Earth Syst. Sci. Data Discuss., doi:10.5194/essd-2015-42, 2016
Manuscript under review for journal Earth Syst. Sci. Data
Published: 19 January 2016

Open Access Earth System Science Data Discussions

An internally consistent data product for the world ocean: the Global Ocean Data Analysis Project, version 2 (GLODAPv2)

A. Olsen[1], R. M. Key[3], S. van Heuven[4], S. K. Lauvset[1,2], A. Velo[5], X. Lin[3], C. Schirnick[6], A. Kozyr[7], T. Tanhua[6], M. Hoppema[8], S. Jutterström[9], R. Steinfeldt[10], E. Jeansson[2], M. Ishii[11], F. F. Pérez[5], and T. Suzuki[12]

[1]Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Allègaten 70, Bergen, Norway
[2]Uni Research Climate, Bjerknes Centre for Climate Research, Allegt. 55, 5007 Bergen, Norway
[3]Atmospheric and Oceanic Sciences, Princeton University, 300 Forrestal Road, Sayre Hall, Princeton, NJ 08544, USA
[4]Royal Netherlands Institute for Sea Research (NIOZ), Marine Geology and Chemical Oceanography, P.O. Box 59, 1790 AB Den Burg, the Netherlands
[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, US 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 Göteborg, 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

Received: 3 December 2015 – Accepted: 10 December 2015 – Published: 19 January 2016

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

Published by Copernicus Publications.

ESSDD
doi:10.5194/essd-2015-42

Abstract

For version 2 of the Global Ocean Data Analysis Project (GLODAPv2) we collated data from 724 scientific cruises covering the global ocean: data assembled in the previous efforts GLODAPv1.1 (Global Ocean Data Analysis Project version 1.1) in 2004, CARINA (CARbon IN the Atlantic) in 2009/10, and PACIFICA (PACIFic ocean Interior CArbon) in 2013, and an additional 168 cruises. Twelve core parameters (salinity, oxygen, macronutrients, seawater CO_2 chemistry parameters and halogenated transient tracers) have been subjected to extensive quality control including systematic evaluation of biases between cruises. The data are available in two formats: (i) as submitted but updated to WOCE exchange format whenever required, and (ii) as a merged and calibrated data product. In the latter, adjustments have been applied to remove significant biases, respecting occurrences of any known or likely time trends. Adjustments determined by previous efforts have been re-evaluated. Hence, GLODAPv2 is not a simple merge of previous collections and some new data, but represents a unique, internally consistent data product.

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 4 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. The latter were obtained either by interpolation of, or by calculation from, measured data. This paper documents the GLODAPv2 history, methods, and products, including a broad overview of the secondary quality control results. The magnitude of and reasoning behind the adjustments are available on a per cruise and parameter basis in an online Adjustment Table.

ESSDD

doi:10.5194/essd-2015-42

1 Introduction

Over the past decade we have witnessed the completion and publication of several ocean carbon data synthesis products as a result of painstaking global community efforts: GLODAPv1.1 (Global Ocean Data Analysis Project version 1.1, Key et al., 2004; Sabine et al., 2005), CARINA (CARbon IN the Atlantic, Key et al., 2010; Tanhua et al., 2009a), SOCAT (Surface Ocean CO_2 Atlas, Pfeil et al., 2013; Bakker et al., 2014) and PACIFICA (PACIFic Interior ocean CArbon, Suzuki et al., 2013) provide easy and open access to uniformly formatted, quality controlled, and well-documented data and have spearheaded major scientific developments in the field. Access to these products allows larger scale, longer term and higher impact science. The collaborative quality control carried out during their preparation helps reveal issues with sampling, measurement and documentation practices. The main goal of GLODAPv2 (Global Ocean Data Analysis Project version 2) was to create a single high-quality internally consistent global data product containing CO_2-relevant ocean interior measurements from ship-based surveys. The need for GLODAPv2 arose from the fact that these data were, at the time GLODAPv2 was initiated, assembled in the three separate and only partially intercalibrated products GLODAPv1.1, CARINA and PACIFICA, and that data from more than 150 cruises had not been included in any of the earlier products. GLODAPv2 forms a foundation that the community can build upon through quality controlling and adding new data on a routine basis.

2 History and goals of GLODAPv2

2.1 Observing programs

Data from the surveys of WOCE/JGOFS, CLIVAR, and GO-SHIP form the GLODAPv2 backbone. The WOCE/JGOFS global ocean survey was carried out in the 1990s and included 64 (mostly) coast-to-coast one-time sections and several addi-

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
ESSDD
doi:10.5194/essd-2015-42
An internally consistent data product for the world ocean
A. Olsen et al.

tional repeat surveys (King et al., 2001). The data enabled an unprecedented mapping of ocean structure and tracer distribution (Sparrow et al., 2005–2011). Notably, the WOCE/JGOFS survey data underlies the current global interior ocean carbon climatology (Key et al., 2004) and through that, most data-based estimates of the global ocean sink for anthropogenic CO_2 (e.g. Khatiwala et al., 2009; Sabine et al., 2004).

The repeat hydrography program within the framework of CLIVAR was instigated in the early 2000s (Feely et al., 2014), aiming for a global repeat survey along selected WOCE sections within a decadal timeframe. This effort evolved into GO-SHIP, a sustained ship-based repeat hydrography program for documenting changes in ocean circulation, structure, heat, freshwater, oxygen, carbon, etc. (Hood et al., 2010). Over 2003–2012 this program carried out the first global reoccupation of selected WOCE hydrographic sections (see Talley et al., 2016, for a review). The next phase of the ongoing survey aims for completion by 2023.

In addition to the data from WOCE/JGFOS and CLIVAR/GO-SHIP, GLODAPv2 contains the data from the large-scale surveys of the 1970s and 1980s: GEOSECS, TTO, and SAVE, and from a multitude of national and regional programs. Examples include the times-series stations KNOT and 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 dedicated coastal surveys and time-series have been included on an opportunistic basis, based on their availability or their existence in any of the three pre-GLODAPv2 products. Examples of time series data *not* included in GLODAPv2 are 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 area *very* well maintained, organized, and distributed already. We strongly encourage dedicated efforts for organizing coastal data in a similar way.

2.2 GLODAPv2 predecessors

GLODAPv2 builds on GLODAPv1.1, CARINA, and PACIFICA (Fig. 1). GLODAPv1.1 evolved from the WOCE/JGOFS survey, and produced the first internally consistent global ocean data set. It enabled accurate and observation-based quantification of the ocean inventory of anthropogenic CO_2 (Sabine et al., 2004). GLODAPv1.1 includes data from 116 cruises carried out between 1972 and 1999, the majority of these are from the last decade of this period. GLODAPv1.1 is not merely a collection of the data from the individual cruises as submitted by PIs. Data were subjected to rigorous quality control (secondary QC) to ensure consistency. The GLODAPv1.1 team had realized that since analytical and standardization methods had evolved dramatically over the almost three decades of data collection, with introduction of new instrumentation (e.g. Johnson et al., 1993) and certified reference material (CRM, Dickson, 2001; Dickson et al., 2003) as prominent achievements, systematic biases were likely to exist in the data. The GLODAPv1.1 team used several approaches to minimize these, including, for total alkalinity, (TAlk), post-cruise calibrations with respect to CRMs and for total dissolved inorganic carbon (TCO_2), results from analysis of replicate samples at C. D. Keeling's mercury manometer instrument at Scripps Institution of Oceanography (SIO). In addition, the team carried out a systematic comparison of data from the different cruises to identify and eliminate any remaining offsets, using multiple linear regression (MLR), isopycnal and crossover analysis. The crossover analysis was carried out using deep water data (> 1500 or 2000 m, depending on region). At these depths a decadal survey can *generally* be considered synoptic, and significant and systematic offsets between two cruises indicate calibration issues. Since each cruise usually overlaps with several others, an assessment of all offsets combined will reveal the biased data.

An unbiased and automated crossover procedure was published by Gouretski and Jancke (2001) and Johnson et al. (2001). They introduced the use of least-squares models (Menke, 1984; Wunch, 1996) to determine the set of corrections required to simultaneously minimize all cruise-by-cruise offsets. Let **G** be the model matrix of size *o* x *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:

$$\mathbf{G} \cdot \mathbf{m} = \mathbf{d} \quad (1)$$

then

$$\mathbf{m} = \mathbf{G}^{\mathrm{T}} \cdot (\mathbf{G} \cdot \mathbf{G}^{\mathrm{T}})^{-1} \cdot \mathbf{d} \quad (2)$$

This model is known as the 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 the uncertainties of the crossover offsets and a priori information on expected measurement accuracy of each cruise into account, while the former only uses the uncertainties of the crossover offsets. These have become the favored models (Tanhua et al., 2010a), and both were used here (Sect. 3.4).

The GLODAPv1.1 team used these procedures to various degrees. For the Atlantic, crossover offsets were evaluated on a case-by-case basis and combined with MLR derived offsets and internal consistency analyses to determine recommended adjustments (Wanninkhof et al., 2003). In the Pacific, manual crossover analysis, MLR, internal consistency, and isopycnal analysis methods were used to determine recommended adjustments to the inorganic carbon chemistry data (Sabine et al., 2002). The final adjustments were based on a subjective combination of the various results. Finally, for the Indian Ocean the crossovers were manually inspected (Sabine et al., 1999).

The GLODAPv1.1 team focused on TCO_2 and TAlk, and the four halogenated transient tracers CFC-11, CFC-12, CFC-113, and CCl_4 (Table 1). They did not themselves quality control the salinity, oxygen, and nutrient data of the cruises, but adopted corrections suggested by three independent efforts, as listed in Table 2. As a result, the GLODAPv1.1 adjustments for these parameters come from different sources – even in the

ESSDD
doi:10.5194/essd-2015-42
An internally consistent data product for the world ocean
A. Olsen et al.

same basin – which is far from ideal. In addition the adjustments for oxygen and nutrients are a mixture of multiplicative and additives ones. As noted by T. Takahashi, LDEO (T. Takahashi, personal communication, 2005), multiplicative nutrient adjustments are much preferred because additive adjustments can seriously bias surface ocean values (Tanhua et al., 2010b).

Apart from said variables, GLODAPv1.1 included data of $\Delta^{14}C$, $\delta^{13}C$, ^{3}H, $\delta^{3}He$, helium, and several derived parameters. The isotopic data were not subjected to secondary quality control. Missing salinity, nutrients and oxygen data were vertically interpolated from existing data at the same station whenever reasonable. All data were flagged using the WOCE scheme (Table 3). An additional flag value of zero, "0", was introduced for data that could have been measured but was somehow approximated – normally through interpolation or seawater CO_2 chemistry calculations.

The CARINA Atlantic Ocean carbon data synthesis effort was based on an initial collaboration instigated by D. Wallace and L. Mintrop, Kiel, in 1999. This effort dwindled in the early 2000s due to lack of funding, but was reinitiated and expanded in scope at an international IOCCP sponsored workshop at Laugarvatn, Iceland in 2006 (IOCCP, 2006), mainly as a result of funding obtained through the EU integrated project (IP) CARBOOCEAN (project 511176 (GOCE)). At this workshop the CARINA goal was set to collate the ample amount of Arctic, Atlantic and Southern Ocean carbon data available but not included in GLODAPv1.1, and subject these to secondary QC.

CARINA evolved into a community effort thanks to CARBOOCEAN funding and lasted for 4 years. Data from 188 cruises and campaigns were assembled, documented, subjected to primary and secondary QC and published as three bias-corrected data product files, one each for the Arctic Mediterranean Seas, the Atlantic Ocean, and the Southern Ocean. All of the original, unadjusted, data were also published in the form of WOCE exchange format files (Swift and Diggs, 2008) through a single web page at CDIAC (http://cdiac.ornl.gov/oceans/CARINA/). Subsequently, CCHDO (Clivar & Carbon Hydrographic Data Office) also imported copies of the individual CARINA cruise files.

ESSDD
doi:10.5194/essd-2015-42

The CARINA group adopted the crossover and inversion analysis as their main secondary QC tool. The group quickly realized that the large data density rendered manual crossover identification and evaluation impractical. Hence, while the data were ingested and subjected to primary control at Princeton University, automated crossover and inversion scripts were developed and applied by the European partners (Tanhua et al., 2010b). This general data flow has been followed in subsequent efforts. The crossover and inversion scripts largely implement the routines of Johnson et al. (2001), or further developments of them. In addition, MLR and internal consistency analyses were used for the secondary QC. The CARINA group also introduced the use of a set of minimum allowable adjustments, respecting expected measurement accuracies of each parameter.

To keep track of the information, the CARINA group developed a web site for organizing output from the secondary QC (Tanhua et al., 2010b). This website (http://carina.geomar.de) has the details of the 4840 crossover locations, including the reasoning behind adjustments that were made. The CARINA group also documented their work in a special issue of the journal Earth System Science Data (Tanhua et al., 2009–2010). This issue includes 20 peer-reviewed papers that present details of the quality control of different parameters in the CARINA regions (Table 4). Besides the 12 core parameters subjected to secondary QC (Table 1), the CARINA product includes data of SF_6, carbon isotopes, He, δ^3He, ^{3}H, and derived and interpolated data more or less as in GLODAPv1.1. No mapped product was prepared, however, numerous subsequent publications with maps, vertical distributions and, modeling have been produced (Manizza et al., 2011; Olsen et al., 2010; Jeansson et al., 2011; Velo et al., 2013; Schmittner et al., 2013; Carter et al., 2014).

PACIFICA is the most recent of the GLODAPv2 precursors. It was initiated in 2006, at the PICES-XV annual meeting and published in 2013 (Suzuki et al., 2013). Altogether, PACIFICA includes data from 306 cruises, covering the time period 1985–2010; 35 of these have previously been published as part of GLODAPv1.1 and/or CARINA, while the remainder are new additions to the data synthesis effort. For the secondary QC the PACIFICA team adopted the crossover and inversion routines developed by the CARINA group. The following parameters were subjected to secondary QC: salinity, oxygen, nitrate, silicate, phosphate, TCO_2, and TAlk (Table 1). All of the PACIFICA recommended adjustments are published at the PACIFICA website (http://pacifica.pices.jp/about.html). Besides its seven core parameters, PACIFICA includes data of SF_6, carbon isotopes, He, δ^3He, ^{3}H, neon, and derived and interpolated data more or less as GLODAPv1.1. No mapped product was prepared.

2.3 GLODAPv2 instigation and goals

In the years 2009–2010 the EU IP CARBOCHANGE, the follow-up project of CARBOOCEAN, was being prepared. By then two of the interior ocean carbon datasets had been completed, GLODAPv1.1 and CARINA, and PACIFICA was well underway. As summarized above, these had been constructed following basically the same principles – albeit with some differences in their execution, and they largely contained data from different cruises. Besides, the first CLIVAR/GO-SHIP global repeat survey was coming to an end and there was an evident need to calibrate and publish these data as an integrated product to promote and support analyses of large scale decadal ocean variability and change. It was obvious that the time was ripe for the synthesis of syntheses, GLODAPv2.

The GLODAPv2 team (Table 5) assumed responsibility for producing a unified bias-corrected global, cruise-based, interior ocean carbon data product by combining all of the data available in GLODAPv1.1, CARINA, and PACIFICA plus any other "new" data, particularly those from the CLIVAR/GO-SHIP survey. The team set the following goals:

- To calibrate the data for the core parameters salinity, oxygen, nitrate, silicate, phosphate, TCO_2, TAlk, pH, and the four halogenated transient tracer species, using procedures developed for CARINA.
- To do the bias correction conservatively to preserve real variability in the final product. Respecting the initial minimum adjustment limits as listed given in Table 6

ESSDD
doi:10.5194/essd-2015-42
An internally consistent data product for the world ocean
A. Olsen et al.

was one guideline in this respect, preforming all crossover and inversion analyses on data from deeper than 2000 m (in some cases 1500 m) and respecting known or likely ocean trends at these depths were the others.

- To include all other commonly occurring parameters beside the 12 core ones.
- To interpolate missing oxygen and nutrient data from existing data at the same station following the CARINA practice.
- To calculate the third seawater CO_2 chemistry parameter (pCO_2 is not included in GLODAPv2, only TCO_2, TAlk, and pH) wherever measured data for two of them were present.
- To prepare an updated mapped global ocean carbon climatology.
- Finally, all original – unadjusted – data were to be made available in WOCE exchange formatted data files at a single access point.

3 GLODAPv2 production

GLODAPv2 was produced through a series of steps. The following strategy was set up early on:

1. Identify and ingest data not included in GLODAPv1.1, CARINA, or PACIFICA due to their recent dates or unavailability at the time of each respective synthesis, and subject these to primary QC. This group of data is referred to as “GLODAPv2 (NEW)” in the following.
2. In parallel, re-evaluate GLODAPv1.1 using the CARINA-developed analysis tools (Tanhua et al., 2010b) to enhance its consistency with respect to CARINA and PACIFICA; this GLODAPv1.2 product, is not to be publicly released, but used internally in step 3.

ESSDD
doi:10.5194/essd-2015-42

3. Combine GLODAPv1.2 with CARINA and PACIFICA to give a global reference data product, 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 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. 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.

By and large this work plan was followed except for a few customizations resulting from the invariable exceptions, delays, surprises and mistakes. All aspects of GLODAPv2 production and outcome are presented here except for the mapped climatology that is described by Lauvset et al. (2015).

3.1 Assembling data from "new" cruises and primary QC

GLODAPv1.1 contained data from cruises carried out until 1998 and CARINA from cruises carried out until 2005. One central GLODAPv2 goal was to add more recent data as well as older data that had not previously been included.

We directly contacted PIs known to have carried out relevant cruises. Additionally, a request letter was circulated to the ocean carbon science community through the IOCCP, and the SOLAS and IMBER core projects of the IGBP. Many PIs were forthcoming leading to submission of significant new data for inclusion in GLODAPv2. Our

ESSDD
doi:10.5194/essdd-2015-42

product contains data from 168 "new" cruises. These cruises are listed in the Supplement. For listings of the cruises from GLODAPv1.1, CARINA, and PACIFICA the reader is referred to the web pages for each product:

- For GLODAP: http://cdiac.ornl.gov/oceans/glodap/
- For CARINA: http://cdiac.ornl.gov/oceans/CARINA/
- For PACIFICA: http://cdiac.ornl.gov/oceans/PACIFICA/

All incoming data were merged as necessary, converted to WOCE exchange format, (Sect. 5.1) and subjected to primary QC to ensure that potential outliers were flagged as questionable (WOCE flag 3) or bad (WOCE flag 4). The primary QC was carried out following routines outlined in Sabine et al. (2005) and Tanhua et al. (2010b), 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, used for individual data points in GLODAPv2, are listed in Table 3.

3.2 Revision of the GLODAPv1.1 data product

Three inconsistencies between GLODAPv1.1, CARINA, and PACIFICA needed consideration, before the global reference data set used in step 3 was constructed:

1. As mentioned above, the evaluation of biases in the parameters salinity, oxygen, nitrate, silicate, and phosphate had not been carried out in GLODAPv1.1 per se, but were extracted from three other analyses (Gouretski and Jancke, 2001; Johnson et al., 2001; Mordy et al., unpublished, confront Table 2) – that contained much of the data as part of other data collections. Hence, these data had not been quality controlled as an entity.

2. No minimum adjustment limits were applied for these 5 parameters, i.e., all corrections suggested by the crossover and inversion analyses carried out by the authors mentioned above were applied to adjust the data (except for the Mordy

et al. (unpublished) adjustments that were filtered before being given to the GLODAPv1.1 team). This is in contrast to the practices of CARINA and PACIFICA, incuring the risk of removing real variations in the data as well as applying corrections beyond the expected level of accuracy of each parameter.

3. In contrast to CARINA and PACIFICA, GLODAPv1.1 used mostly additive adjustments for oxygen and nutrients (Table 2). This is at odds with the nature of likely sources of bias (Tanhua et al., 2010b), and also led to a few highly anomalous surface water concentrations in the final data product.

GLODAPv1.1 was therefore revised to produce GLODAPv1.2. This was carried out by analyzing all of the original – unadjusted – core data in GLODAPv1.1 with the crossover and inversion scripts developed during CARINA (Tanhua et al., 2010b). The crossover scripts identified more crossovers than the $\sim 1^{\circ}$ criterion of GLODAPv1.1 (Sabine et al., 2005), and the inversion scripts suggested more corrections. All suggested corrections were carefully scrutinized and adjustments were applied after, in particular, checking that they respected any expected time trends, natural ocean variability or were not based on previously undetected outliers. GLODAPv1.2 will not be released as an individual product.

3.3 Secondary QC of the core parameters in the GLODAPv2 (NEW) set

The secondary QC of all data collected outside the Arctic Ocean proper (i.e. south of the Fram and Bering straits) was carried out using the secondary quality control toolbox by Lauvset and Tanhua (2015). Every core parameter of each new cruise was subjected to a crossover analysis against a reference database, which was constructed by combining GLODAPv1.2, CARINA, and PACIFICA. Whenever a weighted average offset exceeded the predefined minimum adjustment limit for that parameter, the data were scrutinized and an adjustment applied whenever justified.

The secondary QC of all data collected within the Arctic Ocean proper was carried out by inspection of average property values in its individual basins, and inspection of

ESSDD
doi:10.5194/essd-2015-42

deviations from the values derived using a set of MLR equations specific to the various regions, following the methods by Jutterström et al. (2010).

3.4 Regional crossover and inversion analyses of the merged preliminary product, GLODAPv2 Beta

The regional crossover and inversion analyses were carried out using automated crossover routines as described in Tanhua et al. (2010b). We used the "running cluster" variant for determining crossovers and the WLSQ inversion to determine corrections. This analysis was first carried out on the combined GLODAPv1.2, CARINA, PACIFICA, and adjusted GLODAPv2 (NEW) data, GLODAPv2 Beta, and led to the revision of a fraction of the already vetted adjustments. However, at this stage two issues arose:

Issue 1. For salinity and oxygen two types of submitted data exist. The first type included a single column of parameter values, 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). The second type includes two columns of parameter values, one giving the bottle values and the other giving the CTD values. The first type represents no problem, the submitted data are subjected to primary and secondary QC and bias corrected whenever required, irrespective of them being bottle or CTD values. For the second type, on the other hand, there were in many instances significant differences between the bottle and CTD values, implying that the CTD values had not been properly calibrated with respect to the bottle values before data submission. This had to be dealt with, since in these cases it was normally the CTD values that gave the complete profile, while the (likely accurate) bottle values were sampled more sparsely. We therefore reset all of the existing adjustments for these parameters and introduced a pre-calibration step, identifying seven distinct scenarios, of which the fourth turned out to not occur (it is kept in the list however, to maintain compatibility between this document and material produced during the secondary QC effort):

1. No data are available: no action needed.
2. No bottle values present: use CTD values.
3. No CTD values present: 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 and replace missing bottle values with 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.

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. When something does go wrong with the CTD sensor(s) and this is not discovered until the cruise is over, the cost is catastrophic. The fact also remains that bottle salt/oxygen samples are about the only way to be sure when a sample bottle mis-trips or leaks. Additionally, the cost of analyzing a few expensive tracers (particularly isotopes) on samples that mis-tripped/leaked/etc. quickly exceeds the relatively small cost of shipboard salt/oxygen analysis.

Issue 2. By now we had layer upon layer of adjustments for several datasets, e.g. the CARINA published adjustments + our revision of these, or the PACIFICA published adjustments + our revision of these, or the GLODAPv1.2 derived adjustments + our revision of these. It soon became very challenging to keep track of any systematic

ESSDD
doi:10.5194/essd-2015-42

patterns, such as for example if data from any specific groups of cruises or from specific time periods tended to be offset by some particular magnitude. Moreover, in some cases the justification for the originally proposed adjustment was not entirely clear, this renders the justification of any downstream adjustments unclear as well.

Given these issues we therefore decided to reset the entire database, and carry out a crossover and inversion analysis on the unadjusted data (of which salinity and oxygen data were pre-calibrated as outlined above). As a start, regional analyses were carried out. These were subsequently augmented with customized analyses in order to tease out any underlying patterns in the biases. For example invoking the assumption that cruises from the CLIVAR/GO-SHIP global survey 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 for all these inversions were to take advantage of groups of cruises of known high-quality forming a cohesive grid, against which cruises of unknown quality could be evaluated. This turned out to be a fruitful approach. The most important results are summarized by parameter in Sect. 4.3.

3.5 Quality control of the halogenated transient tracer data

Given their strongly transient nature and low concentration in most deep waters, crossover and inversion analysis is of limited value for halogenated transient tracers (CFC-12, CFC-11, CFC-113, and CCl_4; CFCs for short) data. Further, in the past synthesis products CARINA, GLODAPv1.1, and PACIFICA, the CFCs included 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.

ESSDD
doi:10.5194/essd-2015-42

Hence, the secondary QC of the CFC data was not included in steps 2–4 described above, but was undertaken in a dedicated analysis that focused on the data from the 168 new cruises, and the PACIFICA and Southern Ocean CARINA data. To ensure consistency we also re-evaluated the GLODAPv1.1 CFC data using the same procedure.

The methods included inspection of surface saturation levels, which must be realistic, evaluation of the relationships among the four tracer species for each cruise, and crossover and inversion analysis following CARINA protocols as described by Jeansson et al. (2010) and 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 SF_6 was not carried out because few data were available, however we plan that this will be done in future versions of GLODAP since this parameter is now routinely measured.

3.6 Scale conversion and quality control of the pH data

In the three GLODAPv2 predecessors, data for pH have been dealt with in various ways:

- pH data were not included the GLODAPv1.1 product files *per se*, but 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 (the quality controlled) TCO_2 and TAlk data were also included. The pH data included in the CARINA product files were unified to the Sea-Water Scale (SWS) at 25 °C and surface (0 dbar) pressure.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
ESSDD
doi:10.5194/essd-2015-42
An internally consistent data product for the world ocean
A. Olsen et al.

– PACIFICA included the measured as well as calculated pH data as CARINA, but no secondary QC was preformed (Suzuki et al., 2013). The pH data were reported at the Sea-Water Scale (SWS) at 25 °C and surface (0 dbar) pressure.

For GLODAPv2 we decided to include quality-controlled pH at the total hydrogen ion scale at standard (25 °C and surface (0 dbar) pressure) and in situ (in situ temperature and pressure) conditions in the product. The total scale was preferred since this 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 refitted by Dickson and Millero carbonate dissociation constants (Dickson and Millero, 1987; Merbach et al., 1973), we used the ones of Lueker et al. (2000), which 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, either as observed or as calculated from TCO_2 or TAlk, in order to determine or verify the scale and conditions. This job was somewhat alleviated 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 parameter measured at the cruise in question, or if these measurements 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 labeled as not subjected to full secondary QC (−888, Sect. 4.2).

ESSDD
doi:10.5194/essd-2015-42

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, we performed crossover analysis 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). Recognising this, such data have not been included in the data product unless passing full secondary QC, criteria 2 or 3. Otherwise these were suspended (−666, Sect. 4.2)

4 GLODAPv2 secondary QC results and adjustments

4.1 Preservation of real variability

Throughout secondary quality control we recognized the risk of removing real signals of variability present in the data since the crossover and inversion is an objective method and does not discriminate between real difference and measurement bias. Using data from deeper than 2000 m helped, but in some regions time trends are expected to occur at these dephs over the time scales considered. Therefore, each correction suggested by the crossover and inversion was closely scrutinized. Whenever doubt existed, adjustments were not applied. In particular we were careful about applying adjustments in regions of strong variability (such as the Nordic Seas overflow), or when time trends were detected or suspected. As an example our method of preserving trends, Fig. 2

ESSDD
doi:10.5194/essd-2015-42

shows one type of figure used to evaluate the crossover offsets generated by our analyses. This particular cruise is an occupation of WOCE line AR07W in the Labrador Sea (18HU19960512), 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, upwards to correct the bias of these these data. However, plotting the crossover offsets vs. time (Fig. 2, lower panel), clearly reveals the strong trend in TCO_2 in the area. This shows that the TCO_2 data of 18HU19960512 are not significantly offset from its contemporary cruises, while lower than data from later cruises. This is as expected from the invasion of anthropogenic CO_2 and deep ventilation that occurs in the Labrador Sea (e. g. Yashayaev, 2007). No adjustments were applied to these data.

On the other hand, in a few cases, when the data were very precise and from regions where no deep-ocean changes are expected to occur at the time scales we considered, we applied adjustments somewhat smaller that the limits given in Table 6.

4.2 The Adjustment table

The results of the secondary QC analyses were entered into the online GLODAPv2 Adjustment Table at GEOMAR. This is similar in form and function to the Adjustment Table used in CARINA (Tanhua et al., 2010b). A permanent, non-editable version of this Adjustment Table is available at http://glodapv2.geomar.de. In contrast to the CARINA Adjustment Table, this does not include an entry for each GLODAPv2 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, information is restricted to one entry per cruise including magnitude of any recommended adjustments to the twelve core parameters and comments regarding their justification.

Whenever data are not available for a specific parameter at any given cruise the value −999 has been entered in the adjustment table. Whenever data are available the secondary QC had five potential outcomes:

ESSDD
doi:10.5194/essd-2015-42

1. The data are of good quality, consistent with the rest of the dataset and should not be adjusted. The value 0 (for salinity, TCO_2, TAlk, pH) or 1.00 (for oxygen, nutrients, CFCs) appears in the Adjustment Table.

2. The data are of good quality but are biased: adjust by adding (for salinity, TCO_2, TAlk, pH) or by multiplying (for oxygen, nutrients, CFCs) the number in the Adjustment Table.

3. The data appear of good quality but their nature, being from shallow depths, coastal regions, without crossovers or similar, prohibits full secondary QC, the value −888 appears in the Adjustment Table.

4. The data have not been QC'd, are of uncertain quality, and suspended until full secondary QC has been carried out. The value −666 appear in Adjustment Table and these data are not included in the data product. This option was introduced for non-QC'd NBS pH data in GLODAPv2 (Sect 3.6). For future updates (Sect. 7), all incoming data will be suspended until QC'd.

5. The data are of poor quality and excluded from the data product. The value −777 appears in the Adjustment Table. This was applied both when our QC routines revealed that the data were poor and also when all data had already been flagged questionable or bad by the data provider.

The justifications for our decisions were entered only partially bearing in mind that these quickly become incomprehensible for anyone outside the team – or even for the team members themselves. Besides, the 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. When accessing the table be aware of the following:

– The GEOMAR Adjustment Table gives the dataset source of each cruise, CARINA, PACIFICA, GLODAPv1.2, or GLODAPv2 (NEW).

doi:10.5194/essd-2015-42

- A comment was not always entered when the data appeared unbiased.
- For CARINA cruises we used the CARINA recommended adjustment as the initial value and also included all comments entered during the CARINA QC process, as these were already available to us in the appropriate format. Any comments from before 2011 are thus “CARINA comments”, while any comments from after are “GLODAPv2” revisions and based on the analysis of the GLODAPv2 Beta product, or the re-analysis of the unadjusted data set as described in Sect. 3.4.
- For PACIFICA-sourced cruises we used the PACIFICA recommended adjustments as the initial values. No comments were available with these, hence those that appear in the Adjustment Table are all ours, and justify revisions to the original PACIFICA recommended adjustments, or simply states that these should be maintained, as based on the analysis of the GLODAPv2 Beta product, or the re-analysis of the unadjusted data set as described in Sect. 3.4.
- For GLODAPv1.2 cruises, all adjustment values and comments that appear are based on our analyses. Either from ones described in Sect. 3.2 or from the ones described in Sect. 3.4. Comments from 2012 are typically based on the former, while comments by Steven van Heuven from 2014 are typically based on the latter.
- For GLODAPv2 (NEW) cruises, all adjustment values and comments that appear are based on our analyses. Either from ones described in Sect. 3.3 or from the ones described in Sect. 3.4. Comments by Sara Jutterström or Siv Lauvset typically refer to the former, while comments by Steven van Heuven typically refer to the latter.
- For CFCs the comments are either inherited from CARINA or posted following our analyses described in Sect. 3.5.
- For pH the comments are either inherited from CARINA or posted following our analyses described in Sect. 3.6.

ESSDD
doi:10.5194/essd-2015-42

– Some of the comments refer to workshops where the magnitudes of adjustments were discussed and decided; these are: Bergen, November 2012; Norwich, April 2013; Groningen, October 2013; and Bremen, January 2014. The Bergen workshop dealt with GLODAPv1.2 (Sect 3.2), the secondary QC of the GLODAPv2 (NEW) cruises (Sect. 3.3) was the primary subject of the Norwich workshop, while the analysis of the preliminary product, GLODAPv2 Beta was the subject of the Groningen and Bremen workshops – before the entire product was reset.

As an example of information available at the Adjustment Table we take a closer look at TAlk results for cruise 06MT20040311. Note that familiarity with the crossover and inversion method as described in Tanhua et al. (2010b) is required in order to understand this example and the reader should look up this specific cruise entry in the Adjustment Table while reading. The row for 06MT20040311 is found using the search field in the Adjustment Table, upper right, and the specific summary page for this cruise is opened by clicking at either of the symbols in the leftmost column at the row for this cruise. Once at this cruise's summary page, the figures and comments for TAlk can be accessed by clicking at 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 10 June 2008 and one by Steven van Heuven dated 8 January 2015. In correspondence with the guidelines above, the former was entered during CARINA, while the latter is entered based on analyses described in Sect. 3.4. 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 reset (unadjusted) GLODAPv2 database. While the data were not adjusted in CARINA, since the bias appeared less than the 6 µmol kg^{-1} threshold, during GLODAPv2 we felt that the evidence were convincing enough to apply an adjustment of +4 µmol kg^{-1}. An example of one of the crossovers that supports this adjustment is provided in Fig. 3. 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: sigma-4, potential temperature, and 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., 2010b). 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 (WMO) 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 ± 3.16 µmol kg^{-1}, while it is -3.68 ± 2.48 and -4.83 ± 3.17 µmol kg^{-1} in potential temperature 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. 4. This can also be obtained from the page for this cruise in the Adjustment Table. This shows that the mean offset is -4.5 ± 4.7 µmol kg^{-1}, and cannot be unequivocally 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 µmol kg^{-1} to these data.

Another example is phosphate of 316N20050821. Two comments are provided for phosphate in the Adjustment Table, by Are Olsen 26 February 2014 and by Steven van Heuven 16 June 2014. 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 5 crossover plots and one summary figure (named Xresults.png), uploaded in January 2014 (holding mouse pointer over name gives upload date). These were created during the analyses summarized in Sect. 3.3; indeed this specific cruise is part of the GLODAPv2 (NEW) set. 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 (comments have not always been entered when no adjustment was suggested). However, next two comments revise this, as first mentioned in Are Olsen's comment which summarize a conclusion met for this cruise at the Groningen workshop based on analysis of GLODAPv2 Beta, and then in Steven's authoritative comment following his analysis of the unadjusted data.

These two examples illustrates that it certainly is possible to locate our main evidence for adjustments that we have applied and to backtrack the steps we took to unearth these.

4.3 Secondary QC summary

In this section we provide a brief overview of the outcome of the secondary quality control by parameter. In particular, we summarize noteworthy features and exceptional cases.

Data from altogether 734 cruises were subjected to secondary QC. For ten of these the secondary QC revealed that most if not all of the data were of unacceptable quality. Further, for these ten 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 ten cruises are available in the Cruise Summary Table (CST, cf. 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.

The GLODAPv2 data product 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 7) in

ESSDD
doi:10.5194/essd-2015-42

order to take into account potential changes in personnel, equipment, and procedures during their execution and partly because 4 cruises were adjusted on a per station range basis as a result of obvious bias in one or several parameters 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.

4.3.1 Salinity calibration summary

All of the 780 entries that were analyzed came with some kind of salinity data (Table 8). The outcome of the pre-calibration step is summarized in Table 9. For 295 of the entries only bottle salinity was reported (case 3; or rather, bottle data was *presumably* reported – in many cases these values may have been mislabeled CTD salinity values), while CTD salinity was the only salinity parameter reported for 77 entries (case 2). For the entries that included both bottle and CTD salinity (case 5–7), our analysis revealed significant mismatch between these values at 144 entries (case 6 and 7). This is almost 20 % of all entries. A fraction of this size is simply unacceptable given the complexity of modern climate change issues. We encourage the community to *only* submit calibrated data in the future. The "after-the-fact" linear calibration that we performed will never be as good as what could have been done by the data originators. The CTD salinity values from 141 entries were salvaged by our simple calibration routine (case 6), while CTD salinity from 3 were not (case 7).

For 162 of the entries, full secondary QC could not be carried out (Table 8). Data from 6 entries were deemed to be of too poor quality for inclusion in GLODAPv2. Typically, these showed large and depth dependent offsets and/or unrealistic scatter compared to background data for most parameters. Of the remaining 612 entries, the salinity data from 41 were found clearly biased, warranting an adjustment.

The size distributions of the adjustments that have been applied appear in Table 10 and Fig. 5a. Adjustments smaller than the initial threshold have only been applied to 5

ESSDD
doi:10.5194/essd-2015-42

entries, while the bulk of the adjustments applied are between 0.005 and 0.010. The largest negative and positive adjustment applied are −0.025 and +0.025. For a brief discussion on the resulting improvements of dataset consistency for salinity and subsequent parameters, see Sect. 4.4.

4.3.2 Oxygen calibration summary

Of the 780 entries, 722 had oxygen data (Table 8). For 21 of these only CTD oxygen were present in the file, for 520 only "oxygen" values were present (presumably bottle data, but conceivably these values occasionally were CTD oxygen), while for 181 both types were reported in the file (Table 9). These were handled as described in Sect. 3.4, and the number of cases encountered is summarized in Table 9. For 99 of these entries the CTD oxygen values matched the bottle oxygen data nicely (case 5). For 82 entries the CTD oxygen data had apparently not been calibrated with respect to the corresponding bottle values (case 6 and 7). This is almost 50 % of the data sets submitted with both CTD and bottle oxygen data, hence a quite widespread malpractice, appropriate calibration of any sensor data should of course be carried out at PI level prior to data submission. We encourage a better practice in this respect in the future. Of these 82, 62 could be salvaged (case 6) through a simple linear regression, while 20 were discarded (case 7) as it was not possible to achieve a good fit using our – admittedly simple – procedure.

After this initial calibration step, data of CTD and chemically determined oxygen concentration were merged into a single, "hybrid" variable. Crossover, inversion and subsequent adjustment for bias minimization were performed on this hybrid oxygen. Table 10 and Fig. 5b provide overviews of the applied adjustments. About two thirds of the entries were deemed to be accurate to within our limits, and thus did not require an adjustment. A total of 7 of applied non-zero adjustments were smaller than the threshold of 1 %. These necessarily were cruises with sufficiently high precision that such small bias could be observed beyond doubt. Almost half of the non-zero adjustments were between 1 and 2 %. The other half of the applied non-zero adjustments (99 cruises)

ESSDD
doi:10.5194/essd-2015-42

was greater than 2 %. 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 calibration summary

Nitrate data were available for 709 of the 780 entries (Table 8). 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. Of these, 380 were accepted to be accurate and 150 entries were adjusted (Table 8). Of the applied adjustments, 50 entries (i.e. 33 %) are beneath the initial 2 % limit while 49 % are between 2 and 4 % (Table 10). 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 (see Sect. 4.4). 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. A very 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 calibration summary

Silicate data were available for 678 of the 780 entries (Table 8). The silicate data of 33 entries were found to be of poor quality exhibiting excessive scatter, large offsets or a combination of the two. For 255 entries the silicate data were considered to be accurate to within the uncertainty of our methods, while data from 264 entries were

ESSDD
doi:10.5194/essd-2015-42

adjusted. This is almost 40 % of the entries with silicate, making silicate is the most frequently adjusted parameter 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 our crossover and inversion analyses. US silicate data tended to be 2 % higher than the silicate values recorded at Japanese entries (Fig. 6a and b). 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). Kenso 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). To remedy this rather severe inconsistency we pre-adjusted the Japanese data by +1 % and US data by −1 %. This removed the systematic difference and a clear baseline emerged (Fig. 6c 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 our choice of splitting the difference between the US and Japanese efforts may result in our Pacific Ocean data product to be – 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 of high priority for the nutrient analytical community.

For the South Atlantic and Indian basins, crossover and inversion were performed without notable incidents, and in both basins the dataset consistency has markedly improved (see Sect. 4.4)

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

ESSDD
doi:10.5194/essd-2015-42

themselves in addition to the multiplicative ones our 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 ∼ 50 µmol kg^{-1} tend to be very sensitive to freezing, which can increase the measured concentration by up to 15 % (K. Bakker, NIOZ, personal communication, 2014) while samples with lower silicate are not affected by freezing. Freezing was occasionally suspected (and then generally confirmed) to have been performed on cruises, forcing arbitrary removal of data, and complicating the automated crossover analysis. Although the average offset for silicate at crossovers has been reduced by our efforts 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.

4.3.5 Phosphate calibration summary

688 entries included phosphate data. Of these, data from 59 were found to be of too poor quality for inclusion in the product. Adjustments were applied to 163 entries, Data from 184 entries could not be adequately checked with our routines (Table 8).

Of the 163 adjusted entries, 31 (highly precise) had adjustments smaller than the threshold. 132 entries had larger adjustments (Table 10) with the largest being about ±12 %.

4.3.6 Total dissolved inorganic carbon (TCO_2) calibration summary

TCO_2 was measured at 602 of the 780 entries included in GLODAPv2 (Table 8). During quality control 15 were found to be of too poor quality 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 µmol kg^{-1}. For 8 very precise entries we applied a smaller adjustment (Table 10, Fig. 5f). A few very large adjustments were applied, generally to historic entries (e.g. GEOSECS).

ESSDD
doi:10.5194/essd-2015-42

4.3.7 Total alkalinity (TAlk) secondary QC summary

465 entries had TAlk data. 106 of these could not be subjected to full secondary QC and were set to −888 (Table 8). 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 µmol kg^{-1}. About 75 % of the applied adjustments are equal to or larger than this (Table 10). TAlk is the second most frequently adjusted parameter in GLODAPv2 with 32 % recommended for adjustment. This was the result of a bias identified in Japanese Pacific cruises. Following crossover and inversion analysis (Sect 3.4), a very clear separation was observed between the US entries and (most of) the Japanese Pacific entries. The former had negative corrections recommended (i.e., data being consistently too high) and the latter had positive corrections recommended (i.e., data being consistently too low). These results are illustrated in Fig. 7a and b. The occasional presence of strongly outlying cruises did not obfuscate this result. The offset between US and Japanese labs appears to exist throughout the era of measurements, although in the latest 10 years, too few data exist to be sure. Typically, the available metadata for the Japanese cruises were sparse and did not include information on traceability to CRMs. However, the 6 or so Japanese results after 2005 that are estimated to *not* require an adjustment relative to the US are 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 Dickson (SiO, US), whereas the smaller Japanese lines have used a different method of standardization. The Japanese measurement effort appears to be more internally consistent than that of the USA. 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. Based on the information at hand, in particular the *documented* traceability to CRM for the US cruises, we decided to apply a +6 µmol kg^{-1} adjustment 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. 5g. Following this,

ESSDD
doi:10.5194/essd-2015-42

a crossover and inversion analysis was rerun and gave a clear baseline (Fig. 7c and d). This allowed us to identify adjustments required for individual cruises.

4.3.8 pH data adjustment summary

A total of 259 entries had pH data included, while 521 did not. Out of the ones with data, 59 were found accurate and not bias corrected, while 77 has been adjusted in the product. 67 could not been fully QC'd but we believe their quality is fine (–888), while data from 47 cruises has been suspended from the product as we believe their further QC is required. These are non-QC'd data supplied on the NBS scale (Sect. 3.6).

4.3.9 Halogenated transient tracer 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 8: 273/270 for CFC-11/CFC-12, but only 105 for CFC-113 and 72 for CCl_4). The range of CFC values 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 parameters. 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 8). 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. Hence, while for CFC-113 and CCl_4 about two-thirds of the applied adjustments are larger than 10 % (two times the limit), only about one-third of adjustments for CFC-11 and CFC-12 is of such magnitude (Table 10, and Fig. 5i–l).

ESSDD
doi:10.5194/essd-2015-42

Sulfur hexafluoride, SF_6, is now being measured on some cruises as a replacement for (or in addition to) CFCs. These measurements are included in the data product, but are too sparse for secondary QC.

4.4 Improvement of the internal consistency of the GLODAPv2 data collection

Application of adjustments was done with the aim of reducing the deep-water offsets between the many entries constituting the GLODAPv2 data product. The extent to which we have herein succeeded, we refer to as the "internal consistency improvement", which we express as the decrease in the weighted mean absolute offset between (i) the unadjusted data (after primary QC) and (ii) the adjusted data (after secondary QC). Although this certainly is not the only possible means of quantifying improvement, we believe it to be a good compromise in terms of implementability, understandability and "compactness". 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 we recognize the interesting nature of these details, in Table 11 we only report on the improvements per ocean basin and for the full world ocean. Note that relative improvement for nutrients and TCO_2 and TAlk is higher than that of salinity, accuracy for which was quite high for most cruises already. For all parameters, the internal consistency of the database has been significantly increased

5 GLODAPv2 product description and access

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/

ESSDD
doi:10.5194/essd-2015-42

oceans/GLODAPv2/). The orignal data and product files are described here, while the mapped climatology is described by Lauvset et al. (2015).

5.1 Original data

GLODAPv2 includes original data from 724 cruises. These may consist of several legs, and in some cases multiple related cruises have been merged. Our secondary QC was carried out on a per leg or, in a few rare cases, station range basis, as mentioned above. The CST includes a column that lists individual components of multi-leg cruises analyzed per leg, these also appear in Table 7 (see also Sect. 4.2). All cruise data for the 724 cruises and their associated metadata can be accessed through the CST.

The content of the original data files is as received from the originator, but have been updated to WOCE exchange format, whenever required. File headers, listing essential information on cruise and the analytical procedures were generated for all except the PACIFICA cruises. *No bias adjustments were applied to the data in these files.*

Each cruise and data file is uniquely identified with its GLODAPv2 cruise number and its EXPOCODE. We also include known aliases 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. Since we began this work, the responsibility for platform code assignment has been 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. We generally opted for the older or better-known code in these cases.

Note that for the following time series or campaigns the data have not been segmented into individual cruises, but maintained as collections under a single EXPOCODE, to ease record keeping: EGEE, GIFT, Iceland Sea, Irminger Sea, Kerfix, OWS Mike, and SWITCHYARD time series (assigned EXPOCODES

ESSDD
doi:10.5194/essd-2015-42

are 35A820050607, CARBOGIB2005, IcelandSea, IrmingerSea, 35UCKERIXTS, 58P320011031 and ZZIC2005SWYD), and 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 milliliter or micromol per liter were converted to micromole per kilogram (µmol 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 one atmosphere. The error made by an actual lab temperature deviating up to 5 °C from the assumed 22 °C is insignificant. For the milliliter to micromol conversion for oxygen, we used the factor 44.66, 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 have been applied throughout, and any questionable or bad data identified during our primary QC have been flagged accordingly in these files. However, note that whenever data from an entire cruise were found to be bad following our 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. We urge all users of the individual cruise data 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

ESSDD
doi:10.5194/essd-2015-42

the CST. These two sources can be complementary. For many cruises we also provide access to copies of written cruise reports through the CST and references to relevant scientific publications.

5.2 Product files

The GLODAPv2 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 we have produced four regional subset files. There is one for the Arctic, one for the Atlantic Ocean, one for the Pacific Ocean, and one for the Indian Ocean. The coverage, per decade, of the global GLODAPv2 file is given in Fig. 8, and that of each of the four regional ones in Fig. 9. 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 in the regional subset files, cruise numbers 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 "expocodes" is available, containing all 724 EXPOCODES.

The product files were prepared following the same general procedures as 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 for the 911 records that were associated with Niskin bottle number "0"

ESSDD
doi:10.5194/essd-2015-42

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, we merged any reported CTD and bottle values into one parameter. This was carried out following the scenarios summarized in Sect. 3.4. Basically, whenever both types were reported we replaced the missing bottle values with CTD derived ones, if they were consistent with the bottle values or could be calibrated linearly to achieve this. Any CTD data that were inconsistent with the bottle measurements and could not be calibrated are not included in the product files.

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

4. Whenever 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/US Department of Commerce, 1995), whichever was greater. This parameter is not research quality, but useful for drawing bottom topography for sections.

5. Values flagged 3, 4, or 5 were excluded and their flags set to 9. This applies both to flags set for individual records during the primary QC, and whenever the secondary QC concluded that all measurements of some parameters from a cruise should be discarded.

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

ESSDD
doi:10.5194/essd-2015-42

7. Multi-leg cruises that had been quality controlled on a per-leg basis were combined to single cruises. Table 7 provides an overview of these.

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 Gordon and Garcia (1992). In the few instances only potential temperature values were reported, these values were retained.

10. Whenever sampling pressure or depth was missing this was calculated following UNESCO (1983).

11. GLODAPv2 includes TCO_2, TAlk, and pH data. Generally, whenever only two seawater CO_2 chemistry parameters were reported, the third was calculated. In the final product files some of the cruises thus, invariably, has a mixture of calculated and measured values of specific CO_2 chemistry parameters, i.e. the cruises that had non-collocated measurements of three parameters. This is generally not a problem since the internal consistency of the seawater CO_2 chemistry data at these cruises has been established (Sect. 3.6). However, in some cases so few data were available for the third parameter 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 12 provides an overview of the cruises where measured data were replaced with calculated ones. On the other hand, for those cruises where

ESSDD
doi:10.5194/essd-2015-42

all three had been measured but there were a few holes in the record of each that did not diminish their scientific value, we did not fill these holes with calculated values. For the various constants involved we used the settings as in Velo et al. (2010), but used the Lueker et al. (2000) carbonate dissociation constants instead of the Dickson and Millero (1987) refitted Merbach et al. (1973) ones. The calculations were carried out using CO2SYS (Lewis and Wallace, 1998) for Matlab (van Heuven et al., 2009). Calculated data points are flagged "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 one of the input parameter's flag was "0" the calculated data were given a secondary QC flag of "0" as well. For example, if pH was calculated from unchecked TCO_2 and TAlk, its secondary QC flag have been set to "0", if pH was calculated from unchecked TCO_2 and but checked TAlk, its secondary flag is set to "0", but if the TCO_2 and TAlk data used were both checked, the pH secondary QC flag is set to "1". In the Adjustment Table, we have indicated the records that have calculated CO_2 chemistry parameters in the product 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 data were measured but calculated values are available in the product. And a value of −777c, for instance, means that measured data were bad, and calculated have been inserted in the product file. The Adjustment Table at http://glodapv2.geomar.de can also be exported to an ascii file. In this file a seperate 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 parameter calculated. The adjustment value itself refers to the original data, and should be interpreted as explained in Sect. 4.2.

12. Note also that, similar to GLODAPv1.1, the product files contains some TAlk values that have been calculated from discrete $p\mathrm{CO}_2$ and TCO_2, for cruises were data for only this pair were available. These TAlk data were treated as measured

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.

during our QC 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 (1995), Bu and Warner (1995), Bullister and Wisegarver (1998), and Bullister et al. (2002).

Besides the core parameters, the product includes data for the following: $\Delta^{14}C$, $\delta^{13}C$, ^{3}H, ^{3}He, He, Ne, $\delta^{18}O$, TOC, DOC, DON, TDN, SF_6, and Chl *a*. None of these were subjected to secondary QC. Table 13 specifies the file contents and lists parameter 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. We encourage users to remember that hard-working scientists made these measurements, often under severe conditions. Further, the PIs normally possess insight on the quality and context of the data not known to the GLODAPv2 team. We encourage users to invite individual data-providers to collaborate in scientific investigations that depend on their data. Importantly, this will promote further sharing of data, and better science. In the CST we have included citations to relevant scientific publications for individual cruises, whenever these were known to us. 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, scope, and complexity, errors and mistakes are bound to occur. Besides the product files, we therefore also provide a document that lists known-issues. This will be updated as new errors are found and reported to us by the user community. Cruise specific issues, e.g. errors or data updates are *also* given in field "General Comments for this Cruise" at each cruise's page in the online Adjustment Table.

ESSDD
doi:10.5194/essd-2015-42

7 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, and also revealed particular wide spread 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 (Sect. 4.3.1 and 4.3.2) is an intolerable practice, which can easily be fixed. Lack of documentation is another; 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 (L) or per unit mass (kg) sea water). The lack of universal and certified nutrient standards 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 were used. This can result from missing or excessive 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

ESSDD
doi:10.5194/essd-2015-42

products is an important and warranted activity. GLODAPv2 will form the starting point for routine future interior ocean syntheses, and we plan 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. Future plans also include a Live Access Server for interactive access to the GLODAP data products.

Acknowledgements. GLODAPv2 would not have been possible without the efforts of several hundreds of sea-going scientists who generated funding and dedicated time to collect in situ data, critical to our understanding of the earth system and central to our efforts for projecting its future, and willingly shared their data for the wider benefit of scientific progress. The cruise chief scientists and principal investigators are listed in the Cruise Summary Table at CDIAC, anybody who feels omitted should contact us (preferably CDIAC who maintains this table, kozyra@ornl.gov) so their names can be included.

The GLODAPv2 project itself received support from a number of agencies and projects. Importantly, the EU-IP CARBOCHANGE (FP7 264 878) provided funding for A. Olsen, S. van Heuven, T. Tanhua, and travel support for R. Key as well as the project framework that instigated GLODAPv2. A. Olsen further acknowledges generous support from the FRAM – High North Research Centre for Climate and the Environment, the Centre for Climate Dynamics at the Bjerknes Centre for Climate Research, the EU AtlantOS (grant agreement No 633 211) project and the Norwegian Research Council project SNACS (229 752). R. Key was supported by KeyCrafts grant 2012-001, CICS grants NA08OAR432 0752 and NA14OAR432 0106, NASA grant NNX12AQ22G, NSF grants OCE-082 5163 (with a supplement via WHOI P.O. C119 245) and PLR-142 5989, and Battelle contract #400 0133565 to CDIAC. A. Kozyr was supported by DOE contract DE-AC05-00OR2272 to UT-Battelle, operators of CDIAC under ORNL. S. K. Lauvset and E. Jeansson appreciates support from the Norwegian Research Council (projects DECApH, 214513 and VENTILATE, 229791). The International Ocean Carbon Coordination Project (IOCCP) also supported this activity through the US National Science Foundation grant (OCE- 124 3377) to the Scientific Committee on Oceanic Research. Benjamin Pfeil (University

of Bergen) provided help with some of the data management issues during the preperation of GLODAPv2 and Karel Bakker (NIOZ) shared his invaluable insight with us, helping with the secondary quality control of the nutrient data.

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, Tsukuba, Japan, No. 60, 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.

Bakker, D. C.E, Pfeil, B, Smith, K., Hankin, S., Olsen, A., Alin, S. R., Cosca, C., Harasawa, S., Kozyr, A., Nojiri, Y., O'Brien, K., Schuster, U., Telszewski, M., Tilbrook, B., Wada, C., Alk, J., Barbero, L., Bates, N. R., Boutin, J., Bozec, Y., Cai, W.-J., Caste, R. D., Chavez, F. P., Chen, L., Chierici, M., Currie, K., de Baar, H. J., Evans, W., Feely. R. A., Fransson, A., Gao, Z., Hales, B., Hardman-Mountford, N. J., Hoppema, M., Huang, W.-J., Hunt, C. W., Huss, B., Ichikawa, T., Johannessen, T., Jones, E. M., Jones, S. D., Jutterström, S., Kitidis, V., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefévre, N., Manke, A. B., Mathis, J. T., Merlivat, L., Metzl, N., Murata, A., Newberger, T., Omar, A. M., Ono, T., Park, G.-H., Paterson, K., Pierrot, D., Ríos, A., Sabine, C., Saito, S., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Sieger, R., Skjelvan, I., Steinhoff, T., Sullivan, K., Sun, H., Sutton, A. J., Suzuki, T., Sweeney, C., Takahahsi, T., Tjiputra, J., Tsurushima, N., van Heuven, S. M. A. C.,

Vandermark, D., Vlahos, P., Wallace, D. W. R., Wanninkhof, R., and Watson, A. J.: An update to the Surface Ocean CO_2 Atlas (SOCAT version 2), Earth Syst. Sci. Data, 6, 69–90, 2014.
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, 175–187, 2002.
Carter, B. R., Toggweiler, J. R., Key, R. M., and Sarmiento, J. L.: Processes determining the marine alkalinity and calcium carbonate saturation state distributions, Biogeosciences, 11, 7349–7362, doi:10.5194/bg-11-7349-2014, 2014.
Dickson, A. G.: pH scales and proton-transfer reactions in saline media such as sea water, Geochim. Cosmochim. Ac., 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. I, 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, 191 pp., North Pacific Marine Science Organization, Sidney, BC, Canada, 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.

ESSDD
doi:10.5194/essd-2015-42

Falck, E. and Olsen, A.: Nordic Seas dissolved oxygen data in CARINA, Earth Syst. Sci. Data, 2, 123–131, doi:10.5194/essd-2-123-2010, 2010.

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.

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, Prog. Oceanogr., 48, 337–402, 2001.

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), edited by: Hall, J., Harrison, D. E, and Stammer, D., Venice, Italy, 21–25 September 2009, ESA Publication WPP-306, doi:10.5270/OceanObs09, 2010.

Hoppema, M., Velo, A., van Heuven, S., Tanhua, T., Key, R. M., Lin, X., Bakker, D. C. E., Perez, F. F., Ríos, A. F., Lo Monaco, C., Sabine, C. L., Álvarez, M., and Bellerby, R. G. J.: Consistency of cruise data of the CARINA database in the Atlantic sector of the Southern Ocean, Earth Syst. Sci. Data, 1, 63–75, doi:10.5194/essd-1-63-2009, 2009.

IOCCP: Initial Atlantic Ocean Carbon Synthesis Meeting, Laugarvatn, Iceland 28–30 June, IOCCP Report No 5, Paris, 2006.

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.

Jeansson, E., Olsen, A., Eldevik, T., Skjelvan, I., Omar, A. M., Lauvset, S. K., Nilsen, J. E. Ø., Bellerby, R. G. J., Joahannessen, T., and Falck, E.: The Nordic seas carbon budget: sources, sinks, and uncertainties, Global Biogeochem. Cy., 25, GB4010, doi:10.1029/2010GB003961, 2011.

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.
Johnson, K. M., Wills, K. D., Butler, D. B., Johnson, W. K., and Wong, C. S.: Coulometric total carbon dioxide analysis for marine studies, maximizing the performance of an automated gas extraction system and coulometric detector, Mar. Chem., 44, 167–187, 1993.
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 vs. CFCs, Prog. 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. Cy., 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.
Khatiwala, S., Primeau, F., and Hall, T.: Reconstruction of the history of anthropogenic CO_2 concentrations in the ocean, Nature, 462, 346–349, 2009.
King, B. A., Firing, E., and Joyce, T. M.: Shipboard observations during WOCE, in: Ocean Circulation and Climate, edited by: Siedler, G., Church, J., and Gould, J., Int. Geophys. Ser., Vol. 77, 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.-Meth., 13, 601–608, doi:10.1002/lom3.10050, 2015.
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.
Lo Monaco, C., Álvarez, M., Key, R. M., Lin, X., Tanhua, T., Tilbrook, B., Bakker, D. C. E., van Heuven, S., Hoppema, M., Metzl, N., Ríos, A. F., Sabine, C. L., and Velo, A.: Assessing

the internal consistency of the CARINA database in the Indian sector of the Southern Ocean, Earth Syst. Sci. Data, 2, 51–70, doi:10.5194/essd-2-51-2010, 2010.

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., and Johnson, D. R.: World Ocean Atlas 2009, Volume 1, Temperature, in: NOAA Atlas NESDIS 68, edited by: S. Levitus, US Government Printing Office, Washington D. C., 2010.

Lueker, T. J., Dickson, A. G., and Keeling, C. D.: Ocean pCO_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.

Manizza, M., Follows, M. J., Dutkiewciz, S., Menemenlis, D., McClelland, J. W., Hill, C. N., Peterson, B. J., and Key, R. M.: A model of the Arctic Ocean carbon cycle, J. Geophys. Res., 116, C12020, doi:10.1029/2011JC006998, 2011.

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), 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. and Olsen, A.: Nordic Seas nutrients data in CARINA, Earth Syst. Sci. Data, 2, 205–213, doi:10.5194/essd-2-205-2010, 2010.

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.

Olafsson, J., Olafsdottir, S. R., Benoit-Cattin, A., and Takahashi, T.: The Irminger Sea and the Iceland Sea time series measurements of sea water carbon and nutrient chemistry 1983–2008, Earth Syst. Sci. Data, 2, 99–104, doi:10.5194/essd-2-99-2010, 2010.

ESSDD
doi:10.5194/essd-2015-42

Olsen, A.: Nordic Seas total dissolved inorganic carbon data in CARINA, Earth Syst. Sci. Data, 1, 35–43, doi:10.5194/essd-1-35-2009, 2009a.
Olsen, A.: Nordic Seas total alkalinity data in CARINA, Earth Syst. Sci. Data, 1, 77–86, doi:10.5194/essd-1-77-2009, 2009b.
Olsen, A., Key, R. M., Jeansson, E., Falck, E., Olafsson, J., van Heuven, S., Skjelvan, I., Omar, A. M., Olsson, K. A., Anderson, L. G., Jutterström, S., Rey, F., Johannessen, T., Bellerby, R. G. J., Blindheim, J., Bullister, J. L., Pfeil, B., Lin, X., Kozyr, A., Schirnick, C., Tanhua, T., and Wallace, D. W. R.: Overview of the Nordic Seas CARINA data and salinity measurements, Earth Syst. Sci. Data, 1, 25–34, doi:10.5194/essd-1-25-2009, 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.
Ota, H., Mitsuda, H., Kimura, M., and Kitao, T.: Reference materials for nutrients in seawater: Their development and present homogeneity and stability, in: Comparability of Nutrients in the World's Ocean, edited by: Aoyama, M., Dickson, A. G., Hydes, D. J., Murata, A., Oh, J. R., Roose, P., and Woodward, E. M. S., Mother Tank, Tsukuba, Japan, 2010.
Pfeil, B., Olsen, A., Bakker, D. C. E., Hankin, S., Koyuk, H., Kozyr, A., Malczyk, J., Manke, A., Metzl, N., Sabine, C. L., Akl, J., Alin, S. R., Bates, N., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W.-J., Chavez, F. P., Chen, A., Cosca, C., Fassbender, A. J., Feely, R. A., González-Dávila, M., Goyet, C., Hales, B., Hardman-Mountford, N., Heinze, C., Hood, M., Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., Jones, S. D., Key, R. M., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lourantou, A., Merlivat, L., Midorikawa, T., Mintrop, L., Miyazaki, C., Murata, A., Nakadate, A., Nakano, Y., Nakaoka, S., Nojiri, Y., Omar, A. M., Padin, X. A., Park, G.-H., Paterson, K., Perez, F. F., Pierrot, D., Poisson, A., Ríos, A. F., Santana-Casiano, J. M., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Schneider, B., Schuster, U., Sieger, R., Skjelvan, I., Steinhoff, T., Suzuki, T., Takahashi, T., Tedesco, K., Telszewski, M., Thomas, H., Tilbrook, B., Tjiputra, J., Vandemark, D., Veness, T., Wanninkhof, R., Watson, A. J., Weiss, R., Wong, C. S., and Yoshikawa-Inoue, H.: A uniform, quality controlled Surface Ocean CO_2 Atlas (SOCAT), Earth Syst. Sci. Data, 5, 125–143, doi:10.5194/essd-5-125-2013, 2013.
Pierrot, D., Brown, P., Van Heuven, S., Tanhua, T., Schuster, U., Wanninkhof, R., and Key, R. M.: CARINA TCO_2 data in the Atlantic Ocean, Earth Syst. Sci. Data, 2, 177–187, doi:10.5194/essd-2-177-2010, 2010.

ESSDD
doi:10.5194/essd-2015-42

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, US Departement of Energy, Oak Ridge, Tennessee, 2005.

Sabine, C. L., Key, R. M., Johnson, K. M., Millero, F. J., Sarmiento, J. L., Wallace, D. W. R., and Winn, C. D.: Anthropogenic CO_2 inventory of the Indian Ocean, Global Biogeochem. Cy., 13, 179–198, 1999.

Sabine, C. L., Feely, R. A., Key, R. M., Bullister, J. L., Millero, F. J., Lee, K., Peng, T.-H., Tilbrook, B., Ono, T., and Wong, C. S.: Distribution of anthropogenic CO_2 in the Pacific Ocean, Global Biogeochem. Cy., 16, 1083, doi:10.1029/2001GB001639, 2002.

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. L., Hoppema, M., Key, R. M., Tilbrook, B., van Heuven, S., Lo Monaco, C., Metzl, N., Ishii, M., Murata, A., and Musielewicz, S.: Assessing the internal consistency of the CARINA data base in the Pacific sector of the Southern Ocean, Earth Syst. Sci. Data, 2, 195–204, doi:10.5194/essd-2-195-2010, 2010.

Schmittner, A., Gruber, N., Mix, A. C., Key, R. M., Tagliabue, A., and Westberry, T. K.: Biology and air–sea gas exchange controls on the distribution of carbon isotope ratios ($\delta^{13}C$) in the ocean, Biogeosciences, 10, 5793–5816, doi:10.5194/bg-10-5793-2013, 2013.

Sparrow, M., Chapman, P., and Gould, J.: The World Ocean Circulation Experiment (WOCE) Hydrographic Atlas Series, International WOCE Project Office, Southampton, UK, 2005–2011.

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-

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.

Title Page
Abstract | Instruments
Data Provenance & Structure
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

159, NDP-092, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, 2013.

Swift, J.: Reference quality water sample data: notes on data acquisition, in: The GO-SHIP Repeat Hydrography Manual: A Collection of Expert Reports and Guidelines, edited by: Hood, E. M., Sabine, C. L., and Sloyan, B. M., IOCCP report number 14, ICPO publication series number 134, available at: http://www.go-ship.org/HydroMan.html (last access: 22 December 2015), 2010.

Swift, J. and Diggs, S. C.: Description of WHP exchange format for CTD/Hydrographic data, CLIVAR and Carbon Hydrographic 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, 19.1–19.31, 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, US Department of Energy, Oak Ridge, Tennessee, 2009a.

Tanhua, T., Brown, P. J., and Key, R. M.: CARINA: nutrient data in the Atlantic Ocean, Earth Syst. Sci. Data, 1, 7–24, doi:10.5194/essd-1-7-2009, 2009d.

ESSDD

doi:10.5194/essd-2015-42

Title Page
Abstract | Instruments
Data Provenance & Structure
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc

Printer-friendly Version
Interactive Discussion

Tanhua, T., Olsen, A., Hoppema, M., and Gouretski, V.: CARINA: a consistent carbon-relevant data base for the Arctic, Atlantic and Southern Oceans, Earth Syst. Sci. Data, http://www.earth-syst-sci-data.net/special_issue2.html, 2009–2010.

Tanhua, T., Steinfeldt, R., Key, R. M., Brown, P., Gruber, N., Wanninkhof, R., Perez, F., Körtzinger, A., Velo, A., Schuster, U., van Heuven, S., Bullister, J. L., Stendardo, I., Hoppema, M., Olsen, A., Kozyr, A., Pierrot, D., Schirnick, C., and Wallace, D. W. R.: Atlantic Ocean CARINA data: overview and salinity adjustments, Earth Syst. Sci. Data, 2, 17–34, doi:10.5194/essd-2-17-2010, 2010a.

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, 2010b.

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., Perez, F. F., Brown, P., Tanhua, T., Schuster, U., and Key, R. M.: CARINA alkalinity data in the Atlantic Ocean, Earth Syst. Sci. Data, 1, 45–61, doi:10.5194/essd-1-45-2009, 2009.

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.

Velo, A., Perez, F. F., Tanhua, T., Gilcoto, M., Rios, A. F., and Key, R. M.: Total alkalinity estimation using MLR and neural network techniques, J. Marine Syst., 111–112, 11–18, doi:10.1016/j.jmarsys.2012.09.002, 2013.

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 B, 62, 608–620, 2010.

Wanninkhof, R., Peng, T.-H., Huss, B., Sabine, C. L., and Lee, K.: Comparison of inorganic carbon parameters measured in the Atlantic ocean from 1990–1998 and recommended adjustments, ORNL/CDIAC-140, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, 2003.

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.

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.

Title Page
Abstract Instruments
Data Provenance & Structure
Tables Figures
◂ ▸
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

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, Prog. Oceanogr., 73, 242–276, 2007.

ESSDD

doi:10.5194/essd-2015-42

Title Page

Abstract | Instruments

Data Provenance & Structure

Tables | Figures

◂ | ▸

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. Parameters subjected to secondary QC by the various interior ocean carbon synthesis efforts.

	Sal.	Oxy.	NO_3	Si	PO_4	TCO_2	TAlk	pH	CFC-11	CFC-12	CFC-113	CCl_4
GLODAPv1	–[a]	–[a]	–[a]	–[a]	–[a]	X	X		X	X	X	X
CARINA	X	X	X	X	X	X	X	X	X[b]	X[b]	X[b]	X[b]
PACIFICA	X	X	X	X	X	X	X	–	–	–	–	–
GLODAPv2	X	X	X	X	X	X	X	X	X	X	X	X

[a] The GLODAPv1.1 team used corrections from the sources listed in Table 2, when available in these.
[b] Only CFC data from the Arctic mediterranean seas and Atlantic regions were subjected to secondary QC in CARINA, not the Southern Ocean Data.

ESSDD
doi:10.5194/essd-2015-42

Table 2. Source and nature of earlier GLODAPv1.1 salinity, oxygen, and nutrient adjustments. GJ is Gouretski and Jahnke (2001), J2001 is Johnson et al. (2001) and M refers to unpublished work of Mordy, C., L. Gordon, G. Johnson and A. Ross.

Region/Era	Salinity	Oxygen	Nutrients
Atlantic	GJ2001 Additive	GJ2001 Additive	GJ2001 Additive
Pacific WOCE	J2001 Additive	J2001 Additive	M Multiplicative
Pacific historic	GJ2001 Additive	GJ2001 Additive	GJ2001 Additive
Indian main US WOCE	J2001 Additive	GJ2001 Additive	M Multiplicative
Indian other WOCE and historical	GJ2001 Additive	GJ2001 Additive	GJ2001 Additive

ESSDD
doi:10.5194/essd-2015-42

Table 3. WOCE flags (briefly; for full details see http://geo.h2o.ucsd.edu/documentation/policies/Data_Evaluation_reference.pdf) in GLODAPv2. Data flagged 3 or 4 were not included in the product files.

WOCE Flag Value	Interpretation in GLODAPv2
0	Approximated
1	Not used
2	Good
3	Questionable
4	Clearly bad
5	Value not reported
6	Average of replicate
7	Not used
8	Not used
9	Not measured

ESSDD
doi:10.5194/essd-2015-42

Table 4. Contents of the CARINA Special Issue in Earth System Science Data (Tanhua et al., 2009–2010).

Citation	Paper Name
Falck and Olsen (2010)	Nordic Seas dissolved oxygen data in CARINA
Hoppema et al. (2009)	Consistency of cruise data of the CARINA database in the Atlantic sector of the Southern Ocean
Jeansson et al. (2010)	Nordic Seas and Arctic Ocean CFC data in CARINA
Jutterström et al. (2010)	Arctic Ocean data in CARINA
Key et al. (2010)	The CARINA data synthesis project: introduction and overview
Lo Monaco et al. (2010)	Assessing the internal consistency of the CARINA database in the Indian sector of the Southern Ocean
Olafsson et al. (2010)	The Irminger Sea and the Iceland Sea time series measurements of sea water carbon and nutrient chemistry 1983–2008
Olafsson and Olsen (2010)	Nordic Seas nutrients data in CARINA
Olsen et al. (2009)	Overview of the Nordic Seas CARINA data and salinity measurements
Olsen (2009)	Nordic Seas total dissolved inorganic carbon data in CARINA
Olsen (2009b)	Nordic Seas total alkalinity data in CARINA
Pierrot et al. (2010)	CARINA TCO_2 data in the Atlantic Ocean
Sabine et al. (2010)	Assessing the internal consistency of the CARINA data base in the Pacific sector of the Southern Ocean
Stendardo et al. (2009)	CARINA oxygen data in the Atlantic Ocean
Steinfeldt et al. (2010)	Atlantic CFC data in CARINA
Tanhua et al. (2009)	CARINA: nutrient data in the Atlantic Ocean
Tanhua et al. (2010b)	Quality control procedures and methods of the CARINA database
Tanhua et al. (2010a)	Atlantic Ocean CARINA data: overview and salinity adjustments
Velo et al. (2009)	CARINA alkalinity data in the Atlantic Ocean
Velo et al. (2010)	CARINA data synthesis project: pH data scale unification and cruise adjustments

ESSDD
doi:10.5194/essd-2015-42

Table 5. The GLODAPv2 team in alphabetical order and individual responsibilities.

Who	Institution	Main Tasks
Mario Hoppema	Alfred Wegener Institute, Bremerhaven, Germany	GLODAPv1.2 and Southern Ocean expert
Masao Ishii	Meteorological Research Institute, Japan	PACIFICA and Pacific Ocean expert
Emil Jeansson	Uni Research, Bjerknes Centre for Climate Research, Norway	CFCs and Nordic Seas expert
Sara Jutterström	IVL Swedish Environmental Research Institute, Sweden	Secondary QC of new cruises from the Arctic
Robert M. Key	Princeton University, USA	Data acquisition, primary QC and data product assembly
Alexander Kozyr	Carbon Dioxide Information Analysis Center, USA	Data manager (US), Cruise Summary Table
Siv K. Lauvset	University of Bergen, Bjerknes Centre for Climate Research, Bergen, Norway	Secondary QC of new cruises from outside Arctic and mapped product
Xiaohua Lin	Princeton University, USA	Primary QC and data product assembly.
Are Olsen	University of Bergen, Bjerknes Centre for Climate Research, Norway	Coordinator, GLODAPv1.2 and Nordic Seas expert
Fiz F. Pérez	Instituto de Investigaciones Marinas–CSIC, Spain	CO_2 system and Atlantic Ocean expert
Carsten Schirnick	GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany	Cruise information and Adjustment Table
Reiner Steinfeldt	Universität Bremen, Germany	CFCs and Atlantic Ocean expert
Toru Suzuki	Marine Information Research Center, Japan	PACIFICA and Pacific Ocean expert
Toste Tanhua	GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany	GLODAPv1.2, CFCs and Atlantic Ocean expert
Steven van Heuven	University of Groningen, and Royal Netherlands Institute for Sea Research, the Netherlands	Crossover and inversion analysis of merged and reset product.
Anton Velo	Instituto de Investigaciones Marinas–CSIC, Spain	pH QC and seawater CO_2 chemistry calculations

ESSDD
doi:10.5194/essd-2015-42

Table 6. Initial minimum adjustment limits.

Parameter	Minimum Adjustment
Salinity	0.005
Oxygen	1 %
Nutrients	2 %
TCO_2	4 µmol kg^{-1}
TAlk	6 µmol kg^{-1}
pH	0.005
CFCs	5 %

ESSDD
doi:10.5194/essd-2015-42

Table 7. Multi-leg cruises in GLODAPv2 that received secondary quality control on a per-leg basis but are included as a single entry 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

ESSDD
doi:10.5194/essd-2015-42

Table 8. Summary of secondary QC actions per parameter for the 780 non-dismissed entries.

	Sal.	Oxy.	NO_3	Si	PO_4	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.

ESSDD
doi:10.5194/essd-2015-42

Table 9. Summary of salinity and oxygen calibration needs and actions for the 780 non-dismissed entries subjected to secondary QC.

Case	Description	Salinity		Oxygen	
		Entries	adj./noQC/bad*	Entries	adj./noQCd/bad*
1	No data are available, no action needed	0	–	58	–
2	No bottle values present, use CTD derived values	77	9/33/1	21	7/11/0
3	No CTD values present, use bottle data	295	17/71/4	520	164/95/9
4	(Case not used)	–	–	–	–
5	The CTD values do not deviate significantly from bottle values: replace missing bottle values with CTD values	264	9/14/0	99	22/4/1
6	The CTD values deviate significantly from bottle values: calibrate these using linear fit and replace missing bottle values with calibrated CTD values	141	6/43/1	62	10/11/0
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	0/1/0	20	4/6/0

* adj: adjusted in final data product; noQC: full secondary QC not possible; bad: data were of poor quality and not included in final product. Remainder data were included as is in final product files.

ESSDD
doi:10.5194/essd-2015-42

Table 10. Summary of the distribution of applied adjustments per parameter, in number of adjustments applied for each parameter.

	adj. < limit	limit $\leq$ adj. < $2\cdot$ limit	$2\cdot$ limit $\leq$ adj.
Salinity	5	22	14
Oxygen	7	101	99
NO_3	50	73	27
Si	113	95	56
PO_4	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

ESSDD

doi:10.5194/essd-2015-42

Table 11. 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 absolute offset of the crossover offsets of unadjusted and adjusted data, respectively. *n* is the total number of valid crossovers in the global ocean for the parameter in question.

	ARCTIC			ATLANTIC			INDIAN			PACIFIC			GLOBAL			
	unadj		adj	unadj		adj	unadj		adj	unadj		adj	unadj		adj	*n* (global)
Sal [ppm]	4.1	→	3.8	7.1	→	5.0	2.7	→	1.6	2.4	→	1.9	4.1	→	3.1	~ 12100
Oxy [%]	1.3	→	1.0	1.7	→	0.8	1.4	→	0.7	1.7	→	1.1	1.7	→	0.9	~ 10900
NO_3 [%]	4.2	→	1.6	2.7	→	1.7	1.8	→	1.0	1.0	→	0.8	1.7	→	1.2	~ 9500
Si [%]	8.2	→	3.5	4.8	→	2.7	2.8	→	1.5	1.9	→	0.9	2.8	→	1.7	~ 8300
PO_4 [%]	4.8	→	2.5	4.2	→	2.5	2.7	→	1.1	1.5	→	1.0	2.2	→	1.3	~ 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	~ 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	~ 3400

ESSDD
doi:10.5194/essd-2015-42

Table 12. Cruises where measured carbon parameters have been fully replaced with calculated ones in the GLODAPv2 product files. One of these cruises is from the Alantic (28HE20100404), while the remainder are Pacific cruises.

Parameter	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
	49HO2000601	25	125
	49HO20000621	25	123
	49NZ20041117	277	1049
	49UF20080117	58	117

ESSDD
doi:10.5194/essd-2015-42

Table 13. Parameters 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.

Parameter	Units	Variable Name	Flag Name	Secondary QC Flag Name
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 botttle number		bottle		
Sampling pressure	dbar	pressure		
Sampling depth	m	depth		
Temperature	°C	temperature		
Potential temperature	°C	theta		
Salinity		salinity	G2salinityf	G2salinityqc
Potential density	$kg\,m^{-3}$	sigma0		
Potential density, ref 1000 dbar	$kg\,m^{-3}$	sigma1		
Potential density, ref 2000 dbar	$kg\,m^{-3}$	sigma2		
Potential density, ref 3000 dbar	$kg\,m^{-3}$	sigma3		
Potential density, ref 4000 dbar	$kg\,m^{-3}$	sigma4		
Neutral density	$kg\,m^{-3}$	gamma		
Oxygen	$\mu mol\,kg^{-1}$	oxygen	G2oxygenf	G2oxygenqc
Apparent oxygen utilization	$\mu mol\,kg^{-1}$	aou	G2aouf	
Nitrate	$\mu mol\,kg^{-1}$	nitrate	G2nitratef	G2nitrateqc
Nitrite	$\mu mol\,kg^{-1}$	nitrite	G2nitritef	
Silicate	$\mu mol\,kg^{-1}$	silicate	G2silicatef	G2silicateqc
Phosphate	$\mu mol\,kg^{-1}$	phosphate	G2phosphatef	G2phosphateqc
TCO_2	$\mu mol\,kg^{-1}$	tco2	G2tco2f	G2tco2qc

ESSDD
doi:10.5194/essd-2015-42

Table 13. Continued.

Parameter	Units	Variable Name	Flag Name	Secondary QC Flag Name
TAlk	µmol kg^{-1}	talk	G2talkf	G2talkqc
pH at total scale, 25 °C and zero dbar of pressure		phts25p0	G2phts25p0f	G2phtsqc
pH at total scale, in situ temperature and pressure		phtsinsitutp	G2phtsinsitutpf	G2phtsqc
CFC-11	pmol kg^{-1}	cfc11	G2cfc11f	G2cfc11qc
pCFC-11	ppt	pcfc11		
CFC-12	pmol kg^{-1}	cfc12	G2cfc12f	G2cfc12qc
pCFC-12	ppt	pcfc12		
CFC-113	pmol kg^{-1}	cfc113	G2cfc113f	G2cfc113qc
pCFC-113	ppt	pcfc113		
CCl_4	pmol kg^{-1}	ccl4	G2ccl4f	G2ccl4qc
$pCCl_4$	ppt	pccl4		
SF_6	fmol kg^{-1}	sf6	G2sf6f	
pSF6	ppt	psf6		
$\delta^{13}C$	‰	c13	G2c13f	
$\Delta^{14}C$	‰	c14	G2c14f	
$\Delta^{14}C$ counting error		c14err		
^3H	TU	h3	G2h3f	
^3H counting error	TU	h3err		
δ^3He	%	he3	G2he3f	
^3He counting error	%	he3err		
He	nmol kg^{-1}	he	G2hef	
He counting error	nmol kg^{-1}	heerr		
Ne	nmol kg^{-1}	neon	G2neonf	
Ne counting error	nmol kg^{-1}	neonerr		
$\delta^{18}O$	‰	o18	G2o18f	
Total organic carbon	µmol L^{-1}	toc	G2tocf	
Dissolved organic carbon	µmol kg^{-1}	doc	G2docf	
Dissolved organic nitrogen	µmol kg^{-1}	don	G2donf	
Total dissolved nitrogen	µmol kg^{-1}	tdn	G2tdnf	
Chlorophyll *a*	µg L^{-1}	chla	G2chlaf	

ESSDD
doi:10.5194/essd-2015-42

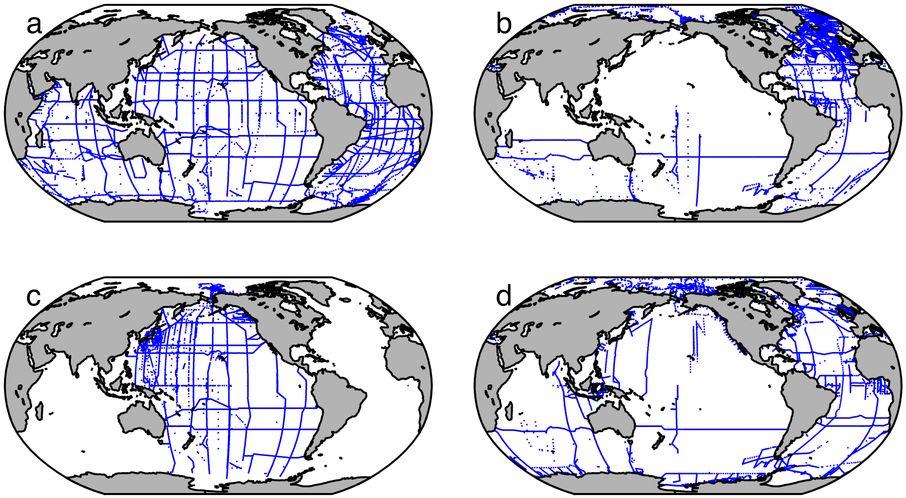


Figure 1. Station locations in **(a)** GLODAPv1.1, **(b)** CARINA, **(c)** PACIFICA, and **(d)** locations of stations in GLODAPv2 new to data synthesis.

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.

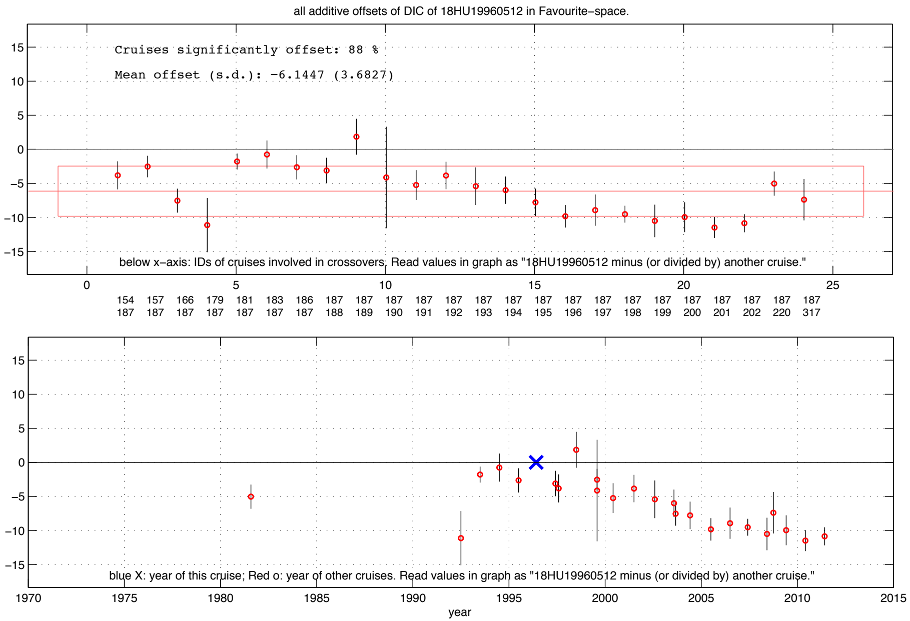


Figure 2. Summary figure used to evaluate TCO_2 crossover offsets of WOCE repeat section AR07W cruise 18HU19960512 in the Labrador Sea. The upper panel shows the 24 crossover offsets that were determined and their mean and standard deviation (note that these cruise numbers do not correspond to the final GLODAPv2 cruise numbers in the Cruise Summary Table and in the product files, in this figure. 18HU19960512 is cruise number 187, while its "official" GLODAPv2 cruise number is 159). The lower panel shows these crossover offsets sorted by time. Favorite-space is the mean of the offsets in sigma-4, pressure and potential density space (cf. Sect 4.2) In both panels, negative values means that 18HU19960512 TCO_2 values are lower than those of the comparison cruise. The lower panel implies a temporal trend, not a measurement bias.

ESSDD
doi:10.5194/essd-2015-42
An internally consistent data product for the world ocean
A. Olsen et al.

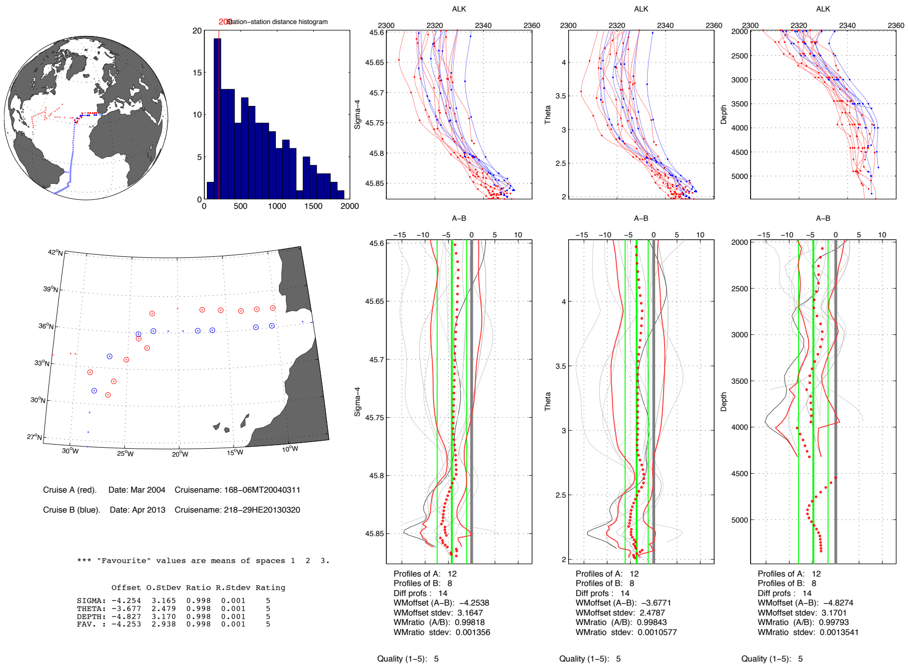


Figure 3. Summary TAlk figure for crossover stations between cruise 06MT20040311 and 29HE20130320, as described in the main text. Note that the cruise numbers given in this figure, 168 and 218, are *not* the GLODAPv2 cruise numbers that are used for our data product. The relative difference here is approximately 4 μmol kg^{-1} with the red (Meteor; 06MT) cruise seeming to have lower abyssal values than the blue (Hesperides; 29HE).

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.

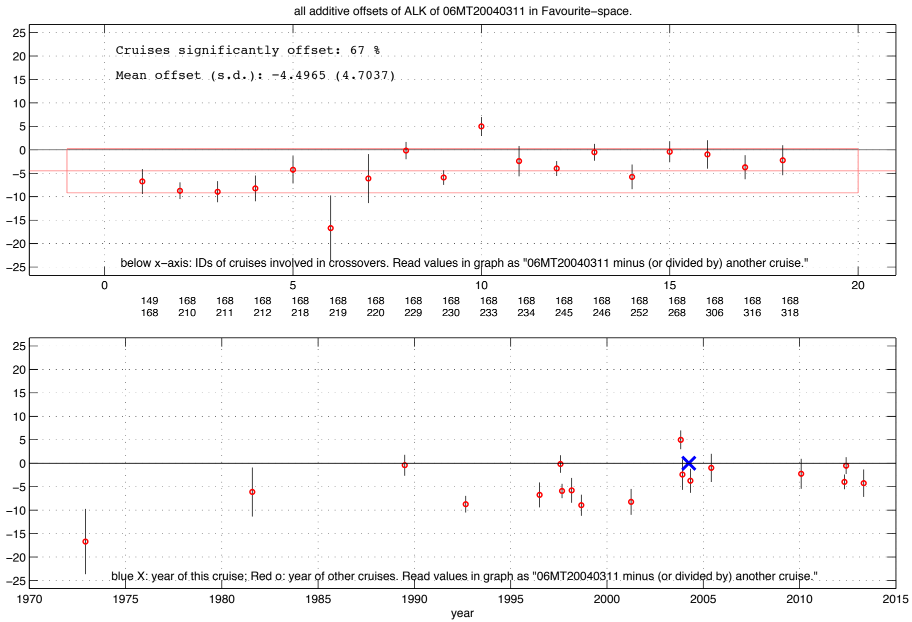


Figure 4. Summary figure for crossover offsets of cruise's 06MT20040311 TAlk data, as in Fig. 2.

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.

Title Page
Abstract | Instruments
Data Provenance & Structure
Tables | Figures
◀◀ | ▶▶
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

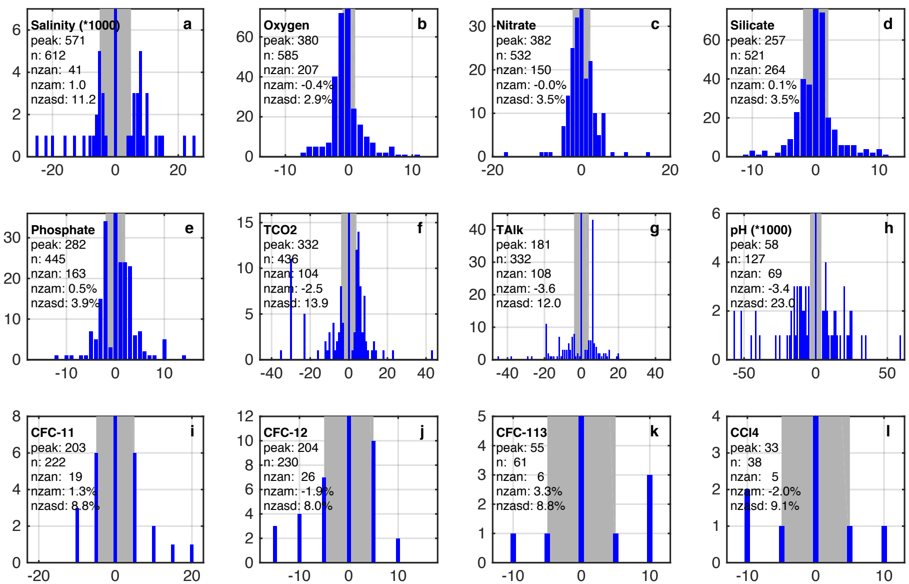
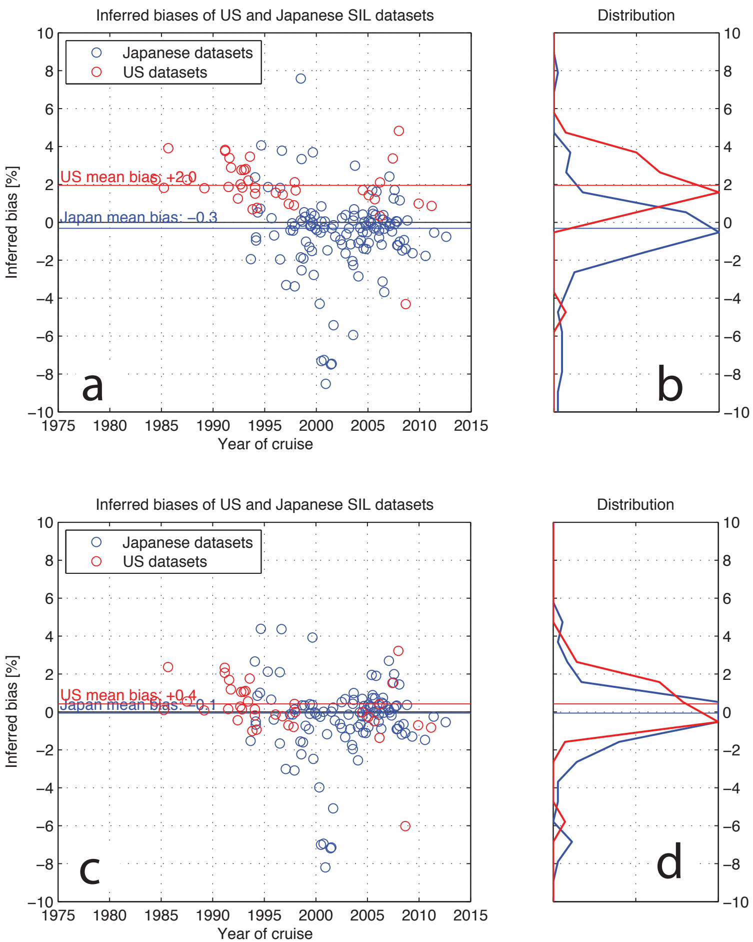


Figure 5. Size distribution of applied adjustments for each core parameter that received secondary QC. In each case the horizontal axis is adjustment magnitude and the vertical axis is cruise count. Gray areas depict the nominal adjustment threshold area. *peak* indicates the number of entries that got no adjustment (or zero adjustment; in all cases this is the peak of the histogram). *n* excludes those entries for which data (i) could not be secondary QC'ed or (ii) was considered of insufficient quality for our product, and thus does not necessarily add up to 780. Measures *nzan*, *nzam* and *nzasd*, are the number, mean and standard deviation of all non-zero adjustments, respectively.

ESSDD
doi:10.5194/essd-2015-42

A. Olsen et al.

Title Page
Abstract
Instruments
Data Provenance & Structure
Tables
Figures
◄
►
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion



ESSDD
doi:10.5194/essd-2015-42

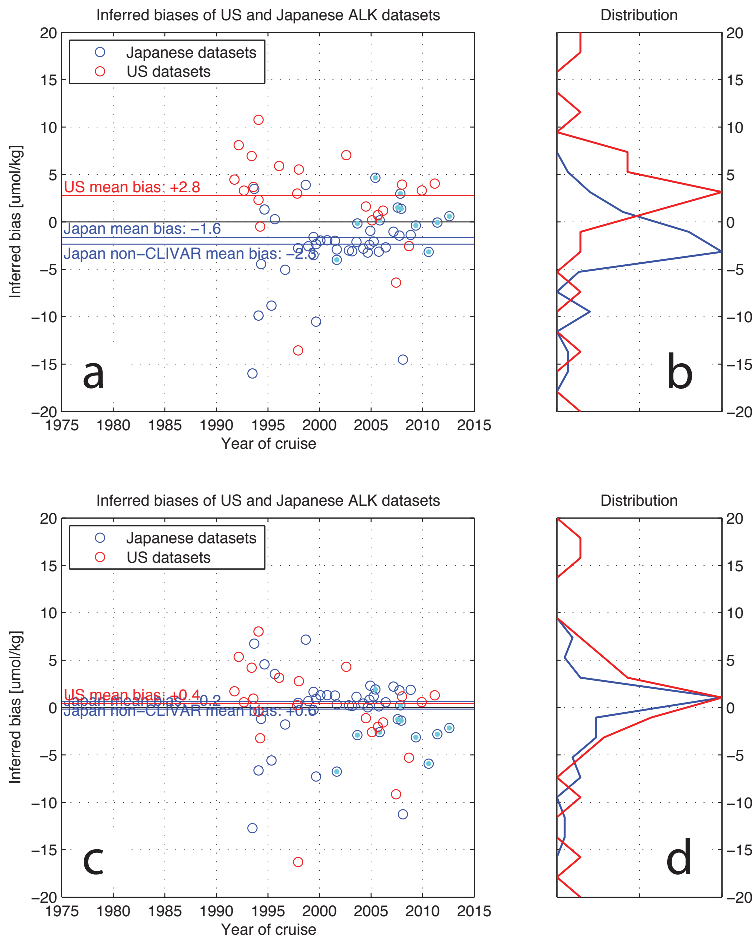
Title Page
Abstract | Instruments
Data Provenance & Structure
Tables | Figures
◀◀ | ▶▶
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Figure 6. Silicate biases between US and Japanese efforts before (**a** and **b**) and after (**c** and **d**) pre-adjustments (US: −1 %; Japan: +1 %) were applied to the data. Circles represent the individual cruise biases (inferred by the GLODAPv2 crossover and inversion method) of Pacific silicate measurements. Data from "Line-P" and many small-scale cruises in the variable Kuroshio region were excluded from this analysis. Red circles: US cruises. Blue circles: Japanese cruises. Red and blue horizontal lines indicate countries' approximate mean offsets. Note the distinct separation between the average correction recommended for the US and Japanese cruises in **(a)** and **(b)**. This indicates a persistent analytical inconsistency. Note that in **(c)** and **(d)** the average recommended corrections are now both approximately 0. Remaining offsets greater than ∼ 2 % (positive or negative) were individually adjusted as part of regular secondary QC, subsequent to this country-specific pre-adjustment. A reasonable assumption would be that one country's data should have received a 2 % correction, but the data are insufficient to determine is this is correct.

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.



ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.

Figure 7. TAlk biases between US and Japanese efforts before (**a** and **b**) and after (**c** and **d**) pre-adjustments were applied to TAlk data of Japanese non-CLIVAR cruises. 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 circles: US cruises. Blue circles: Japanese cruises (CLIVAR cruises are cyan-filled). Cruises of third countries are not shown. Red and blue horizontal lines indicate countries' approximate mean offset. For Japan data the values are split into cruises that were or were not part of CLIVAR. The distinct separation between the average correction recommended for the US and Japanese cruises in **(a)** and **(b)**, indicates a persistent analytical inconsistency. The large amount of Japanese cruises, and the zero-sum constraint of the inversion method force the US cruises towards positive numbers, even though we later decide the US cruise average represents our main reference. Note in **(c)** and **(d)** that the average recommended corrections are now much closer to 0. Note also how adjustments of one half of the total dataset (i.e., Japanese non-CLIVAR) bring both halves of the dataset (i.e., US and Japanese) closer to neutral. Remaining offsets greater than $\sim 4\,\mu mol\,kg^{-1}$ (positive or negative) were individually adjusted as part of regular secondary QC, subsequent to this country-specific pre-adjustment.

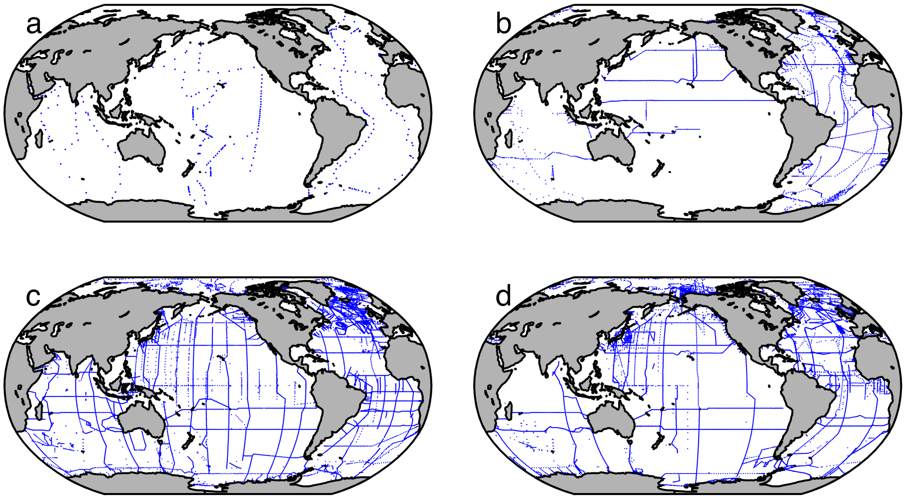


Figure 8. Station locations in the GLODAPv2 data product for data obtained during **(a)** the 1970s, **(b)** the 1980s, **(c)** the 1990s, and **(d)** 2000s and beyond.

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.

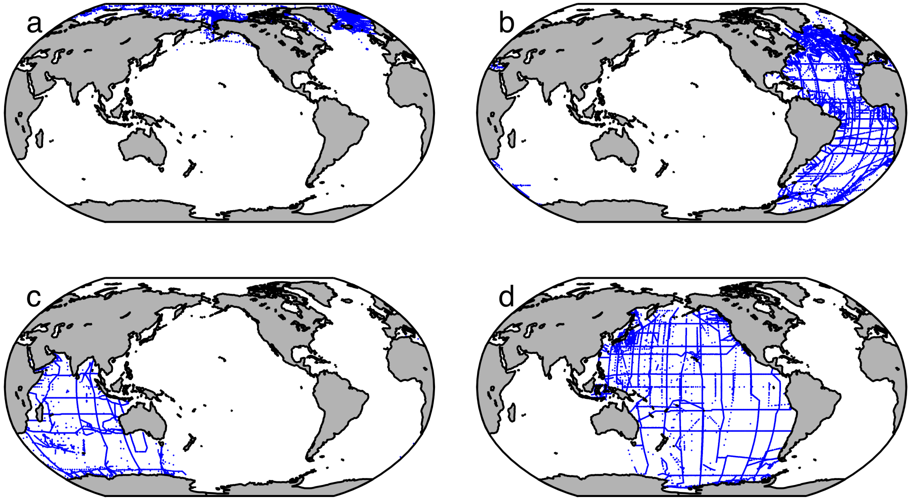


Figure 9. Locations of data included in the **(a)** Arctic, **(b)** Atlantic, **(c)** Indian, and **(d)** Pacific Ocean product files. Note the minor "spillover" near the boundaries.

ESSDD
doi:10.5194/essd-2015-42

An internally consistent data product for the world ocean

A. Olsen et al.