Introduction

This online material is more than just a supplemental material and is really an integral part of our publication that we could not include in our printed paper.

1. 2004GB002397-dcstar.txt In section 1, the ocean general circulation model and simulations that generated the synthetic data are described. In section 2, we provide an in-depth reassessment of the ΔC* method and describe how we applied it to synthetic ocean data. In section 3, multiple linear regression models to predict preformed concentrations are described. Additional details regarding possible errors in the biological correction of the ΔC* method (section 4) and the OMP analysis (section 5) are also provided. The correction for a temporal change in the air-sea CO2 disequilibrium is derived in section 6. In section 7, sources of error related to denitrification are described. Finally, the volume census of the model domain that allows for a global extrapolation of our results is provided in section 8.

2. 2004GB002397-figure1.eps A global multiple linear regression model of ALK^0 as a function of temperature, salinity, and PO. The red diagonal line is a 1:1 line. The dashed lines are 2σ (1σ=9.28 μmol-eq/kg). The correlation coefficient between the predicted and simulated ALK^0 is 0.987.

3. 2004GB002397-figure2.eps An error map of ALK^0 predicted using Equation A1. It shows the difference between predicted and simulated surface ALK in unit of μmol-eq/kg.

4. 2004GB002397-figure3.eps A global multiple linear regression model of DIC_{eq}^0 as a function of temperature, salinity, and alkalinity. The red diagonal line is a 1:1 line. The dashed lines are 2σ (1σ=4.82 μmol-C/kg). The correlation coefficient between the predicted and simulated DIC_{eq}^0 is 0.998. The line of positive deviation outside the 2σ envelope represents predicted DIC_{eq}^0 from one latitude
band in the Arctic that could not be masked out; it makes no substantive contribution in this study.

5. 2004GB002397-figure4.eps A Pacific multiple linear regression model of PO$_4$ as a function of temperature and salinity. The red diagonal line is a 1:1 line. The dashed lines are 2σ (1σ=0.21 μmol-P/kg). The correlation coefficient between the predicted and simulated PO$_4$ is 0.914.

6. 2004GB002397-figure5.eps A solution of the OMP analysis that shows the fractions of the four water-types (a) type 1 NW, (b) type 2 EE, (c) type 3 SW, and (d) type 4 SO.

7. 2004GB002397-figure6.eps Comparison of “true” and estimated preindustrial ∆C* on the σθ=26.8 Pacific isopycnal. (a) The “true” distribution as estimated by Equation 3. (b) The distribution as estimated by using the OMP results shown in Figure 5. (c) The deviation of the estimated from the “true”. (c) The “true” in x-axis versus the estimated in y-axis. Unit is in μmole-C/kg.

8. 2004GB002397-figure7.eps The distributions of dye tracers. (a) The sum of 15 dyes released in the Pacific basin. (b) The sum of 14 dyes released in the Atlantic basin. (c) The sum of 7 dyes released in the Indian basin. The fractions of all 36 dyes sum to 1.

In section 1, the ocean general circulation model and simulations that generated the synthetic data are described. In section 2, we provide an in-depth reassessment of the ΔC* method and describe how we applied it to synthetic ocean data. In section 3, multiple linear regression models to predict preformed concentrations are described. Additional details regarding possible errors in the biological correction of the ΔC* method (section 4) and the OMP analysis (section 5) are also provided. The correction for a temporal change in the air-sea CO₂ disequilibrium is derived in section 6. In section 7, sources of error related to denitrification are described. Finally, the volume census of the model domain that allows for a global extrapolation of our results is provided in section 8.

1. Model Description

The physical model is described in detail by Gnanadesikan et al. [2002] as the KVLOW-AILOW model. The same model is also referred to as the Low-Low or LL model in later publications [Gnanadesikan et al., 2002; Gruber et al., 2001; Jin and Gruber, 2003; Matsumoto and Key, 2004; Mignone et al., 2004].

The biogeochemistry component of the model is based on the BIOTIC code of OCMIP-2 [Najjar and Orr, 1999]. Prognostic variables in the model include potential temperature (T), salinity (S), phosphate (PO₄), dissolved inorganic carbon (DIC), alkalinity (ALK), and oxygen (O₂). Biological export production in the near surface ocean is based on restoring model-simulated PO₄ to the observed monthly climatology [Louanchi and Najjar, 2000]. For all biologically mediated processes, the other biogeochemical variables are linked to PO₄ by a fixed elemental stoichiometry of P:N:C:O₂=1:16:117:-170 [Anderson and Sarmiento, 1994]. The exception is O₂ in near-anoxic waters. Below a threshold of 4 µmol/kg, oxygen consumption halts in the model, as observations and laboratory experiments show that aerobic respiration becomes severely limited at such levels. Remineralization of organic matter to produce PO₄ and DIC continues in the model nevertheless, as nitrate is assumed to take over oxygen as oxidizing agent in the process of denitrification.

In the BIOTIC simulation, the coupled physical-biogeochemical model is first run out with a fixed preindustrial atmospheric CO₂ concentration of 278 ppm until a steady state is reached. Using the equilibrium state as the preindustrial state representing the year 1765, the model is then run forward until year 2000, using the atmospheric time history of CO₂ as a surface boundary condition. Other physical boundary conditions remain unchanged. The difference in the DIC content between this anthropogenic CO₂ run and a control run with constant atmospheric CO₂ directly gives the anthropogenic CO₂ uptake by the model ocean.

In a separate run, two chlorofluorocarbon tracers, CFC-11 and CFC-12, are simulated following the
OCMIP-2 CFC simulation protocols [Orr et al., 1999]. The uptake of CFCs is simulated starting from an initial condition of zero concentration everywhere in the year 1940 and using their atmospheric time histories as surface boundary conditions up to the year 2000.

In both the BIOTIC and CFC historical simulations, the model ocean absorbs the anthropogenic gases, as the increase in their atmospheric concentrations maintains higher partial pressures in the atmosphere relative to the ocean surface. The absorption is strictly abiotic, because there is no biological feedback on the model ocean by these gases.

We also used the same physical model to simulate the distribution of dye tracers as part of our effort to evaluate the OMP analysis as described in section 5.1 of the printed publication. Dye tracers were released from 36 surface regions defined by Gloor et al. [2001]. Because they defined the regions for their purpose of inverting oceanic tracers for air-sea fluxes, the regional breakdown (14 in the Atlantic; 7 in the Indian; and 15 in the Pacific) is not ideal for our application here. The dye tracer distributions were obtained by running the model forward for 3000 years. During this time, the 36 tracers have their concentrations restored to 1 in the surface region that they represent but set to 0 in all other surface regions. In the interior ocean, the tracers are conservative, and so their concentrations can only change by mixing and provide exact fractional contributions of all 36 source water regions.

2. A critical review of the $\Delta C^*$ method and its application to our synthetic data

The measured DIC concentration $DIC_m(t_{obs})$ of a parcel of seawater, collected at time $t=t_{obs}$ from the ocean interior, that has been in contact with the atmosphere since the beginning of the Industrial Revolution (~1765) can contain as much as a few percent of anthropogenic signal $C_{ant}(t_{obs})$. A larger component of $DIC_m(t_{obs})$ is the biological contribution $\Delta DIC_{bio}$ or the amount of DIC added by organic carbon remineralization and CaCO$_3$ dissolution. By far, the largest component would be the preformed, preindustrial DIC concentration $DIC^0(t_0)$, which is what the parcel had when it was last at the surface of the ocean in preindustrial times. The basic premise of the $\Delta C^*$ method is that $\Delta DIC_{bio}$ and $DIC^0(t_0)$ can be reliably estimated with data, and accounting for these components in $DIC_m(t_{obs})$ would give us the anthropogenic CO$_2$ content $C_{ant}$:

$$C_{ant}(t_{obs}) = DIC_m(t_{obs}) - \Delta DIC_{bio} - DIC^0(t_0), \quad (1)$$

The first step in the $\Delta C^*$ method is to estimate the contribution of biology $\Delta DIC_{bio}$. This can be thought of as retracing an interior water parcel’s DIC history back to the location where the parcel was last in contact with the atmosphere. In Figure 1 of our printed paper, water parcel A represents $DIC_m(t_{obs})$. Making the biological correction $\Delta DIC_{bio}$ takes us to point B. The biological correction is given by:

$$\Delta DIC_{bio} = r_{C:O_2}(O_2 - O_2^{sat}) + \frac{1}{2}[Alk - Alk^0 + r_{N:O_2}(O_2 - O_2^{sat})], \quad (2)$$
where $O_2^{sat}$ is the oxygen saturation concentration, and $r_{C:O_2}$ and $r_{N:O_2}$ are the elemental stoichiometry that relates oxygen and nitrate to carbon. The first term on the right hand side is the DIC correction due to the remineralization of organic matter. The second term enclosed by 1/2 is the alkalinity contribution to DIC and contains the $CaCO_3$ and nitrate contributions. In regions where water column denitrification is known to occur, Equation 2 has been extended by a term involving $N^* = NO_3^- - 16 PO_4^{3-} + constant$ [Gruber and Sarmiento, 1997]. As explained by Sabine et al. [1999], this accounts for the fact that denitrification remineralizes organic matter and thus increasing DIC without consuming dissolved oxygen. Since denitrification consumes nitrate, while producing phosphate, denitrification has a strong impact on $N^*$, which can be used to diagnose how much DIC has been added to the water column from this process.

Since we do not have nitrogen as part of our synthetic data (macronutrient in our model is phosphate), we use preformed $PO_4$ ($PO_4^0$) instead of $N^*$ in near-anoxic waters (<4 $\mu$mol-O$_2$/kg). In our analysis, the first sign of denitrification (i.e., $\Delta C^*$ anomaly) becomes evident not at the 4 $\mu$mol-O$_2$/kg level, but already at much higher levels, perhaps 100 $\mu$mol/kg. This is an indication that the anomaly that originates in <4$\mu$mol/kg O$_2$ waters mixes with higher-oxygen waters. We can use the difference between in situ $PO_4$ and $PO_4^0$ concentrations to estimate the full soft-tissue component of $\Delta DIC_{bio}$, because the interior ocean cycling of $PO_4$ and DIC is tightly linked by the model’s elemental stoichiometry of 1:117.

We estimate $PO_4^0$ concentration using a multiple linear regression (MLR) as a function of potential temperature and salinity (see Section 3 below). The biological correction also requires an estimate of preformed alkalinity $ALK^0$. Following Gruber et al. [1996], we estimate $ALK^0$ for our synthetic data by an MLR model as a function of potential temperature, salinity, and a quasi-conservative tracer PO ($PO=O_2+170*PO_4$; see Section 3).

The second and by far the more difficult step in the $\Delta C^*$ method is to estimate the preindustrial, preformed concentration $DIC^0(t_0)$ (the lower horizontal dashed line, Figure 1 of printed paper). Because there is no simple way to estimate $DIC^0(t_0)$, we do in parts. First, we first calculate the expected DIC concentration that would be in equilibrium with the preindustrial atmospheric pCO$_2$, $DIC_{eq}(t_0)$, given surface ocean properties temperature, $T$, salinity, $S$, and preformed alkalinity, $ALK^0$. We then estimate the air-sea CO$_2$ disequilibrium $\Delta C_{diseq}$, which represents the difference between preformed and equilibrium DIC concentrations. The equilibrium and disequilibrium components give what we want (i.e., $DIC^0(t_0)$):

$$DIC^0(t_0) = DIC_{eq}(t_0) + \Delta C_{diseq}(t_0),$$

(3)

For the estimation of the equilibrium DIC concentrations, we linearized our equation using $T$, $S$, and $ALK^0$ following Gruber et al. [1996] ($r^2=0.98$; see Appendix). The MLR models for $DIC_{eq}(t_0)$ and $DIC_{eq}(t)$ (i.e., $DIC_{eq}$ at different times) have the same functional form but with pCO$_2$ dependence calculated for the coefficients of $T$, $S$, and $ALK$ (see Section 3).

To the extent that there is mixing in the ocean, $\Delta C_{diseq}(t_0)$ is a mixture of different disequilibria from different endmember source waters [Gruber, 1998]. Inserting (3) into (1) yields:
\[ C_{ant}(t_{obs}) = DIC_{m}(t_{obs}) - \Delta DIC_{bio} - DIC_{eq}(t) - \Delta C_{diseq}(t_0), \]

\[ = \Delta C^* - \Delta C_{diseq}(t_0), \]  \hspace{1cm} (4)

where \( \Delta C^* \) now represents a quasi-conservative quantity that can be estimated with data [Gruber et al., 1996]. In Figure 1 of our printed paper, \( \Delta C^* \) is the difference between point B and \( DIC_{eq}(t_0) \), indicated by the upper horizontal dashed line. The preindustrial air-sea CO\(_2\) disequilibrium is shown by the distance between the two horizontal dashed lines.

The crux of the \( \Delta C^* \) method is thus estimating the preindustrial air-sea disequilibrium \( \Delta C_{diseq}(t_0) \), and the method uses two mutually exclusive ways to make these estimates. Both assume (1) that ocean mixing occurs primarily along surfaces of equal density, and (2) that the natural carbon cycle in the ocean has remained in steady-state, i.e. that this disequilibrium has not changed over time in response to changes in ocean circulation and/or ocean biology. The first way is applicable to relatively deep isopycnal surfaces, parts of which can be assumed to contain no anthropogenic CO\(_2\). On such surfaces, Equation 4 indicates that the average \( \Delta C^* \) in those parts without anthropogenic carbon represents an average preindustrial disequilibrium:

\[ \Delta C_{diseq}(t_0) = \Delta C^*, \text{ if } C_{ant} = 0 \]  \hspace{1cm} (5)

The second way is applicable to shallower isopycnal surfaces that likely contain anthropogenic CO\(_2\) everywhere on the isopycnals. If we had a reliable estimate of the age of the water parcel \( \tau = t_{obs} - t \) (i.e., the time elapsed since leaving the surface), we would be able to estimate the preformed DIC concentration (point C in Figure 1 of printed paper). Furthermore, we can calculate its surface ocean DIC concentration in equilibrium with atmospheric pCO\(_2\) at times in the past (\( DIC_{eq}(t) \), dot-dashed curve in Figure 1 of printed paper). The difference between the estimated preformed DIC concentration and \( DIC_{eq}(t) \) is the estimate of the air-sea disequilibrium \( \Delta C_{diseq}(t) \). This defines a new quantity \( \Delta C^*_t \) that is identical to \( \Delta C^* \), except that the equilibrium term is referenced to some postindustrial pCO\(_2\): \( t \) instead of the preindustrial pCO\(_2\):

\[ \Delta C^*_t = DIC_{m}(t_{obs}) - DIC_{eq}(t) - \Delta DIC_{bio}, \text{ if } \tau = t_{obs} - t \text{ is known} \]  \hspace{1cm} (6)

This quantity therefore represents air-sea disequilibrium at time \( t \) in the second method. A critical assumption is that the air-sea disequilibrium does not change with time:

\[ \Delta C_{diseq}(t_0) = \Delta C_{diseq}(t) = \text{const} \]  \hspace{1cm} (7)
For the “mid-depth” class of isopycnals (see Sections 3 and 4.1 of the printed paper), we make use of the OMP method to obtain the different water mass mixing ratios. The “observations” that we use to constrain the inverse problem include potential temperature ($\theta$), $S$, $O_2$, $PO_4$, and $\Delta C^*$, directly from the model or calculated. The number of water types that can be solved for depends on the availability of independent observations and how well they constrain the solution. It also depends on the identification and definition of the source water types, a key step in the OMP analysis that is primarily based on $\theta$ and $S$. We note that the OMP analysis is an exercise in inversion that is not strongly constrained by “data”. The choice of weights is somewhat arbitrary, except for the general practice of preferentially weighting the mass conservation equation [Lee et al., 2003; Sabine et al., 2002; Sabine et al., 1999].

3. Multiple linear regression models of $ALK^0$, $DIC_{eq}$, and $PO_4^0$

For our model-derived synthetic data, we developed multiple linear regression (MLR) models of $ALK^0$, DIC in equilibrium with the pre and post industrial atmosphere (i.e., $DIC_{eq}(t_0)$ and $DIC_{eq}(t)$), and preformed PO$_4$ ($PO_4^0$). Only the near surface values of ALK, preindustrial DIC, and PO$_4$ from the three major basins were used; values from the Arctic and marginal seas were removed from the analysis.

A global $ALK^0$ fit is made around near-surface mean values of alkalinity (2271.8 µmol-eq/kg), temperature (14.26 $^\circ$C), salinity (34.77 permil), and the quasi-conservative tracer PO (381.77 µmol-PO/kg):

$$ALK^0 = 2271.8 + (1.636 \pm 0.004) \cdot (T - 14.26) + (70.365 \pm 0.016) \cdot (S - 34.77) + (0.312 \pm 0.000) \cdot (PO - 381.77)$$ (MLR1)

The fit has an $r^2$ of 0.97 (Figure 1, 2004GB002397-figure1.eps). Figure 2 (2004GB002397-figure2.eps) shows the spatial distribution of the MLR-predicted $ALK^0$ error.

A global $DIC_{eq}(t_0)$ fit is made around near-surface mean values of preindustrial DIC$_{eq}$ (1992.6 µmol-C/kg), temperature (14.26 $^\circ$C), salinity (34.77 permil), and alkalinity (2271.8 µmol-eq/kg):

$$DIC_{eq}(t_0) = 1992.6 - (8.208 \pm 0.002) \cdot (T - 14.26) - (6.122 \pm 0.053) \cdot (S - 34.77) + (0.8489 \pm 0.001) \cdot (ALK - 2271.8)$$ (MLR2)

The $r^2$ value for this regression is 0.98 (Figure 3, 2004GB002397-figure3.eps).

For the MLR model for $DIC_{eq}(t)$ (i.e., $DIC_{eq}$ at different pCO$_2$), we calculated the pCO$_2$ dependence of the coefficients of $T$, $S$, and $ALK$ in Equation MLR2. The dependence for the range of pCO$_2$ in the postindustrial times is +0.0081/pCO$_2$ for temperature, +0.0178 for salinity, and +1.328e-4 for ALK.

Finally, for $PO_4^0$, we developed MLR models for each basin, because the global fit was not very good ($r^2=0.72$). Figure 4 (2004GB002397-figure4.eps) shows the Pacific $PO_4^0$ fit, which is made around near-surface mean values of PO$_4$ (0.731 µmol-P/kg), temperature (15.87 $^\circ$C), and salinity (34.62 permil):
\[ P_{04}^I = 0.731 - (0.053 \pm 0.002) \cdot (T-15.87) + (0.065 \pm 0.030) \cdot (S-34.62) \]  

(MLR3)

The \( r^2 \) value for the Pacific fit is 0.83. We do not show the MLR models for the Indian and Atlantic, but their \( r^2 \) values are respectively 0.93 and 0.61. Although the 1 standard deviation of the MLR model is only 0.21 \( \mu \text{mol/kg} \), its effect on DIC is much larger as it needs to be multiplied by the \( r_{PC}=1:117 \).

4. Detailed discussion of the error in the biological correction

This section gives more details to section 2.2 of the printed paper. The reason that the biological correction (Equation 2, printed paper) may be a possible source of errors is that is relies on a number of assumptions: (1) stoichiometric ratios of remineralization of organic matter are constant and known; (2) the preformed concentration of oxygen is correctly estimated by its saturation concentration; and (3) preformed alkalinity concentration \( Alk^0 \) has not changed over time and can be reliably estimated on the basis of a multiple linear regression model.

Gruber [1998] and Sabine et al. [1999] investigated the biases resulting from the violation of the first assumption, particularly the error arising from the use of non-standard carbon to oxygen ratios. They showed that a constant offset in stoichiometry leads to approximately inversely proportional changes in the estimated anthropogenic CO\(_2\), with the magnitude of the changes scaling with the apparent oxygen utilization. For example, when they used a carbon to oxygen ratio that was ~13% lower than their standard ratio of \( r_{C:O2}=0.688 \) (i.e., 117:-170 [Anderson and Sarmiento, 1994]), the anthropogenic carbon inventory increased by about 20% in the Atlantic and about 18% in the Indian. When they used a ratio that was ~13% higher than their standard ratio (i.e., approaching the classical ratio of 106:-138 [Redfield et al., 1963]), the inventory decreased roughly by 15% in the Indian but by only 8% in the Atlantic.

These changes in reconstructed anthropogenic CO\(_2\) are smaller than expected simply by propagating the resulting changes in \( \Delta DIC_{bio} \) onto the estimated \( C_{ant} \). For example, for an AOU of 150 \( \mu \text{mol/kg} \), which is representative for mid thermocline waters, a 13% change in the C:O\(_2\) stoichiometry leads to a change in the estimated \( \Delta DIC_{bio} \) of about 14 \( \mu \text{mol/kg} \), which is of similar magnitude as \( C_{ant} \) in such mid-depth waters. Therefore, if such a bias in \( \Delta DIC_{bio} \) had a proportional impact on the estimation of \( C_{ant} \), the resulting error in \( C_{ant} \) would be of order 1. The reason that this is not the case is because a compensation occurs in the \( \Delta C^* \) method, in which errors introduced in the \( \Delta DIC_{bio} \) terms are offset by the disequilibrium term. This is a result of the disequilibrium being determined from either \( \Delta C^* \) or \( \Delta C_{diseq}^* \), which in the case of a biased estimate of \( \Delta DIC_{bio} \) is biased proportionally. Since anthropogenic CO\(_2\) is ultimately determined by the difference between \( \Delta C^* \) and \( \Delta C_{diseq} \) (Equation 4), errors in \( \Delta DIC_{bio} \) and hence \( \Delta C^* \) or \( \Delta C_{diseq}^* \) tend to cancel out with the errors in \( \Delta C_{diseq} \). In the case of biases in the stoichiometric ratio, this cancellation is only partial because the bias depends on the magnitude of AOU, whereas the cancellation term is constant for a particular water mass on a particular isopycnal surface. As we will see next, the cancellation is nearly complete in the case of errors in preformed oxygen.
With regard to the second assumption, it is well known that winter time oxygen concentrations in high latitudes are often not at saturation. Since preformed concentrations are primarily set by the winter time concentrations in the outcrop regions, any deviation of the preformed oxygen concentration from the saturation concentration is transported into the interior and leads to an error in $\Delta DIC_{\text{bio}}$ according to Equation 8. Ito et al. [2004] recently suggested that the deviation can be more than 50 $\mu$mol kg$^{-1}$ and thus can cause errors in $\Delta DIC_{\text{bio}}$ as large as of 30 $\mu$mol kg$^{-1}$.

Fortunately, as long as this error does not change with time for particular water mass, it cancels out when anthropogenic CO$_2$ is determined. This can be shown easily if, as in Equation 4, $\Delta C^*$ is used to determine $\Delta C_{\text{diseq}}$. Let us say there is a bias in $\Delta DIC_{\text{bio}}$ of magnitude $\Delta\Delta DIC_{\text{bio}}$ in a particular water mass, so that $\Delta DIC'_{\text{bio}} = \Delta DIC_{\text{bio}} + \Delta\Delta DIC_{\text{bio}}$. The prime (') indicates the apparent values. The air-sea CO$_2$ disequilibrium determined for the same water mass using Equations 3 and 4 is then also offset by $\Delta\Delta DIC_{\text{bio}}$ so that $\Delta C'_{\text{diseq}}(t_0) = \Delta C_{\text{diseq}}(t_0) - \Delta\Delta DIC_{\text{bio}}$. Using this biased disequilibrium in (3) leads to no change in the estimated $C_{\text{ant}}$ because the $\Delta\Delta DIC_{\text{bio}}$ terms in $\Delta C'_{\text{diseq}}(t_0)$ and $\Delta DIC'_{\text{bio}}$ cancel out:

$$C_{\text{ant}}(t_{\text{obs}}) = DIC_{\text{m}}(t_{\text{obs}}) - \Delta DIC^*_{\text{bio}} - DIC_{\text{eq}}(t_0) - \Delta C'_{\text{diseq}}(t_0),$$

$$= \Delta C^* - \Delta C_{\text{diseq}}(t_0),$$

(8)

Although it turns out that the deviation of the preformed oxygen from O$_2$ saturation has only a small influence on the estimated anthropogenic CO$_2$, the resulting error in the estimated $\Delta C_{\text{diseq}}$ term would diminish the physical meaning of the disequilibrium term. This would negate a strength of the $\Delta C^*$ method, which is its ability to check each step or term against expected and observed properties.

Biases stemming from the third assumption regarding Alk$^0$ are likely small as well, because there is little evidence that would suggest that preformed alkalinity has changed substantially over time. Although the saturation horizon for aragonite has been affected by the invasion of anthropogenic CO$_2$ [Feely et al., 2004], the reported changes are small. Our estimate of $\Delta DIC_{\text{bio}}$ would only be affected, if the changes were transported from the deeper parts of the thermocline to the surface and alter preformed alkalinity. The only error of concern associated with the estimation of preformed alkalinity is if the linear regression model generally used to estimate preformed alkalinity has biases in certain regions. One possibility that we will later discuss is that most studies used “PO” [Broecker and Peng, 1982] as a conservative tracer in their Alk$^0$ regression models. In regions influenced by denitrification, PO loses its conservative behavior, resulting in a bias in the estimated Alk$^0$. This bias will remain local, however.

In summary, biases associated with the biological correction can be large but have ultimately only a moderate impact on the estimated anthropogenic CO$_2$. This is primarily because biases in $\Delta DIC_{\text{bio}}$ are compensated by their impact on $\Delta C_{\text{diseq}}$. This conclusion is very different from that drawn from an analysis of the random errors, for which there is no equivalent compensation, so that their uncertainties are propagated onto $C_{\text{ant}}$ [Gruber et al., 1996].

5. Illustration and evaluation of the OMP analysis
This section provides more details to sections 4.1 and 5.1 of the printed paper as they relate to the illustration and evaluation of the OMP analysis. In section 4.1, we illustrated the application of the $\Delta C^*$ method to a “mid-depth” isopycnal $\sigma_\theta=26.8$ in the Pacific. Analyses of the T-S properties on this isopycnal surface show that there are four water types or endmembers on this isopycnal surface. The first is the northwest water type (Type 1: NW), which is the freshest and coldest. The second and third water types are the warm and salty waters found in the eastern equatorial region (Type 2; EE) and in the southern hemisphere mid-latitude region adjacent to Australia (Type 3; SW). The fourth water type is the southern outcrop water (Type 4: SO), with intermediate values of temperature and salinity. The fractional contributions of these four water types, as determined by the OMP method are shown in Figure 5 (2004GB002397-figure5.eps). Our analysis shows that water types NW and SO are confined to their respective hemispheres, while water types EE and SW are found in significant fractions in both hemispheres. The highest fractions of each water type are located where we expect them (e.g., nearly 100% of NW and SO is in the outcrop regions).

In section 5.1 of our printed paper, we examined the uncertainty in the OMP analysis in two ways. Our analysis was again for the Pacific isopycnal $\sigma_\theta=26.8$. In one way, we compared the “true” and OMP-predicted $\Delta C^*$ for the preindustrial period, when there were no confounding factors such as anthropogenic carbon and the temporal change in air-sea CO$_2$ disequilibrium. Assuming that the biological correction $\Delta DIC_{bio}$ is accurate, the preindustrial $\Delta C^*$ is a truly conservative tracer. Therefore, if our OMP results are correct, we should be able to reconstruct the preindustrial $\Delta C^*$ distribution using the water type mixing ratios determined by the OMP analysis and their endmember values chosen in exactly the same manner as $\Delta C^*$, to quantify the air-sea CO$_2$ disequilibrium. Figure 6 (2004GB002397-figure6.eps) shows the OMP-derived preindustrial $\Delta C^*$. The derived quantity (Figure 6a) is generally within a few $\mu$mol-C/kg of the target distribution (Figure 6b). An exception is in the northwest Pacific, where there is a strong gradient in the fraction of water type NW (Figure 5a), while the true preindustrial $\Delta C^*$ does not have such a gradient. The error there is as large as 10 $\mu$mol-C/kg (Figure 6c).

In the second way, we used the dye tracers to can calculate $\Delta C_{diseq}$ on the same Pacific isopycnal $\sigma_\theta=26.8$ by a linear combination of the dye (i.e., water mass) fractions and the mean surface $\Delta C^*$, at each of the 36 dye-release regions. The accuracy of the calculated disequilibrium depends on the extent to which: (1) the disequilibrium is constant with time; (2) $\Delta C^*$ truly reflects the disequilibrium; and (3) the dye-release regions are able to resolve important outcrop regions. In the printed paper, we have detailed discussion on the first two conditions, which are identical whether we used the OMP-derived water type fractions or the dye tracer fractions. The success of using the dye tracers to quantify $\Delta C_{diseq}$ therefore rests largely on the third condition. For the Pacific $\sigma_\theta=26.8$, for example, the basin-wide agreement between the “true” and the estimated $C_{air}$ using the dye tracers appears somewhat better than that using the OMP analysis solutions shown in Figure 6b. Locally however, the agreement is much worse in the far north using the dye tracers, because the northwest outcrop region is not resolved by the predefined dye-release regions. The agreement would presumably be improved, if the dye-release regions had finer spatial
An interesting finding of the dye tracer distribution is that waters ventilated in one basin leak out in significant amounts to other basins (Figure 7, 2004GB002397-figure7.eps). For example, the export of Pacific waters into the Indian Ocean via the Indonesian throughflow can be so large that less than 50% of the water in some parts of the \( \sigma_\theta=26.8 \) isopycnal in the Indian Ocean has been ventilated in the same basin. This implies for the \( \Delta C^* \) method that one cannot fully characterize the disequilibrium on an isopycnal surface in any one basin using ocean data from that basin alone (Table 1). This problem is especially pronounced in that the MLR model for \( ALK^0 \) in practice has been estimated independently for each basin and for different sets of data [Lee et al., 2003; Sabine et al., 2002; Sabine et al., 1999]. A remedy would be to use a global MLR model for \( ALK^0 \), as done originally [Gruber, 1998; Gruber et al., 1996] but at the cost of a poorer model fit.

6. Correcting for a change in the air-sea CO\(_2\) disequilibrium

Here we elaborate on the “first improvement” of the \( \Delta C^* \) method as presented in section 5.2 of the printed paper, where we corrected for the change in the air-sea CO\(_2\) disequilibrium over time, \( \Delta\Delta C_{\text{diseq}}(t) \). The correction is applicable to upper ocean where \( \Delta C^*_t \) is used to estimate \( \Delta C_{\text{diseq}}(t) \) (Equation 5), and we suggest a correction to \( \Delta C_{\text{diseq}}(t) \) and hence \( C_{\text{ant}} \) that is approximately proportional to the concentration of the estimated (i.e., uncorrected) \( C_{\text{ant}} \). This comes from our expectation as discussed in our printed paper that the positive bias increases with time and scales roughly with the estimated \( C_{\text{ant}} \):

\[
C_{\text{ant}}^{\text{corr}}(t) = \Delta C^* - \Delta C_{\text{diseq}}(t) - \Delta\Delta C_{\text{diseq}}(t)
\]

\[
\Delta\Delta C_{\text{diseq}}(t) \approx a \cdot C_{\text{ant}}^{\text{uncorr}}(t)
\]

We show that a reasonable first estimate of the correction factor \( a \) can be determined from first principles and data. In order to illustrate the idea, let us assume for the moment that \( \Delta\Delta C_{\text{diseq}}(t) \) is globally uniform, which is not strictly correct, but later we will relax this assumption to account for regionally dependent air-sea disequilibria.

We start our derivation with Equation 4 in the printed paper, which states that the oceanic uptake flux of anthropogenic CO\(_2\) is proportional to the air-sea disequilibrium:

\[
F_{\text{ant}}(t) = k \cdot \gamma_{\text{DIC}} \frac{\Delta \Delta C_{\text{diseq}}(t)}{\alpha},
\]

If we had an independent estimate of the anthropogenic flux, we could solve Equation 4 for \( \Delta\Delta C_{\text{diseq}}(t) \). A very convenient scaling estimate of \( F_{\text{ant}}(t) \) was introduced by Gloor et al. [2003] and later confirmed by Mikaloff-Fletcher et al. [2004] on the basis of forward model simulations. The scaling
argument proposes that over the industrial period, the global mean flux of anthropogenic CO₂ into the ocean at any given time was proportional to the atmospheric $pCO_2$ perturbation, thus:

$$F_{ant}(t) = \beta \cdot pCO_{2,atm}^{ant}(t) \quad (11)$$

Combining Equations 10 and 11 and solving for $\Delta\Delta C_{diseq}(t)$ gives:

$$\Delta\Delta C_{diseq}(t) = \frac{\alpha \cdot \beta}{k \cdot \gamma_{DIC}} pCO_{2,atm}^{ant}(t) \quad (12)$$

We can determine the magnitude of the scaling factor $\beta$ from the integral constraint that the globally and temporally integrated flux must equal the total inventory of anthropogenic CO₂ in the ocean:

$$\int_{t_0}^{t} F_{ant}(t) dt dA = \int C_{ant}(t) dV \quad (13)$$

Solving for $\beta$ by combining Equations 11 and 16 and assuming that the flux $F_{ant}(t)$ is spatially uniform gives:

$$\beta = \frac{\int C_{ant}(t) dV}{A_{oc} \cdot \int_{t_0}^{t} pCO_{2,atm}^{ant}(t) dt} \quad (14)$$

The numerator of Equation 14 for $t=1994$ may be represented by the global $C_{ant}$ inventory of 118±19 Pg C, or (9.83 ±1.58)·10¹⁵ mol, as estimated by Sabine et al. [2004]. We can obtain the denominator by integrating the atmospheric $pCO_2$ perturbation history up to the same year. The integration gives 4535 ppm yr, with a maximum uncertainty of a few 100 ppm yr. $A_{oc}$ is the ice-free area of the surface ocean, which is about 332·10¹² m². The value of $\beta$ calculated with these quantities is 0.0065±0.0012 mol m⁻² yr⁻¹ µatm⁻¹.

Next we would like to replace $pCO_{2,atm}^{ant}(t)$ on the right hand side of Equation 12 with $C_{ant}$, and transform Equation 12 into Equation 9, so that we can determine the factor $a$. Since the amount of $C_{ant}$ in surface waters is often very close to that predicted by assuming that the carbonate system has fully equilibrated with the atmospheric CO₂ perturbation, we can estimate this surface concentration from the buffer factor and $pCO_{2,atm}^{ant}(t)$ [Sarmiento and Gruber, in press]:

$$C_{ant}(t) \approx \frac{\alpha}{\gamma_{DIC}} pCO_{2,atm}^{ant}(t) \quad (15)$$
Solving Equation 15 for $pCO_{2\text{atm}}^\text{ant}(t)$ and inserting the result into Equation 12 gives the final estimate of the magnitude of the temporal change of the air-sea disequilibrium:

$$\Delta\Delta C_{\text{diseq}}(t) = \frac{\beta}{k_{\text{ex}}} C_{\text{ant}}(t)$$

(16)

The globally averaged $k_{\text{ex}}$ for CO$_2$ as determined by Broecker et al. [1985] is about 0.065±0.015 mol m$^{-2}$ yr$^{-1}$ µatm$^{-1}$. Together with our estimated $\beta$ above, this gives a value of for the factor $a$:

$$a = \frac{\beta}{k_{\text{ex}}} = \frac{0.0065 \pm 0.0012}{0.065 \pm 0.015} = 0.10 \pm 0.03$$

(17)

Our analysis thus suggests that we need to apply a roughly 10% downward correction to all $C_{\text{ant}}$ estimates that have been made using $C^\text{*}$, to quantify the air-sea disequilibrium. This estimate of $a$ is accurate as long as Equation 11 holds, that is, the flux of anthropogenic CO$_2$ is proportional to the atmospheric CO$_2$ perturbation. Theoretically this should be the case if atmospheric CO$_2$ rises exponentially, which is mostly true throughout the 20$^{th}$ century. We can then therefore use Equation 11 for the data taken during the 20$^{th}$ century, but this assumption may not hold in the future, particularly if atmospheric CO$_2$ were to follow a stabilization trajectory.

Our global estimate of $a$ (Equation 17) can be tailored to different regions by applying the integral constraint (Equation 14) to individual isopycnal volumes and obtaining regional estimates of $\beta$. By combining these regional estimates of $\beta$ with an estimate of the air-sea gas exchange coefficient, $k_{\text{ex}}$ that is representative for the outcrop region of this isopycnal volume, one would obtain a correction factor $a$ that is unique for each isopycnal volume considered.

7. Sources of error in the $\Delta C^*$ method related to denitrification

We did not have sufficient space in our printed paper to discuss errors in the $\Delta C^*$ method as it relates to denitrification. First, there is likely to be some error in using $PO$ to predict $ALK^0$ and hence the equilibrium DIC concentrations. Denitrification increases PO$_4$ without decreasing O$_2$, and so PO would be biased high. Since the coefficient of PO in the $ALK^0$ MLR model is positive, $ALK^0$ would also be biased high. Its final effect on $C_{\text{ant}}$ would be a positive bias, because the $ALK^0$ coefficient in the MLR model for $DIC_{eq}(t)$ is higher than that for $DIC_{eq}(t_0)$. This source of error has not been identified before but exists in all previous applications of the $\Delta C^*$ method to real ocean data [Gruber, 1998; Lee et al., 2003; Sabine et al., 2002; Sabine et al., 1999], because they all use PO or its variant NO [Lee et al., 2003] to estimate $ALK^0$. However, the magnitude of the positive bias would be very small, because the difference in the $ALK^0$ coefficients in the MLR models for $DIC_{eq}(t)$ and $DIC_{eq}(t_0)$ is about 1%. Second, we can expect some
error from our use of \( PO_4^0 \) in accounting for denitrification in the soft tissue component of \( \Delta DIC_{bio} \) because the MLR model to predict \( PO_4^0 \) is not highly precise (see section 3 above). However, we do not expect from this a bias in the final \( C_{ant} \) estimation. Both errors would be limited regionally, because the denitrification correction is only necessary in waters with low \( O_2 \) content and that show signs of anomalous \( \Delta C^* \).

8. Volume census to make Table 2 of printed paper

To extrapolate our analysis of synthetic data on nine select isopycnal surfaces, we made a volume census of the model ocean, dividing each basin into three volumes corresponding to the three isopycnal classes. In each basin, we identified the deepest isopycnal surface that is free of \( \rho \)CFC ages 30 years and older; this surface serves as the boundary between the “shallow” and “mid-depth” isopycnals. Next we defined the boundary between “mid-depth” and “deep” isopycnals as the deepest isopycnal that is free of \( \rho \)CFC ages 50 years and older. The “mid-depth” and “deep” isopycnals thus can contain waters with ages younger than 30 years, as they should near outcrop regions. The volumes bounded by the identified isopycnals are then used to derive global biases for each isopycnal classes and for the whole ocean.
References


Mignone, B.K., J. Sarmiento, R. Slater, and A. Gnanadesikan, Sensitivity of sequestration efficiency to mixing

Mikaloff-Fletcher, S.E., N. Gruber, A. Jacobson, S. Doney, S. Dutkiewicz, M. Follows, K. Lindsay, D. Menemenlis, and A. Mouchet, Robust estimates of anthropogenic carbon uptake, transport, and storage by the ocean, 2004.


Auxiliary Material Figure 1

A scatter plot showing the relationship between Simulated Preindustrial ALK (umol/kg) and ALK from MLR (umol/kg). The data points are closely aligned along a diagonal line, indicating a strong positive correlation between the two variables.
Auxiliary Material Figure 3

DIC in EQ with 278 ppm (umol/kg)

DIC from MLR (umol/kg)

linearized around mean
Auxiliary Material Figure 4
Auxiliary Material Figure 5

(a) [Map Image]

(b) [Map Image]

(c) [Map Image]

(d) [Map Image]