Marine Chemistry 132–133 (2012) 44–55

Contents lists available at SciVerse ScienceDirect

Marine Chemistry

journal homepage: www.elsevier.com/locate/marchem

Regression-based estimates of the rate of accumulation of anthropogenic CO_2 in the ocean: A fresh look

William Carlisle Thacker *

Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL 33149, USA
Atlantic Oceanographic and Meteorological Laboratory, National Oceanic and Atmospheric Administration, 4301 Rickenbacker Causeway, Miami, FL 33149, USA

ARTICLE INFO

Article history:
Received 28 March 2011
Received in revised form 30 December 2011
Accepted 17 February 2012
Available online 23 February 2012

Keywords:
Carbon dioxide
Regression
MLR
eMLR

ABSTRACT

Regression-based methods used for estimating the rate of increase of anthropogenic CO_2 in the ocean are reviewed and guidelines for improvement are presented. Following these guidelines leads to a local two-regression method, the first regression accounting for changes in oceanic carbon due to natural variability and the second regression associating the remaining systematic temporal variability with the anthropogenic signal to quantify the rate of accumulation. While a formal measure of the accumulation rate's uncertainty is provided by the standard error of the second regression's slope parameter, both the available data's limited ability to characterize carbon's natural variability in the absence of any anthropogenic contribution and the choice of regressors to account for that variability present significant uncertainties that are less easily quantified. An attractive feature of the method is its applicability to data other than those from repeated hydrographic surveys, such as might be provided by appropriately instrumented profiling floats.

1. Introduction

Since first suggested by Wallace (1995) regression-based methods have been widely used for estimating the rate at which anthropogenic CO_2 has been accumulating in the ocean. The data on which these estimates are based come from transoceanic hydrographic surveys which have been repeated at roughly decadal intervals. Simply subtracting corresponding values of carbon observed by the repeated surveys is not sufficient for estimating the increase, as oceanic carbon's natural spatial and temporal variability is larger than the expected anthropogenic signal. The role of regression is to exploit empirical relationships between carbon and other observed environmental variables to improve the signal-to-noise ratio. However, as the ways in which regression has been used for estimating anthropogenic CO_2 are somewhat unorthodox and as previous studies have not been sufficiently clear as to what is signal and what is noise, a fresh look is needed. The purpose of this work is to clarify the role of regression and to suggest improvements to the methodology.

Regression-based methods are attractive, as they need neither assumptions characterizing the rate of exchange of CO_2 through the air–sea interface at any spatial location or at any moment in time nor assumptions about its fate once in the ocean. Instead, the consequences of the air–sea transfer and subsequent transport are reflected in the statistics of the data. And without question, carbon exhibits strong empirical relationships to other observed variables (e.g., Redfield, 1934; Redfield et al., 1963). The issue is only whether the regression-based methods adequately account for these relationships.

Given that the central problem is the possibility of confounding anthropogenic increase with natural spatial and temporal variability, it is useful to consider how this variability might ideally be sampled. For example, suppose profiling floats capable of the same sorts of measurements as obtained by the hydrographic surveys were generously deployed to give data sampled frequently in both space and time. Because such an observing system would capture the natural variability far better than the decadal snapshots confined to particular sections, it is useful to consider how regression might be used if such data were available. Clearly, methods based on decadal differences would have to be modified to accommodate sampling more continuous in time and not confined to survey lines. A start toward improved methods might be possible by focusing on how to estimate the rate of increase of anthropogenic CO_2 in the vicinity of the intersection of two repeated surveys, so that there would be data from a minimum of four different years and variability would be sampled not just in a single plane but in two orthogonal planes to give a richer picture of the empirical relationships in the neighborhood of their intersection. This is in fact the strategy pursued here with data from the region centered on the intersection of lines P06 and P16 in the South Pacific from the stations indicated in Fig. 1.

It is also useful to note that carbon's non-anthropogenic variability results from the variability of physical and biogeochemical processes. Their variability is poorly known and hard to quantify, so it is difficult to judge the extent to which it is adequately characterized by the data from the repeated surveys. Even a fleet of profiling floats would fail to capture near-surface diurnal variability, but they might resolve

* Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL 33149, USA. Tel.: +1 305 361 4323; fax: +1 305 361 4392.
E-mail address: carlisle.thacker@noaa.gov.

0304-4203/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.marchem.2012.02.004

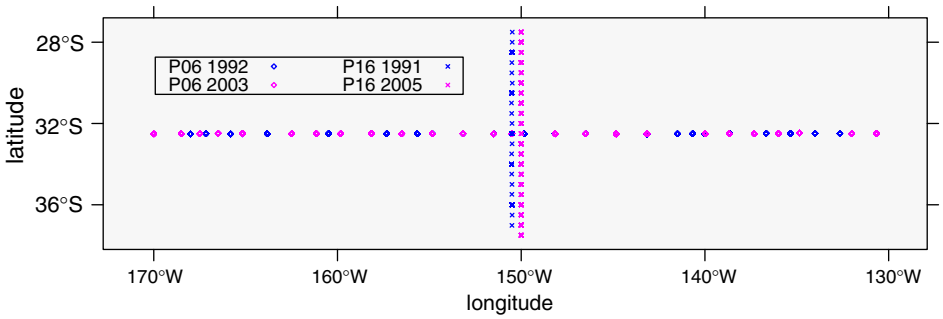

Fig. 1. Locations of stations contributing data to this study.

seasonal variability and variability associated with vertical and horizontal displacements better than the surveys do. Any unsampled variability will not be explained by the regression model and will have the potential of being confounded with the sought anthropogenic signal. In addition, there is the important issue of whether decadal climate changes are altering the empirical relationships between the environmental constituents; this will require a multi-decadal data record to sort out. The focus here is strictly on methodology, but nevertheless the nature of the variability and the adequacy of the sampling should be kept in mind.

2. MLR and eMLR

The method introduced by Wallace (1995) is referred to as the MLR method because it uses multivariate linear regression.[1] Its objective is to compare data from repeated hydrographic surveys, which provide measurements of carbon and concurrent environmental variables. First, a linear regression model is fitted to the earlier data to capture the empirical relationship between carbon and the environmental variables. Then this model is used with data from the later survey to predict what would have been measured if there had been no accumulation of anthropogenic carbon.[2] Difference between the observed and predicted values at each observation point are then interpreted as anthropogenic carbon accumulated during the interval between the surveys.

A flaw in this reasoning can be seen by considering a hypothetical situation where the second survey follows immediately after the first and yields data with exactly the same values as those from the first. By design, there is no possibility of instantaneous accumulation of carbon and none should be inferred from the data. Nevertheless, the MLR method would indicate instantaneous accumulation, positive at some points and negative at others. For this example the differences between observed and predicted values that this method attributes to the accumulation of anthropogenic carbon are simply the residuals of the fit.

Put aside this conceptual difficulty for now, accept that the regression model fitted to the data from the first survey adequately captures the empirical relationships of carbon to other environmental variables, and consider the difference between the second survey's observations and the model's predictions. In addition to the signal of accumulated anthropogenic carbon there is also prediction error, which might be regarded as noise. Rather than trying to estimate the accumulation rate at each point where there is a measurement, a better estimate might be associated with a region containing many measurements. By averaging the differences within the region the noise might be suppressed so that the anthropogenic signal might be better revealed. If the average is over all of the data from the survey, then the problem of hypothetical instantaneous accumulation disappears, because the average of all residuals of the fit is necessarily zero. This suggests that to achieve some degree of localization while separating the anthropogenic signal from noise and simultaneously avoiding the possibility of instantaneous accumulation, the MLR method might attempt to use models fitted to smaller regions.

Now consider the question of whether MLR's model fitted to data from the first survey adequately captures the empirical relationships of carbon to other environmental variables. Are the data from the first survey sufficient to say: "knowing that the other environmental variables are thus and such, then carbon is generally like this"? When the model is fitted to data from widely separated stations and from all depths, as is the usual practice, then the model might know something about how carbon varies from mode water to deep water. However, if the fit would be spatially limited, as suggested above to achieve localization while averaging out noise, then the all of the data in the sample would confirm the general attributes of the water and the regression would seek to refine the details, exploiting regional co-variability to sort out influences from nearby water masses, but the model would only know of spatial co-variability within the fitting region at the time of the first survey. As the second survey provides additional information about the co-variability, fitting the regression model to the combined data from both surveys might be a good idea.

Friis et al. (2005) modified the MLR method, calling their modification the *extended* MLR method, or eMLR for short.[3] This method avoids the problem of instantaneous accumulation by fitting a second model — involving exactly the same environmental variables as the first — to data from the second survey. The difference between the models is regarded as a model for the changes during the interval between the surveys that result from the accumulation of anthropogenic carbon.[4] The difference model is used with the environmental data from the second survey to estimate the changes at each observation point.

Because the eMLR method uses two models, it avoids the problem of possible instantaneous accumulation. However, its use of two models leads to conceptual difficulties relating to the separation of

[1] Among the subsequent studies that have used the MLR method are those of Slansky et al. (1996), Sabine et al. (1999), Ono et al. (2000), McNeil et al. (2001a), McNeil et al. (2001b), Peng et al. (2003), Quay et al. (2007), Levine et al. (2008), Peng and Wanninkhof (2010), and Wanninkhof et al. (2010).

[2] Similarly, Brewer et al. (1995) had suggested using linear regression to predict concentrations of oceanic carbon in the absence of direct observations. They sought better estimates of sound absorption in sea water, which requires improved estimates of pH that might be inferred from the predicted values of carbon. That application is quite different in that it doesn't address the problem of changes in the concentration of carbon.

[3] For examples of studies that have used the eMLR method see those of Olsen et al. (2006), Quay et al. (2007), Sabine et al. (2008), Park et al. (2008), Levine et al. (2008), Brown et al. (2010), Hauck et al. (2010), Peng and Wanninkhof (2010), Wanninkhof et al. (2010), and Goodkin et al. (2011).

[4] This difference model is quite unorthodox, its coefficients being defined by subtracting those of the first model from those of the second.

environmental from anthropogenic variability. Unlike the MLR method, where the model explained environmental variability, leaving the signal together with noise in the residuals, eMLR assumes the signal is entirely in the differences between the two regression models with none in the residuals. In other words, contemporaneous relationships with environmental variables account for all of the anthropogenic carbon together with much of the non-anthropogenic carbon.

Consider a second hypothetical situation. This time the two surveys are well separated in time and there has been substantial accumulation of anthropogenic carbon. However, suppose that carbon exhibits no empirical relationships with any of the environmental variables. Of course this is not the case for actual data, but it might be regarded as the limit of very weak empirical relationships. For this case the model fitted to the data from the earlier survey reduces to the mean of those observed values for carbon. Similarly, the model for the time of the later survey describes carbon as a second mean. Consequently, the eMLR estimate of accumulated anthropogenic carbon is the difference between these two means. This might, at best, be a reasonable estimate for the average accumulation over the entire surveyed region, but it provides no information about the spatial distribution, as eMLR in this case provides exactly the same estimate at all points. To see greater accumulation near the air–sea interface, as might be expected, and less at extreme depths, data in different depth ranges would have to be treated separately.

3. Need for local analyses

While the above discussions of MLR and eMLR both suggest that estimates of accumulation of anthropogenic carbon should be based on more localized analyses, hints at this conclusion can be seen in other studies. For example, Sabine et al. (2008) recognized treating all of the data from repeated trans-Pacific P02 sections along 30°N in a single eMLR analysis to be inappropriate and separated the near-coastal eastern and western data from a central part; nevertheless the central region was still large, extending over 75° of longitude and from surface to bottom. In the same study they also divided the data from the P16 section into three segments, the largest extending from 30°S to 72°S and again over the full depth range. Similarly, Wanninkhof et al. (2010) found that applying the eMLR method within density layers gave noticeably different results than when applying it over the full water column, but their empirical relationships were assumed to be valid for the entire length of the A16 section running from 60°N to 60°S in the Atlantic Ocean; they did not go further and reduce the scope of the empirical relationships both horizontally and vertically.

Plots of hydrographic data provide direct evidence of the need to limit the analysis domain. Consider, for example, the plots in Fig. 2 of data from the relatively limited region surrounding 32.5°S and 150°W but throughout the entire water column: each panel shows

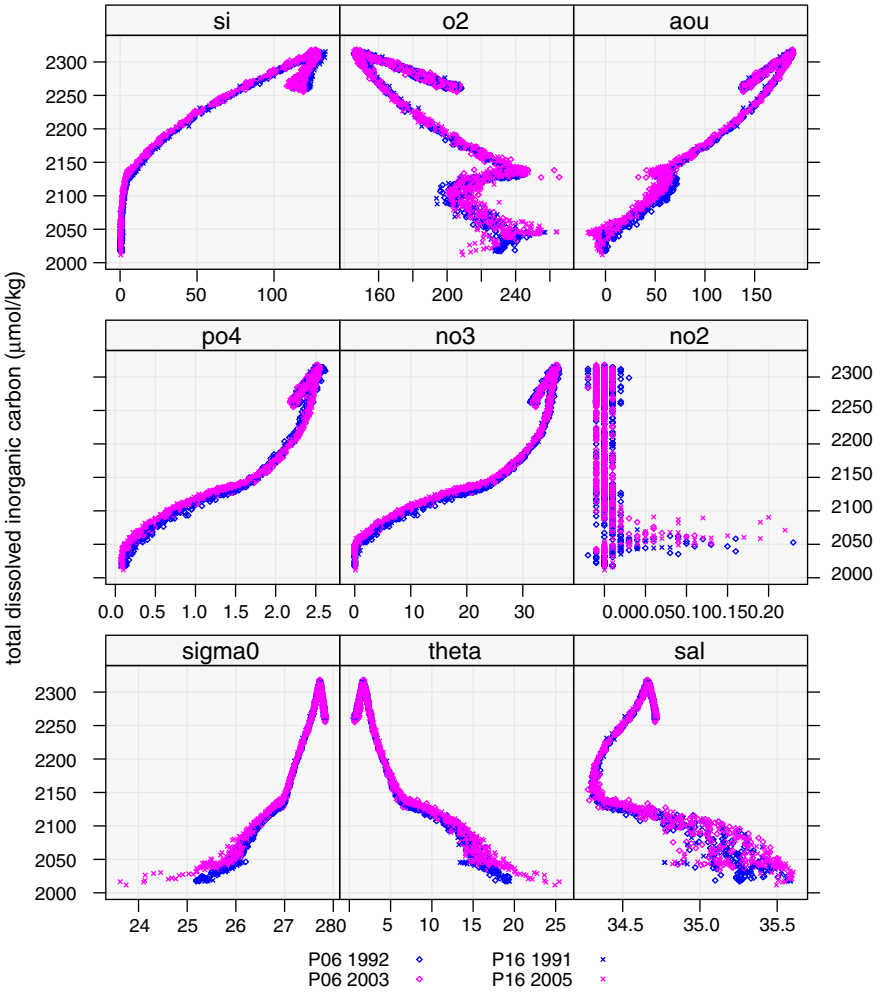

Fig. 2. Co-variability of total dissolved inorganic carbon and nine environmental variables for data at all depths from the stations shown in Fig. 1. From left to right and bottom to top: sigma0 indicates potential density anomaly (kg/m^3) relative to the surface; theta, potential temperature (°C) also relative to the surface; sal, salinity; po4, phosphate (µmol/kg); no3, nitrate (µmol/kg); no2, nitrite (µmol/kg); si, silicate (µmol/kg); o2, oxygen (µmol/kg); and aou, apparent oxygen utilization (µmol/kg). Relationships are clearly not linear over the entire range of the data.

how carbon varies with a potential regressor.[5] First, notice that, even though the common practice of MLR and eMLR is to use linear models, the relationships seen in the data are definitely not linear, with the strongest nonlinearities occurring at the carbon maximum. More important is the fact that the relationships are multivalued.[6] If different models are used above and below the depth of the carbon maximum, then phosphate, nitrate, and silicate become single valued, However. salinity's subsurface minimum makes it unsuitable as a regressor above the carbon maximum unless the domain is again split and different models are used above and below the salinity maximum. Similarly, the plots of oxygen show that the domain above the salinity maximum might need to be split further.

On the other hand, the plots in Figs. 3 and 4 of data limited to intervals 125–275 dbar and 275–525 dbar, respectively, show relationships that are much more suitable for linear modeling. Furthermore, the tendency of the blue points to correspond to lower values of carbon than do the magenta indicates that there might be quantifiable increases in the amount of dissolved inorganic carbon over the decade between the earlier and later surveys.

4. Guidelines

The above considerations suggest several guidelines that might be followed for improving regression-based estimates of rates of accumulation of anthropogenic CO_2. (1) For consistency with the assumption of temporally unchanging background variability, use only one regression model to account for carbon's co-variability with other environmental variables. (2) For the same reason, treat the data symmetrically with respect to time. (3) For consistency with the assumption of linearity, restrict the regression models to smaller regions. (4) Be clear about what is signal and what is noise and about how they are to be separated. (5) Average over sufficient data to insure that estimates of local rates of accumulation are stable. (6) Allow for the use of data other than those from repeated surveys, such as might be provided by autonomous profiling floats. (7) Attempt to obtain reasonable estimates of the errors of the inferred accumulation rates. (8) Use independent data to verify the robustness of the estimated rates.

5. Local two-regression method

The method for estimating the accumulation of anthropogenic carbon, which is presented now, follows the guidelines of the previous section. Its signature attribute is its use of regression twice. The purpose of the first regression is to remove from the data the natural variability that might obscure the anthropogenic signal. Consequently, that signal remains in the residuals of the first regression together with unattributable random variability (noise). The purpose of the second regression is to separate the anthropogenic signal from the noise.

Central to the method is its locality: separate analyses are needed for data from different regions. The extent of the region is governed by two constraints: the background relationships of carbon to the environment should be homogeneous within the region and the data with the region should be sufficient to characterize these relationships. The practice of estimating an accumulation rate at each point with repeated observations is abandoned in favor of an estimate for average rate for the region. This frees the method to consider data from all locations within the region, not just those repeated at the same points, making it capable of handling data that might be provided by a fleet of profiling floats. To illustrate this method two regions are considered: an upper region centering on 200 dbar and a lower region centering on 400 dbar using the data of Figs. 3 and 4, respectively. The likelihood of success of this approach might be anticipated from the separation of blue and magenta in those scatter plots. It is easy to imagine that a linear model fitted to the data might split them with more blue on one side and more magenta on the other and that this bias would show up in the residuals.

The fundamental assumption is that the patterns of natural variability are unchanging over the time interval during which data are available and are unaltered by the increasing anthropogenic component of carbon in the ocean. Consequently, the empirical relationships of carbon to other environmental variables within a limited region can be captured from the temporally distributed data using a local regression model, separating this background from the oceanic carbon resulting from human activity. However, there is a concern that short-term climatic change might actually be altering the nature of these empirical relationships (Goodkin et al., 2011). If climatic changes are shown to alter the relationships among the environmental variables used to characterize the background variability enough to impact the estimate of the rate at which anthropogenic carbon is accumulating, then this method would have to be modified. For now, this fundamental assumption seems reasonable. And as suggested in Section **2**, for the regression model to capture the background variability faithfully, it should be fitted to data from multiple years in order to benefit from exposure to as many situations as possible. If the model successfully describes the co-variability of oceanic carbon with the other environmental variables, any remaining systematic temporal variability can be attributed to the accumulation of anthropogenic CO_2.

A second assumption is that the rate of accumulation is constant over the interval for which there are data. Given the sparsity of temporal sampling, it is unlikely that detecting anything other than a constant rate is possible.[7] The assumption of a constant rate is consistent with a linear temporal trend, so the rate of accumulation of anthropogenic carbon can be determined by fitting a linear function of time to the residuals from the background-regression model. This linear function has two parameters: slope and intercept. The slope determined by the least-squares fit provides the estimate of the rate of accumulation and the standard error of the slope provides a measure of its uncertainty.

5.1. Background regression

The first task is to identify which of the environmental variables to use in the background regression. This choice should reflect the covariability within the local domain being modeled.

Fig. 3 shows that within the shallower layer nitrate and phosphate have strong linear relationships with carbon, so each is an excellent candidate for use as regressor, but as both provide essentially the same information, only one should be used. The same is true of potential density and potential temperature. Likewise, either oxygen or apparent oxygen utilization, which show weaker linear relationships due to the greater scatter about the regression lines, might be used, but not both. Carbon's relationship to silicate is linear with similar scatter, its relationship to salinity is approximately linear with even more scatter, while its relationship to nitrite cannot be called linear at all. The 400 dbar relationships are similar except for carbon to oxygen. Because the boundary between the two layers was chosen near the depth of the first oxygen maximum, carbon increases with increasing oxygen in the shallower layer and decreases in the deeper layer. Note that aou looks like a promising regressor in the upper layer in spite its low values, but because it has a maximum, it is unlikely to provide much useful information about carbon in the deeper layer.

[5] Nitrite, as might be expected, shows little useful relationship to carbon.

[6] Nonlinear regression might be considered over portions of the water column where relationships are single valued.

[7] But there is no fundamental reason that an increasing rate of accumulation might not be considered.

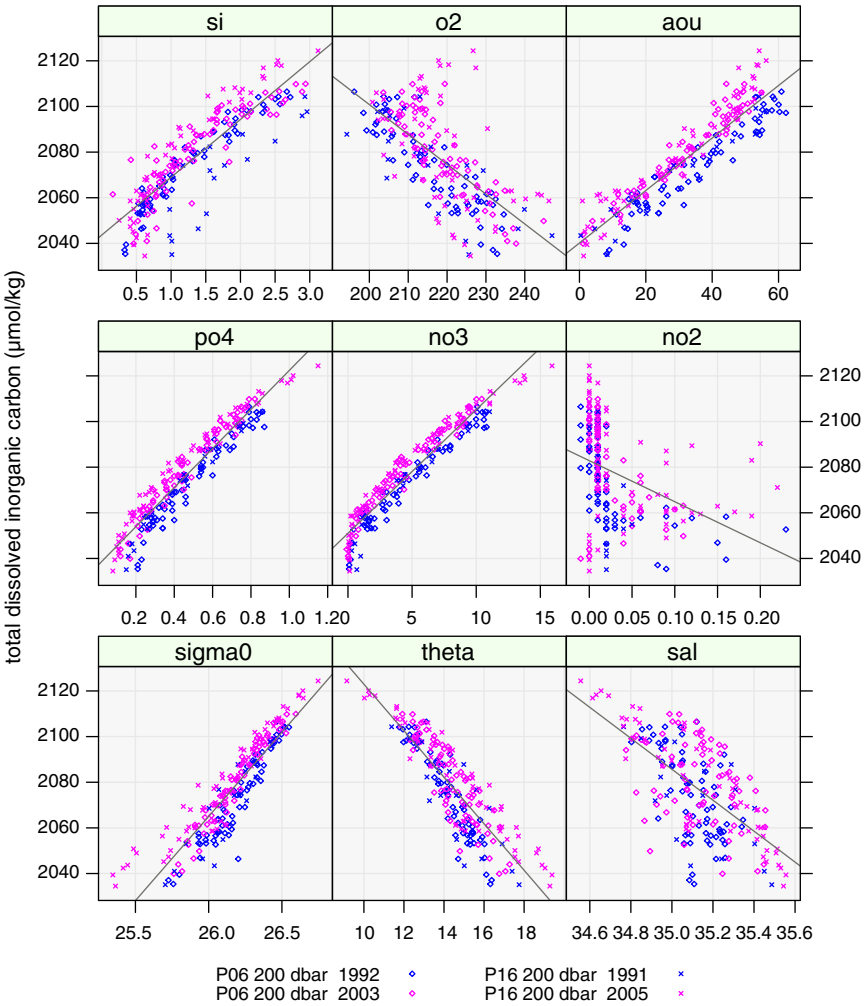

Fig. 3. Co-variability of total dissolved inorganic carbon and nine environmental variables for data restricted to the depth interval centered on 200 dbar. The lines in each panel represent the best univariate fit to the data.

While the scatter plots indicate which variables are the most promising regressors, deciding which and how many to use is not so simple. First, much of the information provided by the candidate regressors is redundant, so different combinations might provide similar estimates. Second, sampling might not be adequate to capture the full range of natural variability. Consequently, an independent sample might lead to different regression coefficients reflecting a different view of what background variability really is and consequently to a different estimate of the accumulation rate. Indeed, independent data might lead to a different opinion concerning the best set of regressors. Thus, the goal is not just to find the best fit to the data but to find a model that would describe an independent sample equally well (e.g. Davis, 1976).

One approach to the choice of variables to characterize the environment is via stepwise regression, adding and/or removing variables until an optimal set is found. The judgment of which set is optimal is generally based on either Mallows (1973) C_p criterion, the Akaike (1974) information criterion (AIC), or the Bayesian information criterion (BIC) (Schwartz, 1978). A drawback is that stepwise regression can be overly permissive, sacrificing robustness by accepting too many regressors. For example, when using stepwise regression with these data, AIC excluded only salinity in the 200-dbar layer, retaining the other eight environmental variables, and in the 400-dbar layer it excluded none. And as stepwise regression makes its choice of regressors from a single sample, it is incapable of judging how well the model might describe independent data.

Because of their mutual correlations, it is unlikely that more than two or three regressors can be used to characterize the environment in either of these two layers without the risk of over fitting.[8] Arguing that the co-variability of environmental variables is largely due to the processes responsible for the distribution of salinity and temperature supporting the ocean's stratification, density should be on the short list of candidate regressors. As temperature and salinity provide little additional information beyond what density provides, for robustness they might be omitted. Nitrate and phosphate both appear to provide useful information about carbon, but given their high correlation due to their Redfield ratio, only one of the two should be used. It is reasonable that the second candidate should be a nutrient to provide information about biochemical variability, so nitrate is included on the short list. Nitrite has already been ruled out. In spite silicate's relatively low levels in the upper ocean, it exhibits relatively strong correlation with carbon, so it is also included in the short list. Although apparent oxygen utilization does not appear useful in the

[8] Over fitting is a consequence of redundancy in the information provided by the regressors, a condition that is sometimes called statistical collinearity. Pairwise redundancy can be identified in scatter plots of pairs of regressors as a tight relationship, while redundancies involving three or more regressors can be approached through the condition number of the matrix of regressor correlation coefficients. The result of over fitting is that minor idiosyncrasies of the fitting data can have undue influence on the regression coefficients, resulting in a model that performs poorly with independent data. It is best avoided by erring on the side of using fewer rather than more regressors.

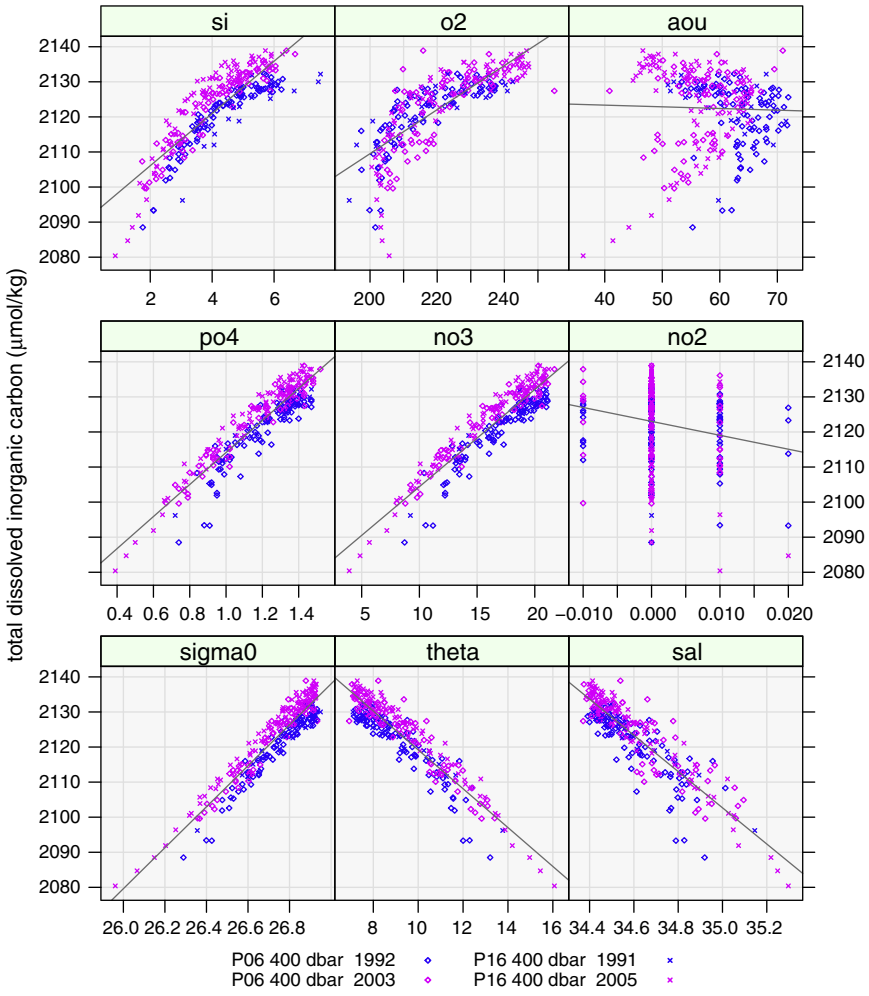

Fig. 4. Co-variability of total dissolved inorganic carbon and nine environmental variables for data restricted to the depth interval centered on 400 dbar. The lines in each panel represent the best univariate fit to the data.

400-dbar layer, it might be useful in the shallower layer, so it has been included as a fourth candidate. An alternative might be to use oxygen, but not both oxygen and apparent oxygen utilization.

To explore the collective utility of these four candidate regressors, six models of background variability were considered. They are listed in Table 1. The default model m0, which has no regressors, ignores the problem of the confounding environmental variability with accumulation of anthropogenic carbon; it just has the intercept term, and when fitted to data that term is evaluated to be the sample mean of the total dissolved inorganic carbon measurements. Model m1 has potential density anomaly as its only regressor; m2 has nitrate as a second regressor; m3 and m4 both retain these two and add silicate and apparent oxygen utilization, respectively, as third regressors. Finally, m5 has all four variables as regressors. Each of the background models is fitted to the data from the 200-dbar layer and separately to the data from the 400-dbar layer, as the empirical relationships are not assumed to be the same in the two layers.[9]

The values for residuals resulting from these fits are connected by lines in Fig. 5 to show how individual observations are corrected by the each model. As each residual is the difference between a measurement of total dissolved inorganic carbon and its modeled counterpart, it is possible to track the impact of adding each additional regressor. Comparing the residuals of m0 to those of the other models shows that failure to account for background variability results in a greater spread in both depth ranges. Comparing residuals of m0 to those of m1 shows that characterizing background variability using only a single variable, density, accounts for most of the reduction in spread. Density is also responsible for a dramatic reordering of some residuals in the shallower layer from low to high and vice versa, which suggests that short-term variability associated with the penetration of the surface mixed layer may be responsible. If so, sampling may be aliasing the long-term relationship between carbon and density. The addition of nitrate further reduces variability in the 200-dbar layer, but the addition of other regressors has less impact, and density alone accounts for essentially all of the reduction in spread in the 400-dbar layer. In both layers the fact that the individual residuals change very little after the initial reduction in spread suggests that the additional regressors provide little additional information about the

Table 1
Regressors are indicated by x for the 6 model types examined in this study.

	Density	Nitrate	Silicate	aou
m0				
m1	x			
m2	x	x		
m3	x	x	x	
m4	x	x		x
m5	x	x	x	x

[9] All computations discussed below were made using R (R Development Core Team, 2011; Venables and Ripley, 2002), which is excellent free software for statistics and data exploration. All graphics were prepared using R's Lattice package (Sarkar, 2008).

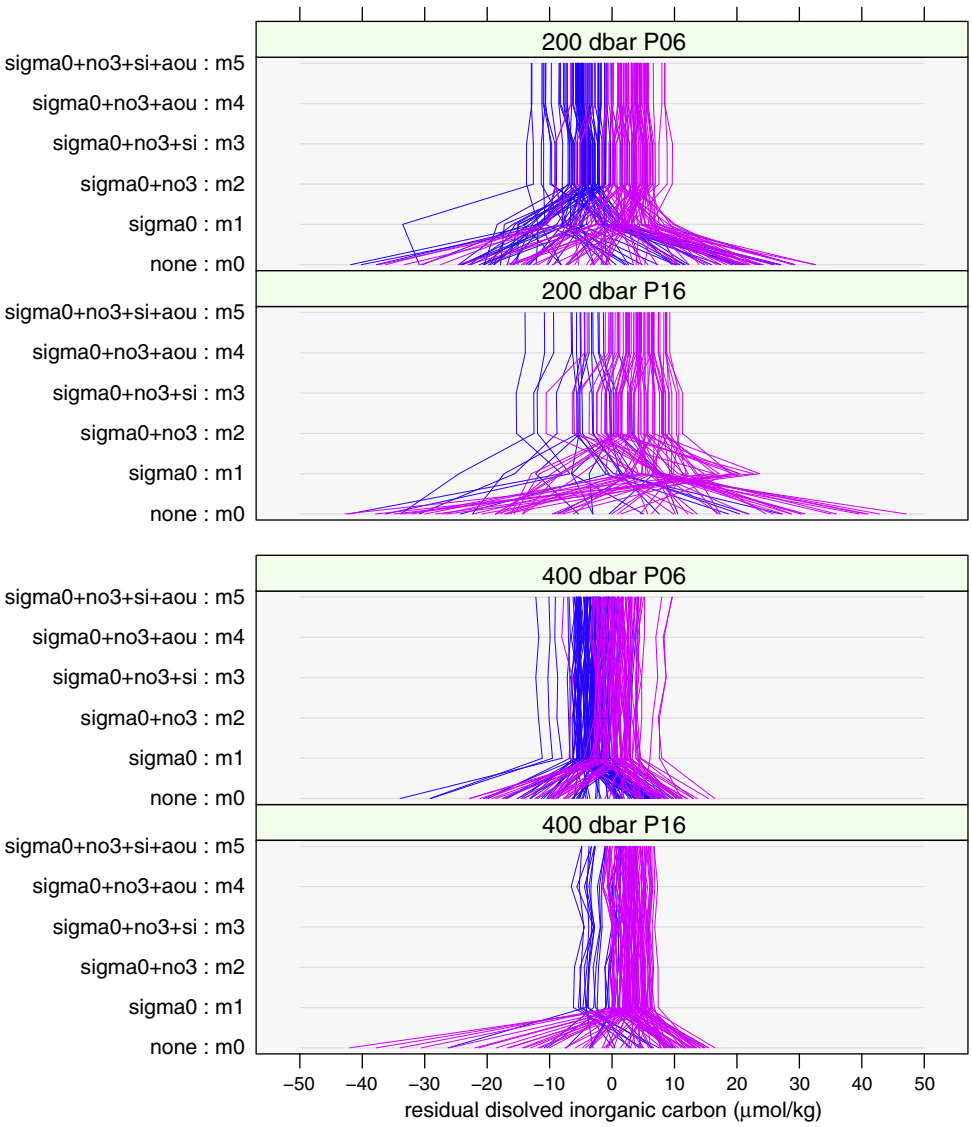

Fig. 5. Residuals of six background-regression models fitted to the data in the 200-dbar (400-dbar) layer are shown in the upper (lower) pair of panels. Residuals are presented separately for data from P06 and P16 so that they can be examined separately. The lines connecting corresponding residuals for the different models indicate that individual observations are only slightly affected as more regressors are added. Blue (magenta) indicates data from the earlier (later) surveys.

environmental variability. As the blue lines indicating data from the earlier surveys are generally to the left of the magenta lines, the anthropogenic signal can be clearly seen. Still, there is a substantial fraction of unexplained spread for blue and magenta lines. If combinations of regressors other than those considered here cannot reduce this remaining spread, it should be regarded as noise.

5.2. Temporal regression

Another view of the residuals is shown in Fig. 6 where they are plotted versus time: they appear as four distinct clusters corresponding to the four surveys providing the data. Again, the spread of residuals at all years is seen to be greatest for the no-regressor case m0 and the greatest reduction in spread is going from m0 to m1. The magenta crosses in Fig. 6 correspond to the means of the residuals for the individual surveys. These means can be regarded as the anthropogenic signal of the individual surveys, and the errors of those means is considerably smaller than the spread of the unattributable noise. Furthermore, the change of the means in time should reflect the rate of accumulation. So the next task is to quantify the rate of accumulation by fitting a linear function of time to the residuals.

The fitted model for each set of residuals is indicated in Fig. 6 by the magenta lines and the slopes of these lines are the estimates of the rate of increase of anthropogenic carbon. Note that the means for the surveys (crosses) are close to the lines. Clearly, regression is serving to average the data as it separates the anthropogenic signal from noise. And as the regression addresses data from the four surveys, it is clear that the accumulation rate as estimated by the slope of the regression line should be regarded an average over the entire space–time region providing the data to which the model has been fitted.

The standard errors of the slope and intercept parameters of the temporal-regression models are analogs of the standard errors of the means for the individual surveys. The standard error of the slope is a particularly useful quantity, as it provides a quantitative measure of the uncertainty of the rate at which anthropogenic carbon has been accumulating within the study region. The rate resulting from each background-regression model together with its standard error is indicated within that model's panel in Fig. 6. The changes in

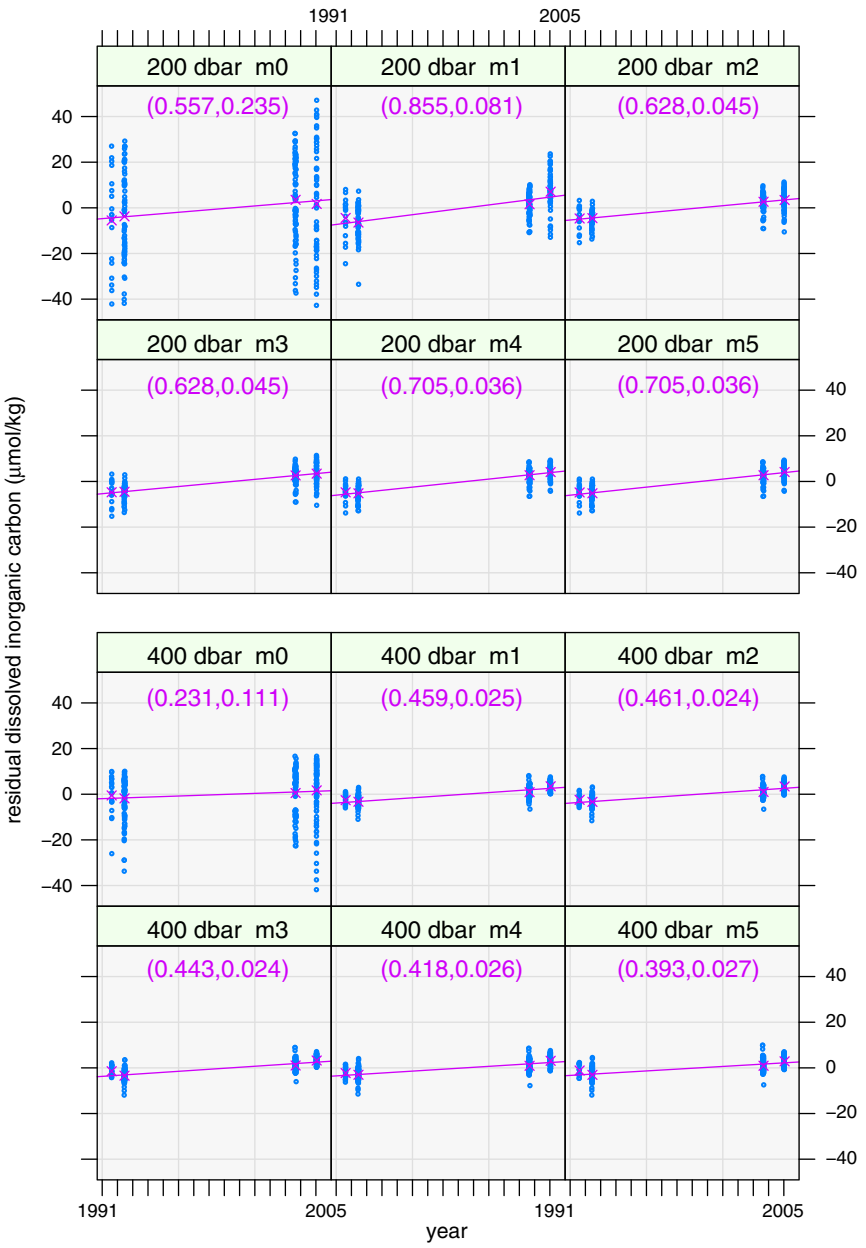

Fig. 6. The plotted points at the date of the observations correspond to residuals of the six background-regression models fitted to data from both P06 and P16 within the two depth intervals as indicated by the panel labels. Crosses indicate the means of the residuals for each survey. Straight lines correspond to the temporal-regression models, and the slopes of these lines are estimates of rates of accumulation of anthropogenic carbon. The pair of numbers displayed in each panel are (accumulation rate, uncertainty of accumulation rate) inferred from the slope of the temporal-regression line and its standard error in units of µmol/kg/yr.

the distribution of residuals from model to model seen in Fig. 5 cause changes in the slopes of the temporal-regression lines and the differences in the estimates of accumulation rates provided by the different models. As expected, estimates of rates for the deeper layer are smaller than those for the layer that is closer to the surface. And as expected, the standard errors of these estimates are smaller than those for the shallower layer, reflecting the reduction of variability with depth.

In both layers, accounting for environmental variability tends to increase the estimated rate of accumulation of anthropogenic carbon with the different environmental models all giving quite similar estimates. But it is important to note that the larger spread in the residuals for m0, which makes no attempt to exploit co-variability of carbon with any environmental variables, causes its estimates to have relatively large uncertainty. In fact, the uncertainty of the 200-dbar layer's null model's estimate is sufficiently large that it spans the accumulation rates inferred from the other five environmental models. However, in the 400-dbar layer where variability is less, the low m0 estimate can be considered distinguishable from the estimates that attempt to account for background variability.

In the 200-dbar range using silicate as a regressor appears to have essentially no impact on either the estimated rate or its uncertainty, as can be seen by comparing results for m3 with those for m2 and results for m5 with those of m4. It appears that the accumulation rate within this depth range is about 0.65 µmol/kg/yr. A better view of the uncertainty of this estimate is given in Section 6 below, where the impact of sampling on the estimate rate is discussed.

In the 400-dbar layer background-models m1, m2, and m3 all give essentially the same estimate for the accumulation rate. The estimate from m4 is a bit smaller, and that from m5 is smaller still. However all have comparable uncertainties as measured by the standard error of the slope parameters of the temporal-regression models. It is easy

to conclude that all provide a better estimate than m0, which ignores the background variability, and considering their uncertainties that they all are in rough agreement with an accumulation rate for this depth range of about 0.4 μmol/kg.

The standard error of the slope parameter doesn't fully describe the uncertainties in the rate at which anthropogenic carbon has been increasing. Observational errors provide an additional source of uncertainty. Another source is the uncertainty of the background regression parameters, which manifests as uncertainty in the residuals to which the temporal regression models are fitted. A third source of uncertainty, which is more difficult to quantify, is the uncertainty in the choice of regressors for quantifying carbon's background variability. The next section addresses a fourth source of uncertainty, that stemming from the available data's limited ability to characterize the full spectrum of variability. Considering all sources of uncertainty, it is clear that the standard error of the slope parameter provides at best an optimistic view of the accuracy of the accumulation rate.

6. Sensitivity to sampling

Two approaches to investigating the sensitivity of the local two-regression estimates of rates of accumulation of anthropogenic CO_2 to the particular data on which they are based are considered. Both involve splitting the data into two sub-samples — one from P06 and the other from P16, but they differ in how the treat the sub-samples. The first follows the same procedure as used for the combined sample, resulting in two independent estimates that can be compared. The second differs in that the background-regression model fitted to data from one sub-sample is applied to the data from the other sample, accounting for environmental variability that had not influenced the fit; so prediction errors rather than residuals determine the slope of the temporal-regression line and thus provide the estimate for the rate of accumulation of anthropogenic carbon.

The first approach is illustrated in Fig. 7. Different colors are used to indicate that some residuals come from fitting to P06 data and the

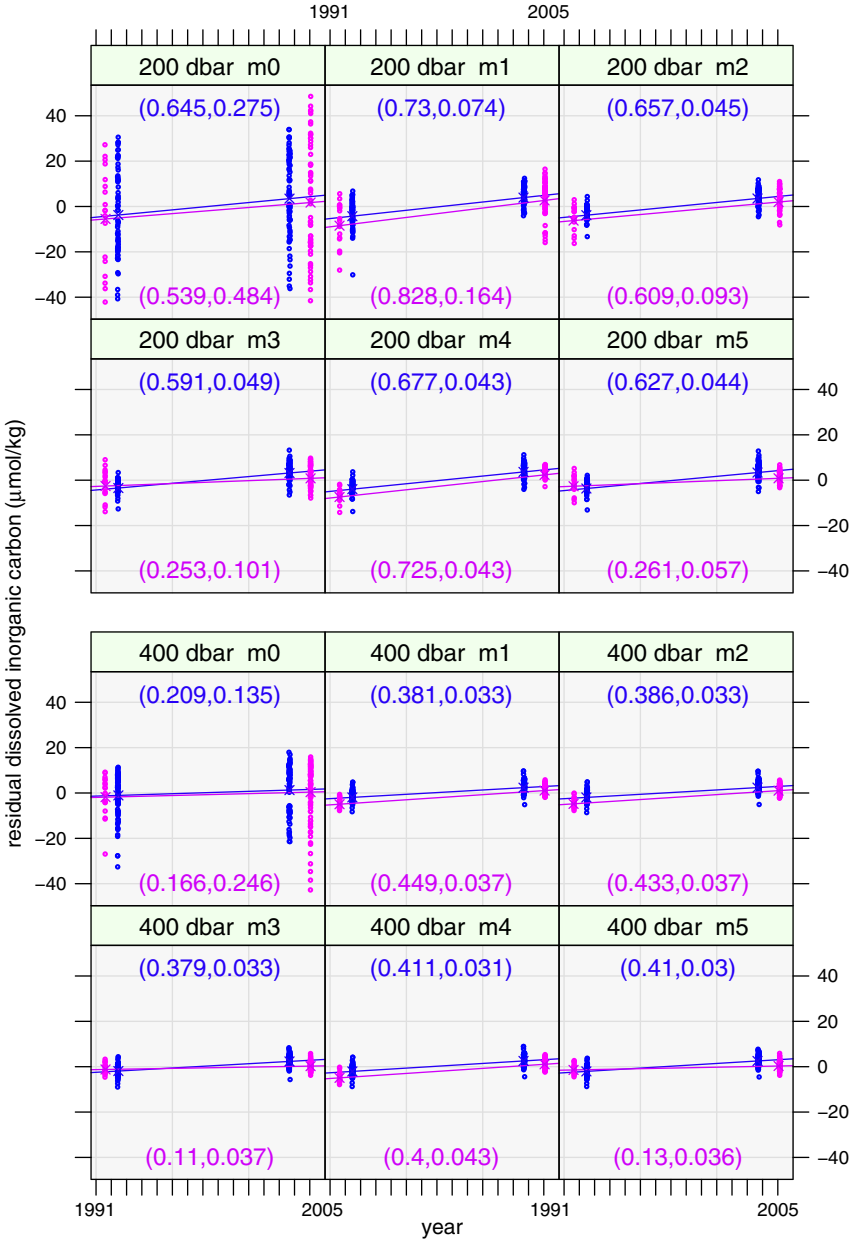

Fig. 7. Like Fig. 6 except that background-regression models have been fitted separately to data from P06 and from P16 in order to explore sensitivity to sampling. Results for P06 are shown in blue and those for P16 in magenta. The two pairs of numbers displayed in each panel are (accumulation rate, standard error of accumulation rate) inferred from the slopes of the two temporal-regression lines in units of μmol/kg/yr.

others come from fitting to P16 data. As data for each is only available from two of the four surveys, the fitted lines pass through the crosses indicating the means for the individual surveys. Thus the estimates for the rates of accumulation benefit only from averaging residuals for each survey but do not benefit from any further smoothing by the temporal regression. Two values of estimated accumulation rates and their standard errors are shown on each panel, once in blue for P06 and again in magenta for P16.

The 200-dbar layer's models m3 and m5, which use silicate, yield different estimates of accumulation rate for the two sections. As silicate showed no effect in Fig. 6, this can be regarded as another indication that it is not a useful regressor for this region and this depth range. On the other hand, models m2 and m4 give results when fitted individually to the two sections that are mutually consistent and also consistent with those based on the combined data. In the 400-dbar layer silicate (m3 and m5) again is associated with the biggest differences between estimates from the separate data sets, and those from m2 and m4 again appear to be mutually consistent and constant with estimates based on the combined data. These results tend to confirm the conclusion drawn from the estimates in Fig. 6 that the accumulation rate is about 0.65 μmol/kg in the 200-dbar layer and about 0.4 μmol/kg in the 400-dbar layer. However, the differences in estimates based on the two subsets of the data suggest that the standard error of the slope parameter underestimates the uncertainty.

It is interesting to note that the estimates for accumulation rates computed individually for the separate sections (Fig. 7) do not necessarily bracket the corresponding estimates based on the combined data (Fig. 6). In the 200-dbar layer they do so only for m0, m2, and m4, and in the 400-dbar layer, only for m6. The reason is that in addition to the slope parameters, which provide the estimates of accumulation rate, the regressions also determine intercept parameters, which reflect the sample means of the variables. In fact, for m4 in the 400-dbar layer, the slopes are essentially the same, but the intercepts are different. Observe that the crosses marking the survey means of the m4 residuals in Figs. 7 and 6 follow different temporal progressions: down–up–up–down versus up–down–down–up. This change in the character of the residuals indicates that accidental trends in the background variables due to inadequate temporal sampling can make a difference. In this case, it manifests very visibly in the intercept estimates, but in general it affects both intercept and slope.

The second approach to checking the sensitivity of the estimates of rates of accumulation of anthropogenic carbon to the peculiarities of the sample is more demanding, as it exposes the fitted models to entirely independent data. It uses the background-regression models that have been fitted to data from one section to predict what the environmental component of the variability in the other section. So, rather than fitting the temporal-regression model to residuals, they are fitted to prediction errors, which being differences from the measured values are similar to residuals but based on independent data. These prediction errors are plotted in Fig. 8 where color is used to indicate the section supplying the verification data. The greater separation between temporal-regression lines compared to what is seen in Fig. 7 suggests that, models fitted to data from one section have a predictive bias when applied to the other section, and that this bias is fairly consistent over the decadal inter-survey interval, leaving the estimates of slope relatively unchanged. The prediction errors also show an increased spread in comparison with the residuals of Fig. 7, and this increase is reflected in somewhat higher uncertainties in the estimates for the rates of accumulation.

Note that, for the null model m0, both approaches compare exactly the same accumulation rates with exactly the same uncertainties. Because the residuals of m0 in Fig. 7 are simply differences between measured values of carbon and their mean and the prediction errors in Fig. 8 are just differences between measured values and the mean of the other sample, corresponding points differ only by an offset. As the points are all shifted by the same amount, the slopes of the regression lines and their standard errors can't differ. However, as expected, the two approaches do produce different estimates for the accumulation rate when accounting for background variability. Sometimes the order of which estimate, blue or magenta, is higher switches, and sometimes it is the same for the two approaches. Taken together, they seem to agree that m3 and m5 might be the most sensitive to sampling in both layers, as each approach produces a pair of quite different estimates, suggesting that silicate is not a useful regressor for these two layers. Model m4 appears to give relatively consistent estimates in Figs. 6, 7, and 8 for both layers. In the 200-dbar layer where apparent oxygen utilization is strongly correlated with carbon, this consistency suggests that it is a useful regressor, but uncertainties render m4 accumulation rates indistinguishable from those of m2. On the other hand, in the 400-dbar layer where apparent oxygen utilization is uncorrelated with carbon, its use as a regressor has relatively little effect on the results.

Overall, the conclusions about sensitivity to sampling are the same whether based on temporal-regression models fitted to background-regression residuals or on those fitted to errors in predicting the background variability of independent data: the available data are sufficient to account for the effects of environmental variations to a reasonable degree of uncertainty, but considering the magnitude of this uncertainty both as characterized by the standard errors of the slope parameters and as indicated by sampling differences, the available data are not sufficient for selecting a best set of regressors.

7. Comparison with previous estimates

Panel-to-panel differences in the uptake rates of Fig. 6 indicate uncertainty associated with the choice of model used to account for background variability, and the within panel differences of Figs. 7 and 8 indicate uncertainties associated with sampling. Clearly, putting all of this information together to get a precise interval estimate of the rate at which anthropogenic carbon has been accumulating in the two layers is difficult. Nevertheless, here is a subjective summary: The accumulation rate within the 200-dbar layer is most likely in the range of 0.5–0.8 μmol/kg/yr and in the 400-dbar layer, 0.35–0.5 μmol/kg/yr.

Murata et al. (2007) use the isopycnal method of Peng et al. (1998) to estimate the accumulation of anthropogenic carbon along P06 using the same data as used in this study. As their results are referenced to potential density rather than to pressure, it is necessary to determine which densities correspond to the two layers of this study. Fig. 3 shows the densities for the 200-dbar layer to be largely in the range of 25.7–26.7 kg/m^3 and Fig. 4 shows the densities for the 400-dbar layer in the range of 26.4–26.9 kg/m^3. In these density ranges Murata et al. (2007) indicate at 150°W in their Fig. 5 accumulations of 6–10 and 8–12 μmol/kg of anthropogenic carbon, respectively, over the 11.2 years between the surveys. These translate into accumulate rates of 0.55–0.9 μmol/kg/yr for the 200-dbar layer, which are roughly in agreement with results found here. However, their rates of 0.7–1.05 μmol/kg/yr for the 400-dbar layer are about twice as high.

Sabine et al. (2008) have estimated the anthropogenic accumulation using data from the two surveys along P16 using eMLR.[10] Their Fig. 4 shows an accumulation over the 13.4 year interval between surveys of about 8–10 μmol/kg at 32.5°S in the depth range of this study's 200-dbar layer and an accumulation of 6.5–8 μmol/kg for the 400-dbar layer. For the 200-dbar layer this amounts to an accumulation rate of 0.6–0.75 μmol/kg/yr, which is quite similar to this study's results. For the 400-dbar layer their accumulation of

[10] They then modify to estimate to correct for changes in apparent oxygen utilization, which is similarly estimated using eMLR, but those corrections are insignificant in the vicinity of 32.5°S.

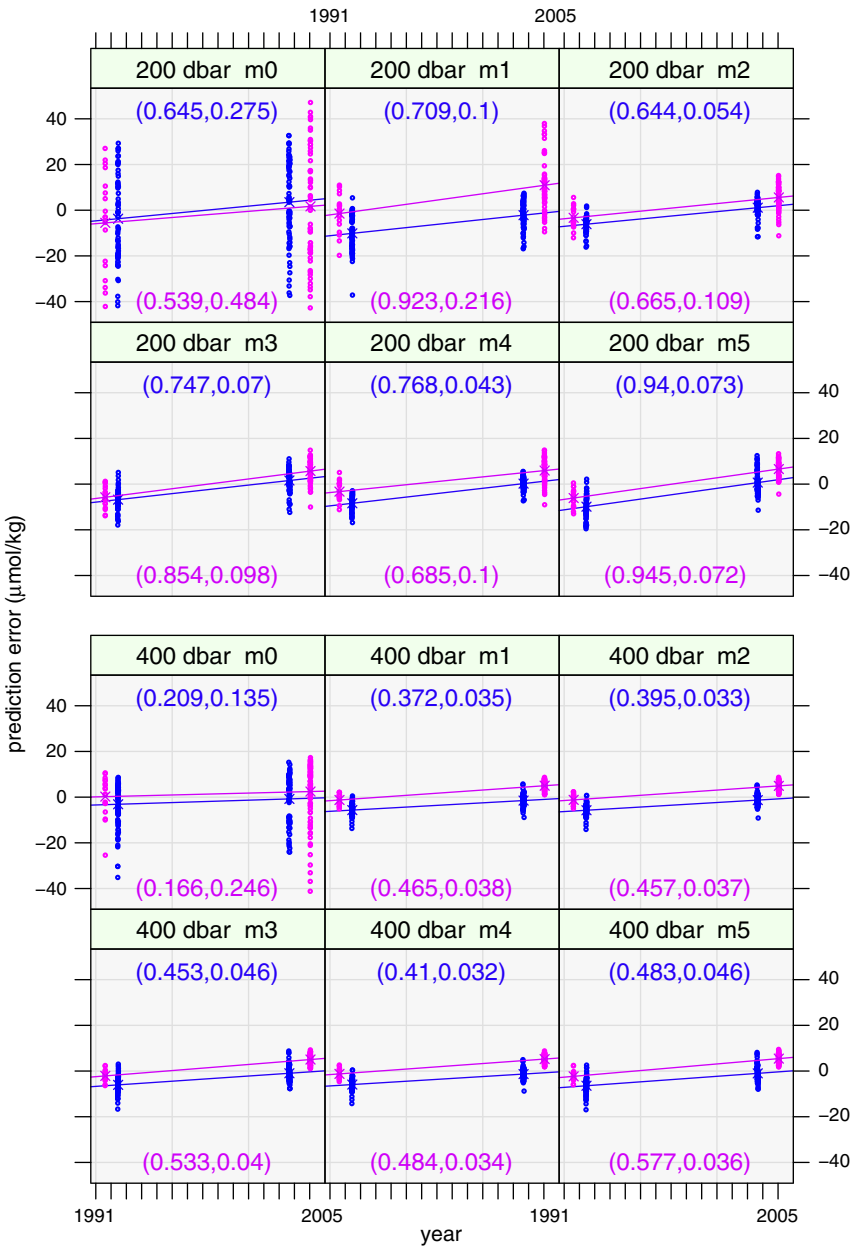

Fig. 8. Like Fig. 7 except that the ordinates now correspond to prediction errors resulting from applying background-regression models to data from the section that had not been used for fitting. Prediction errors for P06 (P16) based on fits to P16 (P06) are shown in blue (magenta). The two pairs of numbers displayed in each panel are (accumulation rate, standard error of accumulation rate) inferred from the slopes of the two temporal-regression lines in units of µmol/kg/yr.

0.5–0.6 µmol/kg/yr is similar to but a bit higher than those found here.

8. Discussion

The first conclusion of this work is that the MLR and eMLR methods suffer from conceptual flaws. Both have problems with the separation of background variability, anthropogenic change, and noise. In addition, both generally treat data over a spatial range that is too large to be consistent with their assumption of linearity, the extent of which can be inferred from scatter plots of the data. And both methods attempt to estimate accumulation rates at individual points without the benefit of spatial averaging to reduce noise. The second conclusion is that, in light of these problems, an improved method is needed, and guidelines for constructing a better method were presented.

The local two-regression method for estimating the accumulation of anthropogenic CO_2 presented in Section 5 provides an alternative to the MLR and eMLR methods that avoids their deficiencies. It directly addresses the problem of separating background variability from anthropogenic signal and noise with the first regression, which identifies the co-variability of oceanic carbon with other environmental variables. A second regression then extracts the anthropogenic signal from the noise. The method is local in that the models are limited to a region within which data exhibit linear relationships, and the rate of accumulation is characteristic of the entire fitting region.

A particularly attractive feature of this approach is that it is suitable for inferring accumulation rates from improved sampling that might be provided by autonomous profiling floats, if they were to be deployed with suitable instruments. Its design allows data from such floats to be used in combination with data from repeated hydrographic surveys and data from fixed moorings. Even with such an

enhanced observing system, an estimate of rate of accumulation of anthropogenic carbon would still require the use of data collected over roughly a decade to be detectable in the residuals from the background regression.

Given the presently available data, the strategy suggested here is to concentrate on the region surrounding the intersection of two repeated surveys so that the temporal sampling is maximized and so that variability can be sampled in orthogonal directions. The question of the horizontal scope of the empirical relationships remains to be explored. It would be nice to confirm that they are sufficiently slowly varying that data from two sets of adjacent intersecting lines can be analyzed jointly to provide estimates within larger regions based on twice as many observations. Analysis within density intervals might be appropriate away from the surface mixed layer and might help extend the horizontal scope of the analysis.

Another conclusion of this work is that any attempt to quantify the uncertainty of estimated rates of accumulation should go beyond formal measures based on how well coefficients have been determined by fitting to the sample at hand. It is important to take into account some measure of the adequacy of the sampling, which can be done by exploring how well the model might explain independent data. And given the well-known problem of over-fitting and the associated difficulty in determine an optimal set of regressors, it is also important to account for the uncertainty associated with their choice.

As the choice of regressors must be made for many layers within each study region, the application of the local two-regression method can be quite laborious. Still, considering the effort that has been made to obtain the data, such labor is not unreasonable. Hopefully, after considering a complete set of layers within several regions where repeating surveys intersect, a pattern will emerge that might guide the model selection.

Acknowledgments

Thanks are due to the Physical Oceanographic Division of the Atlantic Oceanographic and Meteorology for supporting this study. Thanks are also extended to Dennis Hansell, Tsung-Hung Peng, Claudia Schmid, and Rik Wanninkhof for very helpful conversations.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control. 19, 716–723.

Brewer, P., Glover, D.M., Goyet, C., Shafer, D.K., 1995. The pH of the North Atlantic Ocean: improvements to the global model for sound absorption. J. Geophys. Res. 5100, 8761–8776.

Brown, P.J., Bakker, D.C.E., Schuster, U., Watson, A.J., 2010. Anthropogenic carbon accumulation in the subtropical North Atlantic. J. Geophys. Res. 115 C04016, doi:1029/208JC005043.

Davis, R.E., 1976. Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. J. Phys. Oceanogr. 6, 249–266.

Friis, K., Körtizinger, A., Pätsch, J., Wallace, D.W.R., 2005. On the temporal increase of anthropogenic CO_2 in the subpolar North Atlantic. Deep Sea Res. Part I 52, 681–698.

Goodkin, N.F., Levine, N.M., Doney, S.C., Wanninkhof, R., 2011. Impacts of temporal CO_2 and climate trends on the detection of anthropogenic CO_2 accumulation. Global Biogechem. Cycles 25. doi:10.1029/2010GB004009 GB3022.

Hauck, J., Hoppema, M., Bellerby, R.G.J., Volker, C., Wolf-Gladrow, D., 2010. Data-based estimation of anthropogenic carbon and acidification in the Weddell Sea on a decadal timescale. J. Geophys. Res. 115 C03004, doi:10.11029/2009JC005479.

Levine, N.M., Doney, S.C., Wanninkhof, R., Lindsay, K., 2008. Impact of ocean carbon system variability on the detection of temporal increases in anthropogenic CO_2. J. Geophys. Res. 113, C03019. doi:10.1029/2007JC004153.

Mallows, C.L., 1973. Some comments on c_p. Technometrics 15, 661–675.

McNeil, B.I., Matear, R.J., Tilbrook, B., 2001a. Does carbon 13 track anthropogenic CO_2 in the Southern Ocean? Global Biogechem. Cycles 15, 597–614.

McNeil, B.I., Tilbrook, B., Matear, R.J., 2001b. Accumulation and uptake of anthropogenic CO_2 in the Southern Ocean, south of Australia between 1968 and 1996. J. Geophys. Res. 106, 31,431–31,445.

Murata, A., Kumamoto, Y., Watanabe, S., Fukasawa, M., 2007. Decadal increases of anthropogenic CO_2 in the South Pacific subtropical ocean along 32°S. J. Geophys. Res. 112, C05033. doi:10.1029/2005JC003405.

Olsen, A., Omar, A.M., Bellerby, R.G.J., Johannessen, T., Ninnemann, U., Brown, K.R., Olsson, K.A., Olafsson, J., Nondal, G., Kivimae, C., Kringstad, S., Neill, C., Olafsdottir, S., 2006. Magnitude and origin of the anthropogenic CO_2 increase and ^{13}C Suess effect in the Nordic seas since 1981. Global Biogechem. Cycles 20. doi:10.1029/2005GB002669 GB3027.

Ono, T., Watanabe, Y.W., Watanabe, S., 2000. Magnitude and origin of the anthropogenic CO_2 increase and ^{13}C Suess effect in the Nordic seas since 1981. Math. Comput. 72, 317–328.

Park, G.-H., Lee, K., Tishchenko, P., 2008. Sudden, considerable reduction in recent uptake of anthropogenic CO_2 by the East/Japan Sea. Geophys. Res. Lett. 35, L23611. doi:10.1029/2008GL036118.

Peng, T.-H., Wanninkhof, R., 2010. Increase in anthropogenic CO_2 in the Atlantic Ocean in the last two decades. Deep Sea Res. Part I 57, 755–770.

Peng, T.-H., Wanninkhof, R., Bullister, J.L., Feely, R.A., Takahashi, T., 1998. Quantification of decadal anthropogenic CO_2 uptake in the ocean based on dissolved inorganic carbon measurements. Nature 396, 560–563.

Peng, T., Wanninkhof, R., Feely, R.A., 2003. Increase of anthropogenic CO_2 in the Pacific Ocean over the last two decades. Deep Sea Res. Part II 50, 3065–3082.

Quay, P., Sonnerup, R., Stutsman, J., Maurer, J., Kortzinger, A., Padin, X.A., Robinson, C., 2007. Anthropogenic CO_2 accumulation rates in the North Atlantic Ocean from changes in the $^{13}C/^{12}C$ of dissolved inorganic carbon. Global Biogechem. Cycles 21. doi:10.1029/2006GB002761 GB1009.

R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. 3-900051-07-0http://www.R-project.org.

Redfield, A.C., 1934. On the proportions of organic derivatives in sea water and their relation to the composition of plankton. In: Daniel, R.J. (Ed.), James Johnstone Memorial Volume. University Press of Liverpool, pp. 177–192.

Redfield, A.C., Ketchum, B.H., Richards, F.A., 1963. The influence of organisms on the composition of sea water. In: Hill, M.N. (Ed.), The Sea, Volume 2. Interscience Publishers, pp. 26–77.

Sabine, C.L., Key, R.M., Johnson, K.M., Millero, F.J., Poisson, A., Sarmiento, J.L., Wallace, D.W.R., Winn, C.D., 1999. Anthropogenic CO_2 inventory of the Indian Ocean. Global Biogechem. Cycles 13, 179–198.

Sabine, C.L., Feely, R.A., Millero, F.J., Dickson, A.G., Langdon, C., Mecking, S., Greely, D., 2008. Decadal changes in Pacific carbon. J. Geophys. Res. 113 C07021, d07021oi:10.1029/207JC004577.

Sarkar, D., 2008. Lattice: Multivariate Data Visualization with R. Springer, New York978-0-387-75968-5http://lmdvr.r-forge.r-project.org.

Schwartz, G.E., 1978. Estimating the dimension of a model. Ann. Stat. 6, 416–464.

Slansky, C.M., Feeley, R.A., Wanninkhof, R., 1996. The stepwise linear regression method for calculating anthropogenic CO_2 invasion into the North Pacific Ocean. In: Tsanogai, S. (Ed.), Proceedings of the International Marine Science Symposium on Biogeochemical Processes in the North Pacific: Japan Marine Science Foundation, pp. 70–79.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Springer-Verlag, New York.

Wallace, D.W.R., 1995. Monitoring global ocean carbon inventories. OOSDP Background Report Number 5. Ocean Observing System Development Panel. Texas A&M University, College Station, Texas.

Wanninkhof, R., Doney, S., Bullister, J.L., Levine, N.M., Warner, M.J., Gruber, N., 2010. Increase of anthropogenic CO_2 in the interior Atlantic Ocean between 1989 and 2005. J. Geophys. Res. 115. doi:10.1029/2010JC006251 C11028.