Estimating carbonate parameters from hydrographic data for the intermediate and deep waters of the Southern Hemisphere Oceans

H. C. Bostock, S. E. Mikaloff Fletcher, and M. J. M. Williams

National Institute of Water and Atmospheric Research Ltd., Private Bag 14-901, Wellington 6241, New Zealand

Received: 31 January 2013 – Accepted: 25 February 2013 – Published: 2 April 2013

Correspondence to: H. C. Bostock (helen.bostock@niwa.co.nz)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Using GLODAP and CLIVAR ocean carbon data, we have developed several multiple linear regression (MLR) algorithms to estimate alkalinity and dissolved inorganic carbon (DIC) in the intermediate and deep waters of the Southern Hemisphere (south of 25°S) from only hydrographic data (temperature, salinity and dissolved oxygen). A Monte Carlo experiment was used to identify a potential density (σθ) of 27.5 as an optimal break point between the two regimes with different MLR algorithms. The algorithms provide a good estimate of DIC ($R^2 = 0.98$) and alkalinity ($R^2 = 0.91$), and excellent agreement for aragonite and calcite saturation states ($R^2 = 0.99$). Combining the algorithms with the CSIRO Atlas of Regional Seas (CARS), we have been able to map the calcite saturation horizon (CSH) and aragonite saturation horizon (ASH) for the Southern Ocean at a spatial resolution of 0.5°. These maps are more detailed and more consistent with oceanography than the gridded GLODAP data. The high resolution ASH map reveals a dramatic circumpolar shoaling at the Polar Front. North of 40°S the CSH is deepest in the Atlantic (∼4000 m) and shallower in the Pacific Ocean (∼2750 m), while the CSH sits between 3200 and 3400 m in the Indian Ocean.

1 Introduction

Our understanding of the carbonate concentrations and saturation in the oceans has been considerably advanced by the collection of large global datasets such as GEOSECS (1960–1970s), GLODAP (1990’s; Lamb et al., 2002; Key et al., 2004) and more recently CLIVAR/CO$_2$ Repeat Hydrography Program, also known as CARINA/PACIFICA (2000’s; Tanhua et al., 2008; Sabine et al., 2009). However, there are still large gaps, with many areas of the globe that have had little sampling for carbonate parameters. Some of these regions have significant topographic features, such as plateaux and ridges that produce complex currents, thus it is not possible to extrapolate or grid the global datasets across these regions to provide meaningful local estimates.
There may also be temporal variability in these parameters due to seasonal (Feely et al., 1988; Juranek et al., 2009, 2011; McNeil, 2010; McNeil et al., 2011; Alin et al., 2012) or interannual variations caused by phenomenon like El Niño Southern Oscillation (ENSO; McNeil, 2010) or the Pacific Decadal Oscillation (PDO; Kim et al., 2010), which will not have been captured by the one off or decadal repeat transects across the global oceans.

In contrast, there are vast repositories of hydrographic data, which have been compiled to produce ocean climatologies (e.g., World Ocean Atlas (WOA), 2009; CSIRO Atlas of Regional Seas (CARS), 2009). If it was possible to use the detailed hydrographic data to estimate the carbonate parameters in areas where there has been limited sampling for alkalinity and DIC, we could potentially provide detailed carbonate saturation estimates for all areas of the global oceans, with more realistic values in areas where there are complex topography and currents that are not well represented by interpolation of sparse carbon data.

One potential solution would be to develop relationships between carbonate species and hydrographic data using a multiple linear regression (MLR) fit to the existing carbonate data, and then use these relationships and the more widely available oceanographic data to obtain high spatial and temporal resolution information. This MLR approach was first used to predict carbonate species by Wallace (1995), based on observations that carbon exhibited strong correlations with other oceanographic parameters (Brewer et al., 1995). Over the last few years, several publications have used MLR techniques to estimate the carbonate parameters (or carbonate saturation), and changes in anthropogenic carbon uptake from hydrographic measurements (Archer, 1996; Brewer et al., 1997; Millero et al., 1998; Sabine et al., 1999; Lee et al., 2000, 2006; Wallace et al., 2001; McNeil et al., 2001; Friis et al., 2005; Sabine et al., 2008; Juranek et al., 2009, 2011; Kim et al., 2010; McNeil, 2010; Wanninkhof et al., 2010; Peng and Wanninkhof, 2010; Feely et al., 2012). Many of these studies have augmented the hydrographic data with nutrients and/or one of the carbon parameters, hence restricting the usability of these MLR algorithms to regions where this extra data also exists.
Many of the previous efforts to undertake this kind of work have focused on the surface waters of the oceans (e.g. Millero et al., 1998; Lee et al., 2000, 2006). We focus on the intermediate and deep waters and the depth of the calcite (CSH) and aragonite saturation horizons (ASH), as our interest is in understanding how carbonate ion concentrations of bottom waters, which interact with the seabed, affect the distribution of deep water carbonate organisms and sediments. For example, the global depth distribution of deep sea habitat-forming, aragonitic, stony (Scleractinian) corals have been found to be controlled by the aragonite saturation state, with 95% of the organisms found above the ASH (Guinotte et al., 2006). With the ASH currently shoaling at 1–2 m yr\(^{-1}\) in the Southern Hemisphere oceans (Feely et al., 2012), and predicted to shoal considerably in the Southern Ocean by 2100 (Orr et al., 2005; McNeil and Matear, 2008), this could significantly reduce the habitat for these aragonitic organisms.

In this paper, we develop MLR algorithms to estimate carbonate parameters for intermediate and deep waters of the Southern Hemisphere, using hydrographic data (temperature, salinity and dissolved oxygen) and carbonate data (alkalinity and dissolved inorganic carbon – DIC) collected during the WOCE/GLDAP and CLIVAR campaigns from the Southern Ocean. After testing the MLR algorithms with detailed CTD hydrographic data from the WOCE transects, we then apply the algorithms to the CARS (2009) climatology database to produce detailed maps of the CSH and ASH for the Southern Hemisphere oceans.

2 Oceanography

The oceanography of the Southern Hemisphere oceans is dominated by the Southern Ocean (Orsi et al., 2005). The Southern Ocean links all the large ocean basins of the world via the Antarctic Circumpolar Current (ACC), and acts as an important biogeochemical exchange venue, orchestrating the transfer of nutrients between deep and surface waters and the exchange of carbon between the ocean and atmosphere via the biological pump and direct sea-air exchange. The ACC homogenises the water
masses south of the SAF, although there are subtle differences in the water chemistry at depth due to the influence of deep waters entering the Southern Ocean and the formation of intermediate waters that flow north into the adjacent Pacific, Atlantic and Indian basins.

Subantarctic Mode Waters (SAMW) and the low salinity Antarctic Intermediate Waters (AAIW) have high oxygen and low alkalinity, as they are formed by the subduction of Antarctic surface waters (AASW) at the SAF (Fig. 1; e.g. McCartney, 1977, 1982; Talley, 1996, 1999; Sloyan et al., 2010). The primary region of formation is the southeast Pacific, but there is subduction at several other locations in the Southern Ocean (Sallée et al., 2010; Bostock et al., 2013). These cool AASW are subducted between potential density ($\sigma_\theta$) surfaces of 26.9 and 27.3 and have low alkalinity and intermediate DIC (Gonzalez-Davila et al., 2011; Bostock et al., 2013; Fig. 1). The aragonite saturation horizon (ASH) sits at the base of the AAIW in the majority of the South Pacific basin (Bostock et al., 2013).

Below the intermediate waters is the Circumpolar Deep Water (CDW). The CDW is commonly split into lower (lCDW, $\sigma_\theta > 27.75$) and upper (uCDW, $\sigma_\theta = 27.3$ to 27.75) components. The uCDW has a low oxygen signature, with high alkalinity and DIC in the Pacific Sector due to the influence of old Pacific Deep Water (PDW; Bostock et al., 2011; Fig. 1). These deep waters upwell at the Polar Front (PF; Fig. 1), where old carbon rich waters release CO$_2$ back to the atmosphere. The lCDW also has high alkalinity and DIC (Fig. 1) but has higher salinity and oxygen than the overlying uCDW due to the influence of Antarctic Bottom Water (AABW) (Fig. 1; Orsi et al., 1999). CDW also has a large influence from North Atlantic Deep Water (NADW), especially in the South Atlantic, with its slightly higher salinity of 34.72–34.73 (Warren, 1981) and lower alkalinity and DIC (Gonzalez-Davila et al., 2011).

There is some variability in the intermediate and deep waters below the subtropical gyres in the different ocean basins. In the southeastern Pacific, there is significant upwelling in the Peru/Chile Current, which extends from 43° S to 10° S. This results in the shoaling of the ASH to < 500 m (Feely et al., 2004; Bostock et al., 2013). These
waters overlie the low salinity, well oxygenated AAIW, which has recently formed in the southeast Pacific (McCartney, 1977).

The circulation in the southwest Pacific is complex due to the topography. There are multiple types of AAIW in the southwest Pacific (Tomczak and Godfrey, 1994; Hamilton, 2006; Bostock et al., 2013). Beneath the AAIW in the central South Pacific sits the highly corrosive PDW, with very high alkalinity and DIC, which causes the CSH to shoal to 2700 m between 130° W to 180° (Bostock et al., 2011). These waters have excess alkalinity due to the dissolution of carbonate (Feely et al., 2002). The PDW is also the main water mass that enters the South Fiji and New Caledonia Basins in the southwest Pacific, as the complex topography of this region blocks the flow from the south, and the only entrance is via the Solomon Sea in the north (Sokolov and Rintoul, 2000). In contrast, the Tasman basin only receives CDW, with lower alkalinity and DIC, directly from the Southern Ocean. Therefore the deep waters in the Tasman basin have a deeper CSH at ~3100 m (Bostock et al., 2011).

The AAIW flows into the South Atlantic along the western boundary under the Malvinas Current at 500–1200 m, but then continues north under the western boundary Brazil Current (Tomczak and Godfrey, 1994). A tongue of uCDW lies immediately below the AAIW, but is only found as far north as 22° S (Stramma and England, 1999; Gonzalez-Davila et al., 2011). NADW, which has lower DIC and alkalinity than other deep waters, dominates the deep waters of the South Atlantic (Stramma and England, 1999). The NADW mainly flows down the Southwest Atlantic, with smaller amounts in the Southeast Atlantic (Stramma and England, 1999). Below the NADW are a number of weak tentacles of northward flowing AABW, all of which are topographically constrained by seafloor features, especially in the southeast Atlantic where AABW is prevented from flowing north of 25° S by the Walvis Ridge (Tomczak and Godfrey, 1994; Stramma and England, 1999).

In the Indian Ocean the AAIW, with its conspicuous salinity and alkalinity minimum, overlies the Indian Deep Water (IDW). The IDW is characterised by a salinity maximum, intermediate alkalinity and DIC. These IDW have been influenced by mixing with
NADW, especially in the deep western boundary current (Ganacharud et al., 2000). In the east, the CDW/AABW enters via the Australian-Antarctic Discordance and then moves north along the western boundary of the Ninety East Ridge (Tomczak and Godfrey, 1994).

3 Data

This study uses all the WOCE/GLODAP, CLIVAR voyages for which alkalinity and DIC were measured on bottle samples, and the associated hydrographic data were also collected (temperature, salinity, dissolved oxygen) (Table 1; Fig. 2). We used all the data south of 25°S and deeper than 200 m, as the intermediate and deep water masses formed in the Southern Ocean lie beneath the subtropical gyres. In total there are ~20 000 bottle sample data for alkalinity and DIC. The sampling and analyses of the hydrographic and carbonate parameters were performed following standard protocols. Total alkalinity was determined by potentiometric titration and calculated using the Mehrbach et al. (1973) carbonate constants as refitted by Dickson and Millero, 1987, while DIC was analysed using coulometric titration; U.S. Department of Energy (DOE), 1994) and using certified reference material (CRM; Dickson, 2001). The accuracy of the carbonate parameters is estimated at ±3 µmol kg⁻¹ for DIC and ±5 µmol kg⁻¹ for alkalinity (Lamb et al., 2002; Dickson et al., 2007). A secondary quality control is performed by comparing the deep water (> 2000 m) hydrographic, nutrient and carbonate data where they cross other transects, and most of the data is within measurement error (Key et al., 1996, 2002; Lamb et al., 2002; Sabine et al., 2009; Tanhua et al., 2010). We excluded one voyage (P15S 1996), due to a larger offset in the alkalinity data (~5 µmol kg⁻¹; Key et al., 2002).

We undertook a further quality control of the carbonate and hydrographic data using two approaches. Any data that was larger or smaller than typical oceanographic values were removed (Table 2). The data were then plotted against each other, e.g. temperature vs salinity, DIC vs oxygen. These were plotted against data of similar potential
density to determine any obvious outliers greater than 3 standard deviations from the mean (Pearson, 2002). Outliers were not removed if there was a cluster of points from one voyage, or in one region, that could represent a real feature. For example, a cluster (24 data points) of very low oxygen samples exists immediately offshore the west coast of South America on WOCE line P06E. This is an oxygen minimum zone, thus these data points represent a real oceanographic feature and have been retained in the dataset. In total ∼150 samples (<1%) were removed from the overall dataset of ∼20,000 data points. Most of the outliers were due to erroneous alkalinity and DIC measurements, but several were due to high oxygen and salinity data.

The data was collected over several decades 1990s and 2000s. We found that there were subtle variations in the MLR algorithm coefficients when only the 1990’s GLODAP data is used compared to the 2000’s CLIVAR data and therefore we have not made any corrections for the uptake of anthropogenic carbon in the intermediate and deep waters. The anthropogenic signal is considered to be minor in the intermediate and deep waters, with only 5–10 µmol kg^{-1} changes in DIC estimated between 200–1000 m (Sabine et al., 2008; Sallée et al., 2012). This is small relative to the natural variability in the oceans and of similar order of magnitude to the combined uncertainty associated with the measurements (±2 µmol kg^{-1} when using a CRM) and our algorithms. Furthermore, any approach we might take to correct for the anthropogenic signal would introduce its own errors.

Once we had developed and tested the algorithms using the GLODAP/CLIVAR data, we applied them to the CSIRO Atlas of Regional Seas (CARS, 2009) climatology (Dunn and Ridgway, 2002; Ridgway et al., 2002). This database provides temperature, salinity and dissolved oxygen that have been subjected to consistent quality control and have been interpolated taking into account both bathymetry and land. CARS provides mean values and annual and semi-annual cycles, but does not provide information about temporal variability over longer timescales. The annual and semi-annual modes allow monthly climatological fields to be derived. In the Southern Ocean, research vessel logistics and operations have led to a summer bias in oceanographic measurements,
so here we use only the CARS data temporally interpolated for January (Austral Summer).

4 Approach

In this approach, we treat the observed DIC and alkalinity (Alk), from the GLO-DAP/CLIVAR cruises as a linear combination of temperature, \( T \), the salinity, \( S \), with a background salinity of 35 removed, oxygen, \( O_2 \), and a constant offset, \( C \),

\[
\text{DIC} = \alpha T + \beta (S - 35) + \gamma O_2 + C \\
\text{Alk} = \alpha T + \beta (S - 35) + \gamma O_2 + C,
\]

where \( T \) has units of \( ^\circ \)C and DIC, Alk, and \( O_2 \) have units of \( \mu \text{mol kg}^{-1} \). Then, a Singular Value Decomposition (SVD) is used to determine the best values of the parameters \( \alpha \), \( \beta \), \( \gamma \), and \( C \) to fit the observations.

Initially, this optimization was undertaken with all of the data below the surface layer (depth > 200 m). Analysis of the residuals, e.g. the difference between the modelled and observed values, as a function of potential density (\( \sigma_\theta \)) revealed that there were likely to be two distinct regimes with different relationships between carbonate parameters and \( T \), \( S \), and \( O_2 \). We were able to substantially improve the fit by dividing the data into two different density regimes and fitting them separately.

The optimal break point between the two regimes was determined objectively using a Monte Carlo simulation. We generated a random normal distribution of 5000 possible break points centred around \( \sigma_\theta = 27.6 \) with a standard deviation of 0.1 and calculated the best fits to equations 1 and 2 for each of these 5000 \( \sigma_\theta \) break points. Then, we evaluated the goodness of fit using the reduced \( \chi^2 \). The results of this Monte Carlo simulation show a clear, fairly flat minimum in the reduced \( \chi^2 \) over a \( \sigma_\theta \) range of about 27.45 to 27.55 for Alkalinity and 27.4 to 27.6 for DIC (figure 2). At \( \sigma_\theta \) higher than 27.65, the reduced \( \chi^2 \) begins to increase steeply.

6233
For this analysis, we chose to use a break point of $\sigma_\theta = 27.5$ throughout, because it represents a minimum reduced $\chi^2$ for both DIC and Alkalinity. This is also fits with the oceanography as it sits at the boundary between intermediate and deep waters. In order to avoid discontinuities at the density boundary between the two regimes when calculating values of DIC and Alkalinity from these regressions, we take $\sigma_\theta$ averaged weighting of the two fits for data over the density range of $\sigma_\theta = 27.45–27.55$.

The residual standard error (RSE) for the MLR fit is $\pm 7.3 \mu\text{mol kg}^{-1}$ for DIC and $\pm 9.8 \mu\text{mol kg}^{-1}$ for alkalinity, based on all the data used in the optimisation. If we remove the alkalinity and DIC measured-estimated residuals greater than 3 standard deviations from the mean (about 1 % of the total data), considered outliers (Pearson, 2002), there is a minor improvement in the RSE, from $\pm 7.3 \mu\text{mol kg}^{-1}$ to $\pm 6.8 \mu\text{mol kg}^{-1}$ for DIC and $\pm 9.8 \mu\text{mol kg}^{-1}$ to $\pm 9.2 \mu\text{mol kg}^{-1}$ for alkalinity. In general, 70 % of the DIC and alkalinity estimates are within 1 standard deviation of the data, and the $R^2$ of the measured vs estimated alkalinity and DIC was greater than 0.9 (Table 3).

The MLR picks out temperature and oxygen as the main controllers of DIC (Table 3). This is not surprising as DIC is affected by organic matter formation (or break down) in the water column which produces (or requires oxygen). Lee et al. (2000) previously determined that surface DIC can be determined from temperature and nitrate (the latter closely related to oxygen in the oceans). The temperature factor may also be acting as a proxy for depth, with an increased $\alpha$ coefficient for deeper waters as oxygen decreases (Table 3). Alkalinity, on the other hand, is primarily influenced by salinity and temperature (Table 3; Millero et al., 1998; Lee et al., 2006). Alkalinity is also affected by carbonate dissolution and precipitation. Thus the measured vs estimated correlation for alkalinity ($R^2 = 0.91$) is slightly lower than DIC ($R^2 = 0.98$), as none of the parameters ($T$, $S$ or dissolved $O$) are directly affected by carbonate dissolution. Temperature and oxygen may be playing a proxy role in predicting depth, as the coefficients $\alpha$ and $\gamma$ increase for estimating alkalinity in the deep waters.

The largest residuals between the measured and estimated DIC and alkalinity are in the lower mixed layer/thermocline waters ($\sigma_\theta = 25.5$ to 26.5; 200–500 m) and the
deep waters ($\sigma_\theta = 27.7$ to $28; >3000 \text{ m}; \text{Fig. 4}$). While both DIC and alkalinity are over and underestimated in the thermocline, the estimates for the deep waters are predominantly lower than the measured values. The high residuals are most commonly found in the PDW (1800–3500 m in P16S; Fig. 4). This is a region of excess alkalinity due to carbonate dissolution by these old corrosive PDW (Feely et al., 2002). A separate algorithm could be developed for the PDW, but this was not undertaken as the water mass could not be defined by potential density. It would therefore have to be defined by oxygen content and another parameter, increasing the complexity and usefulness of this deliberately simple approach.

South of the PF, the MLR overestimates alkalinity (and DIC to a lesser extent) in the upper 1000 m (Fig. 4). These very cold, dense waters are estimated by the deep alkalinity MLR algorithm as $\sigma_\theta > 27.5$. This region adjacent to Antarctica is controlled primarily by the cold temperatures, and it appears that a separate MLR algorithm may be required for the region between the PF and Antarctica.

There are anomalous residuals in both DIC and alkalinity between 145° W and 180° along P06 2003 (not shown). This seems to be related to lower oxygen values along this section of the P06 transect. This is especially evident in the DIC (which is more affected by oxygen) with residuals throughout the water column, while the alkalinity residuals are only higher in the deeper waters (where oxygen plays a larger role in the MLR; Table 3). This highlights the requirement of using well calibrated hydrographic data.

Previous researchers have used the MLR technique to directly determine aragonite saturation states (Juranek et al., 2009; Kim et al., 2010; Alin et al., 2012). Here we calculate the other carbonate parameters ([CO$_3^{2-}$], $\Omega$ aragonite, $\Omega$ calcite) using CO2sys software (Lewis and Wallace, 1998) and the same temperature, salinity, pressure, alkalinity and DIC data. Silica and phosphate were assumed to be 0 in the CO2sys calculation as these parameters have a minimal effect on the calculation. A comparison of the [CO$_3^{2-}$] calculated from the measured alkalinity and DIC with the estimated alkalinity and DIC gives an $R^2 = 0.96$, RSE $= \pm 4 \mu$mol kg$^{-1}$, while $\Omega$ aragonite and $\Omega$calcite give
an $R^2 = 0.99$, RSE = ±0.05 and ±0.08 respectively (Fig. 5). This is just greater than the overall uncertainties of the aragonite and calcite saturation state calculations of ±0.03 and ±0.05, respectively (Mucci, 1983; Millero, 1995; Feely et al., 2012). The high $R^2$ for $\Omega$ aragonite and $\Omega$ calcite suggests that the errors in the alkalinity and DIC are offset in the calculation of the saturation states.

5 Application and discussion

The development of the algorithm to determine alkalinity and DIC from hydrographic observations allows us to estimate the distribution of alkalinity and DIC where only hydrographic data has been collected. It also allows us to utilise oceanographic climatologies to provide more extensive coverage of the ocean.

Alkalinity and DIC values south of 25°S were estimated from CARS temperature, salinity and dissolved oxygen. From these, $\Omega$ calcite and $\Omega$ aragonite were determined using the CO2sys program. This provided saturation states at each 0.5° in latitude and longitude, and a vertical resolution of 250 m between 2000 and 5000 m, 100 m between 1000 and 2000 m, and 10–50 m in the upper 1000 m of the water column. At each grid point in both latitude and longitude, the vertical profiles in $\Omega$ were linearly interpolated to determine the density where $\Omega = 1$ (Fig. 6) and then the depth of the CSH and ASH (Fig. 7).

The CSH is at highest densities in the South Atlantic, where the NADW has high salinities, and in the cold, dense waters south of the PF (Fig. 6). The ASH also shows the highest densities in these regions, with a dramatic increase at the PF. The ASH generally sits at the base of the AAIW at $\sigma_\theta = 27.3$, but sits at a higher density (and is much more variable) in the South Atlantic, probably due to mixing between the higher alkalinity, denser NADW and the lower alkalinity, less dense AAIW. The lowest density ASH is found off the west coast of South America due to the upwelling into the thermocline waters and the southward flow of EqPIW, with high DIC (Bostock et al., 2010).
The depth of the ASH and CSH show similar patterns to previous global compilations (Feely et al., 2004) but have significantly more detail. Figure 7 shows the ASH and CSH maps derived from GLODAP gridded alkalinity and DIC data. This was supplemented with the temperature and salinity from CARS subsampled to the coarser (both vertical and horizontal) GLODAP grid and followed the same procedure followed to determine the depth where $\Omega = 1$. For CSH, the coarser grid in the GLODAP data resulted in significant masking from the lower resolution bathymetry. The gridded GLODAP data appears to have been interpolated without taking into account bathymetry. This has created patterns in the ASH and CSH that are inconsistent with known ocean circulation. For example, there is evidence for a shoaling of the CSH in the eastern Tasman Sea – a flow over effect from the data in the main South Pacific basin east of New Zealand – yet the only flow of deep waters into this basin is from the south due to the topographic highs such as Lord Howe Rise (Bostock et al., 2011). In contrast, CARS reduces distortion in the gridded fields by taking into account both bathymetry and landmasses (Dunn and Ridgway, 2002). Thus we are able to derive maps of ASH and CSH that have higher resolution and are more consistent with the known oceanographic circulation.

South of 45°S (approximately the location of the STF), the CSH in the Southern Ocean is fairly uniform around 3100–3400 m. There is a subtle deepening of the CSH between the SAF and the PF throughout the Southern Ocean. This is not an artefact of the CARS data, as it is also evident in the raw data from the individual transects (e.g. Fig. 5). This deepening of the CSH is coincident with a shoaling of the $\Omega$ calcite = 1.4 to 1.6, which suggests it is related to the physical oceanography perhaps the upwelling of PDW at the PF also results in the minor depression of the CSH.

The old corrosive PDW causes the CSH to be considerably shallower (<3000 m) in most of the South Pacific Basin. The influence of the PDW into the South Fiji and the New Caledonia basins to the north of New Zealand is readily apparent, while the Tasman Sea has a deeper CSH as it is fed from the Southern Ocean by CDW, with no influence of PDW (Bostock et al., 2011). The deep western boundary current, which is
predominantly made up of CDW and flows east of New Zealand around the Campbell Plateau, Chatham Rise and then north along the Kermadec ridge, is also evident with a deepening of the CSH along its flow path compared to the rest of the South Pacific Basin.

The NADW influence is clearly shown with a deeper CSH (> 3500 m) in the South Atlantic and in the western Indian Ocean, to the west of the Madagascar Ridge. The eastern Indian Ocean has the largest difference in the CSH between the interpolated GLODAP data and the estimates from the CARS climatology. The alkalinity data from GLODAP is patchy for this region and thus the calculated CSH is highly variable, which is then interpolated across the ocean. In the eastern Indian Ocean, the deep waters are made up of CDW, and therefore we would expect the depth of the CSH to be similar to the Southern Ocean.

The ASH depth estimated from the CARS dataset shows a significant shoaling between the SAF and the PF with the shallowest ASH at the PF due to the upwelling of corrosive uCDW. During January (Austral Summer), the ASH shoals to < 1000 m in the South Atlantic and Indian sectors of the Southern Ocean and < 750 m water depth in the Pacific region of the Southern Ocean (Fig. 7). This detail is not evident in the GLODAP interpolations (Fig. 7; Feely et al., 2004). Similar to the GLODAP gridded data, the ASH is deeper, but highly variable in depth in the South Atlantic, probably due to the mixing of the NADW and the AAIW, while off the west coast of South America there is considerable shoaling of the ASH to < 500 m due to the upwelling. These more detailed maps of the estimated ASH and CSH from the CARS climatology will be used for predictive benthic habitat and carbonate sediment models.

While these MLR algorithms work extremely well for estimating the alkalinity, DIC, \([\text{CO}_3^{2-}]\), \(\Omega\) aragonite and \(\Omega\) calcite for the intermediate and deep waters of the Southern Hemisphere open oceans, slight improvements can be made when the South Pacific, South Atlantic and Indian basin data are run separately. Work in the coastal zone of northeast Pacific (Oregon) (Juranek et al., 2009) and the Sea of Japan (Kim et al., 2010), has shown that separate, localised, algorithms are required. Hence, although we
found our algorithm to work well in the open ocean, more testing is needed to determine if is useful in marginal seas. South of the PF our MLR overestimates the alkalinity and DIC in the upper 1000 m (Fig. 5), suggesting that waters around Antarctica may be sufficiently distinct as to benefit from a separate routine.

With continued uptake of anthropogenic CO$_2$ the relationship between DIC and alkalinity and hydrographic parameters will alter, as demonstrated by McNeil et al. (2001). Previous studies have used differences between MLR fits in different time periods, in a technique called extended MLR (eMLR), to determine the uptake of anthropogenic CO$_2$ over the past couple of decades (McNeil et al., 2001b; Peng et al., 2003; Friis et al., 2005; Sabine et al., 2008; Wanninkhof et al., 2010; Peng and Wanninkhof, 2010; Feely et al., 2012). This is an especially important problem in the Southern Ocean, where the AASW are taking up a large proportion of anthropogenic CO$_2$. These waters then subduct to form AAIW, changing its carbonate chemistry, and storing large concentrations of CO$_2$ in AAIW (Sabine et al., 2004; Murata et al., 2007, 2008, 2010; Feely et al., 2012). The small differences in the algorithms for the 1990’s and 2000’s that we found suggest that the anthropogenic carbon bias is relatively small at present, but in the future, new data will be needed to determine how the algorithms evolve in response to anthropogenic uptake.

With the potential routine deployment of dissolved oxygen sensors on future Argo floats, algorithms like these will allow routine determination of carbonate parameters from Argo floats (Juranek et al., 2011). Giving the potential for both higher spatial and temporal resolution of carbonate saturation estimates for the intermediate and deep waters of the Southern Hemisphere oceans. This would help to determine if there are seasonal, interannual (Feely et al., 1988; Juranek, et al., 2009; McNeil et al., 2011; Alin et al., 2012), or decadal (Kim et al., 2010) variability in intermediate and deep waters. This could be significant in areas of upwelling like the west coast of South America. If future Argo floats will also incorporate direct in-situ carbonate measurements (Byrne and Yao, 2008), these algorithms could also be a useful tool for quality control of the data.
MLR algorithms such as these, and other methods such as neural networks (Velo et al., 2013; Sasse et al., 2013), do not abrogate the need for ongoing sampling to directly measure DIC and alkalinity. Instead they give us the tools to maximise the value of direct measurements by allowing meaningful extrapolation to larger areas. Potentially they will allow identification of specific regions of interest, where processes can only be understood with dedicated sampling programmes.

6 Conclusions

Using the global carbonate datasets of GLODAP (1990s) and CLIVAR (2000s), we have developed MLR algorithms to accurately estimate the DIC, alkalinity, and saturation states of calcite and aragonite for intermediate and deep waters of the Southern Hemisphere open oceans. When used in conjunction with the CARS (2009) database this provides new detailed maps of the CSH and ASH, which are oceanographically consistent. These will be used for future benthic habitat mapping and carbonate sediment models. With the deployment of oxygen sensors on Argo floats, these algorithms can be used to look at seasonal and interannual changes in the carbonate saturation states in the intermediate and deep waters.

There are larger errors south of the PF and further testing is required for marginal seas. These algorithms do not abolish the need for future sampling and measurement of DIC and alkalinity, as the coefficients of the MLR algorithms will change with future anthropogenic CO₂ uptake by the oceans. They, however, may help to identify specific regions and depths of interest for targeted future sampling.

Acknowledgements. The authors would like to acknowledge all the funding agencies, voyage leaders and participants and that contributed towards the global databases of WOCE/GLODAP and CLIVAR these are an invaluable resource. Funding for this work has come from the New Zealand Ministries for Primary Industries (formerly the Ministry of Fisheries) (ZBD2010/41), and Business Innovation and Employment (CO1X1226) and NIWA as part of its government funded core research.
References


Wallace, D. W. R.: Monitoring global ocean carbon inventories, Ocean Observing System Development Panel, Texas A&M University, College Station, TX, 54 pp., 1995.


Table 1. GLODAP/CLIVAR voyages used for this study.

<table>
<thead>
<tr>
<th>WOCE transect number</th>
<th>Dates</th>
<th>Research Ship</th>
<th>Voyage Leaders and Principal Scientists responsible for carbon analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pacific</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P15S 2001</td>
<td>24 May–8 Jun 2001</td>
<td>R/V Franklin</td>
<td>Wijffels, Tilbrook</td>
</tr>
<tr>
<td>Southern</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR03 2001</td>
<td>29 Oct–22 Nov 2001</td>
<td>R/V Aurora Australis</td>
<td>Rintoul, Tilbrook</td>
</tr>
<tr>
<td>Indian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I06S 2008</td>
<td>5 Feb–8 Mar 2008</td>
<td>R/V Revelle</td>
<td>Speer, Feely</td>
</tr>
<tr>
<td>I06S 1996</td>
<td>20 Feb–22 Mar 1996</td>
<td>R/V Marion Dufresne</td>
<td>Poisson</td>
</tr>
<tr>
<td>I08 2007</td>
<td>4 Feb–17 Mar 2007</td>
<td>R/V Revelle</td>
<td>Swift, Feely</td>
</tr>
<tr>
<td>I05 2002</td>
<td>10 Mar–16 Apr 1995</td>
<td>R/V Knorr</td>
<td>Talley, Winn</td>
</tr>
<tr>
<td>Atlantic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A10 2003</td>
<td>6 Nov–5 Dec 2003</td>
<td>R/V Mirai</td>
<td>Yoshikawa, Murata</td>
</tr>
<tr>
<td>A10 1993</td>
<td>27 Dec–31 Jan 1993</td>
<td>R/V Meteor</td>
<td>Siedler, Wallace</td>
</tr>
</tbody>
</table>
**Table 2.** Initial quality control on the GLODAP and CLIVAR data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Data removed if:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salinity</td>
<td>&lt; 33.5 or &gt; 36</td>
</tr>
<tr>
<td>Oxygen</td>
<td>&gt; 350 µmol kg$^{-1}$</td>
</tr>
<tr>
<td>Alkalinity</td>
<td>&lt; DIC or &lt; 2000 or &gt; 2600 µmol kg$^{-1}$</td>
</tr>
<tr>
<td>DIC</td>
<td>&lt; 1800 or &gt; 2500 µmol kg$^{-1}$</td>
</tr>
</tbody>
</table>
Table 3. Coefficients and constants, RSE and $R^2$ for the MLR algorithms.

<table>
<thead>
<tr>
<th>Depth</th>
<th>Parameter</th>
<th>Coefficients for $T$, $S - 35$, dissolved O respectively</th>
<th>RSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Alk</td>
<td>$-7.418$</td>
<td>$96.957$</td>
<td>$-0.079$</td>
</tr>
<tr>
<td>(200 m to $&lt; 27.5 \pm 0.05 \sigma_\theta$)</td>
<td>DIC</td>
<td>$-14.866$</td>
<td>$53.682$</td>
<td>$-0.569$</td>
</tr>
<tr>
<td>Deep</td>
<td>Alk</td>
<td>$-17.027$</td>
<td>$100.25$</td>
<td>$-0.663$</td>
</tr>
<tr>
<td>(&gt; $27.5 \pm 0.05 \sigma_\theta$)</td>
<td>DIC</td>
<td>$-23.154$</td>
<td>$13.524$</td>
<td>$-1.017$</td>
</tr>
</tbody>
</table>
Fig. 1. P16 2005 data. (A) Temperature (°C), (B) Salinity, (C) Dissolved Oxygen (µmol/kg) from the CTD data with a vertical resolution of 2 m. (D) Alkalinity (µmol kg⁻¹), (E) Dissolved Inorganic Carbon (DIC) (µmol kg⁻¹) are measured on bottle samples. Overlain are the potential density ($\sigma_\theta$) isopycnals with $\sigma_\theta = 27.5$ highlighted in bold.
Fig. 2. Chi squared ($\chi^2$) of the best fit for all the data from the Monte Carlo (5000 iterations) run on potential density surfaces ($\sigma_\theta$). (A) Alkalinity data, (B) DIC data. $\sigma_\theta = 27.5$ was chosen as the optimal breakpoint as it has a minimum $\chi^2$ for both Alkalinity and DIC. The area shaded in grey is the overlap ($\sigma_\theta = 27.5 \pm 0.05$) where both algorithms are used and the value is averaged.
Fig. 3. Measured alkalinity and DIC versus estimated from the MLR algorithm. Residuals plotted against potential density and water depth. The area in the circle is data from P06 2003 from 145° W to 180°.
Fig. 4. P16S transect. Predicted alkalinity and DIC using the CTD data from CLIVAR transect P16 2005 using the MLR algorithms. (A) Alkalinity (with $\Omega$ calcite overlain), (B) residuals between (A) and the measured bottle data (Fig. 1d), (C) DIC (with $\Omega$ aragonite overlain), (D) residuals between (C) and the measured bottle data (Fig. 1e).
Fig. 5. Carbonate parameters $[\text{CO}_3^{2-}]$, $\Omega$ aragonite and $\Omega$ calcite calculated using CO2SYS (Lewis and Wallace, 1998), using measured versus estimated DIC and alkalinity.
Fig. 6. Potential density of the (A) CSH (contours at $\sigma_\theta = 0.02$ intervals) and (B) ASH (contours at $\sigma_\theta = 0.2$ intervals) predicted from the CARS 2009 data (January). Potential density of (C) CSH and (D) ASH from the bottle data from GLODAP and CLIVAR.
Fig. 7. (A) Depth of the CSH and (B) ASH estimated from the CARS 2009 data (January). (C) Depth of the CSH and (D) ASH from the GLODAP gridded data. Contours are at 250 m intervals.