# The Global Ocean Data Analysis Project version 2 (GLODAPv2) – an internally consistent data product for the world ocean

**Are Olsen**[1], **Robert M. Key**[2], **Steven van Heuven**[3], **Siv K. Lauvset**[1,4], **Anton Velo**[5], **Xiaohua Lin**[2], **Carsten Schirnick**[6], **Alex Kozyr**[7], **Toste Tanhua**[6], **Mario Hoppema**[8], **Sara Jutterström**[9], **Reiner Steinfeldt**[10], **Emil Jeansson**[4], **Masao Ishii**[11], **Fiz F. Pérez**[5], and **Toru Suzuki**[12]

[1]Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Allègaten 70, 5007 Bergen, Norway
[2]Atmospheric and Oceanic Sciences, Princeton University, 300 Forrestal Road, Sayre Hall, Princeton, NJ 08544, USA
[3]Royal Netherlands Institute for Sea Research (NIOZ), Marine Geology and Chemical Oceanography, P.O. Box 59, 1790 AB Den Burg, the Netherlands
[4]Uni Research Climate, Bjerknes Centre for Climate Research, Nygårdsgaten 112, 5007 Bergen, Norway
[5]Instituto de Investigaciones Marinas – CSIC, Eduardo Cabello 6, 36208 Vigo, Spain
[6]GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany
[7]Carbon Dioxide Information Analysis Center, Environmental Sciences Division, Oak Ridge National Laboratory, U.S. Department of Energy, Building 4500N, Mail Stop 6290, Oak Ridge, TN 37831-6290, USA
[8]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bussestrasse 24, 27570 Bremerhaven, Germany
[9]IVL Swedish Environmental Research Institute, Ascheberggatan 44, 411 33 Gothenburg, Sweden
[10]University of Bremen, Institute of Environmental Physics, Otto-Hahn-Allee, 28359 Bremen, Germany
[11]Oceanography and Geochemistry Research Department, Meteorological Research Institute, Japan Meteorological Agency, 1-1 Nagamine, Tsukuba, 305-0052, Japan
[12]Marine Information Research Center, Japan Hydrographic Association, 1-6-6-6F, Hanedakuko, Ota-ku, Tokyo, 144-0041, Japan

*Correspondence to:* Are Olsen (are.olsen@gfi.uib.no)

**Abstract.** Version 2 of the Global Ocean Data Analysis Project (GLODAPv2) data product is composed of data from 724 scientific cruises covering the global ocean. It includes data assembled during the previous efforts GLODAPv1.1 (Global Ocean Data Analysis Project version 1.1) in 2004, CARINA (CARbon IN the Atlantic) in 2009/2010, and PACIFICA (PACIFic ocean Interior CArbon) in 2013, as well as data from an additional 168 cruises. Data for 12 core variables (salinity, oxygen, nitrate, silicate, phosphate, dissolved inorganic carbon, total alkalinity, pH, CFC-11, CFC-12, CFC-113, and $CCl_4$) have been subjected to extensive quality control, including systematic evaluation of bias. The data are available in two formats: (i) as submitted but updated to WOCE exchange format and (ii) as a merged and internally consistent data product. In the latter, adjustments have been applied to remove significant biases, respecting occurrences of any known or likely time trends or variations. Adjustments applied by previous efforts were re-evaluated. Hence, GLODAPv2 is not a simple merging of previous products with some new data added but a unique, internally consistent data product. This compiled and adjusted data product is believed to be consistent to better than 0.005 in salinity, 1 % in oxygen, 2 % in nitrate, 2 % in silicate, 2 % in phosphate, $4\,\mu\mathrm{mol\,kg^{-1}}$ in dissolved inorganic carbon, $6\,\mu\mathrm{mol\,kg^{-1}}$ in total alkalinity, 0.005 in pH, and 5 % for the halogenated transient tracers.

The original data and their documentation and doi codes are available at the Carbon Dioxide Information Analysis Center (http://cdiac.ornl.gov/oceans/GLODAPv2/). This site also provides access to the calibrated data product, which is provided as a single global file or four regional ones – the Arctic, Atlantic, Indian, and Pacific oceans – under the doi:10.3334/CDIAC/OTG.NDP093_GLODAPv2. The product files also include significant ancillary and approximated data. These were obtained by interpolation of, or calculation from, measured data. This paper documents the GLODAPv2 methods and products and includes a broad overview of the secondary quality control results. The magnitude of and reasoning behind each adjustment is available on a per-cruise and per-variable basis in the online Adjustment Table.

## 1 Introduction

Over the past few years increasing evidence for substantial anthropogenic ocean change has emerged. The ocean is warming (Levitus et al., 2012), becoming more acidic (Lauvset et al., 2015), and losing oxygen (Helm et al., 2011). As climate change progresses, these changes will aggravate (Bopp et al., 2013) and may cause significant changes to ocean circulation, ecosystems, and harvestability. Documentation and understanding of ocean change and variability are to a large extent provided through global repeat hydrography programs, with extensive coordination of sampling and measurements of physical and biogeochemical properties (Talley et al., 2016). The data collected during the WOCE/JGOFS (A list of abbreviations appears in Appendix C) global hydrographic survey of the 1990s were combined in the data product GLODAPv1.1 (Sabine et al., 2005; Key et al., 2004) following extensive quality control. By providing easy and open access to internally consistent and properly documented integrated data this product spearheaded major scientific developments, including the first observational estimate of the global ocean anthropogenic $CO_2$ inventory (Sabine et al., 2004). In 2009 GLODAPv1.1 was followed by CARINA (CARbon IN the Atlantic ocean; Key et al., 2010; Tanhua et al., 2009), which combined hydrographic and biogeochemical data from the Arctic, Atlantic, and Southern oceans into a consistent product. Recently, a dedicated synthesis of Pacific Ocean scientific cruise data, PACIFICA (PACIFic Interior ocean CArbon), was published (Suzuki et al., 2013). These two latter data syntheses include a significant amount of data from national projects, ensuring their availability and consistency with global repeat hydrography data.

However, a simple merging of these three products does not give an updated global and fully consistent data product. This is primarily because somewhat different variables were subjected to secondary QC for each product and also because the methods used for the secondary QC have been slightly altered from product to product. Since, in addition, a relatively large amount of new data had become available, in particular those from the CLIVAR/GO-SHIP global repeat survey, GLODAPv2 was instigated to prepare an updated, unified, bias-corrected interior ocean data product, which would

- include data from GLODAPv1.1, CARINA, PACIFICA, and any new data (more recent as well as older, previously unavailable);

- have calibrated and bias-corrected data for the core variables salinity, oxygen, nitrate, silicate, phosphate, dissolved inorganic carbon ($TCO_2$), total alkalinity (TAlk), pH, and the four halogenated transient tracer species, based on consistent secondary QC procedures;

- preserve actual variability and trends;

- include other commonly measured variables;

- contain interpolated values for missing salinity, oxygen, and nutrient data whenever possible;

- include calculated values for the third seawater $CO_2$ chemistry variable ($pCO_2$ is not included in GLODAPv2, only $TCO_2$, TAlk, and pH) wherever measured data for two of them were present.

In addition, an updated mapped global ocean carbon climatology based on the data product was to be prepared, and all original – unadjusted – data were to be made available as WOCE exchange formatted data files at a single access point.

This paper summarizes sources of data for GLODAPv2 (Sect. 2), describes the primary and secondary quality control (QC) procedures (Sect. 3) and results (Sect. 4), introduces the GLODAPv2 data products and access (Sect. 5), provides recommendations for use (Sect. 6), and concludes with a summary of lessons learned during the preparation of this product (Sect. 7). The global ocean mapped climatology is presented in Lauvset et al. (2016).

## 2 Data sources

GLODAPv2 includes all data in GLODAPv1.1, CARINA, and PACIFICA, as well as data from 168 new cruises. The new data originate both from recent cruises, completed after production of the previous data syntheses, and from less recent cruises for which the data have only become available recently. Their sampling locations are shown alongside the

sampling locations of GLODAPv1.1, CARINA, and PACI-FICA cruises in Fig. 1. The new data were obtained by directly contacting principal investigators known to have carried out relevant cruises and by circulating a request letter to the ocean carbon science community through the IOCCP, as well as the SOLAS and IMBER core projects of the IGBP. All of the new data are listed in the Supplement. For the cruises from GLODAPv1.1, CARINA, and PACIFICA the reader is referred to the web pages for each product at CDIAC.

Altogether, GLODAPv2 includes data from 724 cruises. Data from the surveys of WOCE/JGOFS (King et al., 2001; Sabine et al., 2005), CLIVAR, and GO-SHIP (Feely et al., 2014; Hood et al., 2010; Talley et al., 2016) form the backbone. In addition, data from the large-scale surveys of the 1970s and 1980s – GEOSECS, TTO, and SAVE – and from a multitude of national and regional programs have been included. Examples include the time series stations KNOT, K2 (e.g., Wakita et al., 2010), and Line P (e.g., Wong et al., 2007) in the Pacific; the Indian Ocean INDIGO (e.g., Mantisi et al., 1991) and OISO (e.g., Metzl, 2009) programs; the Irminger and Iceland Sea time series data (Olafsson et al., 2009); and several Arctic Ocean (e.g., Jutterström and Anderson, 2005; Giesbrecht et al., 2014) and Nordic Seas data (e.g., Jutterström et al., 2008; Olsen et al., 2010).

GLODAPv2 is primarily an open-ocean data product. Data from a few coastal surveys and time series have been included on an opportunistic basis. Time series data not included in GLODAPv2 include BATS (Steinberg et al., 2001) and HOT (Dore et al., 2003). The rationale is that the large amount of data from these time series would tend to bias the GLODAPv2 data product without improving its spatial detail, and the fact that these data are well maintained, organized, and readily available.

## 3 GLODAPv2 methods

### 3.1 Primary quality control

All individual cruise data files used for GLODAPv1.1, CARINA, and PACIFICA existed in the required WOCE exchange format and had been subjected to primary QC during the preparation of these products. All of the new data were merged as necessary, converted to WOCE exchange format, and also subjected to primary QC. The primary QC was carried out following routines outlined in Sabine et al. (2005) and Tanhua et al. (2010), primarily by inspecting property–property plots. Outliers showing up in two or more different property–property plots were generally flagged as such. The WOCE QC flags are listed in Table 1. As with previous products, a reduced flag set was used for the data product, while the full set was used for the individual cruise data files.

### 3.2 Secondary quality control

#### 3.2.1 Merging of sensor and bottle data for salinity and oxygen

For salinity and oxygen, two types of submitted data exist. Data files may have a single column of values for each, being either from analyses of water samples (in the following referred to as bottle values/bottle salinity/bottle oxygen) or derived from CTD sensor pack data (in the following referred to as CTD values/CTD salinity/CTD oxygen). Otherwise, data files may include two columns of values, one containing the bottle values and the other the CTD values. For GLODAPv2 production the first type of data was subjected to crossover and inversion analysis (see Sect. 3.2.2) and bias-corrected whenever required, irrespective of them being bottle or CTD values.

For the data files including both CTD and bottle values, it was normally the CTD values that gave the complete profile, while the (likely more accurate) bottle values were sampled more sparsely. These data were therefore merged into single "hybrid" salinity and oxygen prior to the crossover and inversion analyses. The consistency between CTD and bottle data from the same cruise was evaluated in this step. When significant offsets existed, the CTD data were corrected using a simple linear fit to the bottle data.

Altogether, seven possible scenarios were defined. The fourth never occurred, but it is included to maintain consistency with material produced during the secondary QC:

1. No data are available: no action needed.

2. No bottle values: use CTD values.

3. No CTD values: use bottle values.

4. Did not occur, case not used.

5. The CTD values do not deviate significantly from bottle values: replace missing bottle values with CTD values.

6. The CTD values deviate significantly from bottle values: calibrate CTD values using linear fit with respect to bottle data and replace missing bottle values with the so-calibrated CTD values.

7. The CTD values deviate significantly from bottle values, and no good linear fit can be obtained for the cruise: use bottle values and discard CTD values.

The number of cases encountered for each scenario is summarized with the other secondary QC results for salinity and oxygen in Sects. 4.3.1 and 4.3.2. This merger step results in the GLODAPv2 data product having only a single column for salinity and a single column for oxygen. The original individual cruise files contain salinity and oxygen (CTD and/or bottle) data as submitted by the data originator.
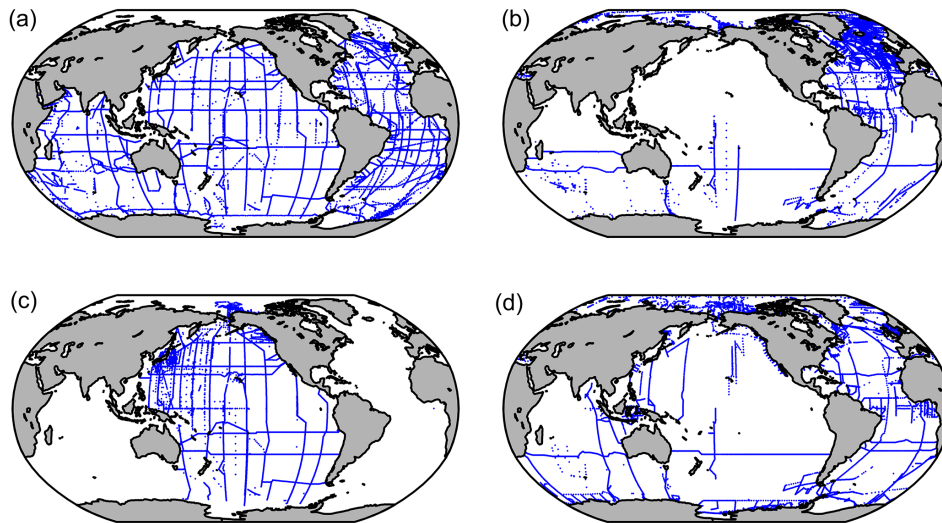
**Figure 1.** Station locations in **(a)** GLODAPv1.1, **(b)** CARINA, and **(c)** PACIFICA, as well as **(d)** locations of stations in GLODAPv2 new to data synthesis.

**Table 1.** WOCE flags in GLODAPv2 exchange format original data files and in product files (briefly for exchange files; for full details see http://geo.h2o.ucsd.edu/documentation/policies/Data_Evaluation_reference.pdf).

| WOCE flag value | Interpretation in original data/product files |
| --- | --- |
| 0 | Not used/interpolated or calculated value |
| 1 | Data not received/not used[a] |
| 2 | Acceptable/acceptable |
| 3 | Questionable/not used[b] |
| 4 | Bad/not used[b] |
| 5 | Value not reported/not used[a] |
| 6 | Average of replicate/not used[c] |
| 7 | Manual chromatographic peak measurement/not used[c] |
| 8 | Irregular digital peak measurement/not used[b] |
| 9 | Sample not drawn/no data |

[a] Flag set to 9 in product files [b] Data are not included in the GLODAPv2 product files and their flags are set to 9. [c] Data are included but flag is set to 2.

### 3.2.2 Crossover and inversion analysis of salinity, oxygen, nutrients, $TCO_2$, and TAlk

The secondary quality control of salinity, oxygen, nutrients, $TCO_2$, and TAlk was carried out through crossover and inversion analyses. This two-step procedure was introduced by Gouretski and Jancke (2001) and Johnson et al. (2001). First, crossover analysis is used to determine cruise-by-cruise offsets by comparing data where two different cruises cross or come close to each other. Next, possible corrections to data are determined in the inversion step. This uses least-squares models (Menke, 1984; Wunch, 1996) to calculate the set of corrections required to simultaneously minimize all cruise-by-cruise offsets. Let $\mathbf{G}$ be the model matrix of size $o \times n$, where $o$ is number of crossovers and $n$ number of cruises, $d$ is the $o$ crossover offsets, and $m$ is the $n$ corrections such that

$$\mathbf{G} \times m = d, \tag{1}$$

then

$$m = \mathbf{G}^T \times (\mathbf{G} \times \mathbf{G}^T)^{-1} \times d. \tag{2}$$

This model is known as simple least squares (SLSQ). Johnson et al. (2001) also introduced the weighted least squares (WLSQ) and weighted damped least squares (WDLSQ) models. The latter takes into account the uncertainties of the crossover offsets and a priori information on expected measurement accuracy of each cruise, while the former only uses the uncertainties of the crossover offsets.

The crossover offsets can be determined in various ways. For GLODAPv1.1 crossover offsets were calculated from stations within $1°$ ($\sim 100$ km) of each other. During CARINA, more elaborate and automated crossover methods

**Table 2.** Initial minimum adjustment limits introduced by CARINA and subsequently used for PACIFICA and GLODAPv2.

| Variable | Minimum adjustment |
|----------|--------------------|
| Salinity | 0.005 |
| Oxygen | 1 % |
| Nutrients | 2 % |
| $TCO_2$ | $4 \, \mu mol \, kg^{-1}$ |
| TAlk | $6 \, \mu mol \, kg^{-1}$ |
| pH | 0.005 |
| CFCs | 5 % |

were developed, for example the "running-cluster" crossover routine, which determines the difference profiles for all station pairs within 200 km from each other. The crossover offset and its standard deviation are then calculated as the weighted mean and standard deviation of all difference profiles. This is highly advantageous for comparing data from repeat sections (Tanhua et al., 2010).

The bias correction of the data included in GLODAPv2 is based on crossover and inversion analyses of the entire unadjusted database. The crossovers offsets were calculated using the running-cluster crossover routine (Tanhua et al., 2010), with data from beneath 2000 m to minimize effects of real variations. Only a fraction of the corrections determined by the crossover and inversion analyses were actually applied to adjust the data. For example, corrections lower than the expected measurement precision – or minimum adjustment limits (introduced by CARINA, Table 2) – were usually not applied, unless the data were very precise and evidence unequivocal. Time trends in the data also give rise to corrections that should not be applied. All corrections were therefore manually evaluated; those that were actually applied are called adjustments. While regional WLSQ inversions of the crossover offsets were used as a first step, they were usually subsequently augmented with customized analyses to determine any underlying patterns and the final adjustments: for example, invoking the assumption that cruises from the WOCE and CLIVAR surveys are of superior quality and may be used as core cruises in a WDLSQ inversion, or carrying out analyses on a subset of data from a given region. An overall strategy was to use of a group of cruises with known high quality to form a cohesive grid against which cruises of unknown quality could be evaluated. Usually only one adjustment per cruise/leg was allowed for each variable – i.e., the underlying assumption for these analyses is that any bias is constant over the duration of the entire cruise/leg. In cases of obvious and significant drift or excessive scatter, all data for the variable at the cruise in question were usually excluded from the product.

In addition to the analyses of the entire and unadjusted dataset, several preliminary analyses were carried out. In particular, (1) GLODAPv1.1 was re-evaluated using the CARINA developed crossover and inversion tools, producing GLODAPv1.2 (not publicly released), and (2) all new data were evaluated on an individual basis using crossovers against a preliminary global reference consisting of GLODAPv1.2 (i.e., re-evaluated GLODAPv1.1), CARINA, and PACIFICA combined, using a software package documented in Lauvset and Tanhua (2015). This is more extensively documented in Appendix A. Familiarity with these preliminary analyses can be useful when accessing the documentation in the GLODAPv2 online Adjustment Table, which is described in Sect. 4.2.

For the Arctic Ocean, crossover and inversion analyses were used in combination with secondary QC procedures described by Jutterström et al. (2010), because of the sparse data and heterogeneous conditions. These include inspection of average property values in individual basins, and inspection of deviations from the values derived using a set of multiple linear regression (MLR) equations specific to the various regions.

### 3.2.3 Quality control of the halogenated transient tracer data

Given the strongly transient nature and low concentration of halogenated transient tracers (CFC-11, CFC-12, CFC-113, and $CCl_4$; CFCs for short) in most deep waters, crossover and inversion analysis is of limited value for these variables. Further, in the previous synthesis products the included CFCs had been subjected to quality control of varying extent:

- In GLODAPv1.1 they were subjected to full primary and secondary QC.

- In CARINA, the CFC data were subjected to full primary and secondary QC in the Arctic and Atlantic regions, but not in the Southern Ocean region.

- No secondary QC was carried out for the PACIFICA CFC data.

Here, secondary QC of the CFC data focused on the 168 new cruises as well as the PACIFICA and Southern Ocean CARINA data. To ensure consistency, the GLODAPv1.1 CFC data were re-evaluated using the same procedures.

The CFC methods included inspection of surface saturation levels, evaluation of the relationships among the tracers from each cruise, and crossover and inversion analysis, all following CARINA protocols (Jeansson et al., 2010; Steinfeldt et al., 2010). Adjustments to CFC-113 and $CCl_4$ data have only been suggested in a few cases as their potential loss by decomposition in the water column renders secondary QC a questionable task. Secondary QC of sulfur hexafluoride ($SF_6$) was not possible because few data were available.

### 3.2.4 Scale conversion and quality control of the pH data

In the three GLODAPv2 predecessors, pH data were treated in various ways:

- pH data were not included in the GLODAPv1.1 product files per se but were used in combination with $TCO_2$ to calculate TAlk whenever that was missing and pH available. The TAlk data were then subjected to secondary QC.

- In CARINA, pH data were subjected to secondary QC and included in the regional product files (Velo et al., 2010). pH calculated from (quality-controlled) $TCO_2$ and TAlk data were also included. The pH data included in the CARINA product files were unified to the seawater scale (SWS) at 25 °C and surface (0 dbar) pressure.

- PACIFICA included measured as well as calculated pH data, like CARINA, but no secondary QC was preformed (Suzuki et al., 2013). The pH data were reported on the total hydrogen ion scale at 25 °C and surface (0 dbar) pressure.

For GLODAPv2 it was decided to include quality-controlled pH on the total hydrogen ion scale at both standard (25 °C and surface (0 dbar) pressure) and in situ (temperature and pressure) conditions. The total hydrogen ion scale was preferred, which has been recommended by Dickson et al. (2007) and by Dickson (2010).

Scale conversion of reported pH was carried out using the procedures of Velo et al. (2010), with the exception that, instead of the Merbach carbonate dissociation constants refitted by Dickson and Millero (Dickson and Millero, 1987; Merbach et al., 1973), the ones of Lueker et al. (2000) were used. These are based on the measurements of Merbach et al. (1973) but made consistent with the total hydrogen ion scale. While the thermodynamic calculations themselves are easily performed with the CO2SYS toolbox (Lewis and Wallace, 1998; van Heuven et al., 2011) with the proper settings, missing or wrong information on scale and/or temperature and pressure conditions of reported data is not infrequent, which makes the scale conversion a challenging task. Hence, all reported pH data were compared with surrounding data for each cruise, as either observed or calculated from $TCO_2$ and TAlk, in order to determine or verify the scale and conditions. This job was somewhat simplified as the pH scale of data from the CARINA and PACIFICA data syntheses has already been determined (Velo et al., 2010; Suzuki et al., 2013).

Crossover analysis of pH was not possible because data only exist for a small fraction of the cruises. Instead, one of three options was selected (in order of increasing complexity):

1. If pH was the only seawater $CO_2$ chemistry variable measured at the cruise in question, or if the measure-

ments had not been carried out at the same stations and/or depths as the other $CO_2$ chemistry data, the pH values were inspected for spread. If this appeared acceptable, then the data were kept but were labeled as not subjected to full secondary QC (−888; see Sect. 4.2).

2. If the pH data were accompanied by (unbiased or bias-corrected) $TCO_2$ *and* TAlk data, the internal consistency of the measurements was evaluated and used to adjust (or in some cases discard) the pH data if these appeared offset.

3. If the pH data were accompanied by (unbiased or bias-corrected) $TCO_2$ *or* Talk (allowing calculations of TAlk *or* $TCO_2$) and collocated with (unbiased or bias-corrected) measured data of TAlk *or* $TCO_2$ of other cruises, crossover analysis was preformed between calculated and measured data of respective cruises. If the calculated TAlk (or $TCO_2$) values were offset from the measured values of the other cruise, the pH data of the cruise of interest were adjusted to minimize this offset (provided that the scatter in the pH data was acceptable; otherwise, they were discarded).

The NBS scale for pH measurements has large inherent uncertainties (Dickson, 1984). Recognizing this, such data have not been included in the data product unless passing full secondary QC, criteria 2 or 3.

## 4 GLODAPv2 secondary QC results and adjustments

### 4.1 Preservation of real variability

The risk of removing real signals of variability present in the data was recognized throughout secondary quality control, in particular because the crossover and inversion is an objective method that does not discriminate between real difference and measurement bias. By only using data deeper than 2000 m for crossover analyses, this risk was reduced, but in some regions deep-water time trends are expected to occur over the decadal timescales considered. Therefore, each correction suggested by the crossover and inversion analysis was scrutinized. Whenever doubt existed, adjustments were not applied, in particular in regions of strong variability (such as the Nordic Seas overflow), or when time trends were detected or suspected. As an example of a method of preserving trends, Fig. 2 shows one type of figure used to evaluate the crossover offsets. This particular cruise (18HU19960512) is an occupation of WOCE line AR07W in the Labrador Sea, and the crossover offsets indicate a bias in $TCO_2$ of $-6 \, \mu mol \, kg^{-1}$, and the inversion suggested a correction of the same magnitude. However, plotting the crossover offsets vs. time as in Fig. 2 clearly reveals the strong $TCO_2$ trend. The gradual decrease in the offsets implies a temporal $TCO_2$ increase at depth rather than a negative bias (as

implied by the mean of the offsets). This is consistent with anthropogenic $CO_2$ uptake and the deep mixing that occurs in this region (Yashayaev, 2007). Cruise 18HU19960512 is not appreciably offset from contemporaneous cruises. No adjustments were applied to these data.

## 4.2 The adjustment table

The results of the secondary QC analyses were entered into the online GLODAPv2 Adjustment Table hosted at GEO-MAR in Kiel, Germany. This is similar in form and function to the Adjustment Table used in CARINA (Tanhua et al., 2010). A permanent, non-editable version of this Adjustment Table is available at http://glodapv2.geomar.de. Table 3 summarizes the type of entries in the Adjustment Table. In contrast to CARINA, the GLODAPv2 Adjustment Table does not include an entry for each crossover; the large number of crossover locations made this unmanageable. Even though at many locations either of the involved cruises may not have the required deep, high-quality data, the number of successfully assessed crossovers ranges from ~ 3400 for TAlk to ~ 12 100 for salinity. Hence, there is one entry per cruise, providing access to summary figures from the crossover analysis and the magnitude and justification of any recommended adjustments. Further details are provided in Appendix B.

## 4.3 Secondary QC summary

Data from 734 cruises were subjected to secondary QC. For 10 of these the secondary QC revealed that most if not all of the data were of unacceptable quality. Further, for these 10 cruises, better quality data from the same region were available, and they were therefore not included in the final product files. The original data from these 10 cruises are available through the Cruise Summary Table (CST, see Sect. 5.1) at CDIAC, at the very end of the CST. They have been assigned cruise number 9999 and secondary QC results are not included in the summaries below.

GLODAPv2 thus includes data from 724 cruises. These were split into a total of 780 cruises/legs/station ranges during secondary QC. This is partly because most cruises consisting of individual legs were analyzed on a per-leg basis (Table 4) in order to take into account potential changes in personnel, equipment, and procedures during their execution and partly because four cruises were adjusted on a per-station range basis as a result of obvious bias in one or several variables for specific parts of, but not the entire, cruise (these are 74AB20050501, 316N19831007, 06GA20000506, and 06AQ19920521). Respecting this distinction, we therefore refer in the following summary to analyzed "entries" instead of cruises, where an entry is an entire cruise (the large majority), leg, or station range.

Application of adjustments was done with the aim of reducing the deep-water offsets between the many entries. A measure of this reduction is given by the "internal consis-

tency improvement". This is the decrease in the weighted mean of the absolute offsets of all crossovers between (i) the unadjusted data (after primary QC) and (ii) the adjusted data (after secondary QC) (Tanhua et al., 2010). This is not the only means of quantifying improvement, but it is a good compromise in terms of implementation, clarity, and brevity. Certainly, improvement will be different between geographical regions, vessels, labs, and countries, with smallest improvements generally observed between the large hydrographic repeat surveys. Conversely, appreciable local improvements are observed for smaller cruises run by groups without a primary focus of delivering climate-quality data (e.g., biological process studies). While the interesting nature of these details is recognized, Table 5 only provides the improvements per ocean basin and for the full world ocean. The relative improvement for nutrients, $TCO_2$, and TAlk is higher than for salinity. Salinity accuracy was quite high for most cruises already. The internal consistency of all variables subjected to secondary QC has been increased significantly.

Summaries of the secondary QC actions are presented in Tables 6 and 7. Figure 3 summarizes the distribution of the adjustments that were applied. Details on the secondary QC results are presented per variable in the following subsections.

### 4.3.1 Salinity merging and adjustment summary

All 780 entries came with salinity data (Table 6). Prior to the crossover and inversion analyses, the CTD and bottle salinity values were merged as described in Sect. 3.2.1. The different actions in this respect are summarized in Table 8.

After the data were merged, they were subjected to crossover and inversion analyses. For 162 of the entries, full secondary QC could not be carried out, and data from 6 entries were deemed to be of too poor a quality for inclusion in GLODAPv2 (Table 6). Typically, these showed large and depth-dependent offsets and/or unrealistic scatter compared to background data. Of the remaining 612 entries, the salinity data from 41 were found clearly biased, warranting an adjustment (Table 6).

Adjustments smaller than the initial threshold have only been applied to five entries, while the bulk of the adjustments applied are between 0.005 and 0.010 (Table 7, Fig. 3g). The largest negative and positive adjustments applied are −0.025 and +0.025. Application of the adjustments increased the global consistency of the salinity data from 0.0041 to 0.0031 as evaluated from the weighted mean of the absolute crossover offsets (Table 5).

### 4.3.2 Oxygen merging and adjustment summary

Of the 780 entries, 722 had oxygen data (Table 6). Data of CTD and chemically determined oxygen concentration were merged into a single, "hybrid" variable using procedures in Sect. 3.2.1, with results summarized in Table 8. Crossover,
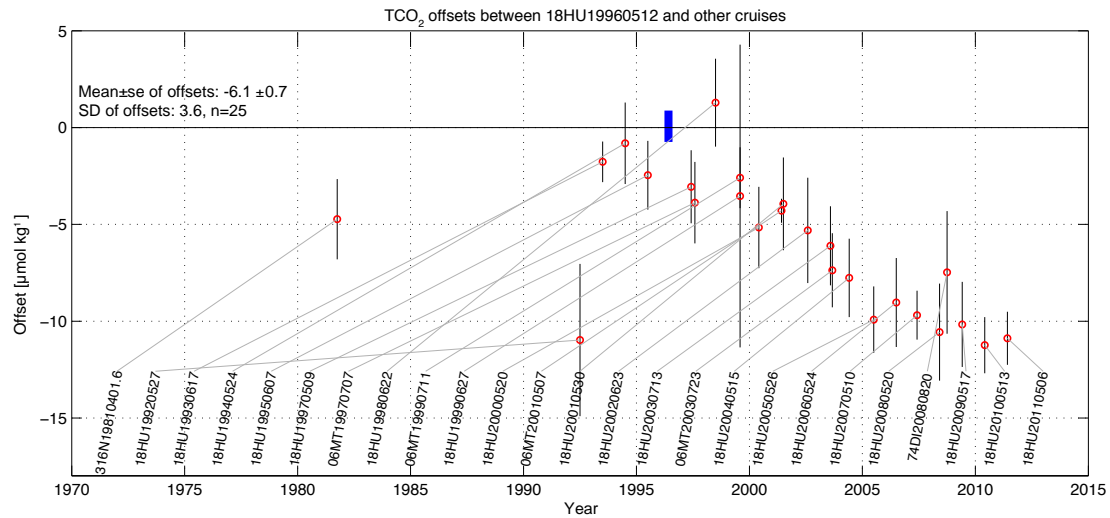
**Figure 2.** Summary figure used to evaluate $TCO_2$ crossover offsets of WOCE repeat section AR07W cruise 18HU19960512 in the Labrador Sea. The figure shows the 25 crossover offsets that were determined, sorted by time. 18HU19960512 is indicated by the blue line. Negative values mean that 18HU19960512 $TCO_2$ values are lower than those of the comparison cruise.

**Table 3.** Possible values in the Adjustment Table and their interpretation.

| Value | Interpretation |
| --- | --- |
| −999 | No data exist for this variable for the cruise in question. |
| −888 | The data appear to be of good quality but their nature, being from shallow depths, without crossovers or similar, prohibits full secondary QC. |
| −777 | The data are of poor quality and excluded from the data product. |
| −666 | The data have not been quality-controlled, are of uncertain quality, and suspended until full secondary QC has been carried out. They are not included in the product. |
| 0/1* | The data are of good quality, consistent with the rest of the dataset and should not be adjusted. |
| Any other number | The data are of good quality but are biased: adjust by adding (for salinity, $TCO_2$, TAlk, pH) or multiplying (for oxygen, nutrients, CFCs) the number in the Adjustment Table. |

* The value of 0 is used for variables with additive adjustments (salinity, $TCO_2$, TAlk, pH) and 1 for variables with multiplicative adjustments (for oxygen, nutrients, CFCs). This is mathematically equivalent to "no adjustment" in each case.

inversion, and subsequent adjustment for bias minimization were performed on this hybrid oxygen. A total of 378 of the entries were deemed to be accurate to within the minimum adjustment limits, and thus did not require an adjustment (Table 6). A total of seven applied non-zero adjustments were smaller than the threshold of 1 % (Table 7, Fig. 3f). These necessarily were cruises with sufficiently high precision so that such small bias could be observed beyond doubt. Almost half of the non-zero adjustments were between 1 and 2 %, while the other half of the applied non-zero adjustments (99 cruises) was greater than 2 % (Table 7, Fig. 3f). The largest adjustments applied were −7.2 and +11 %. This rather tight distribution is testimony to the high accuracy generally achieved in oxygen measurements.

### 4.3.3 Nitrate adjustment summary

Nitrate data were available for 709 of the 780 entries (Table 6). Of these, data from 42 were of insufficient quality for inclusion, data from 137 could not be fully quality-controlled, and data from 530 received successful secondary QC (Table 6). Of these, 380 were accepted to be accurate and 150 entries were adjusted. Of the applied adjustments, 50 (i.e., 33 %) are beneath the initial 2 % limit, while 49 % are between 2 and 4 % (Table 7, Fig. 3b). The high fraction receiving small adjustments illustrates the high precision commonly attained with nitrate analysis. The secondary nitrate QC was performed without notable peculiarities. Secondary QC markedly increased the internal consistency of the nitrate data (Table 5). This suggests (i) that the nitrate data are generally highly precise (while not necessarily accurate), and (ii) that our assumption that each entry suffered from not more than one, constant bias is generally valid. Very

**Table 4.** Multi-leg cruises in GLODAPv2 that received secondary quality control on a per-leg basis but are included as a single cruise in the product files.

| Cruise number | EXPOCODE | Expocodes of individual legs |
| --- | --- | --- |
| 102 | 18DD19940906 | 18DD19940906; 18DD19941013 |
| 236 | 316N19720718 | 316N19720718.1; 316N19720718.2; 316N19720718.3; 316N19720718.4; 316N19720718.5; 316N19720718.6; 316N19720718.7; 316N19720718.8; 316N19720718.9 |
| 237 | 316N19810401 | 316N19810401; 316N19810416; 316N19810516; 316N19810619; 316N19810721; 316N19810821; 316N19810923 |
| 238 | 316N19821201 | 316N19821201; 316N19821229; 316N19830130 |
| 242 | 316N19871123 | 316N19871123.1; 316N19871123.2; 316N19871123.3; 316N19871123.4; 316N19871123.5; 316N19871123.6 |
| 243 | 316N19920502 | 316N19920502; 316N19920530; 316N19920713 |
| 255 | 316N19950829 | 316N19950829; 316N19950930 |
| 257 | 316N19951202 | 316N19951202; 316N19951230 |
| 268 | 318M19730822 | 318M19730822; 318M19730915; 318M19731007; 318M19731031; 318M19731204; 318M19740102; 318M19740205; 318M19740313; 318M19740412; 318M19740513 |
| 269 | 318M19771204 | 318M19771204; 318M19771216; 318M19780128; 318M19780307; 318M19780404 |
| 273 | 318M20091121 | 318M20091121; 318M20100105 |
| 298 | 325019850330 | 325019850330; 325019850504 |
| 319 | 32MW19890206 | 32MW19890206; 32MW19890309; 32MW19890402 |
| 338 | 33MW19930704 | 33MW19930704.1; 33MW19930704.2 |
| 370 | 35MF19850224 | 35MF19850224; 35MF19860401; 35MF19870114 |
| 439 | 49HH19910813 | 49HH19910813; 49HH19910917 |
| 486 | 49NZ20030803 | 49NZ20030803; 49NZ20030909 |
| 497 | 49NZ20051031 | 49NZ20051031; 49NZ20051127 |
| 507 | 49NZ20090410 | 49NZ20090410; 49NZ20090521 |

**Table 5.** Improvements resulting from the GLODAPv2 quality control split out per basin and for the global dataset. The numbers in the table are the weighted mean of the absolute offsets of all crossovers of unadjusted and adjusted data, respectively. $n$ is the total number of valid crossovers in the global ocean for the variable in question.

| | Arctic | | Atlantic | | Indian | | Pacific | | Global | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | unadj | adj | unadj | adj | unadj | adj | unadj | adj | unadj | adj | $n$ (global) |
| Salinity [ppm] | 4.1 | => 3.8 | 7.1 | => 5.0 | 2.7 | => 1.6 | 2.4 | => 1.9 | 4.1 | => 3.1 | $\sim 12\,100$ |
| Oxygen [%] | 1.3 | => 1.0 | 1.7 | => 0.8 | 1.4 | => 0.7 | 1.7 | => 1.1 | 1.7 | => 0.9 | $\sim 10\,900$ |
| Nitrate [%] | 4.2 | => 1.6 | 2.7 | => 1.7 | 1.8 | => 1.0 | 1.0 | => 0.8 | 1.7 | => 1.2 | $\sim 9500$ |
| Silicate [%] | 8.2 | => 3.5 | 4.8 | => 2.7 | 2.8 | => 1.5 | 1.9 | => 0.9 | 2.8 | => 1.7 | $\sim 8300$ |
| Phosphate [%] | 4.8 | => 2.5 | 4.2 | => 2.5 | 2.7 | => 1.1 | 1.5 | => 1.0 | 2.2 | => 1.3 | $\sim 8800$ |
| $TCO_2$ [µmol kg$^{-1}$] | 6.1 | => 3.5 | 4.4 | => 2.9 | 4.5 | => 2.2 | 4.0 | => 2.3 | 4.4 | => 2.6 | $\sim 5800$ |
| TAlk [µmol kg$^{-1}$] | 8.2 | => 3.5 | 7.5 | => 3.5 | 5.2 | => 3.3 | 3.4 | => 2.2 | 5.8 | => 2.8 | $\sim 3400$ |

few exceptions were encountered that exhibited either strong instrumental drift or strong station-to-station variability.

The southeastern corner of the Pacific (30–90° S, 120–70° W) is a region of particular uncertainty for nitrate. The data do not form a cohesive network with an unambiguous "baseline". An important source of uncertainty here is drift of the nitrate measurements from 33RO20071215 and/or 31DS19940124.

### 4.3.4 Silicate adjustment summary

Silicate data were available for 678 of the 780 entries (Table 6). The silicate data of 33 entries were found to be of poor quality, exhibiting excessive scatter, large offsets, drift, or a combination of these. For 255 entries the silicate data were considered to be accurate to within the uncertainty of our methods, while data from 264 entries were adjusted (Table 6). This is almost 40 % of the entries, making silicate the
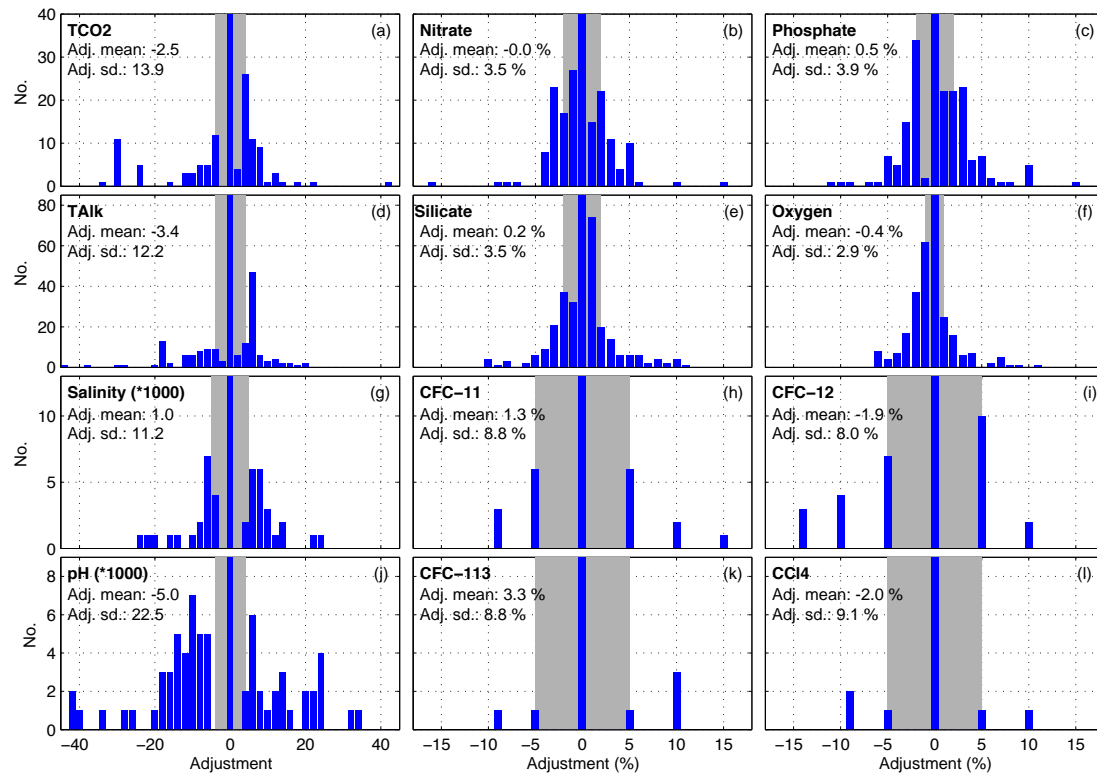
**Figure 3.** Size distribution of applied adjustments for each core variable that received secondary QC. Gray areas depict the initial minimum adjustment limit. Entries for which data could not be secondary quality-controlled or were considered of insufficient quality for our product are excluded from this figure.

**Table 6.** Summary of secondary QC actions per variable for the 780 non-dismissed entries.

| | Salinity | Oxygen | Nitrate | Silicate | Phosphate | TCO$_2$ | TAlk | pH | CFC-11 | CFC-12 | CFC-113 | CCl$_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| With data | 780 | 722 | 709 | 678 | 688 | 602 | 465 | 259 | 273 | 270 | 105 | 72 |
| No data | 0 | 58 | 71 | 102 | 92 | 178 | 315 | 521 | 507 | 510 | 675 | 708 |
| Unadjusted[a] | 571 | 378 | 380 | 255 | 282 | 332 | 180 | 59 | 208 | 207 | 57 | 33 |
| Adjusted[b] | 41 | 207 | 150 | 264 | 163 | 104 | 150 | 77 | 26 | 19 | 6 | 5 |
| −888[c] | 162 | 127 | 137 | 126 | 184 | 151 | 106 | 67 | 30 | 30 | 15 | 14 |
| −666[d] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 47 | 0 | 0 | 0 | 0 |
| −777[e] | 6 | 10 | 42 | 33 | 59 | 15 | 29 | 9 | 9 | 14 | 27 | 20 |

[a] The data are included in the data product file as is, with a secondary QC flag of 1 (Sect. 5.2). [b] The adjusted data are included in the data product file with a secondary QC flag of 1 (Sect. 5.2). [c] Data appear of good quality but have not been subjected to full secondary QC. They are included in data product with a secondary QC flag of 0 (Sect. 5.2). [d] Data are of uncertain quality and suspended until full secondary QC has been carried out; they are excluded from the data product. [e] Data are of poor quality and excluded from the data product.

most frequently adjusted variable in GLODAPv2. The single reason for this is that the silicate data of a large fraction of Pacific entries were adjusted to remove an average 2 % offset in silicate observed between the US and Japanese entries from this region. This systematic "country-specific" bias was revealed by the crossover and inversion analyses. Figure 4a and b present silicate biases between US and Japanese – uncorrected – cruises in the Pacific, from a dedicated inversion analysis of these data. It is evident from these that US silicate data tend to be approximately 2 % higher than the Japanese values. This systematic bias has been hinted at by results from laboratory comparison exercises (Aoyama et al.,

2010; S. Becker, Scripps, personal communication, 2014; K. Bakker, NIOZ, personal communication, 2014). Kanso reference material for nutrients in seawater (RMNS) samples were analyzed on several of the cruises involved, but the results were not consistently used for correction since the assigned values had not yet been certified (S. Becker, personal communication, 2014). While it would appear reasonable to assume that one country's data should receive a 2 % correction, the data and evidence are inadequate to determine which. The biases determined by the inversion provide no information in this respect: the lower mean bias of Japanese cruises is just a consequence of the zero-sum constraint of

**Table 7.** Summary of the distribution of applied adjustments per variable, in number of adjustments applied for each variable.

|  | adj. < limit | limit ≤ adj. < 2 × limit | 2 × limit ≤ adj. |
|---|---|---|---|
| Salinity | 5 | 22 | 14 |
| Oxygen | 7 | 101 | 99 |
| Nitrate | 50 | 73 | 27 |
| Silicate | 113 | 95 | 56 |
| Phosphate | 31 | 92 | 40 |
| $TCO_2$ | 8 | 51 | 45 |
| TAlk | 37 | 76 | 37 |
| pH | 0 | 25 | 52 |
| CFC-11 | 0 | 17 | 9 |
| CFC-12 | 0 | 12 | 7 |
| CFC-113 | 0 | 2 | 4 |
| $CCl_4$ | 0 | 2 | 3 |

**Table 8.** Summary of salinity and oxygen merger actions for the 780 non-dismissed entries subjected to secondary QC.

| Case | Description | Salinity | Oxygen |
|---|---|---|---|
| 1 | No data are available, no action needed. | 0 | 58 |
| 2 | No bottle values present: use CTD-derived values. | 77 | 21 |
| 3 | No CTD values present: use bottle data. | 295 | 520 |
| 4 | (Case not used) | – | – |
| 5 | The CTD values do not deviate significantly from bottle values: replace missing bottle values with CTD values. | 264 | 99 |
| 6 | The CTD values deviate significantly from bottle values: calibrate CTD values using linear fit with respect to bottle data and replace missing bottle values with the so-calibrated CTD values. | 141 | 62 |
| 7 | The CTD values deviate significantly from bottle values, and no good linear fit can be obtained for the cruise: use bottle values and discard CTD values. | 3 | 20 |

the inversion combined with the larger number of Japanese cruises. The sum of all corrections suggested by the inversion has to be zero and the crossover and inversion tends to conclude that the most frequently measured value is the least biased one. In this case it is the deep silicate measured at Japanese cruises, since there are more Japanese than US cruises; hence, these come out with smaller mean bias than the US cruises, while the "true" value is unknown. Thus, to remedy this inconsistency the Japanese data were pre-adjusted by +1 % and US data by −1 %. This removed the systematic difference and a clear baseline emerged (Fig. 4c and d).

After this pre-adjustment, the set of Pacific silicate data were subjected to regular crossover and inversion analysis to obtain the total required correction. Note that the choice of splitting the difference between the US and Japanese efforts may result in the Pacific Ocean data product being – at least – between −1 and +1 % biased against the "true" level. However, for the purposes of this data product such

residual, systemic, bias between the Pacific and the other major ocean basins (Atlantic, Arctic, Indian) is currently not seen as problematic. Nonetheless, reconciliation between the Japanese and US results should be a high priority for the nutrient analytical community.

For the South Atlantic and Indian basins, crossover and inversion were performed without notable incidents.

Bias minimization of silicate was rather challenging in the North Atlantic Ocean, where silicate values may range from near zero at the ocean surface to well over $50 \, \mu mol \, kg^{-1}$ at depth. At the low end of that range, additive calibration biases manifest themselves in addition to the multiplicative ones the methods were designed to deal with (e.g., residual silicate in the "nutrient-free" seawater used for standards preparation). Additionally, samples with nominal silicate values over $\sim 50 \, \mu mol \, kg^{-1}$ tend to be very sensitive to freezing, which can decrease the measured concentration by up to 15 % due to polymerization (Karel Bakker, NIOZ, personal communication, 2014). Samples with lower silicate concen-
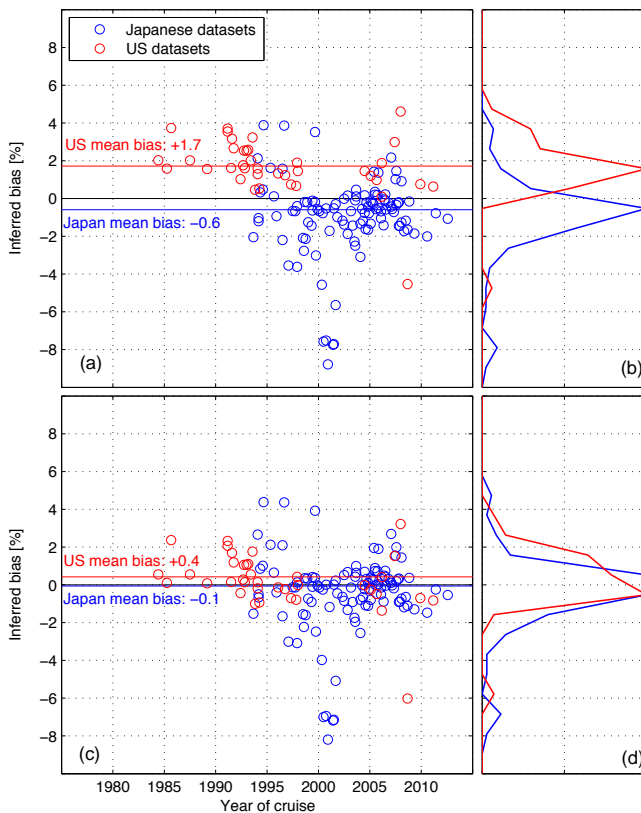
**Figure 4.** Silicate biases between US and Japanese efforts before **(a, b)** and after **(c, d)** pre-adjustments (US: $-1\%$; Japan: $+1\%$) were applied to the data (see main text for details). Data from "Line P" and many small-scale cruises in the variable Kuroshio region were excluded from this analysis. Red and blue horizontal lines indicate countries' approximate mean offsets.

tration seem to not be affected by freezing. Freezing was occasionally suspected (and then generally confirmed) to have been used on cruises, forcing arbitrary removal of data, and complicating the automated crossover analysis. Although the average offset for silicate at crossovers has been reduced in the North Atlantic Ocean, the solution there is not particularly satisfying and a more thorough assessment is expected to be able to substantially improve our results locally.

Overall, the application of the adjustments improved the global consistency of the silicate data by more than a percent, but with regional differences. In the Arctic the secondary QC has been in particular effective; consistency has been improved from 8.2 to 3.5 % (Table 5). This may also be due to removal of obviously poor data from some cruises in this region (see Fig. 7 for regional distributions of data).

### 4.3.5 Phosphate adjustment summary

A total of 688 entries included phosphate data (Table 6). Of these, data from 59 were found to be of too poor a quality for inclusion in the product. Adjustments were applied to 163 entries. Data from 184 entries could not be adequately checked with our routines.

Of the 163 adjusted entries, 31 (highly precise) received adjustments smaller than the threshold; 132 entries had larger adjustments (Table 7, Fig. 3c), with the largest being about $\pm 12\%$.

### 4.3.6 TCO$_2$ adjustment summary

TCO$_2$ was measured on 602 of the 780 entries (Table 6). The quality of 15 were too poor to be retained. Data from 151 were not fully quality-controlled, and of the remainder, 332 entries were accurate within the uncertainty of our methods and 104 were adjusted. The minimum TCO$_2$ adjustment was initially set to $4\,\mu mol\,kg^{-1}$. For eight very precise entries a smaller adjustment was applied (Table 7, Fig. 3a). A few very large adjustments were applied, generally to historic entries (e.g., GEOSECS).

Globally the consistency has improved by $1.8\,\mu mol\,kg^{-1}$, and by much more in some regions (Table 5). The largest improvement is observed for data from the Arctic region.

### 4.3.7 TAlk adjustment summary

In total, 465 entries had TAlk data; 106 of these could not be subjected to full secondary QC and were set to $-888$ (Table 6). Of the remainder, 29 were deemed too poor for inclusion, 180 were of good quality and unbiased, and 150 needed adjustment. The initial minimum allowable adjustment was $6\,\mu mol\,kg^{-1}$ (Table 2). About 75 % of the applied adjustments are equal to or larger than this (Table 7, Fig. 3d). TAlk is the second most frequently adjusted variable in GLODAPv2 with 32 % recommended for adjustment. This was the result of a bias identified in Japanese Pacific cruises. Following crossover and inversion analysis a very clear separation was observed between the US entries and (most of) the Japanese Pacific entries. This is illustrated in Fig. 5a and b. Japanese data appear consistently too low, while US data appear consistently high. The offset between US and Japanese labs appears to exist throughout the era of measurements, although too few data exist in the latest 10 years to be sure. Typically, the available metadata for the Japanese cruises were sparse and did not include information on traceability to CRMs. However, the six or so Japanese results after 2005 that were found to *not* require an adjustment relative to the US were all CLIVAR/GO-SHIP lines. A possible explanation is that these Japanese CLIVAR/GO-SHIP measurements have been standardized against the certified reference material provided by A. G. Dickson (Dickson, 2001; Dickson et al., 2003), whereas the smaller Japanese lines have used a
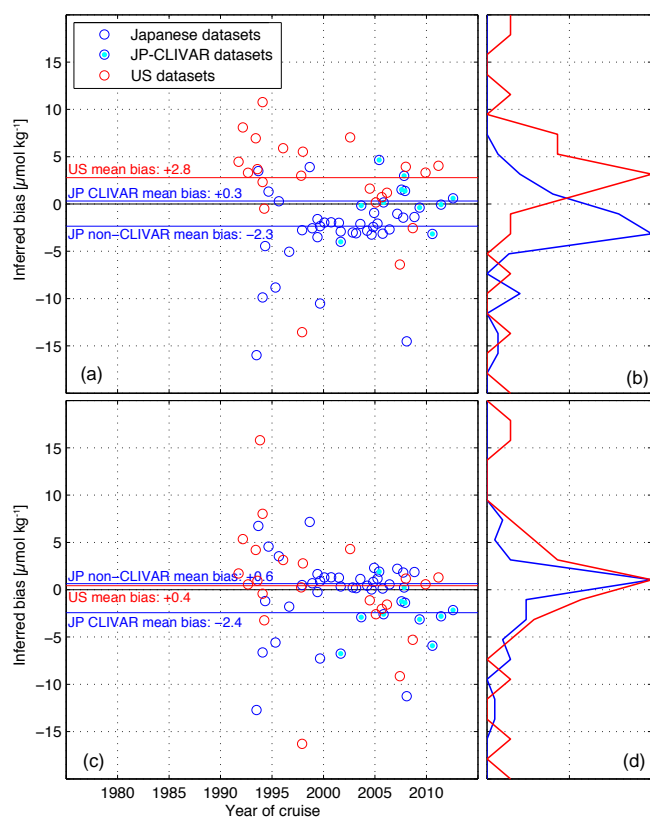
**Figure 5.** TAlk biases between US and Japanese efforts before **(a, b)** and after **(c, d)** pre-adjustments were applied to TAlk data of Japanese non-CLIVAR cruises (see main text for details). Circles represent the biases (inferred by the GLODAPv2 inversion method) of TAlk measurements of individual cruises in the Pacific Ocean. Data of "Line P" and many small-scale cruises in the variable Kuroshio region were excluded from this analysis. Red and blue horizontal lines indicate countries' approximate mean offset. For Japanese data the values are split into cruises that were or were not part of CLIVAR.

different method of standardization. Some additional information may be gleaned from assessing these results on a per-ship or per-lab basis rather than per-country, but such analyses have not been performed. As has been noted earlier, the crossover and inversion method does not provide any information about which set of data is the correct one. For silicate the difference was split between Japanese and US cruises, in the absence of additional information. In this case, based on the *documented* traceability to CRM for the US cruises, a $+6 \, \mu mol \, kg^{-1}$ pre-adjustment was applied to the TAlk of the Japanese non-CLIVAR Pacific cruises. This is the reason for the peak in the distribution of applied adjustments visualized in Fig. 3d. The consistency of US and Japanese Pacific TAlk data after the pre-adjustment is shown in Fig. 5c and d. Note that the average corrections are now essentially identical for both countries and close to zero.

After this pre-adjustment, the set of Pacific TAlk data were subjected to regular crossover and inversion analysis to obtain the total required correction.

### 4.3.8 pH adjustment summary

A total of 259 entries included pH data (Table 6). Of these, 59 were found accurate, while 77 were adjusted; 67 could not be fully quality-controlled but are thought to be accurate (−888). Data from 47 cruises were suspended and further QC is required. These are all data supplied on the NBS scale (Sect. 3.2.4).

### 4.3.9 CFC-11, CFC-12, CFC-113, and CCl$_4$ adjustment summary

During WOCE and CLIVAR, CFC-11 and CFC-12 were commonly measured, whereas data for CFC-113 and CCl$_4$ are less abundant. This is reflected in the number of entries with CFC data available in GLODAPv2 (Table 6: 273/270 for CFC-11/CFC-12, but only 105 for CFC-113 and 72 for CCl$_4$). The range of CFC concentrations in deep water spans about 2 orders of magnitude ($\sim 0.01$ to $1.0 \, pmol \, kg^{-1}$). Areas with higher concentration are often subject to temporal variability, as they are close to the deep-water formation areas. In regions with less temporal variability, CFC concentrations are low, and a relative error of $\sim 10\%$ might still be smaller than the accuracy of the data. Consequently, data adjustment is more difficult than for other variables. The threshold for adjustment was set to 5% as in CARINA. As a result, only about 10% of the CFC data have been corrected, less than for the other quantities (Table 6). Quality control of CFC-113 and CCl$_4$ is even more difficult. For these two, adjustments have only been applied if repeat cruises from the same area were available and the data from these repeats were clearly inconsistent. For applied CFC-113 and CCl$_4$ adjustments, about 65% are larger than 10%, or 2 times the limit. Only about 35% of adjustments for CFC-11 and CFC-12 are that large (Table 7 and Fig. 3h, i, k, l).

## 5 GLODAPv2 product access and description

GLODAPv2 consists of three components: the original data, the bias-corrected product files, and the mapped climatology. They are available at CDIAC (http://cdiac.ornl.gov/oceans/GLODAPv2/). The original data and product files are described here, while the mapped climatology is described by Lauvset et al. (2016).

### 5.1 Original data

GLODAPv2 includes original data from 724 cruises, and access and documentation for individual cruise files are provided through the CST at the GLODAPv2 web page at CDIAC. The 724 cruises may consist of several legs, and

in a few cases multiple cruises have been merged. Among other things, the CST includes a column that lists individual components of multi-leg cruises analyzed per leg (Sect. 4.3).

The content of the original data files is as received from the originator, but the files have been updated to WOCE exchange format (Swift and Diggs, 2008) whenever required. File headers, listing essential information on cruises and the analytical procedures, were generated for all except the PACIFICA cruises. No bias adjustments were applied to the data in these files, and they also contain the oxygen and salinity data as submitted – i.e., no merged bottle and CTD values are included.

Each cruise and data file is uniquely identified with its GLODAPv2 cruise number and its EXPOCODE. Known aliases are also specified in the CST. The GLODAPv2 cruise numbers were assigned sequentially after sorting by EXPOCODE. EXPOCODES were constructed by combining the NODC platform code (http://www.nodc.noaa.gov/General/NODC-Archive/platformlist.txt) with the sailing date of the cruise in the format YYYYMMDD. In a few cases when the sailing date could not be determined, the date of the first sampling was used. After the inception of GLODAPv2, the responsibility for platform code assignment was assumed by the ICES data center (http://ices.dk/marine-data/vocabularies/Pages/default.aspx). A few differences exist between the two sets of codes; the older or better-known code was preferred in these cases.

Note that, for the following time series or campaigns, the data have not been segmented into individual cruises but instead maintained as collections under a single EXPOCODE, to ease record keeping: the EGEE, GIFT, Iceland Sea, Irminger Sea, Kerfix, OWS Mike, and SWITCHYARD time series (assigned EXPOCODES are 35A820050607, CARBOGIB2005, IcelandSea, IrmingerSea, 35UCKERFIXTS, 58P320011031, and ZZIC2005SWYD), and the OMEX-1 Nordic Seas, OMEX-1 North Atlantic, and OMEX-2 North Atlantic campaigns (assigned EXPOCODES are OMEX1NS, OMEX1NA, and OMEX2NA).

All concentration units are those set for WOCE and used in earlier data products. In particular, any oxygen and nutrient concentrations reported in milliliters or micromoles per liter were converted to micromoles per kilogram ($\mu$mol kg$^{-1}$). The default procedure for nutrients was to use seawater density at reported salinity, an assumed lab temperature of 22 °C, and a pressure of 1 atm. The error made by an actual lab temperature deviating up to 5 °C from the assumed 22 °C is insignificant. For the milliliter to micromole conversion for oxygen, the factor 44.66 was used, derived using the ideal gas law at standard temperature and pressure, corrected for the non-ideal behavior of oxygen, while for the per-liter to per-kilogram conversion potential density was used whenever draw temperatures were unavailable.

Note also that the original TTO-NAS data file contains the potentiometrically measured TCO$_2$ and non-adjusted TAlk, while the data product contains the adjusted and calculated TAlk and TCO$_2$ derived using the recommendations by Tanhua and Wallace (2005).

WOCE quality flags (Table 1) have been applied throughout. Any questionable or bad data identified during primary QC are included and flagged accordingly in these files. However, note that whenever data from an entire cruise were found to be bad following secondary QC, they have not necessarily been flagged as such in the individual data files. However, this may be noted in the metadata, and is definitely noted in the Adjustment Table at GEOMAR. The Adjustment Table record for each specific cruise can be directly accessed via the hyperlink that appears in the rightmost column in the CST. All users of the individual cruise data files are encouraged to respect the WOCE flags that have been applied and also to consult the notes in the Adjustment Table and all available metadata before any analyses are carried out. Metadata for each cruise is usually contained in the header of each exchange file and/or in the "Metadata" link in the CST. These two sources can be complementary. For many cruises, access to copies of written cruise reports is provided through the CST, as well as references to relevant scientific publications.
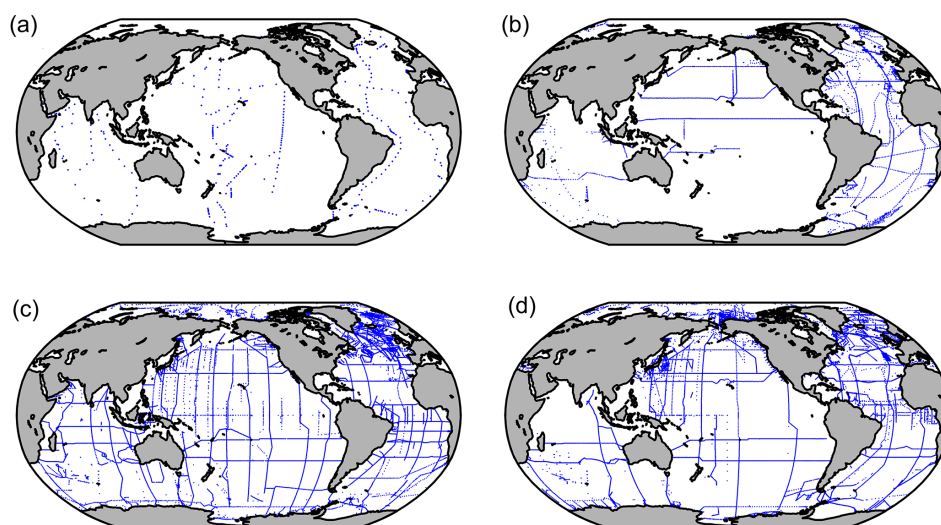
## 5.2　Product files

The GLODAPv2 data product is available as one global file containing all 724 cruises, with bias minimization adjustments applied to the data. Cruises are in alphabetical order of EXPOCODES. In addition, four regional subset files have been produced: one each for the Arctic, Atlantic, Pacific, and Indian oceans. The global decadal coverage of GLODAPv2 is given in Fig. 6, and that of each regional file in Fig. 7. The files are available as comma-separated ASCII files (*.csv) and as binary MATLAB format files (*.mat; MATLAB, 2015).

There is no data overlap in the regional files – i.e., a single cruise can only appear in one of the regional files even though some cruises cover multiple basins. In the product files each cruise is identified using its unique GLODAPv2 cruise number to avoid text strings in the data files, i.e., EXPOCODES are not included. In the global file, cruise numbers increase consecutively, while cruise numbers in the regional subset files increase but are not consecutive. A lookup table is provided with the data files to facilitate matching of cruise number and EXPOCODE. In the MATLAB version of the product files, a structure array named "expocodes" is available, containing all 724 EXPOCODES.

The product files were prepared following the same general procedures used for GLODAPv1.1 (Key et al., 2004; Sabine et al., 2005) and CARINA (Key et al., 2010) and are only summarized here:

1. If temperature was missing, then all data for that record were set to −9999/NaN and their flags to 9. The same was done when pressure/depth was missing, except for the 911 records that were associated with Niskin bottle

**Figure 6.** Station locations in the GLODAPv2 data product for data obtained during **(a)** the 1970s, **(b)** the 1980s, **(c)** the 1990s, and **(d)** 2000s and beyond.
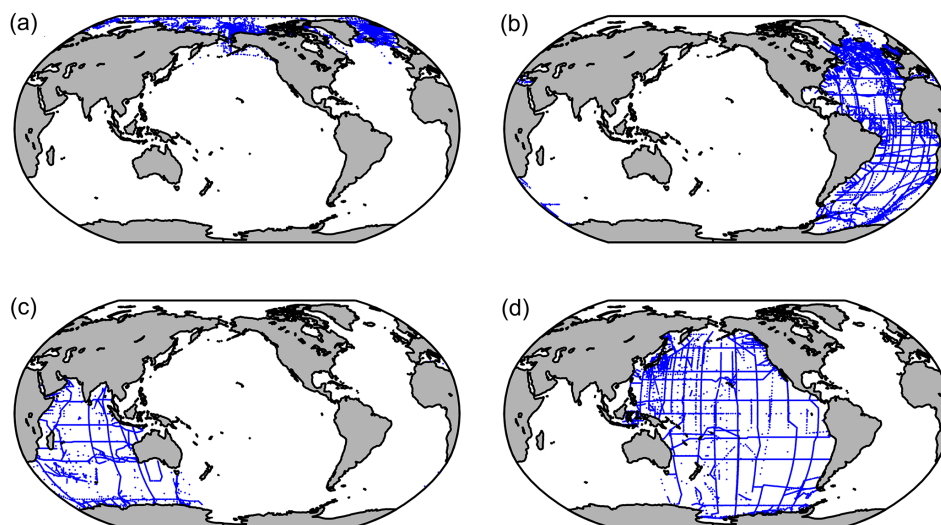


**Figure 7.** Locations of data included in the **(a)** Arctic, **(b)** Atlantic, **(c)** Indian, and **(d)** Pacific ocean product files. Note the minor "spillover" near the boundaries.

number "0" and had actual data. These were considered to be surface samples collected at station and were retained. Their pressure and depth were set to 0.

2. For both oxygen and salinity, any reported CTD and bottle values were merged following procedures summarized in Sect. 3.2.1.

3. In some cases nitrate plus nitrite was reported instead of nitrate. Whenever explicit nitrite concentrations were reported, these were subtracted to get the nitrate values; otherwise, $NO_3 + NO_2$ was simply renamed to $NO_3$.

4. When bottom depths were not given, they were approximated as the deepest sample pressure $+ 10$ or extracted from the bathymetry of the TerrainBase (National Geophysical Data Center/NESDIS/NOAA/U.S. Department of Commerce, 1995), whichever was greater. This variable is not research-quality, but it is useful for drawing bottom topography for sections.

5. All data with quality flags 3, 4, 5, or 8 were excluded from the product files and their flags set to 9. Hence, in the product files a flag 9 can indicate not measured (as is also the case for the original exchange formatted data files) or excluded from product; in any case, no data value appears.

6. The 12 core variables were calibrated using adjustments from the GEOMAR Adjustment Table. For these variables the data product also contains secondary QC flags, indicating by cruise and variable whether ("1") or not ("0") data successfully received secondary QC. A "0" flag here generally means that data were too shallow or geographically too isolated for crossover analysis. Flag "0" corresponds to a "−888" adjustment value in the Adjustment Table.

7. Multi-leg cruises that had been quality-controlled on a per-leg basis (Table 4) were combined to single cruises.

8. To ensure that as many carbon data as possible were accompanied by supporting biogeochemical data, missing salinity, oxygen, nitrate, silicate, and phosphate values were vertically interpolated whenever practical, using a quasi-Hermetian piecewise polynomial. "Whenever practical" means that interpolation was limited to the vertical data separation distances given in Table 4 in Key et al. (2010). Interpolated values are flagged 0.

9. Values for potential temperature; potential densities referenced to 0, 1000, 2000, 3000, and 4000 dbar; neutral density; and apparent oxygen utilization were calculated using Fofonoff (1977), Bryden (1973), UNESCO (1981), and Garcia and Gordon (1992). In the few instances in which only potential temperature values were reported, these values were retained.

10. Whenever sampling pressure or depth was missing this was calculated following UNESCO (1981).

11. GLODAPv2 includes $TCO_2$, TAlk, and pH data. Generally, whenever only two seawater $CO_2$ chemistry variables were reported, the third was calculated. In the final product files some of the cruises thus, invariably, have a mixture of calculated and measured values of specific $CO_2$ chemistry variables, i.e., the cruises that had non-collocated measurements of three variables. This is generally not a problem since the internal consistency of the seawater $CO_2$ chemistry data at these cruises has been established (Sect. 3.2.4). However, in some cases so few data were available for the third variable that the internal consistency could not be established (typically when the instrument had been brought along for testing or training purposes). The few measured data points were then replaced with calculated ones. Table 9 provides an overview of the cruises where measured data were replaced with calculated ones. On the other hand, for those cruises where all three variables had been measured but there were a few holes in the record of each that did not diminish their scientific value, these holes were not filled with calculated values. For the various constants involved, the same as those in Velo et al. (2010) were used, except for carbonate dissociation constants, where the ones of Lueker et al. (2000) were used instead of the constants of Merbach et al. (1973) refitted by Dickson and Millero (1987). The calculations were carried out using CO2SYS (Lewis and Wallace, 1998) for MATLAB (van Heuven et al., 2011). Calculated data points are assigned WOCE flag "0" and thus easily excluded from any analysis if desired. Note also that the secondary QC flags of the measured carbon data have been carried through to the calculated ones, and if at least one of the input variable's flag was "0" the calculated data were given a secondary QC flag of "0" as well.

12. Note also that, similar to GLODAPv1.1, the product files contain some TAlk values that have been calculated from discrete $pCO_2$ and $TCO_2$, for cruises where data for only this pair were available. These TAlk data were treated as measured during the secondary QC analyses and are not indicated as calculated in the Adjustment Table. They do have WOCE flag 0 in the product files, though.

13. Partial pressures for CFC-11, CFC-12, CFC-113, $CCl_4$, and $SF_6$ were calculated using the solubilities by Warner and Weiss (1985), Bu and Warner (1995), Bullister and Wisegarver (1998), and Bullister et al. (2002).

Besides the core variables, the product contains data for the following: $\Delta^{14}C$, $\delta^{13}C$, $^3H$, $\delta^3He$, He, Ne, $\delta^{18}O$, total organic carbon (TOC), dissolved organic carbon (DOC), dissolved organic nitrogen (DON), total dissolved nitrogen (TDN), $SF_6$, and chlorophyll $a$ (Chl $a$). None of these were subjected to secondary QC. Table 10 specifies the file contents and lists variable names used. Missing data are set to −9999 in the csv files and NaN in the MATLAB files.

## 6 Recommendations for data use

GLODAPv2 is freely accessible and can be used without any fees, login requirements, or other restrictions. Whenever GLODAPv2 is used, this paper here should be cited in any publication. We also ask users to remember that hardworking scientists made these measurements, often under severe conditions. Further, the principal investigators normally possess insight on the quality and context of the data not known to the GLODAPv2 team. Hence, inviting individual data providers to collaborate in scientific investigations that depend on their data is considered good and fair practice. Importantly, this will promote further sharing of data and will be beneficial to science. In the CST, citations to relevant scientific publications for individual cruises have been provided whenever these were known. GLODAPv2 users are encouraged to cite these papers. Data providers are encouraged to supply additional references to specific cruise data by contacting CDIAC directly. Finally, in a product of this size,

**Table 9.** Cruises where measured carbon variables have been fully replaced with calculated ones in the GLODAPv2 product files. One of these cruises is from the Atlantic (29HE20100405), while the remainder are Pacific cruises.

| Variable | Cruise | Number of measured values removed | Number of calculated values added |
|---|---|---|---|
| TCO$_2$ | 29HE20100405 | 40 | 953 |
| TAlk | 49HO19980718 | 71 | 337 |
| | 49XK19960617 | 69 | 317 |
| pH | 49EW19981003 | 23 | 53 |
| | 49HG19930413 | 75 | 188 |
| | 49HO20000601 | 25 | 125 |
| | 49HO20000621 | 25 | 123 |
| | 49NZ20041117 | 277 | 1049 |
| | 49UF20080117 | 58 | 117 |

scope, and complexity, errors and mistakes are bound to occur. Besides the product files, a document that lists known issues is provided at CDIAC. This will be updated as new errors are found and reported by the user community. Cruise-specific issues, e.g., errors or data updates, are also given in the field "Annotations for this cruise in GLODAPv2" at each cruise's page in the online Adjustment Table.

## 7   Conclusions, lessons learned, and outlook

Over the past 30–40 years, the scope, quality, and frequency of earth system observations have increased in response to awareness of human pressures on our planet. These observations are gathered as part of a multitude of programs, with various requirements for data quality and handling. Global coordination exits in the form of WOCE, CLIVAR, IOCCP, GO-SHIP, etc., but its influence is far from uniform. As a result, data are stored in various places, in various formats, and with inconsistent documentation. Quite often, different versions of the same data are available. Such issues restrict integrated use of data for large-scale and/or long-term assessments. In the worst case it will limit data usability for future generations. GLODAPv2 and its predecessors have attempted to deal with this issue. We believe that we have been largely successful in our undertaking. For example, for TCO$_2$ measurements, Dickson et al. (2007) sets the target within-cruise precision to 1.5 µmol kg$^{-1}$ and the between-cruise range of bias to 4 µmol kg$^{-1}$. For TAlk the targets are 3 and 6 µmol kg$^{-1}$, respectively. The internal consistency improvement (Table 5) indicates that the analyses and adjustments carried out for these two variables have increased the overall consistency from larger or slightly smaller than to clearly smaller than the between-cruise bias targets. In fact, for TAlk, the adjusted data now appear consistent to the within-cruise precision target.

GLODAPv2 has also revealed particular widespread sampling and measurement issues that must be tackled by the community. The frequently occurring sloppy routines for calibrating oxygen and salinity data retrieved from the CTD package are an intolerable and widespread practice. Out of the 780 entries with salinity data, 144, or almost 20 %, contained CTD data that had clearly not been calibrated with regard to the bottle measurements (Table 8). For oxygen the fraction was somewhat less, 11 % (Table 8). However, looking only at the data files that included both CTD and bottle oxygen, significant offsets between the two were found in almost 50 % of the files (Table 8). Given the complexity of modern climate change issues, this is simply unacceptable. Only carefully calibrated CTD values should be submitted. The "after-the-fact" linear calibrations that we performed will never be as good as what could have been done by the data originators.

It should be noted that the practice of measuring salinity and oxygen on only a fraction of samples with the aim of calibrating the CTD sensor has become more common. Although this practice is strongly discouraged by GO-SHIP, some programs persist. The arguments given are that running salt/oxygen on every Niskin bottle is too expensive or that calibration of the CTD does not require that many samples. The latter is generally, but not always, true. However, when something does go wrong with the CTD sensor(s) and this is not discovered until the cruise is over, the cost is very high. The fact also remains that bottle salt/oxygen samples are about the only way to be sure when a sample bottle mistrips or leaks. Additionally, the cost of analyzing a few expensive tracers (particularly isotopes) on samples that mistripped, leaked, etc. quickly exceeds the relatively small cost of shipboard salt/oxygen analysis.

Inadequate documentation is another widespread issue; metadata may be completely missing or lack information on important details, such as method, calibration material and practices, or even reporting scale (e.g., whether data were reported as per unit volume (liters) or per unit mass (kilograms) sea water). The lack of universal and certified nutrient stan-

**Table 10.** Variables in the GLODAPv2 comma separated (csv) product files, their short names, and units, in order of appearance. In the MATLAB product files that are also supplied, a "G2" has been added to every variable name.

| Variable | Units | Variable name | WOCE flag name[a] | Secondary QC flag name[b] |
|---|---|---|---|---|
| GLODAPv2 assigned sequential cruise number | | cruise | | |
| Station | | station | | |
| Cast | | cast | | |
| Year | | year | | |
| Month | | month | | |
| Day | | day | | |
| Hour | | hour | | |
| Minute | | minute | | |
| Latitude | | latitude | | |
| Longitude | | longitude | | |
| Bottom depth | m | bottomdepth | | |
| Pressure of the deepest sample | dbar | maxsampdepth | | |
| Niskin bottle number | | bottle | | |
| Sampling pressure | dbar | pressure | | |
| Sampling depth | m | depth | | |
| Temperature | °C | temperature | | |
| Potential temperature | °C | theta | | |
| Salinity | | salinity | salinityf | salinityqc |
| Potential density anomaly | $kg\,m^{-3}$ | sigma0 | (salinityf) | |
| Potential density anomaly, ref 1000 dbar | $kg\,m^{-3}$ | sigma1 | (salinityf) | |
| Potential density anomaly, ref 2000 dbar | $kg\,m^{-3}$ | sigma2 | (salinityf) | |
| Potential density anomaly, ref 3000 dbar | $kg\,m^{-3}$ | sigma3 | (salinityf) | |
| Potential density anomaly, ref 4000 dbar | $kg\,m^{-3}$ | sigma4 | (salinityf) | |
| Neutral density anomaly | $kg\,m^{-3}$ | gamma | (salinityf) | |
| Oxygen | $\mu mol\,kg^{-1}$ | oxygen | oxygenf | oxygenqc |
| Apparent oxygen utilization | $\mu mol\,kg^{-1}$ | aou | aouf | |
| Nitrate | $\mu mol\,kg^{-1}$ | nitrate | nitratef | nitrateqc |
| Nitrite | $\mu mol\,kg^{-1}$ | nitrite | nitritef | |
| Silicate | $\mu mol\,kg^{-1}$ | silicate | silicatef | silicateqc |
| Phosphate | $\mu mol\,kg^{-1}$ | phosphate | phosphatef | phosphateqc |
| $TCO_2$ | $\mu mol\,kg^{-1}$ | tco2 | tco2f | tco2qc |
| TAlk | $\mu mol\,kg^{-1}$ | talk | talkf | talkqc |
| pH at total scale, 25 °C and 0 dbar of pressure | | phts25p0 | phts25p0f | phtsqc |
| pH at total scale, in situ temperature and pressure | | phtsinsitutp | phtsinsitutpf | phtsqc |
| CFC-11 | $pmol\,kg^{-1}$ | cfc11 | cfc11f | cfc11qc |
| pCFC-11 | ppt | pcfc11 | (cfc11f) | |
| CFC-12 | $pmol\,kg^{-1}$ | cfc12 | cfc12f | cfc12qc |
| pCFC-12 | ppt | pcfc12 | (cfc12f) | |
| CFC-113 | $pmol\,kg^{-1}$ | cfc113 | cfc113f | cfc113qc |
| pCFC-113 | ppt | pcfc113 | (cfc113f) | |
| $CCl_4$ | $pmol\,kg^{-1}$ | ccl4 | ccl4f | ccl4qc |
| $pCCl_4$ | ppt | pccl4 | (ccl4f) | |
| $SF_6$ | $fmol\,kg^{-1}$ | sf6 | sf6f | |
| pSF6 | ppt | psf6 | (sf6f) | |
| $\delta^{13}C$ | ‰ | c13 | c13f | |
| $\Delta^{14}C$ | ‰ | c14 | c14f | |
| $\Delta^{14}C$ counting error | ‰ | c14err | | |
| $^{3}H$ | TU | h3 | h3f | |
| $^{3}H$ counting error | TU | h3err | | |
| $\delta^{3}He$ | % | he3 | he3f | |
| $\delta^{3}He$ counting error | % | he3err | | |
| He | $nmol\,kg^{-1}$ | he | hef | |
| He counting error | $nmol\,kg^{-1}$ | heerr | | |
| Ne | $nmol\,kg^{-1}$ | neon | neonf | |
| Ne counting error | $nmol\,kg^{-1}$ | neonerr | | |
| $\delta^{18}O$ | ‰ | o18 | o18f | |
| Total organic carbon | $\mu mol\,L^{-1\,c}$ | toc | tocf | |
| Dissolved organic carbon | $\mu mol\,L^{-1\,c}$ | doc | docf | |
| Dissolved organic nitrogen | $\mu mol\,L^{-1\,c}$ | don | donf | |
| Total dissolved nitrogen | $\mu mol\,L^{-1\,c}$ | tdn | tdnf | |
| Chlorophyll $a$ | $\mu g\,kg^{-1\,c}$ | chla | chlaf | |

[a] The only derived variable assigned a separate WOCE flag is AOU as it depends strongly on both temperature and oxygen (and less strongly on salinity). For the other derived variables, the applicable WOCE flag is given in parentheses. [b] Secondary QC flags are used to indicate whether data have been subjected to full secondary QC (1) or not (0), as described in Sect. 5.2. [c] Units have not been checked; some data in $\mu mol\,kg^{-1}$ (for TOC, DOC, DON, TDN) or $\mu g\,L^{-1}$ (for Chl $a$) most likely occur.

dards had particularly strong ramifications for Pacific silicate data (Sect. 4.3.4). This issue is recognized by the community and being addressed with the introduction of certified reference material (Aoyama et al., 2012). It is important that this material is used widely and consistently in the future. Our analyses have demonstrated that biases can occur, even if certified material was used. This can result from missing or replicated density corrections (i.e., conversion of data from $\mu$mol $L^{-1}$ to $\mu$mol $kg^{-1}$ twice), or from more fundamental problems.

In light of these brief considerations, it is our firm belief that scientist-driven data synthesis, generating well-documented, quality-controlled, and internally consistent data products is an important and warranted activity. GLODAPv2 will form the starting point for routine future interior ocean syntheses; the plan is to produce updated versions on a routine basis in the years to come. Incoming cruises not yet subjected to QC or included in the product file will be listed at the end of the CST at CDIAC. When the number of incoming cruises warrants an update of our product, their consistency with GLODAPv2 will be checked on a case-by-case basis using crossover routines. New versions of the product files will then be generated with these data added. Any "known issues" (Sect. 6) will also be remedied in these updated versions.

## 8  Data availability

All data presented in this contribution have been deposited at CDIAC, and made available at http://cdiac.ornl.gov/oceans/GLODAPv2/ in formats described above. The doi assigned to the product is doi:10.3334/CDIAC/OTG.NDP093_GLODAPv2. Each of the 724 individual original cruise data files has also been assigned a separate doi, these are not listed here, but can be retrieved through the CST at the GLODAPv2 web page at CDIAC.

## Appendix A: Initial strategy and actual workflow

The initial strategy was to carry out production of GLO-DAPv2 in a series of semi-parallelized steps:

1. Identify and ingest data not included in GLODAPv1.1, CARINA, or PACIFICA and subject these to primary QC. These were grouped into the dataset GLODAPv2 (NEW).

2. In parallel, re-evaluate GLODAPv1.1 using the CARINA-developed analysis tools (Tanhua et al., 2010) to enhance its consistency with respect to CA-RINA and PACIFICA; this GLODAPv1.2 product was not to be publicly released but was used internally in step 3.

3. Combine GLODAPv1.2 with CARINA and PACIFICA to give a global reference data product, and analyze the consistency of the GLODAPv2 (NEW) data with respect to this product using crossovers.

4. Assemble a preliminary product, GLODAPv2.beta, from the four data sources GLODAPv1.2, CARINA, PACIFICA, and GLODAPv2 (new) and carry out regional crossover and inversion analyses to ensure global consistency of GLODAPv2.

5. In parallel, analyze consistency of halogenated transient tracer data using specialized methods and software.

6. In parallel, convert reported pH data to common scale (total hydrogen scale at 25°C and surface (0 dbar) pressure, and also at in situ conditions) and quality-control these data using specialized methods and software.

7. Prepare the GLODAPv2 bias-corrected data product and the mapped climatology.

This strategy was largely followed. GLODAPv1.2 was prepared. All new data were subjected to primary QC, and secondary QC against the internal merged GLODAPv1.2, CARINA, and PACIFICA product. However, at step 4, during the consistency analysis of the GLODAPv2.beta product, it became very difficult to fully track the justification of the adjustments as they had been determined in multiple analyses, e.g., the CARINA published adjustments plus our revision of these, or the PACIFICA published adjustments plus our revision of these, or the GLODAPv1.2 derived adjustments plus our revision of these. In addition, the presence of non-calibrated salinity and oxygen CTD sensor data for a fairly large number of cruises was discovered. The entire database was therefore reset, and crossover and inversion analysis was conducted on the unadjusted data as described in Sect. 3.2.2. Steps 5–7 were conducted as intended and described in the main text (Sects 3.2.3, 3.2.4, and 5.2).

## Appendix B: Guide to the adjustment table

The content of the Adjustment Table was added sequentially as work progressed. Hence, comments frequently pertain to revisions of existing adjustments, and in some cases the entire history of the development of a specific adjustment can be extracted from the comments in the table. Some of the comments may also refer to workshops where the magnitudes of the adjustments were discussed and decided; these workshops are listed in Table B1. When accessing the table be aware of the following:

– A comment was not always entered when the data appeared unbiased.

– The GEOMAR Adjustment Table gives the dataset source of each cruise, CARINA, PACIFICA, GLO-DAPv1.2 (i.e., the re-evaluated GLODAPv1.1), or GLODAPv2 (NEW), the last of which being the new cruises.

– For CARINA cruises the CARINA recommended adjustment was used as the initial value and all comments entered during the CARINA QC process have been included, as these were already available in the appropriate format. Any comments from before 2011 are thus "CARINA comments", while any comments from after are "GLODAPv2" revisions, based on either the analysis of the beta version or the reset, unadjusted, database.

– For PACIFICA-sourced cruises the PACIFICA recommended adjustments were used as initial values. No comments were available with these. Those that appear in the Adjustment Table are from GLODAPv2, either based on the analysis of the beta version, or of the reset, unadjusted, database, and justify revisions to the original PACIFICA adjustments, or simply state that these should be maintained.

– For GLODAPv1.2 cruises, all adjustment values and comments that appear are based on our analyses. For either the preliminary revision of GLODAPv1.1 to GLO-DAPv1.2 or the analyses of the beta product or the reset database. Comments from 2012 are typically based on the first, while comments by Steven van Heuven from 2014 are typically based on the last two.

– For GLODAPv2 (NEW) cruises, all adjustment values and comments are based on our analyses. They are either from the preliminary analyses of each cruise against the global intermediate reference dataset (GLO-DAPv1.2, CARINA and PACIFICA) or based on the analyses of the beta product or the reset database. Comments by Sara Jutterström or Siv Lauvset typically refer to the first, while comments by Steven van Heuven typically refer to the final two.

**Table B1.** Workshops conducted during GLODAPv2 production and their main topics.

| Place | Time | Topic |
|---|---|---|
| Bergen, Norway | Nov 2012 | Revision of GLODAPv1.1 |
| Norwich, UK | Apr 2013 | Preliminary QC of new data |
| Groningen, the Netherlands | Oct 2013 | Secondary QC of full dataset |
| Bremen, Germany | Jan 2014 | Secondary QC of full dataset |

– For CFCs the comments are either inherited from CARINA or posted following our analyses described in Sect. 3.2.3.

– For pH the comments are either inherited from CARINA or posted following our analyses described in Sect. 3.2.4.

As an example of information available at the Adjustment Table, results of evaluation of TAlk from cruise 06MT20040311 are presented in the following. Note that familiarity with the crossover and inversion method as described in Tanhua et al. (2010) is advantageous. The QC results for 06MT20040311 can be found using the search field in the Adjustment Table, upper right, and the specific summary page for this cruise is opened by clicking on either of the symbols in the leftmost column in the row for this cruise. Once at this cruise's summary page, the figures and comments for TAlk can be accessed by clicking on the row "Alkalinity [+]" in the table to the left. The summary page for this cruise can alternatively be accessed through the link in the rightmost column at this cruise's row in the CST (Sect. 5.1). The TAlk data of 06MT20040311 was evaluated in CARINA and re-evaluated in GLODAPv2. There are two comments for TAlk in the Adjustment Table, one by Fiz Perez and Anton Velo dated 2008-06-10 and one by Steven van Heuven dated 2015-01-08. The former was entered during CARINA, while the latter was entered as part of GLODAPv2 QC. There are a total of 27 crossover figures available; by holding the mouse pointer over these, their upload time appears. It then becomes evident that the ones named "Xover__******.png" were uploaded in 2008 and are Anton Velo's figures, while those named "unadjusted_*****.pdf" were uploaded in 2014, generated during Steven van Heuvens analysis of the unadjusted GLODAPv2 database. While the data were not adjusted in CARINA, since the bias appeared less than the $6\,\mu mol\,kg^{-1}$ threshold, during GLODAPv2 the evidence was convincing enough to apply an adjustment of $+\,4\,\mu mol\,kg^{-1}$. An example of one of the crossovers that supports this adjustment is provided in Fig. B1. The three panels to the left are a map with the station locations of the two cruises, a histogram of the distances between the stations involved in the crossover, and a map of the stations involved in the actual crossover. The next three pairs of panels show the actual data compared (upper) and the difference profiles (lower) in three spaces: potential density anomaly referenced to 4000 dbar pressure ($\sigma_4$, axis label: "Sigma-4"), potential temperature ($\Theta$, axis label: "Theta"), and depth (axis label: "Depth"). These difference profiles were determined by comparing station pairs in the crossover that were separated by less than 200 km, in accordance with the "running cluster" procedure (Tanhua et al., 2010). In the difference plots the light curves in the background are the individual difference profiles, the red dotted and solid lines are the average difference and standard deviation (with depth), and the solid green vertical lines are the calculated weighted mean offset and standard deviation. These numbers are also printed in the summary table beneath each difference panel, along with the number of profiles involved from each cruise. In this case, where TAlk is analyzed, the additive offset is the appropriate one to consider. In $\sigma_4$ space this is $-4.25 \pm 3.16\,\mu mol\,kg^{-1}$, while it is $-3.68 \pm 2.48$ and $-4.83 \pm 3.17\,\mu mol\,kg^{-1}$ in $\Theta$ and depth space, respectively. This figure leaves little doubt that the 06MT20040311 TALK values are lower than those of 29HE20130320.

The results from all 18 crossovers identified for TAlk for the 06MT20040311 cruise are presented in Fig. B2. This can also be obtained from the page for this cruise in the Adjustment Table. This shows that the mean offset is $-4.5 \pm 4.7\,\mu mol\,kg^{-1}$, and cannot be ascribed to the presence of a trend in the data. The magnitude of the bias was confirmed by the inversion calculation. We therefore applied an adjustment of $+4\,\mu mol\,kg^{-1}$ to these data.

Another example is phosphate of 316N20050821. Two comments are provided for phosphate in the Adjustment Table, one by Are Olsen for 2014-02-26 and another by Steven van Heuven for 2014-06-16. There are also two sets of crossover figures; one set that can be traced back to Siv (Lauvset) while the other set is associated with Steven (van Heuven). Siv's figures have five crossover plots and one summary figure (named Xresults.png), uploaded in January 2014 (holding the mouse pointer over a name gives the upload date). This cruise is new to data synthesis (its source is GLODAPv2 (NEW)) and Siv's figures were created during the preliminary analyses of new data. The summary figure gives a mean offset of 0.986; since this is too small to warrant an adjustment (given the 2 % threshold for nutrients), none was suggested during this analysis of the GLODAPv2 (NEW) data, and no comment was entered. However, the next two comments revise this, based on the final analyses of all original data combined.
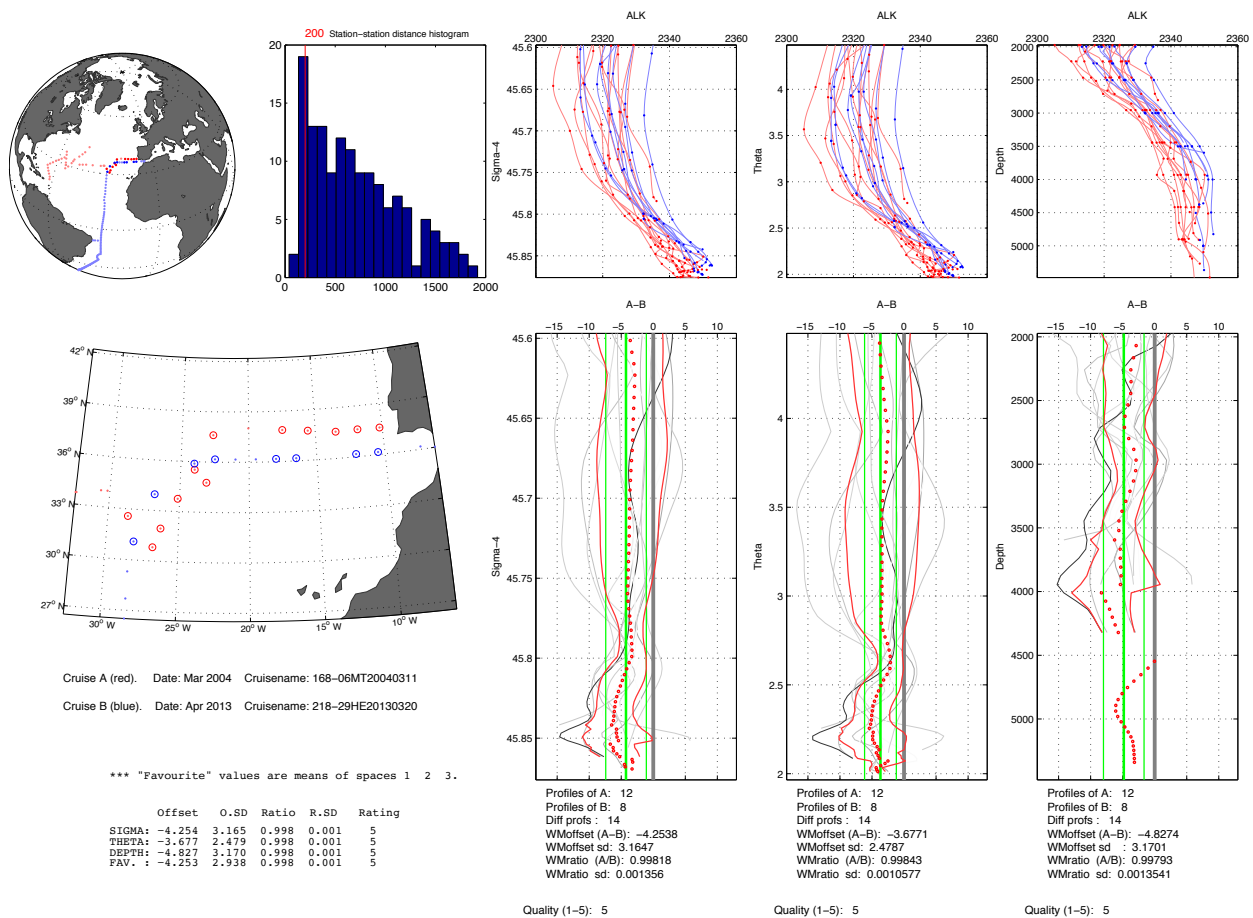
**Figure B1.** Summary TAlk figure for crossover stations between cruise 06MT20040311 and 29HE20130320. Note that the figure is as it appears in the Adjustment Table and from before final cruise numbers were assigned; hence, the cruise numbers given in this figure, 168 and 218, are *not* the GLODAPv2 cruise numbers that are used for our data product. The relative difference here is approximately $4\,\mu\mathrm{mol\,kg^{-1}}$, with the red (*Meteor*; 06MT) cruise seeming to have lower abyssal values than the blue (*Hesperides*; 29HE). The 06MT data were adjusted by $+4\,\mu\mathrm{mol\,kg^{-1}}$ for the final product file.

These two examples illustrate that it is certainly possible to locate the main evidence for adjustments that have been applied and to backtrack the steps taken to unearth these.

Finally, note that in the Adjustment Table the records that have calculated $CO_2$ chemistry variables in the product have been indicated by adding a "c" to their adjustment value. The adjustment value itself applies to any measured data. For instance, for pH a value of "$-999c$" means that no measured data are available, but calculated values are part of the product. Furthermore, a value of $-777c$, for instance, means that measured data were bad, and the calculated data have been inserted into the product file. The Adjustment Table at http://glodapv2.geomar.de can also be exported to an ASCII file. In this file a separate column indicates presence of calculated values in the product; this takes values of "1" for $TCO_2$, "2" for Talk, "3" for pH, and "0" for no carbon variable calculated.
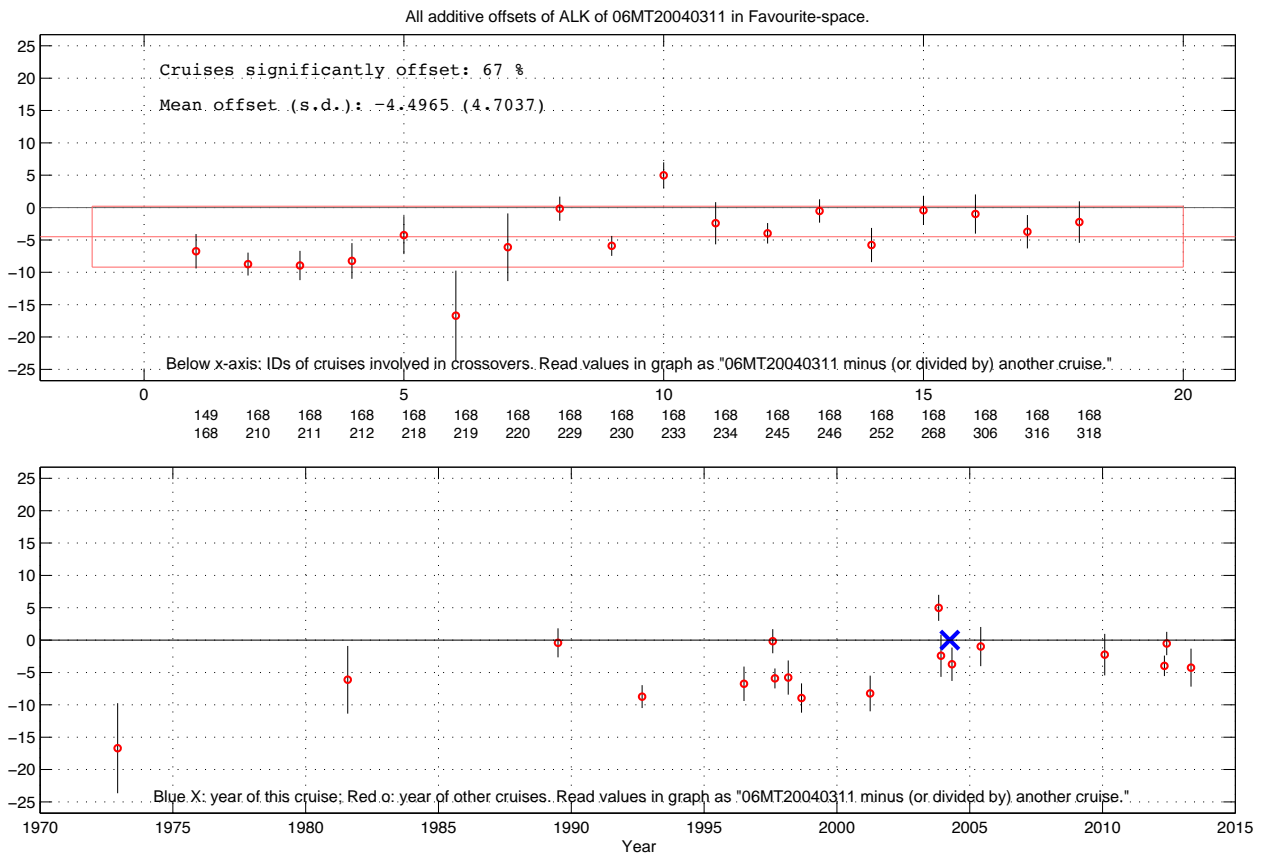
**Figure B2.** Summary figure for crossover offsets of cruise's 06MT20040311 TAlk data. The upper panel shows the 18 crossover offsets that were determined, as well as their mean and standard deviation. $X$-axis labels are cruise numbers for the crossover pairs (note that the figure is as it appears in the Adjustment Table and from before final cruise numbers were assigned; hence, the cruise numbers do not correspond to the final GLODAPv2 cruise numbers in the CST and in the product files – in this figure 06MT20040311 is cruise number 168, while its "official" GLODAPv2 cruise number is 58). The lower panel shows these crossover offsets sorted by time. "Favourite-space" is the mean of the offsets in $\Theta$, pressure, and $\sigma_4$ space. In both panels, negative values mean that 06MT20040311 TAlk values are lower than those of the comparison cruise. The lower panel shows offsets sorted with time. The 06MT data were adjusted by $+4\,\mu\mathrm{mol\,kg^{-1}}$ for the final product file.

## Appendix C: List of abbreviations (excluding variable names)

| | |
|---|---|
| BATS | Bermuda Atlantic Time-series Study |
| CARINA | Carbon in the Atlantic Ocean |
| CDIAC | Carbon Dioxide Information Analysis Center |
| CLIVAR | Climate and Ocean: Variability, Predictability and Change |
| EGEE | Etude de la circulation océanique et du climat dans le Golfe de Guinée |
| GEOSECS | Geochemical Ocean Sections |
| GIFT | Gibraltar Fixed Time Series |
| GLODAP | Global Data Analysis Project |
| GO-SHIP | Global Ocean Ship-based Hydrographic Investigations Panel |
| CRM | Certified reference material |
| CST | Cruise Summary Table |
| CTD | Conductivity–temperature–depth (profiler) |
| HOT | Hawaiian Ocean Time-series |
| ICES | International Council for the Exploration of the Sea |
| IGBP | International Geosphere-Biosphere Programme |
| IOCCP | International Ocean Carbon Coordination Project |
| INDIGO | Indien Geochimie Ocean |
| IMBER | Integrated Marine Biogeochemistry and Ecosystem Research |
| JGOFS | Joint Global Ocean Flux Study |
| Kerfix | Kerguelen Point Fixe |
| KNOT | Kyodo North Pacific Ocean Time-series |
| K2 | Japanese time series station at 47° N, 160° E |
| NESDIS | National Environmental Satellite, Data, and Information Service |
| NIOZ | Royal Netherlands Institute for Sea Research |
| NOAA | National Oceanic and Atmospheric Administration |
| NODC | National Oceanographic Data Center |
| OISO | Océan Indien Service d'Observation |
| OMEX | Ocean Margin Exchange |
| OWS | Ocean weather station |
| PACIFICA | Pacific Ocean Interior Carbon |
| QC | Quality control |
| SAVE | South Atlantic Ventilation Experiment |
| SLSQ | Simple least squares |
| SOLAS | Surface Ocean – Lower Atmosphere Study |
| TTO-NAS | Transient Tracers in the Ocean–North Atlantic Study |
| WDLSQ | Weighted damped least squares |
| WLSQ | Weighted least squares |
| WOCE | World Ocean Circulation Experiment |

**The Supplement related to this article is available online at doi:10.5194/essd-8-297-2016-supplement.**

# References

Aoyama, M., Anstey, C., Barwell-Clarke, J., Baurand, F., Becker, S., Blum, M., Coverly, S. C., Czobik, E., d'Amico, F., Dahllöf, I., Dai, M., Dobson, J., Pierre-Duplessix, O., Duval, M., Engelke, C., Gong, G.-C., Grosso, O., Hirayama, A., Inoue, H., Ishida, Y., Hydes, D. J., Kasai, H., Kerouel, R., Knockaert, M., Kress, N., Krogslund, K. A., Kumagai, M., Leterme, S. C., Mahaffey, C., Mitsuda, H., Morin, P., Moutin, T., Munaron, D., Murata, A., Nausch, G., Ogawa, H., van Ooijen, J., Pan, J., Paradis, G., Payne, C., Prove, G., Raimbault, P., Rose, M., Saito, K., Saito, H., Sato, K., Schmidt, C., Schütt, M., Shammon, T. M., Olafsdottir, S., Sun, J., Tanhua, T., Weigelt-Krenz, S., White, L., Woodward, E. M. S., Worsfold, P., Yoshimura, T., Youénou, A., and Zhang, J. Z.: 2008 Inter-laboratory comparison study of a reference material for nutrients in seawater, Technical Reports of the Meteorological Research Institute, No. 60, Tsukuba, Japan, 134 pp., 2010.

Aoyama, M., Ota, H., Kimura, M., Kitao, T., Mitsuda, H., Murata, A., and Sato, K.: Current Status of Homogeneity and Stability of the Reference Materials for Nutrients in Seawater, Anal. Sci., 28, 1–6, 2012.

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., and Vichi, M.: Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models, Biogeosciences, 10, 6225–6245, doi:10.5194/bg-10-6225-2013, 2013.

Bryden, H. L.: New polynomials for thermal expansion, adiabatic temperature gradient and potential temperature of sea water, Deep-Sea Res., 20, 401–408, 1973.

Bu, X. and Warner, M. J.: Solubility of chlorofluorocarbon-113 in water and seawater, Deep-Sea Res. Pt. I, 42, 1151–1161, 1995.

Bullister, J. L. and Wisegarver, D. P.: The solubility of carbon tetrachloride in water and seawater, Deep-Sea Res. Pt. I, 45, 1285–1302, 1998.

Bullister, J. L., Wisegarver, D. P., and Menzia, F. A.: The solubility of sulfur hexafluoride in water and seawater, Deep-Sea Res. Pt. I, 49, 175–187, 2002.

Dickson, A. G.: pH scales and proton-transfer reactions in saline media such as sea water, Geochim. Cosmochim. Acta, 48, 2299–2308, 1984.

Dickson, A. G.: Reference materials for oceanic $CO_2$ measurements, Oceanography, 14, 21–22, 2001.

Dickson, A. G.: The carbon dioxide system in seawater: equilibrium chemistry and measurements, in: Guide to best practices for ocean acidification research and data reporting, edited by: Riebesell, U., Fabry, V. J., Hanson, L., and Gattuso, J.-P., Publication Office of the European Union, Luxemburg, 2010.

Dickson, A. G. and Millero, F. J.: A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media, Deep-Sea Res. Pt. A, 34, 1733–1743, 1987.

Dickson, A. G., Afghan, D., and Anderson, G. C.: Reference materials for oceanic $CO_2$ analysis: A method for the certification of total alkalinity, Mar. Chem., 80, 185–197, 2003.

Dickson, A. G., Sabine, C. L., and Christian, J. R. (Eds.): Guide to best practices for ocean $CO_2$ measurements. PICES special publication 3, North Pacific Marine Science Organisation, Sidney, BC, Canada, 191 pp., 2007.

Dore, J. E., Lukas, R., Sadler, D. W., and Karl, D. M.: Climate-driven changes to the atmospheric $CO_2$ sink in the subtropical North Pacific Ocean, Nature, 424, 754–757, 2003.

Feely, R. A., Talley, L. D., Bullister, J. L., Carlson, C. A., Doney, S. C., Fine, R. A., Firing, E., Gruber, N., Hansell, D. A., Johnson, G. C., Key, R. M., Langdon, C., Macdonald, A., Mathis, J. T., Mecking, S., Millero, F. J., Mordy, C. W., Sabine, C. L., Smethie, W. M., Swift, J. H., Thurnherr, A. M., Wanninkhof, R.

H., and Warner, M. J.: The US Repeat Hydrography CO$_2$/Tracer Program (GO-SHIP): Accomplishments from the first decadal survey. A US CLIVAR and OCB Report, 2014-15, US CLIVAR Project Office, 2014.

Fofonoff, N. P.: Computation of potential temperature of seawater for an arbitrary reference pressure, Deep-Sea Res., 24, 489–491, 1977.

Garcia, H. and Gordon, L. I.: Oxygen solubility in seawater: Better fitting equations, Limnol. Oceanogr., 37, 1307–1312, 1992.

Giesbrecht, K. E., Miller, L. A., Davelaar, M., Zimmermann, S., Carmack, E., Johnson, W. K., Macdonald, R. W., McLaughlin, F., Mucci, A., Williams, W. J., Wong, C. S., and Yamamoto-Kawai, M.: Measurements of the dissolved inorganic carbon system and associated biogeochemical parameters in the Canadian Arctic, 1974–2009, Earth Syst. Sci. Data, 6, 91–104, doi:10.5194/essd-6-91-2014, 2014.

Gouretski, V. V. and Jahnke, K.: Systematic errors as the cause for an apparent deep water property variability: global analysis of the WOCE and historical hydrographic data, Progr. Oceanogr., 48, 337–402, 2001.

Helm, K. P., Bindoff, N. L., and Church, J. A.: Observed decreases in oxygen content of the global ocean, Geophys. Res. Lett., 38, L23602, doi:10.1029/2011GL04953, 2011.

Hood, M., Fukasawa, M., Gruber, N., Johnson, G. C., Körtzinger, A., Sabine, C., Sloyan, B., Stansfield, K., and Tanhua, T.: Ship-based repeat hydrography: A strategy for a sustained global program, in: Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009, edited by: Hall, J., Harrison, D. E., and Stammer, D., ESA Publication WPP-306, doi:10.5270/OceanObs09, 2010.

Jeansson, E., Olsson, K. A., Tanhua, T., and Bullister, J. L.: Nordic Seas and Arctic Ocean CFC data in CARINA, Earth Syst. Sci. Data, 2, 79–97, doi:10.5194/essd-2-79-2010, 2010.

Johnson, G., Robbins, P. E., and Hufford, G. E.: Systematic adjustments of hydrographic sections for internal consistency, J. Atmos. Ocean. Tech., 18, 1234–1244, 2001.

Jutterström, S. and Anderson, L. G.: The saturation of calcite and aragonite in the Arctic Ocean, Mar. Chem., 94, 101–110, 2005.

Jutterström, S., Jeansson, E., Anderson, L. G., Bellerby, R., Jones, E. P., Smethie Jr, W. M., and Swift, J. H.: Evaluation of anthropogenic carbon in the Nordic Seas using observed relationships of N, P, and C versus CFCs, Progr. Oceanogr., 78, 78–84, 2008.

Jutterström, S., Anderson, L. G., Bates, N. R., Bellerby, R., Johannessen, T., Jones, E. P., Key, R. M., Lin, X., Olsen, A., and Omar, A. M.: Arctic Ocean data in CARINA, Earth Syst. Sci. Data, 2, 71–78, doi:10.5194/essd-2-71-2010, 2010.

Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R. A., Millero, F. J., Mordy, C., and Peng, T.-H.: A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP), Global Biogeochem. Cyc., 18, GB4031, doi:10.1029/2004GB002247, 2004.

Key, R. M., Tanhua, T., Olsen, A., Hoppema, M., Jutterström, S., Schirnick, C., van Heuven, S., Kozyr, A., Lin, X., Velo, A., Wallace, D. W. R., and Mintrop, L.: The CARINA data synthesis project: introduction and overview, Earth Syst. Sci. Data, 2, 105–121, doi:10.5194/essd-2-105-2010, 2010.

Key, R. M., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M.,

Jutterström, S., Steinfeldt, R., Jeansson, E., Ishi, M., Pérez, F. F., and Suzuki, T.: Global Ocean Data Analysis Project, Version 2 (GLODAPv2), ORNL/CDIAC-162, NDP-093, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, doi:10.3334/CDIAC/OTG.NDP093_GLODAPv2, 2016

King, B. A., Firing, E., and Joyce, T. M.: Shipboard observations during WOCE, in: Ocean circulation and climate, Int. Geophys. Ser., vol. 77, edited by: Siedler, G., Church, J., and Gould, J., Academic Press, San Diego, US, 2001.

Lauvset, S. K. and Tanhua, T.: A toolbox for secondary quality control on ocean chemistry and hydrographic data, Limnol. Oceanogr., 13, 601–608, doi:10.1002/lom3.10050, 2015.

Lauvset, S. K., Gruber, N., Landschützer, P., Olsen, A., and Tjiputra, J.: Trends and drivers in global surface ocean pH over the past 3 decades, Biogeosciences, 12, 1285–1298, doi:10.5194/bg-12-1285-2015, 2015.

Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the 1° × 1° GLODAP version 2, Earth Syst. Sci. Data, 8, 325–340, doi:10.5194/essd-8-325-2016, 2016.

Levitus, S., Antonov, J. I., Boyer, T. P. Baranova, O. K., Garcia, H. E., Locarnini, R. A., Mishonov, A. V., Reagan, J. R., Seidov, D., Yarosh, E. S., and Zweng, M. M.: World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010, Geophys. Res. Lett., 39, L10603, doi:10.1029/2012GL051106, 2012.

Lewis, E. and Wallace, D. W. R.: Program developed for CO$_2$ system calculations, ORNL/CDIAC-105, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN, 1998.

Lueker, T. J., Dickson, A. G., and Keeling, C. D.: Ocean $p$CO$_2$ calculated from dissolved inorganic carbon, alkalinity, and equations for $K_1$ and $K_2$: validation based on laboratory measurements of CO$_2$ in gas and seawater in equilibrium, Mar. Chem., 70, 105-119, 2000.

Mantisi, F., Beauverger, C., Poisson, A., and Metzl, N.: Chlorofluoromethanes in the western Indian Sector of the Southern Ocean and their relations with geochemical data, Mar. Chem., 35, 151–167, 1991.

MATLAB, MAT-file format, available at: https://www.mathworks.com/help/pdf_doc/matlab/matfile_format.pdf, last access: 21 December 2015.

Menke, W.: Geophysical Data Analysis: Discrete Inversion Theory, Academic Press, Amsterdam, the Netherlands, 1984.

Merbach, C., Culberson, C., Hawley, J., and Pytkowicz, R.: Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure, Limnol. Oceanogr., 18, 897–907, 1973.

Metzl, N.: Decadal increase of oceanic carbon dioxide in Southern Indian Ocean surface waters (1991–2007), Deep-Sea Res. Pt. II, 56, 607–619, 2009.

Olafsson, J., Olafsdottir, S. R., Benoit-Cattin, A., Danielsen, M., Arnarson, T. S., and Takahashi, T.: Rate of Iceland Sea acidification from time series measurements, Biogeosciences, 6, 2661–2668, doi:10.5194/bg-6-2661-2009, 2009.

Olsen, A., Omar, A. M., Jeansson, E., Anderson, L. G., and Bellerby, R. G. J.: Nordic seas transit time distributions

and anthropogenic $CO_2$, J. Geophys. Res., 115, C05005, doi:10.1029/2009JC005488, 2010.

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T.-H., Kozyr, A., Ono, T., and Ríos, A. F.: The oceanic sink for anthropogenic $CO_2$, Science, 305, 367–371, 2004.

Sabine, C., Key, R. M., Kozyr, A., Feely, R. A., Wanninkhof, R., Millero, F. J., Peng, T.-H., Bullister, J. L., and Lee, K.: Global Ocean Data Analysis Project (GLODAP): Results and Data, ORNL/CDIAC-145, NDP-083, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U. S. Departement of Energy, Oak Ridge, Tenesse, 2005.

Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., and Knap, A. H.: Overview of the US JGOFS Bermuda time-series study (BATS): a decade-scale look at ocean biology and biogeochemistry, Deep-Sea Res. Pt. II, 48, 1405–1447, 2001.

Steinfeldt, R., Tanhua, T., Bullister, J. L., Key, R. M., Rhein, M., and Köhler, J.: Atlantic CFC data in CARINA, Earth Syst. Sci. Data, 2, 1–15, doi:10.5194/essd-2-1-2010, 2010.

Stendardo, I., Gruber, N., and Körtzinger, A.: CARINA oxygen data in the Atlantic Ocean, Earth Syst. Sci. Data, 1, 87–100, doi:10.5194/essd-1-87-2009, 2009.

Suzuki, T., Ishii, M., Aoyama, A., Christian, J. R., Enyo, K., Kawano, T., Key, R. M., Kosugi, N., Kozyr, A., Miller, L. A., Murata, A., Nakano, T., Ono, T., Saino, T., Sasaki, K., Sasano, D., Takatani, Y., Wakita, M., and Sabine, C. L.: PACIFICA Data Synthesis Project, ORNL/CDIAC-159, NDP-092, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U. S. Department of Energy, Oak Ridge, Tennessee, 2013.

Swift, J. and Diggs, S. C.: Description of WHP exchange format for CTD/Hydrographic data, CLIVAR and Carbon Hydrographic Data Office, UCSD Scripps Institution of Oceanography, San Diego, Ca, US, 2008.

Talley, L. D, Feely, R. A., Sloyan, B. M., Wanninkhof, R., Baringer, M. O., Bullister, J. L., Carlson, C. A., Doney, S. C., Fine, R. A., Firing, E., Gruber, N., Hansell, D. A., Ishii, M., Johnson, G. C., Katsumata, K., Key, R. M., Kramp, M., Langdon, C., Macdonald, A. M., Mathis, J. T., McDonagh, E. L., Mecking, S., Millero, F. J., Mordy, C. W., Nakano, T., Sabine, C. L., Smethie, W. M., Swift, J. H., Tanhua, T., Thurnherr, A. M., Warner, M. J., and Zhang, J -Z.: Changes in ocean heat, carbon content and ventilation: Review of the first decade of global repeat hydrography (GO-SHIP), Ann. Rev. Mar. Sci., 8, 185–215, doi:10.1146/annurev-marine-052915-100829, 2016.

Tanhua, T. and Wallace, D. W. R.: Consistency of TTO-NAS inorganic carbon data with modern measurements, Geophys. Res. Lett., 32, L14619, doi:10.1029/2005GL023248, 2005.

Tanhua, T., Olsen, A., Hoppema, M., Jutterström, S., Schirnick, C., van Heuven, S., Velo, A., Lin, X., Kozyr, A., Alvarez, M., Bakker, D. C. E., Brown, P., Falck, E., Jeansson, E., Lo Monaco, C., Olafssson, J., Perez, F. F., Pierrot, D., Rìos, A. F., Sabine, C. L., Schuster, U., Steinfeldt, R., Stendardo, I., Andeson, L. G., Bates, N. R., Bellerby, R. G. J., Blindheim, J., Bullister, J. L., Gruber, N., Ishii, M., Johannessen, T., Jones, E. P., Köhler, J., Körtzinger, A., Metzl, N., Murata, A., Musielewicz, S., Omar, A. M., Olsson, K. A., de la Paz, M., Pfeil, B., Rey, F., Rhein, M., Skjelvan, I., Tilbrook, B., Wanninkhof, R, Mintrop, L., Wallace, D. W. R., and Key, R. M.: CARINA Data Synthesis Project, ORNL/CDIAC-157, NDP-091, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U. S. Department of Energy, Oak Ridge, Tennessee, 2009.

Tanhua, T., van Heuven, S., Key, R. M., Velo, A., Olsen, A., and Schirnick, C.: Quality control procedures and methods of the CARINA database, Earth Syst. Sci. Data, 2, 35–49, doi:10.5194/essd-2-35-2010, 2010.

UNESCO (United Nations Educational, Scientific and Cultural Organization): Tenth report of the joint panel on oceanographic tables and standards, UNESCO Technical Paper in Marine Science, 36, 13–21, 1981.

van Heuven, S., Pierrot, D., Rae, J. W. B, Lewis, E., and Wallace, D. W. R, MATLAB program developed for $CO_2$ system calculations, ORNL/CDIAC-105b, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, TN, doi:10.3334/CDIAC/otg.CO2SYS_MATLAB_v1.1, 2011.

Velo, A., Pérez, F. F., Lin, X., Key, R. M., Tanhua, T., de la Paz, M., Olsen, A., van Heuven, S., Jutterström, S., and Ríos, A. F.: CARINA data synthesis project: pH data scale unification and cruise adjustments, Earth Syst. Sci. Data, 2, 133–155, doi:10.5194/essd-2-133-2010, 2010.

Wakita, M., Watanabe, S., Murata, A., Tsurushima, N., and Honda, M.: Decadal change of dissolved inorganic carbon in the subarctic western North Pacific Ocean, Tellus, 62B, 608–620, 2010.

Warner, M. J. and Weiss, R. F.: Solubilities of chlorofluorocarbons 11 and 12 in water and seawater, Deep-Sea Res. Pt. I, 32, 1485–1497, 1985.

Wong, C. S., Xie, L. S., and Hsieh, W. W.: Variations in nutrients carbon and other hydrographic parameters related to the 1976/77 and 1988/89 regime shifts in the sub-arctic Northeast Pacific, Prog. Oceanogr., 75, 326–342, 2007.

Wunch, C.: The Ocean Circulation Inverse Problem, Cambridge University Press, 1996.

Yashayaev, I.: Hydrographic changes in the Labrador Sea, 1960–2005, Progr. Oceanogr., 73, 242–276, 2007.