parameters except  $\text{rN:P}$  influences [TA] through the  $\text{CaCO}_3$  cycle,  $\text{rN:P}$  influences [TA] via POM fluxes. An increase in  $\text{rN:P}$  tends to erode the vertical gradient of [TA] through increased organic nitrogen export.

[39] One important feature of the TA distribution is that different subsets of parameters have sensitivity patterns with very similar shapes. For example,  $S_{\text{rC:P}}^{TA}$ ,  $S_{\alpha}^{TA}$ ,  $S_R^{TA}$  and  $S_d^{TA}$  have similar structures.  $S_{\sigma}^{TA}$  and  $S_{1/\kappa}^{TA}$  also show similar patterns. This leads to correlations among the parameters and has important implication for our optimization results as we discuss in section 4.3.

### 3.5. S Patterns for DIC ( $S^{DIC}$ )

[40] The  $S^{DIC}$  patterns (Figure 6) show how DIC is redistributed in response to changes in the parameters. In addition to being affected by changes in biological production and remineralization occurring within the ocean, the DIC distribution is also affected by carbon gained or lost through air-sea gas exchange. Parameter changes that enhance the soft-tissue pump cause the uptake of  $\text{CO}_2$  from the atmosphere into the ocean by reducing the surface DIC concentration relative to the deep ocean. On the other hand, an enhancement of the carbonate pump causes outgassing of carbon by reducing the surface [TA] and consequently increasing the oceanic  $\text{pCO}_2$ . The resulting carbon exchange is the combined effect of both changes in the soft-tissue pump and the carbonate pump. The decomposed  $S^{DIC}$  patterns presented in Figure 7 further elucidate the parameter change impact on carbon distribution through the three pumps. In the following three subsections, we describe the  $S^{DIC}$  pattern for each parameter.

#### 3.5.1. Parameters $\text{rC:P}$ and $\text{rN:P}$

[41] Increasing  $\text{rC:P}$  enhances the vertical and basin-to-basin contrasts of DIC (first row in Figure 6). It produces a prominent depletion of DIC in newly formed NADW and a maximum enrichment in the oldest water of the North Pacific. This is because both the soft-tissue and carbonate pumps become stronger in response to the increased C:P ratio (first row in Figure 7). With an increase in  $\text{rC:P}$  the amount of organic carbon consumed and remineralized in the deeper ocean is increased for a fixed amount of new production. Albeit smaller in magnitude, increasing  $\text{rC:P}$  also increases  $\text{CaCO}_3$  fluxes through the dependency of  $\text{CaCO}_3$  formation on the export of POC. The enhanced soft-tissue and carbonate pumps however have opposing impacts on air-sea carbon exchange. The ocean tends to take up  $\text{CO}_2$  with the enhanced soft-tissue pump while the enhanced carbonate pump tends to release  $\text{CO}_2$  to the atmosphere. The gas exchange component shows the net uptake, which implies that the impact of the enhanced soft-tissue pump

**Figure 5.** The zonally averaged  $S^{TA}$  patterns along the Atlantic (right column) are decomposed into contributions from the soft-tissue pump (left column) and the carbonate pump (middle column). From top to bottom,  $S_{\text{rC:P}}^{TA} [\mu\text{eq kg}^{-1} (\text{molC/molP})^{-1}]$ ,  $S_{\text{rN:P}}^{TA} [\mu\text{eq kg}^{-1} (\text{molN/molP})^{-1}]$ ,  $S_{\sigma}^{TA} [\mu\text{eq kg}^{-1}]$ ,  $S_{1/\kappa}^{TA} [\mu\text{eq kg}^{-1} \text{ year}^{-1}]$ ,  $S_{\alpha}^{TA} [\mu\text{eq kg}^{-1}]$ ,  $S_R^{TA} [\mu\text{eq kg}^{-1}]$  and  $S_d^{TA} [\mu\text{eq kg}^{-1} \text{ km}^{-1}]$ . The same color scale is applied to all components for each parameter, but contour intervals are adjusted to better show detailed structure of each component.

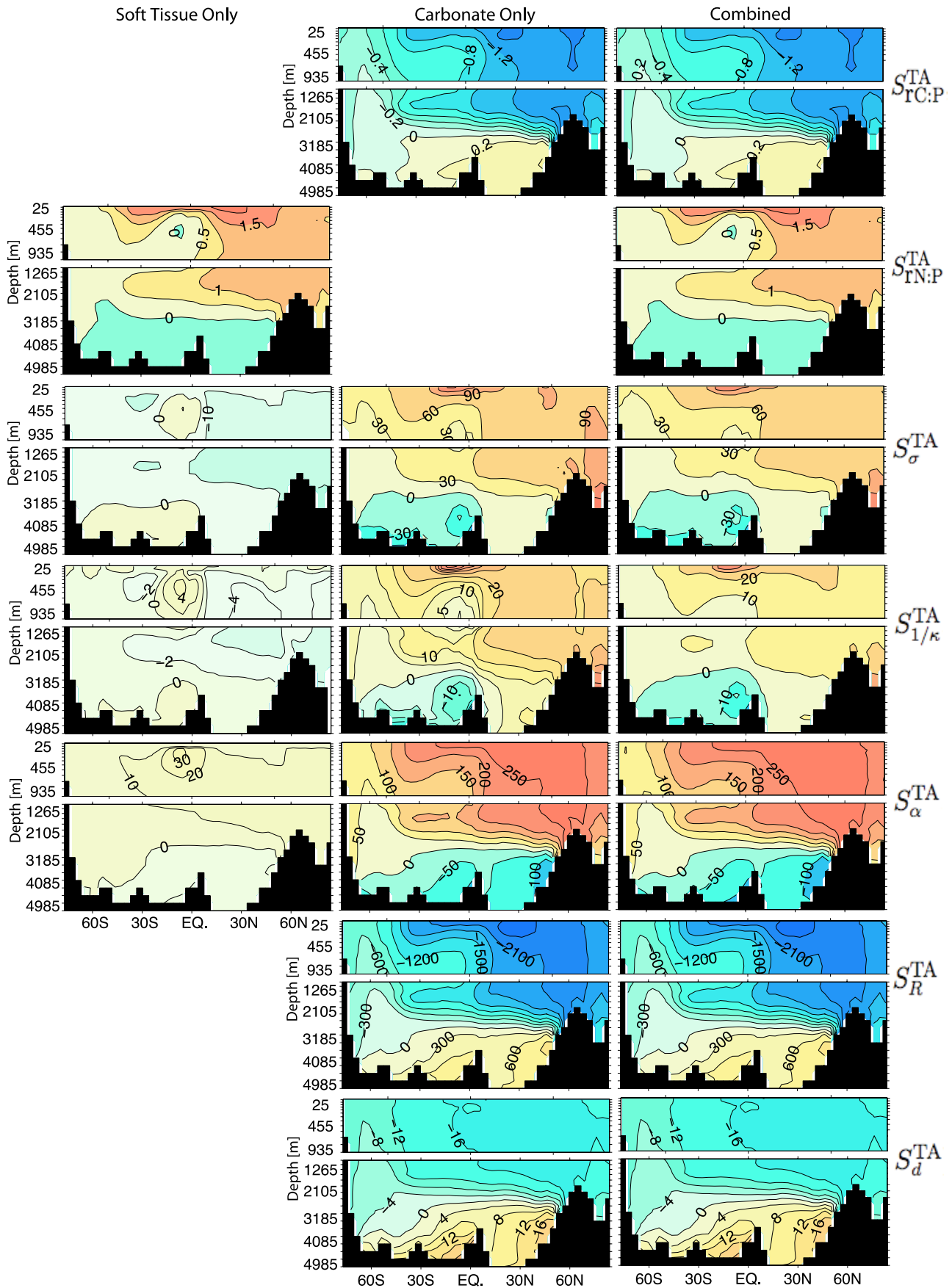


Figure 5

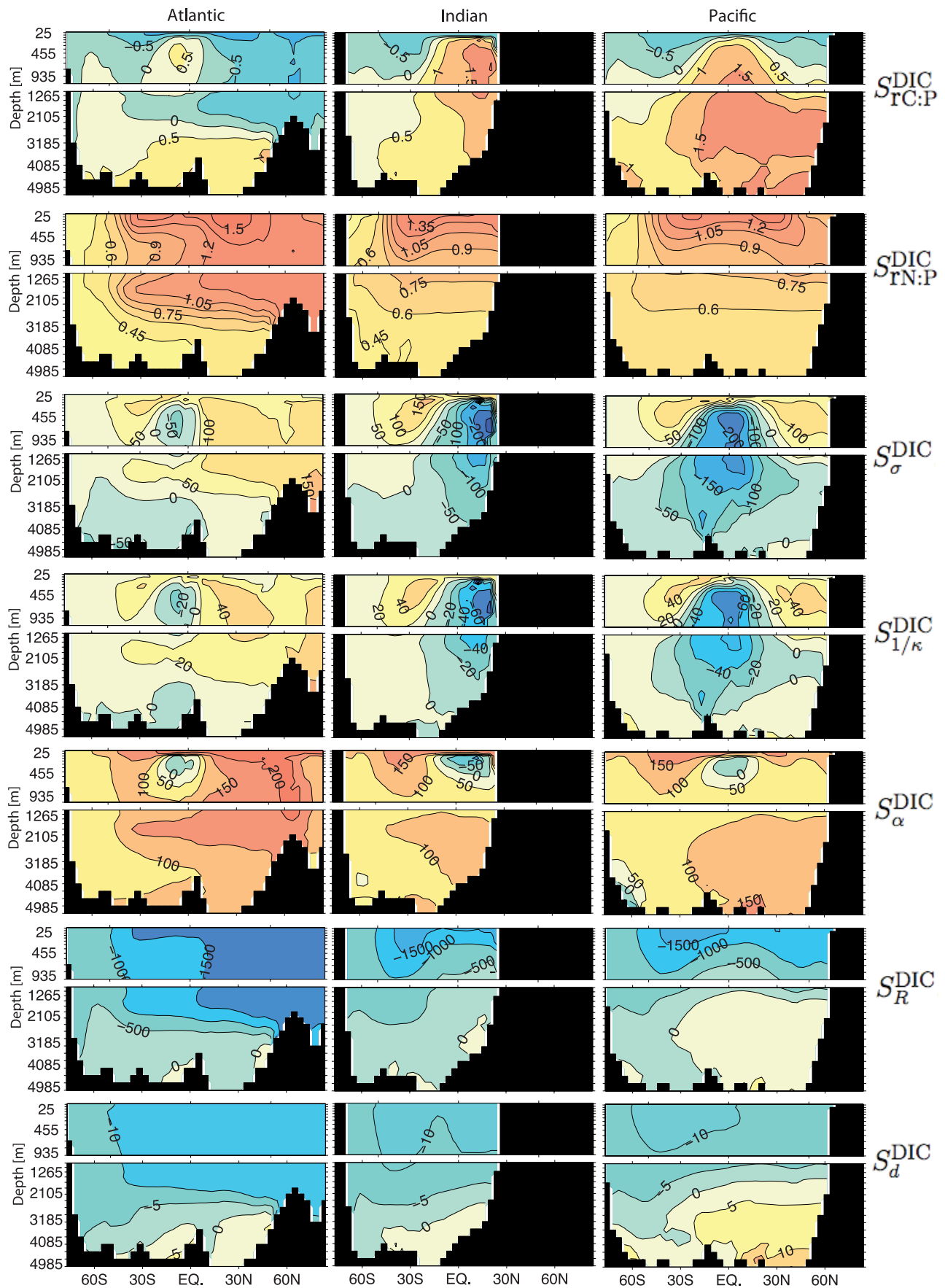
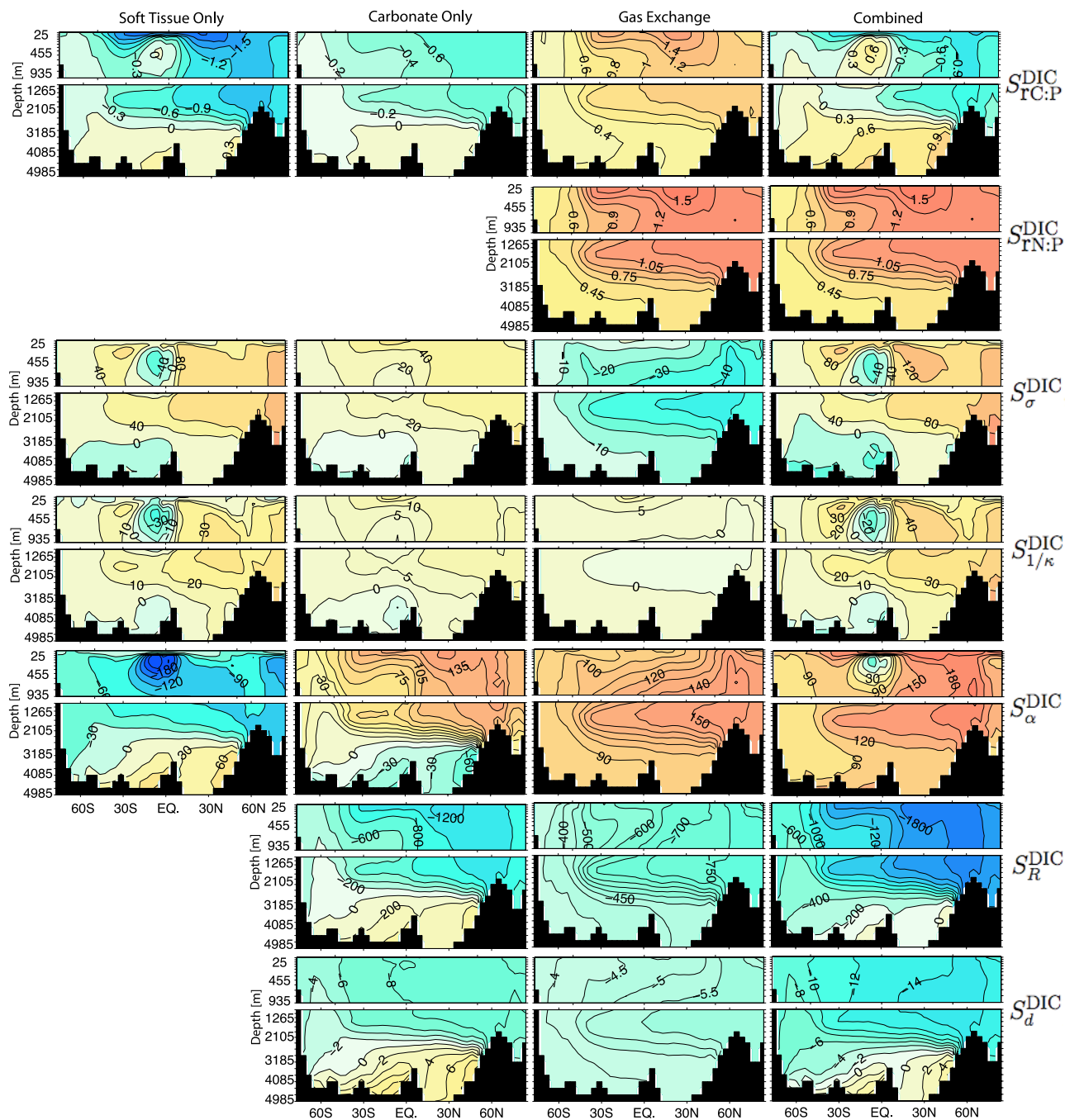


Figure 6



**Figure 7.** The zonally averaged  $S^{DIC}$  patterns along the Atlantic (fourth column) are decomposed into contributions from the soft-tissue pump (first column), the carbonate pump (second column), and the gas exchange pump (third column). From top to bottom,  $S_{TC:P}^{DIC}$  [ $\mu\text{mol kg}^{-1} (\text{molC/molP})^{-1}$ ],  $S_{TN:P}^{DIC}$  [ $\mu\text{mol kg}^{-1} (\text{molN/molP})^{-1}$ ],  $S_{\sigma}^{DIC}$  [ $\mu\text{mol kg}^{-1}$ ],  $S_{1/\kappa}^{DIC}$  [ $\mu\text{mol kg}^{-1} \text{ year}^{-1}$ ],  $S_{\alpha}^{DIC}$  [ $\mu\text{mol kg}^{-1}$ ],  $S_R^{DIC}$  [ $\mu\text{mol kg}^{-1}$ ] and  $S_d^{DIC}$  [ $\mu\text{mol kg}^{-1} \text{ km}^{-1}$ ]. The same color scale is applied to all components for each parameter, but contour intervals are adjusted to better show detailed structure of each component.

**Figure 6.**  $S^{DIC}$  patterns showing the sensitivity of the equilibrium solution of DIC to change in the parameter values. From top to bottom,  $S_{TC:P}^{DIC}$  [ $\mu\text{mol kg}^{-1} (\text{molC/molP})^{-1}$ ],  $S_{TN:P}^{DIC}$  [ $\mu\text{mol kg}^{-1} (\text{molN/molP})^{-1}$ ],  $S_{\sigma}^{DIC}$  [ $\mu\text{mol kg}^{-1}$ ],  $S_{1/\kappa}^{DIC}$  [ $\mu\text{mol kg}^{-1} \text{ year}^{-1}$ ],  $S_{\alpha}^{DIC}$  [ $\mu\text{mol kg}^{-1}$ ],  $S_R^{DIC}$  [ $\mu\text{mol kg}^{-1}$ ] and  $S_d^{DIC}$  [ $\mu\text{mol kg}^{-1} \text{ km}^{-1}$ ]. The zonally averaged DIC sensitivities in the Atlantic (left column), Indian (middle column) and Pacific (right column) are shown. Warm colors correspond to positive sensitivities and cold colors correspond to negative sensitivities.

dominates over the impact of the enhanced carbonate pump when controlling air-sea carbon fluxes.

[42] For this OCMIP-2 type model, the rN:P parameter influences DIC only through air-sea CO<sub>2</sub> exchange. Increasing rN:P causes an increase in surface [TA] in all basins as shown by  $S_{rN:P}^{TA}$  in Figure 4. This shifts the surface water CO<sub>2</sub> system equilibrium to favor a net uptake of CO<sub>2</sub> from the atmosphere. Consistent with the surface [TA] change,  $S_{rN:P}^{DIC}$  (second row in Figures 6 and 7) shows the highest sensitivities in the subtropical surface oceans. The sensitivity decreases gradually with depth because the carbon originates from the atmosphere.

### 3.5.2. Parameters $\sigma$ and $1/\kappa$

[43] Zonally averaged  $S_{\sigma}^{DIC}$  and  $S_{1/\kappa}^{DIC}$  show similar patterns (third and fourth rows in Figure 6). The maximum enrichment of DIC is observed near the subtropical thermocline in the three basins. In contrast, DIC is mostly depleted in the thermocline of the Indian and Pacific tropical regions. The  $S_{\sigma}^{DIC}$  and  $S_{1/\kappa}^{DIC}$  patterns are best captured by their soft-tissue components indicating the dominant role played by organic carbon fluxes in the net DIC distribution when  $\sigma$  or  $1/\kappa$  is perturbed (third and fourth rows in Figure 7). Increasing  $\sigma$  or  $1/\kappa$  causes an increase in the global inventory of DOM. This leads to an increase in the amount of DOC that is transported away from production zones by surface currents before being remineralized. This lateral transport of carbon causes [DIC] to decrease in areas where surface currents are divergent such as in the equatorial oceans and causes [DIC] to increase in areas where surface currents are convergent such as in the subtropical gyres. The lateral contrasts are clearly visible in  $S_{1/\kappa}^{DIC}$  and  $S_{\sigma}^{DIC}$  shown in Figures 6 and 7. An increase in either  $\sigma$  or  $1/\kappa$  tends to decrease the export production of POC (Figure 8 in KP06) and also of CaCO<sub>3</sub>. The reduced CaCO<sub>3</sub> fluxes erode the vertical gradient of [DIC]. The similar biological effects of the parameter changes yield nearly identical patterns in the soft-tissue and carbonate components for  $S_{\sigma}^{DIC}$  and  $S_{1/\kappa}^{DIC}$ .

[44] Despite gross similarities in the sensitivity patterns for  $\sigma$  and  $1/\kappa$ , the net effect on air-sea carbon exchange is different for each parameter. Increasing  $\sigma$  decreases the capacity of the ocean to store carbon, while increasing  $1/\kappa$  increases it. This can be seen in the gas exchange component of  $S_{\sigma}^{DIC}$  and  $S_{1/\kappa}^{DIC}$  shown in Figure 7. The opposing effect is due to the different response of the soft-tissue component of [DIC] in the surface ocean with changes in either  $1/\kappa$  or  $\sigma$ . With an increase in  $\sigma$  less DIC is exported as POC so that the surface [DIC] increases. An increase in  $1/\kappa$  on the other hand allows a larger fraction of the DOC to remineralize below the surface because of its longer lifetime. As a result, the soft-tissue component of [DIC] decreases in some parts of the surface ocean. The differences in surface [DIC] then influences the air-sea carbon exchange explaining the different sensitivities. The overall effect is that negative sensitivities associated with  $\sigma$  are dominant for the gas exchange component of  $S_{\sigma}^{DIC}$  whereas the weaker negative sensitivities for  $S_{1/\kappa}^{DIC}$  are present only in the water masses ventilated through the northern North Atlantic and North Pacific, not shown, but that are visible when zonal averages for individual basins are plotted.

### 3.5.3. Parameter $\alpha$

[45] The  $S_{\alpha}^{DIC}$  pattern is mostly positive in all three basins (fifth row in Figure 6). By increasing  $\alpha$ , a larger fraction of POC reaches the deep ocean. The strength of the soft-tissue pump is thus enhanced as can be seen in the fifth row of Figure 7. The resulting decrease in surface [DIC] is most pronounced in the thermocline of the equatorial oceans. The negative sensitivities shown in the equatorial thermocline is a consequence of subsurface convergence and upwelling which can return nutrients remineralized in the upper thermocline back to the euphotic layer. A deeper remineralization of POC leads to less reentrainment of carbon from below, causing [DIC] to decrease considerably in that area. The increased downward flux of POC couples with the Atlantic-to-Pacific deepwater flow to produce the large increase in [DIC] in the bottom waters of the Pacific.

[46] The carbonate component of  $S_{\alpha}^{DIC}$  contoured in Figure 7 shows the global effects of reduced production of CaCO<sub>3</sub> resulting from deeper sequestration of nutrients. The reduced production and dissolution of CaCO<sub>3</sub> increase [DIC] in the upper ocean while decreasing [DIC] in the deeper ocean. We can infer from the strengthened soft-tissue pump and the weakened carbonate pump that the oceanic pCO<sub>2</sub> would decrease significantly because of changes in biologically driven [DIC] and [TA] near the surface. The gas exchange component indeed shows that the entire ocean basins favor uptake of CO<sub>2</sub> as manifested by widespread positive signs of  $S_{\alpha}^{DIC}$ . We also note that North Atlantic Deep Water mass responds most sensitively to an increase in  $\alpha$  in terms of carbon uptake.

### 3.5.4. Parameters $R$ and $d$

[47] The  $S_R^{DIC}$  and  $S_d^{DIC}$  patterns (sixth and seventh rows in Figures 6 and 7) show the combined effects from the carbonate pump and the gas exchange pump. An increase in  $R$  increases the export of CaCO<sub>3</sub> from the euphotic layer. This produces an enhanced vertical and basin-to-basin contrast in DIC. A deeper dissolution of CaCO<sub>3</sub> also causes a larger fraction of CaCO<sub>3</sub> to be pumped into the deeper part of the ocean and so also produces an enhancement in the vertical and basin-to-basin contrasts of DIC. Compared to that of  $S_R^{DIC}$ , the carbonate component of  $S_d^{DIC}$  is characterized by contour lines concentrated at depths below ~2000 m (also seen in  $S_d^{TA}$  in Figure 4). This is because  $d$  affects [DIC] only through dissolution processes of which the reference e-folding length scale is 3500 m. In comparison,  $R$  influences [DIC] through both production and dissolution. As a result, the upper ocean also shows considerable sensitivities with respect to a change in  $R$ . An increase in either  $R$  or  $d$  raises the oceanic pCO<sub>2</sub> and thus the ocean releases carbon as can be seen in the gas exchange component of  $S_R^{DIC}$  and  $S_d^{DIC}$ .

### 3.6. General Remarks on $S^{DIC}$ Patterns

[48] The sensitivity of the three pump components shows how changes in biological processes couple with different aspects of the ocean circulation (Figure 7). The soft-tissue component shows coupling with both the wind-driven and meridional overturning circulations. The parameters  $\sigma$  and  $1/\kappa$  interact most strongly with the wind-driven circulation while rC:P and  $\alpha$  interact predominantly with the overturning circulation. The carbonate component shows that

**Table 3.** Percentage of the Spatial Variance in the Tracer Fields Captured by the Model Before and After the Optimization<sup>a</sup>

Data	OCMIP-2	OPT	OPT(rN:P = 16)
Phosphate	61	70	70
Total alkalinity	49	71	71
DIC	49	75	75

<sup>a</sup>Units are percent.

changes in CaCO<sub>3</sub> fluxes are coupled with the deep circulation below the thermocline. The gas exchange component is tightly coupled with the global ventilation patterns. In particular, NADW responds most sensitively to changes in  $\alpha$ ,  $\sigma$ ,  $R$ , and  $d$ . New water is formed and transported into the interior over the highly productive North Atlantic Ocean, giving NADW high sensitivities to changes in production and remineralization. For changes in rC:P, rN:P and  $1/\kappa$ , high sensitivities are also evident in subtropical thermocline water masses.

[49] One important feature of the  $S^{DIC}$  patterns is that they have distinct patterns across the parameters. The combination of the three sensitivity components results in more or less unique patterns of the sensitivity. The distinct sensitivity pattern allows the parameters to be constrained better using DIC data than TA data. We discuss the implication of the distinct patterns for parameter constraint in section 4.3.

## 4. Parameter Optimization

### 4.1. Optimization Method and Data

[50] To optimize the biogeochemical model parameters, we form a cost function which provides an objective measure of the model-data misfit. The cost function is the weighted sum of the fraction of unexplained variance of PO<sub>4</sub>, TA and DIC integrated over the global ocean given by

$$\begin{aligned}
 C(\mu) &= w_1 \frac{\langle ([\text{PO}_4] - [\text{PO}_4]_{obs})^2 \rangle}{\text{Var}([\text{PO}_4]_{obs})} + w_2 \frac{\langle ([\text{TA}] - [\text{TA}]_{obs})^2 \rangle}{\text{Var}([\text{TA}]_{obs})} \\
 &\quad + w_3 \frac{\langle ([\text{DIC}] - [\text{DIC}]_{obs})^2 \rangle}{\text{Var}([\text{DIC}]_{obs})} \\
 &\equiv w_1 C^{\text{PO}_4}(\mu) + w_2 C^{\text{TA}}(\mu) + w_3 C^{\text{DIC}}(\mu), \quad (44)
 \end{aligned}$$

where the angle brackets are operators that compute the global mean of tracer concentrations as defined in equation (26) and where the subscript *obs* denotes the observed tracer field. The denominators are the spatial variances of observed tracer distributions, i.e.,  $\text{Var}([\text{PO}_4]_{obs}) \equiv \langle ([\text{PO}_4]_{obs} - \langle [\text{PO}_4]_{obs} \rangle)^2 \rangle$ ,  $\text{Var}([\text{TA}]_{obs}) \equiv \langle ([\text{TA}]_{obs} - \langle [\text{TA}]_{obs} \rangle)^2 \rangle$  and  $\text{Var}([\text{DIC}]_{obs}) \equiv \langle ([\text{DIC}]_{obs} - \langle [\text{DIC}]_{obs} \rangle)^2 \rangle$ . The weights,  $w_1$ ,  $w_2$  and  $w_3$  can be adjusted to take into account the relative uncertainties of PO<sub>4</sub>, TA and DIC measurements. The  $C^{\text{PO}_4}(\mu)$ ,  $C^{\text{TA}}(\mu)$  and  $C^{\text{DIC}}(\mu)$  represent

the PO<sub>4</sub>, TA and DIC components of the cost function, respectively. We chose  $w_1 = w_2 = w_3 = 1$  and discuss the sensitivity of our results to this particular choice of weights in section 4.2.

[51] We use the globally gridded data sets of PO<sub>4</sub> from World Atlas 2001 [Conkright *et al.*, 2002] and of TA and DIC from GLODAP [Key *et al.*, 2004] to constrain the parameters. The cost function is computed by interpolating the observations to our model grid. While atmospheric pCO<sub>2</sub> is fixed at the preindustrial level of 278 ppm in our steady state solver, the DIC data is constructed on the basis of the 1990s observations [Key *et al.*, 2004]. We estimate the anthropogenic DIC distribution and add it to our preindustrial equilibrium DIC field before minimizing the model-data difference. To estimate the cumulative amount of anthropogenic carbon taken up by the ocean until the year 1990, we time-integrate the carbon model starting from the reference preindustrial equilibrium distribution with atmospheric CO<sub>2</sub> prescribed to follow the observed record from 1765 to 1990 according to the OCMIP-2 protocol (<http://www.ipsl.jussieu.fr/OCMIP/>). The difference field between the DIC distribution in 1990 and that in the preindustrial time is then added to our preindustrial equilibrium DIC field when evaluating the model-data misfit.

[52] We optimize all seven parameters simultaneously by minimizing the discretized version of equation (44). In the discretized version of the cost function, the square of model-data difference at each grid point is weighted by the volume of the grid box. For convenience and ease of use, we adopted Matlab's FMINSEARCH routine to minimize the cost function. This routine uses the Nelder-Mead simplex algorithm [see Lagarias *et al.*, 1998, and references therein]. For our seven parameter cost function the minimum was found after several hundred iterations. The Nelder-Mead simplex algorithm is a robust minimization algorithm but it does not utilize the cost function gradient during its search. Given the smoothness of our cost function we expect that a factor of 10 decrease in the number of function evaluations could have been achieved had we adopted a quasi-Newton method for the optimization itself.

### 4.2. Optimized Parameters

[53] Our optimization significantly improves the agreement between the model and observations (Table 3). Before optimizing the parameters, the simulation with OCMIP-2 parameters captures 61%, 49% and 49% of the spatial variance of the PO<sub>4</sub>, TA and DIC fields. The simulation using the optimized parameter set increases the variance captured by the model to 70%, 71% and 75%, respectively.

[54] The optimized parameter values are given in Table 2. To determine how well each parameter is constrained, we also present in Figure 8 the contour plots of the cost function in the proximity of its minimum. The plots are

**Figure 8.** Contour plots of  $100 \times (C(\mu) - C(\mu^{opt}))$  showing the shape of the cost function for the combined data of PO<sub>4</sub>, TA and DIC with variations in parameters, where  $C(\mu^{opt}) = 0.30 + 0.29 + 0.25 = 0.84$ . The parameters indicated on the *x* axis and the *y* axis are varied in the neighborhood of their minima (Table 2) with the other parameters being held constant at their optimal values. A cross marks the location of the minimum of the cost function and the deviations from the minimum value are contoured. We arbitrarily draw contour lines at 3, 10 and increments of 10 afterward. The contour labeled 3 corresponds to an average 1% decrease in the fraction of the spatial variance in the observed tracer data captured by the model.



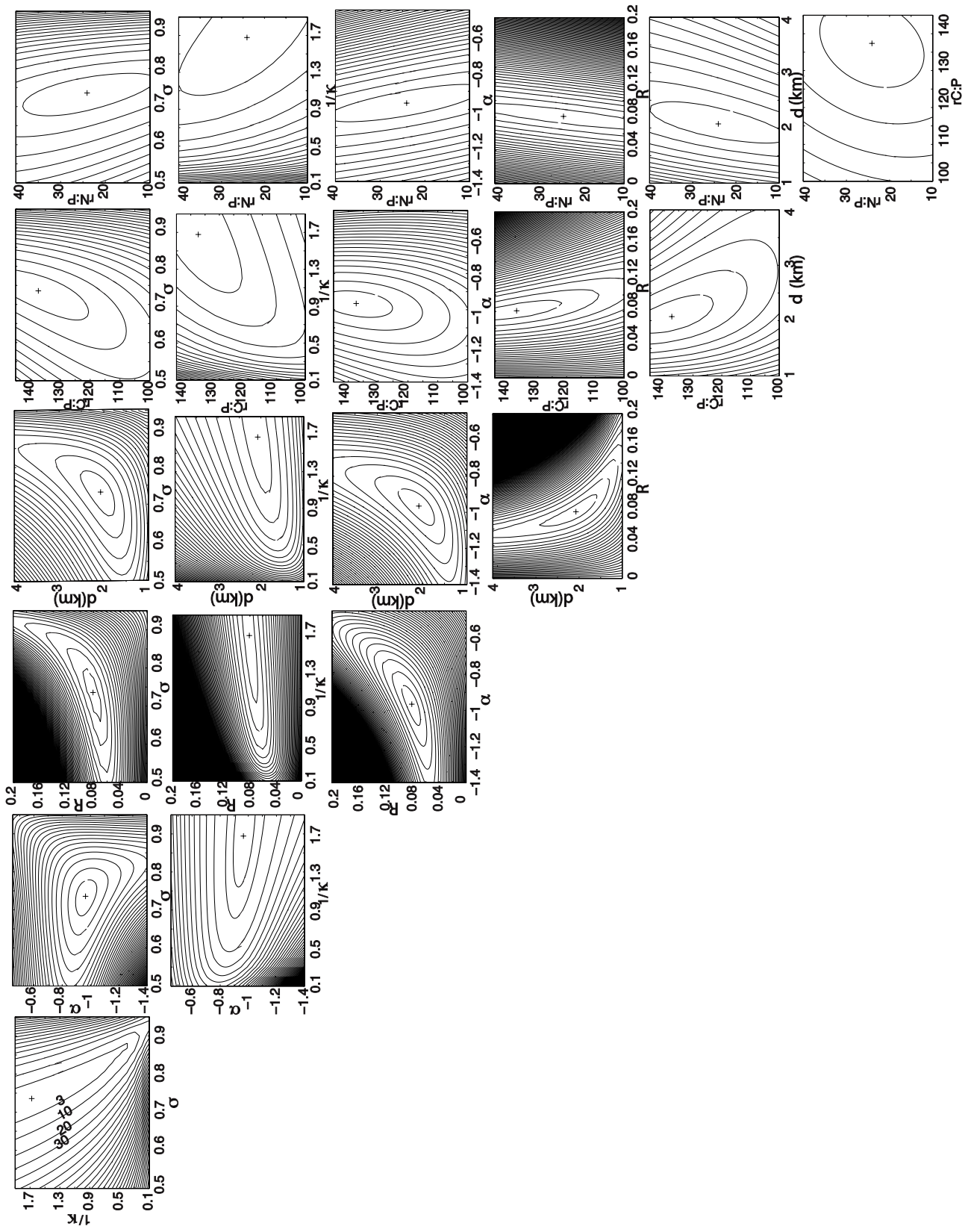


Figure 8

two-dimensional cross sections through the cost function in parameter space. The cost function is evaluated by computing steady state solutions on a grid of parameter values in the neighborhood of the optimal parameter set. In Figure 8 we discretized the parameter ranges using equally spaced points with 10 points for  $\sigma$  and  $1/\kappa$ , 16 points for  $d$ , rC:P and rN:P, 20 points for  $\alpha$  and 21 points for  $R$ . A cross marks the location of the minimum of the cost function multiplied by 100, which is  $100 \times 0.84 = 84$ , and the deviations from the minimum value are contoured. Thus the contour line of “3” indicates an average 1% decrease in the fraction of the spatial variance in the observed tracer data captured by the model. From the shape of the contour lines, we can extract information on the extent to which the parameters are constrained either individually or in combination.

#### 4.2.1. Stoichiometry Ratios: rC:P and rN:P

[55] After optimization we obtain C:N:P ratios of 137:24:1. These are higher than  $117 \pm 14/16 \pm 1/1$ , which Anderson and Sarmiento [1994] estimate on the basis of a nonlinear inverse model and nutrient data. Large values for the C:P ratio have been reported in previous studies [e.g., Takahashi et al., 1985] but our estimate of the N:P ratio is greater than the reference value by 50%. Such a large value has not been reported elsewhere.

[56] A contour plot of the cost function as a function of rC:P and rN:P in Figure 8 shows how rC:P is relatively well constrained in comparison to rN:P. For example, the increase in the cost function by 3 results in a  $\pm 8\%$  change in rC:P, i.e., 126–149, when all parameters except rC:P are held constant at their optimal values. However, for the same increase in the cost function rN:P changes by  $\pm 46\%$ , i.e., 13–36. The relatively weak sensitivity of the cost function to rN:P results in a relatively larger uncertainty for rN:P compared to rC:P.

[57] We speculate that the poor constraint on rN:P is because the OCMIP-2 model does not explicitly simulate important processes associated with the nitrogen cycle [Najjar et al., 2007]. In the OCMIP-2 model, the rN:P is used only to account for the contribution of nitrate fluxes to TA and plays a role in controlling DIC distribution only through gas exchange processes. Thus, by using model formulations resolving the nitrogen cycle and combining with nitrogen data, it might be possible to better constrain rN:P.

#### 4.2.2. Parameters Controlling CaCO<sub>3</sub> Fluxes Only: R and d

[58] We obtain the CaCO<sub>3</sub> production ratio  $R = 0.081$  which is larger than the value of  $R = 0.07$  used as the reference OCMIP-2 value. Our value is slightly larger than the range  $R = 0.06–0.08$  suggested by Yamanaka and Tajika [1996] but falls within the range of  $0.06 \pm 0.03$  given by Sarmiento et al. [2002] and within the range of  $R = 0.07–0.10$  given by Jin et al. [2006].

[59] Our optimized value  $d = 2100$  m is 40% smaller than  $d = 3500$  m obtained from Yamanaka and Tajika [1996] and

used as the reference OCMIP-2 value. This large change in the value of  $d$  is significant because of the relatively high sensitivity of the cost function to changes in  $d$  (Figure 8). An increase in the cost function by 3 suggests the optimal range of  $d = 1800–2400$  m. By changing the value of  $d$  from the reference to optimal value with the other parameters being held at their optima, the cost function decreases by as much as 50. The e-folding length scale of  $d = 2100$  m produces the profile of CaCO<sub>3</sub> downward fluxes that gradually decrease throughout the entire water column, indicating that a significant fraction of exported CaCO<sub>3</sub> dissolves at depths shallower than the global mean calcite saturation horizon (See Figure 11 and discussion in section 4.4). Our optimization result thus provides an alternative line of evidence in support for the existence of processes leading to CaCO<sub>3</sub> dissolution at depths shallower than the calcite saturation horizon in the Atlantic, Indian and Pacific oceans obtained using TA\* [Chung et al., 2003; Sabine et al., 2002; Feely et al., 2002]. That we obtain a similar result using a different approach is important because the recent study of Friis et al. [2006] has suggested that the studies using the TA\* tracer have not properly taken into account the influence of transport.

#### 4.2.3. Parameters $\sigma$ , $1/\kappa$ , and $\alpha$

[60] In a previous study (KP06), we reported results from optimization of  $\sigma$ ,  $1/\kappa$  and  $\alpha$  using only PO<sub>4</sub> data. In the present study, we find that the optimization using the combined data set of PO<sub>4</sub>, TA and DIC leads to only minor changes in the optimal parameter values of  $\sigma$ ,  $1/\kappa$  and  $\alpha$ . The optimized value for  $\sigma$  is identical within 2 significant figures to our previous value. Our new optimized value of  $\alpha = -0.97$  is only slightly larger than the value  $\alpha = -1.0$  we obtained previously. Our new optimized value of  $1/\kappa = 1.7$  years lies in between 1.0 and 2.8 years we obtained previously. The shape of the cost function based on the equally weighted combination of DIC, TA and PO<sub>4</sub> presented in the  $\sigma - 1/\kappa$  plot of Figure 8 is also similar to the cost function based on only PO<sub>4</sub> shown in KP06. This indicates that  $1/\kappa$  and to a lesser extent  $\sigma$  remain relatively poorly constrained even with the addition of DIC and TA data as can be deduced from the strong correlation between the sensitivity patterns for  $S^{TA}$  and  $S^{DIC}$  for both  $\sigma$  and  $1/\kappa$ .

[61] In our previous study (KP06), we attributed a large fraction of model improvement to changes in the two parameters,  $\sigma$  and  $1/\kappa$  that control DOM production and remineralization, because of their effect on the nutrient trapping problem in the eastern equatorial oceans. We find that this same effect also improves the agreement between the model and the observations for DIC and TA as we discuss in section 4.5.

#### 4.3. Relative Importance of DIC and TA for Constraining Parameters

[62] We investigate the degree to which DIC or TA alone constrains the parameters by plotting two-dimensional con-

**Figure 9.** The TA component of Figure 8,  $100 \times (C^{TA}(\mu) - C^{TA}(\mu^{opt}))$  where  $C^{TA}(\mu^{opt}) = 0.29$ . The parameters indicated on the x axis and the y axis are varied in the neighborhood of their minima (Table 2) with the other parameters being held constant at their optimal values. A cross marks the location of the minimum of the cost function and the deviations from the minimum value are contoured. We arbitrarily draw contour lines at 1, 5 and increments of 5 afterward. The contour labeled 1 corresponds to a 1% decrease in the fraction of the spatial variance in the observed TA data captured by the model.

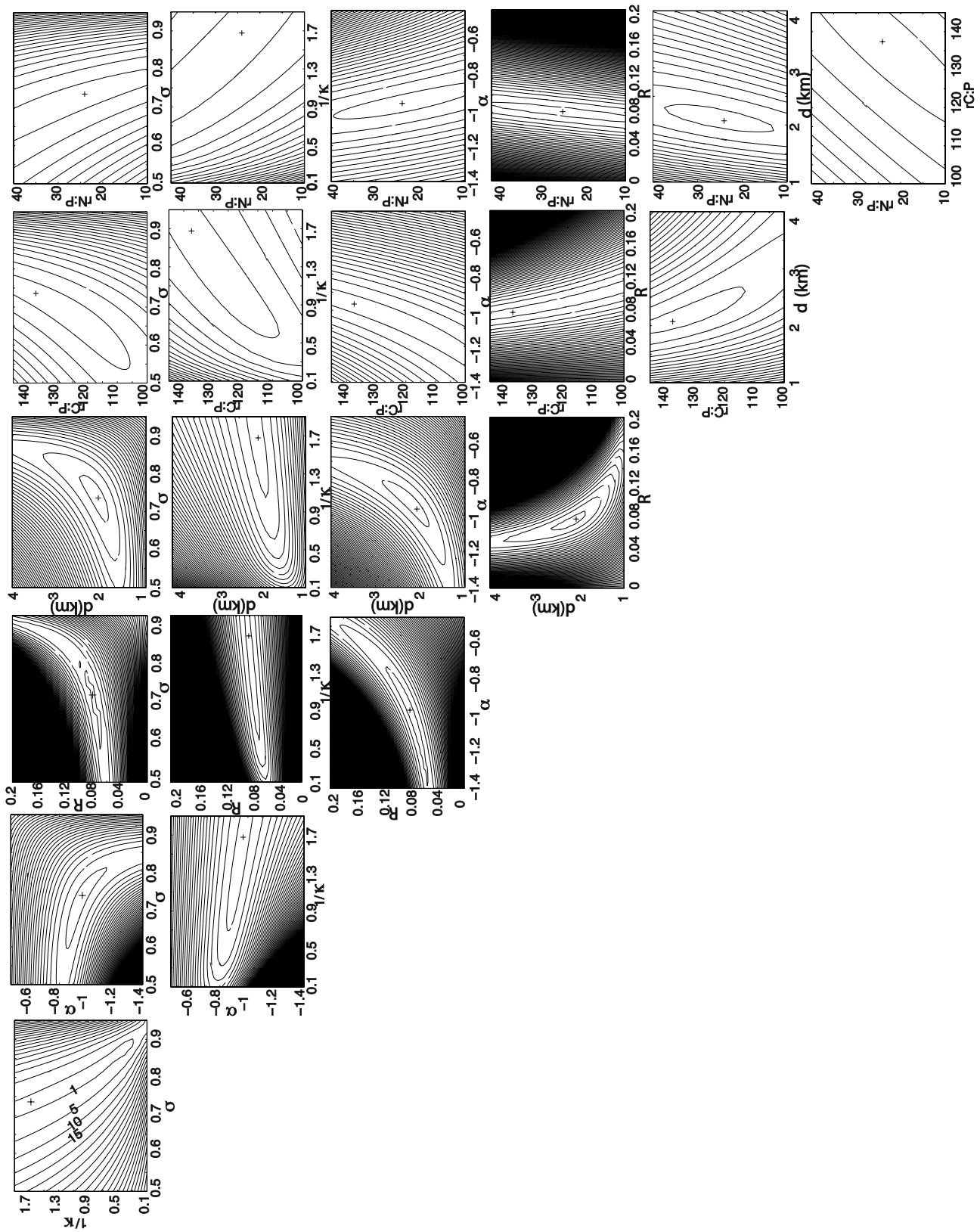


Figure 9

four plots of the DIC and TA cost-function components in Figures 9 and 10. The minimum of the total cost function, indicated by the cross, approximately coincides with the minimum of each component. This suggests that the optimum parameter set is insensitive to the choice of the weights in the cost function (i.e.,  $w_1$ ,  $w_2$  and  $w_3$  in equation (44)). Figures 9 and 10 do show however that the two components of the cost function have different shapes, indicating that uncertainty estimates would depend on the particular choice of weights. For example, the rC:P-R plot for the TA component of the cost function shows that the parameters are constrained mostly to a one-dimensional family in which rC:P and  $R$  are negatively correlated whereas the DIC component of the cost function shows little correlation between each parameter. To summarize the similarities in the  $S$  patterns, we show the correlation coefficient for each pair of  $S^{TA}$  patterns (lower triangle in Table 4) and for each pair of  $S^{DIC}$  patterns (upper triangle in Table 4). As one expect from a visual inspection of the  $S^{TA}$  patterns shown in Figure 4, all pairs of the  $S^{TA}$  patterns are highly correlated. We see correlation coefficients that exceed  $|r| = 0.8$  in all of the pairs except for rN:P- $d$ , which has  $|r| = 0.74$ . In contrast, the  $S^{DIC}$  patterns are relatively less correlated. Only the pairs  $\sigma - 1/\kappa$ ,  $R - d$ ,  $1/\kappa - R$  and  $\alpha$ -rC:P have the correlation coefficients  $|r| > 0.8$ . As a result, these parameter pairs (in particular,  $\sigma - 1/\kappa$  and  $d$ - $R$ ) are not well constrained independently using either TA data or DIC data.

[63] Combining multiple tracers for the optimization is particularly complimentary for pairs of parameters that have  $S$  patterns with correlation coefficients of opposing signs. This is the case for  $\alpha$ -rC:P and rC:P-rN:P where the combination of TA and DIC data greatly improves the ability of the inversion to resolve individual parameters.

#### 4.4. Export of Organic Carbon and CaCO<sub>3</sub>

[64] The postoptimization globally integrated POC export at the base of the euphotic layer is 14.8 GtC/a. The preoptimization value using the OCMIP-2 reference parameters was 15.5 GtC/a. Interestingly, the estimated POC export production estimated in KP06 after having optimized only the parameters associated with the phosphorous cycle was also 14.8 GtC/a despite having used a C:P stoichiometry ratio rC:P = 117. The effect of optimizing all the model parameters using the combined PO<sub>4</sub>, DIC and TA data has been to lower the net export of particulate organic phosphorus in comparison to the optimization of KP06 so that the implied POC export remained fixed even though rC:P increased from 117 to 137.

[65] In comparison to the OCMIP-2 reference value, the globally integrated CaCO<sub>3</sub> export production from the euphotic layer increases from 1.09 GtC/a to 1.20 GtC/a. In spite of the greater export production at the depth of 75 m, the vertical profiles of CaCO<sub>3</sub> fluxes (Figure 11) show that smaller amounts of CaCO<sub>3</sub> are exported below a

depth of  $\sim 700$  m for the optimized solution. This can be attributed to the substantial reduction in the dissolution length scale of CaCO<sub>3</sub> after optimization. As a result the optimized model suggests that a greater fraction of CaCO<sub>3</sub> is regenerated in the upper ocean above  $\sim 2600$  m. The optimized solution is consistent with the study of *Feely et al.* [2004] which estimated that up to 60% of total water column dissolution occurs above a depth of 2000 m. For our model with optimized parameters 67% of the CaCO<sub>3</sub> export is dissolved in the upper 2000 m. Our model is unable to distinguish between dissolution in the water column and in the sediments since it does not include any sedimentation processes or inputs of CaCO<sub>3</sub> from rivers.

[66] We find that the combined data sets of PO<sub>4</sub>, TA and DIC constrain well the export production of POC and CaCO<sub>3</sub>. Figures 12a and 12b show the globally integrated CaCO<sub>3</sub> export production as a function of  $\sigma$  and  $1/\kappa$  and as a function of  $\alpha$  and rC:P. We overlay the contour line of the 1% decrease in the fraction of spatial variance captured by the model, which is taken from Figure 8, over the contour plots of the CaCO<sub>3</sub> export production. The minimum and maximum of the globally integrated CaCO<sub>3</sub> export production within the parameter space defined by a 1% decrease in the fraction of spatial variance captured by the model is 1.1 GtC/a and 1.3 GtC/a. With the rain ratio of  $R = 0.081$ , the POC export production is constrained to  $15 \pm 1$  GtC/a. Our study suggests that the spatial distributions of PO<sub>4</sub>, TA and DIC are highly sensitive to changes in the export of POC and CaCO<sub>3</sub>. In other words, having a right amount of downward export of POC and CaCO<sub>3</sub> provides a necessary condition for minimizing the model-data misfit. We also point out that the vertical export of biogenic material depends sensitively on aspects of the ocean circulation that are responsible for bringing nutrients to the euphotic layer [e.g., *Gnanadesikan et al.*, 2002]. This implies that the optimized value of the export production of CaCO<sub>3</sub> shown in Figures 12a and 12b would shift toward higher values if we used transport models with higher vertical mixing regimes and lower values if we used transport models with lower vertical mixing regimes. The fraction of CaCO<sub>3</sub> that dissolves above a depth of 2000 m is a function of  $d$  only and constrained to  $67 \pm 4\%$  of which  $\pm 4$  is obtained by allowing a 1% decrease in the fraction of spatial variance captured by the model (Figure 12c).

#### 4.5. Equilibrium Solutions Based on Optimized Parameters

[67] We present the equilibrium solutions based on the optimized parameters listed in Table 2. The vertically and zonally averaged TA and DIC fields are presented in the bottom panels of Figures 1 through 3. The top and middle panels show respectively the solutions obtained using the OCMIP-2 reference parameters and the observed TA and

**Figure 10.** The DIC component of Figure 8,  $100 \times (C^{DIC}(\mu) - C^{DIC}(\mu^{opt}))$  where  $C^{DIC}(\mu^{opt}) = 0.25$ . The parameters indicated on the x axis and the y axis are varied in the neighborhood of their minima (Table 2) with the other parameters being held constant at their optimal values. A cross marks the location of the minimum of the cost function and the deviations from the minimum value are contoured. We arbitrarily draw contour lines at 1, 5 and increments of 5 afterward. The contour labeled 1 corresponds to a 1% decrease in the fraction of the spatial variance in the observed DIC data captured by the model.

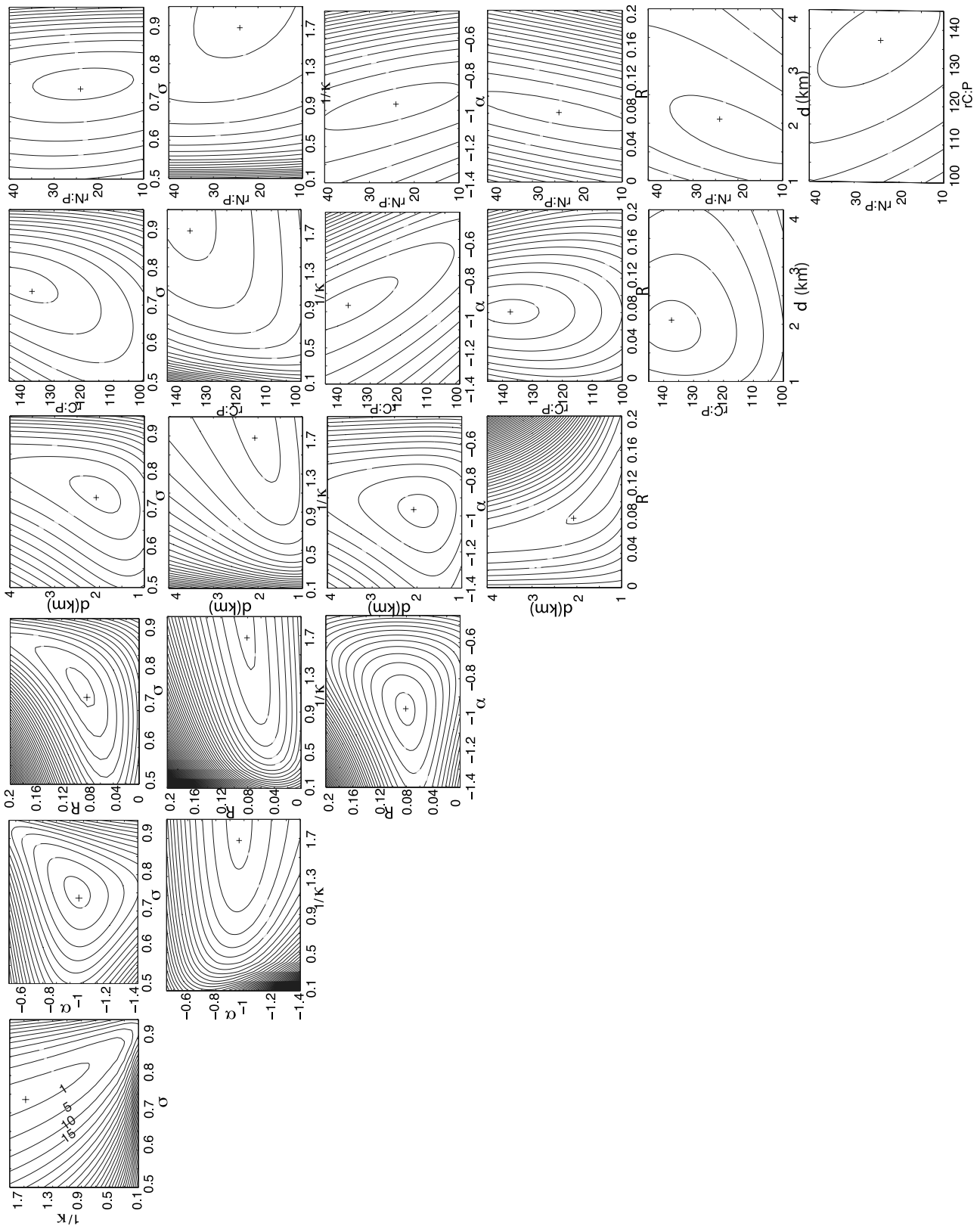


Figure 10

**Table 4.** Correlation Coefficients Between Pairs of  $S^{TA}$  Patterns (Lower Triangle) and Between Pairs of  $S^{DIC}$  Patterns (Upper Triangle)

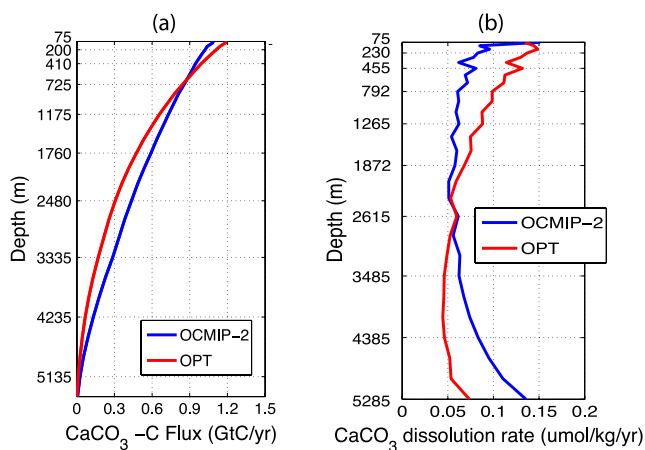
Parameter	$\sigma$	$1/\kappa$	$\alpha$	$R$	$d$	rC:P	rN:P
$\sigma$		0.95	-0.32	-0.72	-0.62	-0.59	0.21
$1/\kappa$	1.00		-0.08	-0.84	-0.75	-0.35	0.45
$\alpha$	0.97	0.97		-0.24	-0.17	0.83	0.74
$R$	-0.99	-0.99	-0.99		0.94	0.00	-0.79
$d$	-0.88	-0.86	-0.93	0.91		-0.05	-0.74
rC:P	-0.99	-0.99	-0.99	1.00	0.91		0.60
rN:P	0.91	0.92	0.88	-0.91	-0.74	-0.91	

DIC fields. The simulation with the OCMIP-2 reference parameters was discussed in section 2.4.

[68] The most noticeable improvement is made in the equatorial Pacific Ocean, where column-averaged [TA] and [DIC] are significantly reduced (Figure 1). Approximately 50% of the excess [TA] and approximately 64% of the excess [DIC] in the reference simulation are removed in the optimized solution. This is attributable to increased dissolved organic matter in high-productivity upwelling regions in the equatorial oceans. As discussed in section 4.2.3, the increased DOM pools are mainly due to increased  $\sigma$  and  $1/\kappa$ . Dissolved forms of nutrients and carbon are transported laterally by the ocean currents while being remineralized. The lateral transport of DOM then causes a decrease in the production of organic matter and  $\text{CaCO}_3$  in the upwelling region. As a result, part of the nutrient trapping problem is reduced.

[69] Zonally averaged [TA] and [DIC] distributions also show better agreements in terms of the vertical and basin-to-basin contrasts (Figures 2 and 3). With the optimized parameters, the simulated [TA] and [DIC] are increased in NADW and decreased in deep waters in the North Pacific Ocean. The reduced vertical and basin-to-basin contrasts are consistent with the weakened carbonate pumps after optimization. As shown in Figure 11 and discussed in section 4.4, the larger fraction of  $\text{CaCO}_3$  that dissolves in the upper ocean leads to weakening of the carbonate pump.

[70] It is important to point out that some discrepancies between the simulated and observed fields remain in spite of the improvements made after optimization. For example, the excess [TA] and [DIC] in high-productivity equatorial Pacific Ocean are not completely removed. The discrepancies in the AABW in the Atlantic are also barely

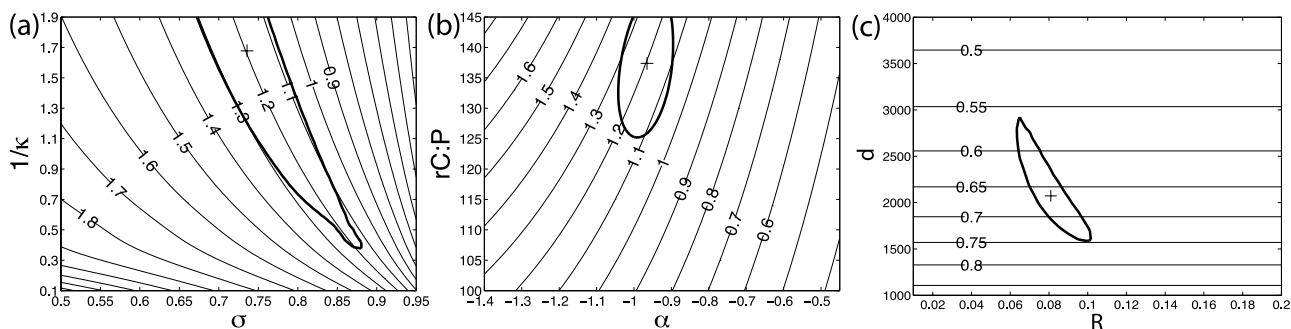


**Figure 11.** (a) Globally integrated  $\text{CaCO}_3$ -C fluxes with depth in  $\text{GtC a}^{-1}$  obtained using the OCMIP-2 reference parameters (blue curve) and using the optimized parameter set (red curve), (b) Globally averaged  $\text{CaCO}_3$  dissolution rates with depth in  $\mu\text{mol kg}^{-1} \text{a}^{-1}$ .

improved for both TA and DIC fields. Except for a slight improvement shown in the TA field, our simulation still overestimates both [TA] and [DIC] near the bottom of the Atlantic. Our optimized parameter set also fails to capture high concentrations of TA and DIC in the thermocline of the North Pacific Ocean. As we discussed in KP06 and in section 2.4 these errors are likely due to circulation errors, but the need to add additional complexity to the biogeochemistry model cannot be ruled out.

## 5. Discussion and Conclusions

[71] Our sensitivity and optimization study has been made computationally feasible by the use of an implicit solver based on Newton's method to obtain steady state solutions of phosphate ( $\text{PO}_4$ ), dissolved organic phosphorous (DOP), total alkalinity (TA) and dissolved inorganic carbon (DIC). Our study uses global DIC, TA and  $\text{PO}_4$  tracer data to optimize the parameters in Table 1, which control the tracer distributions through stoichiometric ratios and the production and remineralization of organic and inorganic carbon. We also compute sensitivity patterns ( $S$  patterns) by taking



**Figure 12.** Globally integrated export production  $\text{CaCO}_3$ -C in  $\text{GtC a}^{-1}$  (a) as a function of  $\sigma$  and  $1/\kappa$  and (b) as a function of  $\alpha$  and rC:P (c) the fraction of  $\text{CaCO}_3$  that dissolves above a depth of 2000 m as a function of  $R$  and  $d$ . The 1% uncertainty contour lines of the cost function from Figure 8 are overlaid for the three panels.

partial derivatives of equilibrium solutions with respect to the parameters.

[72] Our sensitivity analysis is useful not only to help understand the parameter optimization result but also to elucidate the mechanisms that impact tracer distributions when parameter values are changed. We introduce a novel decomposition method that allows us to interpret the  $S$  patterns through changes in the soft-tissue, carbonate and gas exchange pumps. The decomposed  $S$  patterns illustrate quantitatively how the carbonate pump couples to the deep branch of the overturning circulation through  $d$ , how the soft-tissue pump couples to the wind-driven circulation through  $\sigma$  and  $1/\kappa$  and to the overturning circulation through  $rC:P$  and  $\alpha$ , and how the gas exchange pump couples strongly with the newly ventilated NADW through  $\alpha$ ,  $\sigma$ ,  $R$  and  $d$ .

[73] The  $S$  patterns for TA show similar patterns for all parameters because it is dominated by the contribution from the carbonate pump. The  $S$  patterns for DIC show more distinct structures across parameters because the three pumps are more equally important to redistribute carbon in the model's ocean-atmosphere system. The distinct patterns imply that the parameters can be better constrained individually when the DIC data is used in conjunction with the TA data.

[74] For the most part, the present work has focused on the parameter sensitivity and optimization as a means of improving simulations. But our parameter optimization can be thought of as the solution to an inverse problem that seeks to constrain uncertain biogeochemical parameters using observations. For example, our optimization can be thought of as an alternative to tracer based methods such as the method used by *Anderson and Sarmiento* [1994] for determining  $rN:P$  and  $rC:P$  or to the TA\* method used by *Chung et al.* [2003], *Sabine et al.* [2002], and *Feely et al.* [2002] to determine the depth of  $CaCO_3$  dissolution. From this point of view the circulation model provides prior information about the ocean circulation that helps reduce the indeterminacy due to transport in tracer inversion. The study of *Jin et al.* [2006] uses a global circulation model in a similar spirit to estimate export fluxes.

[75] As a parameter inversion, our study is incomplete because we have not provided probabilistic uncertainty estimates for our optimal parameters. Formal uncertainty estimates can be obtained in a straightforward way from the Hessian of the cost function [e.g., *Thacker*, 1989], except that without having included an estimate of the spatial covariance of the model-data residuals, the apparent number of degrees of freedom in the combined DIC, TA and  $PO_4$  data set is so large ( $n = 163,884$ ) that it produces unacceptably small error estimates. That the number of degrees of freedom in the residuals is in fact much less is easy to see from the smoothness of the model output and of the WOA01 and GLODAP data sets. We are currently working on a procedure in which we estimate the spatial covariance matrix for the residuals as part of the inversion and hope to report on this work elsewhere.

[76] The sensitivity of the cost function indicates that our optimized parameters  $rC:P = 137$ ,  $\sigma = 0.74$ ,  $\alpha = -0.97$ ,  $R = 0.081$  and  $d = 2100$  m are relatively well constrained whereas  $rN:P = 24$  and  $1/\kappa = 1.7$  years are less well constrained in this study. The poor constraint on  $rN:P$  and

$1/\kappa$  is thought to be due to the biogeochemistry model that does not resolve important processes controlling the parameters. For example, *Anderson and Sarmiento* [1994] point out the importance of considering denitrification processes when constraining the  $N:P$  ratio. This suggests that the relatively poor constraint on  $rN:P$  might be improved by explicitly resolving the nitrogen cycle so that nitrogen data might be used to further constrain  $rN:P$ . The parameter  $1/\kappa$  could be better constrained if we resolved a seasonal cycle of DOM and if we used DOM data.

[77] Nonetheless the higher  $C:P$  ratio and the significantly smaller dissolution length scale of  $CaCO_3$  that we obtain in this optimization compared to their reference values seem robust. Because our estimate of  $rN:P$  did not appear robust and seemed unacceptably large we repeated the optimization holding the  $N:P$  ratio fixed at its reference value of  $rN:P = 16$ . The resulting optimized values  $rC:P = 141$ ,  $\sigma = 0.73$ ,  $1/\kappa = 1.8$  years,  $\alpha = -0.96$ ,  $R = 0.073$  and  $d = 2100$  m all remained within the 1% cost-function sensitivity ranges (Table 2). This gives us additional confidence that with the exception of  $rN:P$  our optimization results are robust.

[78] The strong coupling of ocean physical and biogeochemical dynamics as shown in the  $S$  patterns implies that the behavior of the model's carbon system would differ if the circulation and ventilation fields changed. The high sensitivity of model solution to circulation has been pointed out in several previous studies [e.g., *Najjar et al.*, 2007]. This highlights the importance of conducting sensitivity and optimization studies similar to ours in several different circulation models and seasonally varying circulations. Such studies will help better constrain model parameters and better quantify their uncertainty so as to improve the predictive skill of simulations of the future carbon cycle.

[79] **Acknowledgments.** We thank all those who contributed to constructing the global-scale database for DIC, TA and phosphate data. We thank N. Gruber for useful suggestion related to an earlier version of our manuscript. We also thank E. Kwon's thesis committee members, E. Druffel and K. Moore, who continuously provide valuable comments and help. Constructive comments from two anonymous reviewers are also greatly appreciated. Support from the National Science Foundation grants OCE 02215156 and OCE 0623647 is also acknowledged.

## References

- Amestoy, P. R., I. S. Duff, J. Koster, and J.-Y. L'Excellent (2001), A fully asynchronous multifrontal solver using distributed dynamic scheduling, *SIAM J. Matrix Anal. Appl.*, 3(1), 15–41.
- Anderson, L. A., and J. L. Sarmiento (1994), Redfield ratios of remineralization determined by nutrient data analysis, *Global Biogeochem. Cycles*, 8(1), 65–80.
- Aumont, O., J. C. Orr, P. Monfray, G. Madec, and E. Maier-Raimer (1999), Nutrient trapping in the equatorial Pacific: The ocean circulation solution, *Global Biogeochem. Cycles*, 13(2), 351–369.
- Brewer, P. G., G. T. F. Wong, M. P. Bacon, and D. W. Spencer (1975), An oceanic calcium problem?, *Earth Planet. Sci. Lett.*, 26, 81–87.
- Chung, S.-N., K. Lee, R. A. Feely, C. L. Sabine, F. J. Millero, R. Wanninkhof, J. L. Bullister, R. M. Key, and T.-H. Peng (2003), Calcium carbonate budget in the Atlantic Ocean based on water column inorganic carbon chemistry, *Global Biogeochem. Cycles*, 17(4), 1093, doi:10.1029/2002GB002001.
- Conkright, M. E., et al. (2002), *World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures*, CD-ROM documentation, 17 pp., Natl. Oceanogr. Data Cent., Silver Spring, Md.
- Feely, R. A., et al. (2002), In situ calcium carbonate dissolution in the Pacific Ocean, *Global Biogeochem. Cycles*, 16(4), 1144, doi:10.1029/2002GB001866.
- Feely, R. A., C. L. Sabine, K. Lee, W. Berelson, J. Kleyvas, V. J. Fabry, and F. J. Millero (2004), Impact of anthropogenic  $CO_2$  on the  $CaCO_3$  system in the oceans, *Science*, 305, 362–366.

- Friedrichs, M. A. M., R. R. Hood, and J. D. Wiggert (2006), Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data, *Deep Sea Res., Part II*, 53, 576–600, doi:10.1016/j.dsr2.2006.01.026.
- Friis, K., R. G. Najjar, M. J. Follows, and S. Dutkiewicz (2006), Possible overestimation of shallow-depth calcium carbonate dissolution in the ocean, *Global Biogeochem. Cycles*, 20, GB4019, doi:10.1029/2006GB002727.
- Gent, P. R., and J. C. McWilliams (1990), Isopycnal mixing in ocean circulation models, *J. Phys. Oceanogr.*, 20, 150–155.
- Gnanadesikan, A., R. D. Slater, N. Gruber, and J. L. Sarmiento (2002), Oceanic vertical exchange and new production: A comparison between models and observations, *Deep Sea Res., Part II*, 49(1–3), 363–401.
- Gruber, N., and J. L. Sarmiento (2002), Biogeochemical/physical interactions in elemental cycles, in *The Sea: Biological-Physical Interactions in the Oceans*, vol. 12, edited by A. R. Robinson, J. J. McCarthy, and B. J. Rothschild, pp. 337–399, John Wiley, Hoboken, N. J.
- Holzer, M., and F. Primeau (2006), The diffusive ocean conveyor, *Geophys. Res. Lett.*, 33, L14618, doi:10.1029/2006GL026232.
- Holzer, M., and F. Primeau (2008), The path-density distribution of oceanic surface-to-surface transport, *J. Geophys. Res.*, 113, C01018, doi:10.1029/2006JC003976.
- Jin, X., N. Gruber, J. P. Dunne, J. L. Sarmiento, and R. A. Armstrong (2006), Diagnosing the contribution of phytoplankton functional groups to the production and export of particulate organic carbon, CaCO<sub>3</sub>, and opal from global nutrient and alkalinity distributions, *Global Biogeochem. Cycles*, 20, GB2015, doi:10.1029/2005GB002532.
- Kelly, C. T. (2003), *Solving Nonlinear Equations With Newton's Method*, *SIAM Fundamentals of Algorithms*, vol. 1, 104 pp., Soc. for Appl. and Ind. Math., Philadelphia.
- Key, R. M., A. Kozyr, C. L. Sabine, K. Lee, R. Wanninkhof, J. L. Bullister, R. A. Feely, F. J. Millero, C. Mordy, and T.-H. Peng (2004), A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP), *Global Biogeochem. Cycles*, 18, GB4031, doi:10.1029/2004GB002247.
- Kwon, E. Y., and F. Primeau (2006), Optimization and sensitivity study of a biogeochemistry ocean model using an implicit solver and in situ phosphate data, *Global Biogeochem. Cycles*, 20, GB4009, doi:10.1029/2005GB002631.
- Lagarias, J. C., J. A. Reeds, M. H. Wright, and P. E. Wright (1998), Convergence properties of the Nelder-Mead simplex method in low dimensions, *SIAM J. Optim.*, 9(1), 112–147.
- Large, W. C., J. C. McWilliams, and S. C. Doney (1994), Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization, *Rev. Geophys.*, 32, 363–403.
- Martin, J. H., G. A. Knauer, D. M. Karl, and W. W. Broenkow (1987), VERTEX: Carbon cycling in the northeast Pacific, *Deep Sea Res.*, 34, 267–285.
- Matear, R. J., and G. Holloway (1995), Modeling the inorganic phosphorus cycle of the North Pacific using an adjoint data assimilation model to assess the role of dissolved organic phosphorus, *Global Biogeochem. Cycles*, 9(1), 101–119.
- Najjar, R. G., et al. (2007), Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from Phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2), *Global Biogeochem. Cycles*, 21, GB3007, doi:10.1029/2006GB002857.
- Primeau, F. (2005), Characterizing transport between the surface mixed layer and the ocean interior with a forward and adjoint global ocean transport model, *J. Phys. Oceanogr.*, 35(4), 545–564, doi:10.1175/JPO2699.1.
- Primeau, F., and M. Holzer (2006), The ocean's memory of the atmosphere: Residence-time and ventilation-rate distributions of water masses, *J. Phys. Oceanogr.*, 36, 1439–1456.
- Sabine, C. L., R. A. Feely, R. M. Key, J. L. Bullister, F. J. Millero, K. Lee, T. H. Peng, B. Tilbrook, T. Ono, and C. S. Wong (2002), Distribution of anthropogenic CO<sub>2</sub> in the Pacific Ocean, *Global Biogeochem. Cycles*, 16(4), 1083, doi:10.1029/2001GB001639.
- Sarmiento, J. L., J. Dunne, A. Gnanadesikan, R. M. Key, K. Matsumoto, and R. Slater (2002), A new estimate of the CaCO<sub>3</sub> to organic carbon export ratio, *Global Biogeochem. Cycles*, 16(4), 1107, doi:10.1029/2002GB001919.
- Takahashi, T., W. S. Broecker, and S. Langer (1985), Redfield ratio based on chemical data from isopycnal surfaces, *J. Geophys. Res.*, 90, 6907–6924.
- Thacker, W. C. (1989), The role of the Hessian matrix in fitting models to measurements, *J. Geophys. Res.*, 94, 6177–6196.
- Volk, T., and M. I. Hoffert (1985), Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO<sub>2</sub> changes, in *The Carbon Cycle and Atmospheric CO<sub>2</sub>: Natural Variations Archean to Present*, *Geophys. Monogr. Ser.*, vol. 32, edited by E. T. Sundquist and W. S. Broecker, pp. 99–110, AGU, Washington, D. C.
- Wanninkhof, R. (1992), Relationship between wind speed and gas exchange over the ocean, *J. Geophys. Res.*, 97, 7373–7382.
- Yamanaka, Y., and E. Tajika (1996), The role of the vertical fluxes of particulate organic matter and calcite in the ocean carbon cycle: Studies using an ocean biogeochemical general circulation model, *Global Biogeochem. Cycles*, 10(2), 361–382.
- Yamanaka, Y., and E. Tajika (1997), Role of dissolved organic matter in the marine biogeochemical cycle: Studies using an ocean biogeochemical general circulation model, *Global Biogeochem. Cycles*, 11(4), 599–612.

E. Y. Kwon, Atmospheric and Oceanic Sciences Program, Princeton University, 300 Forrester Road, Sayre Hall, Princeton, NJ 08544, USA. (ekwon@princeton.edu)

F. Primeau, Department of Earth System Science, University of California, Croul Hall, Irvine, CA 92697, USA. (fprimeau@uci.edu)