ORNL/CDIAC-162
NDP-093

GLOBAL OCEAN DATA ANALYSIS PROJECT, VERSION 2

(GLODAPv2)

By the GLODAPv2 group:

Robert M. Key[1], Are Olsen[2,3,], Steven van Heuven[4], Siv K. Lauvset[2,5], Anton Velo[6], Xiaohua Lin[1], Carsten Schirnick[7], Alex Kozyr[8], Toste Tanhua[7], Mario Hoppema[9], Sara Jutterström[10], Reiner Steinfeldt[11], Emil Jeansson[5], Masao Ishii[12], Fiz F. Perez[6], and Toru Suzuki[13].

[1]Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA
[2]Geophysical Institute, University of Bergen, Norway.
[3]Bjerknes Centre for Climate Research, Bergen, Norway
[4] Royal Netherlands Institute for Sea Research, Texel, The Netherlands
[5]Uni Research Climate, Bjerknes Centre for Climate Research, Bergen, Norway.
[6]Instituto de Investigaciones Marinas de Vigo, CSIC, Vigo, Spain
[7] GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany
[8]Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, U.S.A.
[9]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany
[10]IVL Swedish Environmental Research Institute, Göteborg, Sweden
[11]University of Bremen, Bremen, Germany
[12]Geochemical Research Department, Meteorological Research Institute, Japan Meteorological Agency, Japan
[13]Marine Information Research Center, Japan Hydrographic Association, Japan

Date Published: December 14, 2015

Prepared for the Climate Change Research Division
Office of Biological and Environmental Research
U.S. Department of Energy
Budget Activity Numbers KP 12 04 01 0 and KP 12 02 03 0

Prepared by the
Carbon Dioxide Information Analysis Center
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6335
managed by
UT-BATTELLE, LLC
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-00OR2272

CONTENTS

ABBREVIATIONS AND ACRONYMS

CARINA	Carbon in the Atlantic Ocean
CLIVAR	Climate and Ocean: Variability, Predictability and Change
CDIAC	Carbon Dioxide Information Analysis Center
CST	Cruise Summary Table
DOE	Department of Energy
EU	European Union
GLODAP	Global Ocean Data Analysis Project
ICES	International Council for the Exploration of the Sea
ORNL	Oak Ridge National Laboratory
NCEI	National Centers for Environmental Information
NODC	National Oceanographic Data Center (now NCEI)
PACIFICA	PACIFic ocean Interior CArbon
PICES	North Pacific Marine Science Organization
QC	Quality Control
WOCE	World Ocean Circulation Experiment

ABSTRACT

Key, R.M., A. Olsen, S. van Heuven, S. K. Lauvset, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, S. Jutterström, R. Steinfeldt, E. Jeansson, M. Ishii, F. F. Perez, T. Suzuki. 2015. Global Ocean Data Analysis Project, Version 2 (GLODAPv2). ORNL/CDIAC-162, NDP-093. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Dept. of Energy, Oak Ridge, Tennessee. doi: 10.3334/CDIAC/OTG.NDP093_GLODAPv2

Here we report the completion of a uniformly calibrated open ocean data product on inorganic carbon and carbon-relevant variables that we call Global Ocean Data Analysis Project version 2 or simply GLODAPv2. The product includes data from approximately one million individual seawater samples collected from over 700 cruises during the years 1972-2013. Extensive quality control and subsequent calibration were carried out for salinity, oxygen, nutrient, carbon dioxide, total alkalinity, pH and chlorofluorocarbon data. Following calibration, the data were used to produce global climatology maps for many of the parameters. In addition to the data products and the mapped distributions, all of the original data files without alteration other than formatting and unification of units are made freely available along with whatever metadata was collected. An on-line cruise summary table provides data access and additional information including references to publications that have used data from specific cruises.

The GLODAPv2 database is freely available as a numeric data package (NDP-93) at the Carbon Dioxide Information Analysis Center (CDIAC) http://cdiac.ornl.gov/oceans/GLODAPv2. The NDP consists of the original cruise data files, adjusted data product, and this documentation, which describes the GLODAPv2 project.

Keywords: GLODAPv2, carbon dioxide, dissolved inorganic carbon, total alkalinity, pH, oxygen, nutrients, quality control, CLIVAR, Atlantic Ocean, Pacific Ocean, Indian Ocean and Arctic Ocean.

1. INTRODUCTION

GLODAPv2 (Global Ocean Data Analysis Project version 2) is an international data synthesis project for interior ocean inorganic carbon data and related variables for the global ocean. GLODAPv2 was officially instigated with the EU integrated project CARBOCHANGE in 2010. However, its history can be traced back much further than that as we began assembling a collection of data immediately in the wake of the release of the first version of GLODAP, GLODAPv1.1 (Key et al., 2004; Sabine et al., 2005) in anticipation of sufficient data, momentum and funding to assemble the next version. Meanwhile, the CARINA synthesis of carbon-relevant data from the Arctic, Atlantic and Southern Ocean was completed (Key et al., 2010, Tanhua et al. 2009) as was the Pacific effort, PACIFICA (Suzuki et al., 2013). Combined with GLODAPv1.1, these products formed the natural basis for GLODAPv2. The aim of GLODAPv2 was to unify the data of the first version of GLODAP (GLODAPv1.1) with the data from CARINA and PACIFICA, add any new data that were made available to us and fully re-evaluate all of these data to produce a global calibrated data product and mapped climatology. This document provides a brief overview of GLODAPv2. Complete and extensive documentation is provided in Olsen et al. (2015) and Lauvset et al. (2015).

2. DATA COLLECTION, CONVERSION AND PRIMARY QUALITY CONTROL

GLODAPv2 includes data from all ocean areas of the globe (Figure 1). Data from altogether 556 unique cruises were available from GLODAPv1.1, CARINA and PACIFICA, and data from 168 new cruises were added to the collection. Note that the GLODAPv2 definition of "cruise" is imprecise. In some cases, what were originally referred to as “legs” are now called cruises. In others, what may have been described as multiple cruises have been merged. In a few instances, our QC analysis was carried out on subsets (strings of stations) of a single leg. None of these distinctions are particularly important. Therefore, in this and other GLODAPv2 discussions, “cruise” will be used to refer to a numbered item in the Cruise Summary Table (CST). The CST provides summary information for each cruise included in GLODAPv2, including its constituent legs, GLODAPv2 cruise number, EXPOCODE, known aliases and links to relevant scientific publications and our secondary QC (see section 3) results. The EXPOCODES have been generated by combining the NODC platform code (http://www.nodc.noaa.gov/General/NODC-Archive/platformlist.txt) with the YYYYMMDD formatted starting date of the cruise. During our work the ICES data center took over the responsibility for assigning platform codes; in the few cases where the NODC and ICES assigned codes differed, we generally preferred the older or better known code. The GLODAPv2 cruise numbers were assigned sequentially after sorting all cruises in order of increasing (numerically and alphabetically) EXPOCODE.

All of the original data files were reformatted to WOCE exchange formatted files (Swift and Diggs, 2008), which includes a header that provides the essential metadata for each variable, whenever this could be obtained. The unit of concentration for nutrients and oxygen was converted to $\mu mol\ kg^{-1}$, if not reported in this unit already. This was carried out assuming a lab temperature of 22°C for nutrients. For oxygen we used the factor 44.66 for any milliliter to μmol conversion and potential density for liter to kg conversion, as the draw temperature was generally not available.

Primary QC was carried out following Sabine et al. (2005) and Tanhua et al. (2010), by inspecting date, time, sampling positions and property-property plots. Any outliers observed in two or more property-property plots were flagged as such using the flagging scheme of WOCE (Swift and Diggs, 2008).

3. SECONDARY QUALITY CONTROL

Secondary QC was performed to allow us to identify measurement biases in the data, and correct for these in the GLODAPv2 product files. The secondary QC was carried out using crossover and inversion procedures as described by Johnson et al. (2001), Gouretski and Jancke (2001), Tanhua et al. (2010) and Lauvset and Tanhua (2015). Data for the following sea water constituents were subjected to secondary QC: Salinity, Oxygen, Nitrate, Silicate, Phosphate, TCO_2, TAlk, pH, CFC-11, CFC-12, CFC-113 and CCl_4. The secondary QC was carried out in a series of steps as described in detail by Olsen et al. (2015), and only its most important principles are summarized here. Comparison of data from different cruises was carried out using data from beneath 2000 m only, since the time scale covered by GLODAPv2 can generally be considered synoptic at these depths and within the cut-off limits (defined by the GLODAPv2 group) for adjustment application. In the areas where data cannot be considered synoptic at the time scale in question, we took great care to not detrend any data with our adjustments. Hence, GLODAPv2 includes trends in both oxygen and TCO_2 in known deep-water formation areas such as the Labrador Sea and Greenland Sea. All adjustments derived in GLODAPv1.1, CARINA and PACIFICA were reevaluated within the context of the entire GLODAPv2 dataset and revised where necessary. The results of the secondary QC are provided on a per cruise and parameter basis in the online Adjustment Table at https://glodapv2.geomar.de. Access to this table is also available via various links at the primary CDIAC web site (http://cdiac.ornl.gov/oceans/GLODAPv2/). Adjustments for salinity, TCO_2, TAlk and pH are additive while the adjustments for oxygen, the nutrients and CFCs are multiplicative. A positive additive adjustment or a multiplicative adjustment greater than 1 indicates that the original data were biased low. A negative adjustment or a multiplicative adjustment less than 1 indicates that the original data were biased high. The secondary QC results are described in detail by Olsen et al. (2015).

4. GLODAPv2 DATA AND PRODUCT

GLODAPv2 data are served as (1) the original cruise data sets that were subjected to reformatting and primary QC as described in Sect. 2, and (2) a merged data product that was

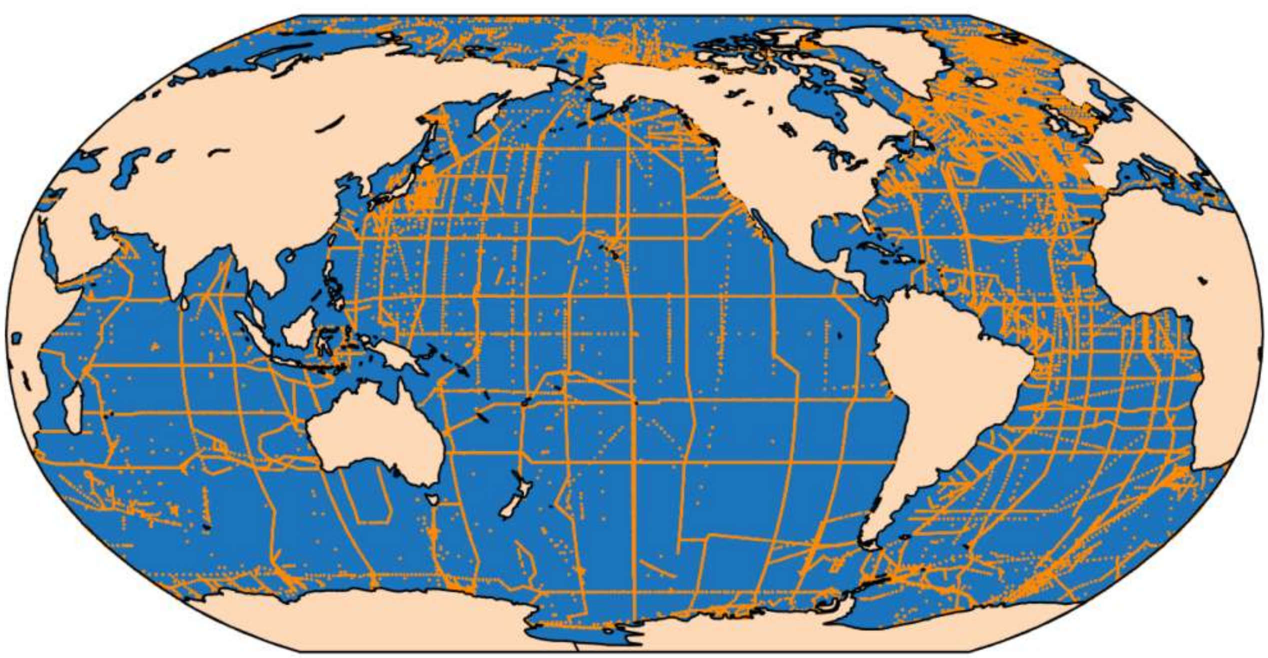

Figure 1: GLODAPv2 station locations.

prepared following the secondary QC as described in Sect 3. In addition, we have prepared a mapped climatology, which is briefly described in Sect. 5 of this document and more extensively by Lauvset et al. (2015).

The original data files (as obtained from the data originator) and documentation can be accessed through the GLODAPv2 CST. *None of the adjustments that we have derived have been applied to the data in the original files.*

The bias-corrected data product is available as a global product file that includes data from all ocean areas, and additionally as four regional files, one each for the Arctic, Atlantic, Indian and Pacific Oceans. These altogether 5 collections are available as comma separated ASCII files (*.csv) and as MATLAB binary data files (*.mat). There is no data overlap among the regional files. A few cruises covered more than one region, but these were never divided nor included in more than one regional file. *All of the adjustments determined and vetted by the GLODAPv2 group have been applied on the data in the product files.* In addition, we have interpolated missing oxygen and nutrients from station data whenever possible following Key et al. (2010), and calculated the third inorganic carbon parameter (TCO_2, TAlk or pH), whenever only two were present in the data file. For some cruises, which had just a few data points for the third inorganic carbon parameters measured for instrument testing or personnel training purposes for instance, the data for this third parameter were replaced with calculated ones. Details of these procedures are provided by Olsen et al. (2015). All pH data have been converted to the total scale at surface pressure (0 dbar) and 25ºC (variable phts25p0) and to the total scale at *in situ* pressure and temperature (variable phtsinsitutp). A full list of parameters included in the product files is provided in Olsen et al. (2015). In the discrete product files, each parameter name is preceded by "G2" (i.e. G2phts25p0 for pH on the total hydrogen scale at 25˚C and surface pressure, etc.).

5. GLODAPv2 MAPPED CLIMATOLOGY

The GLODAPv2 mapped product includes global fields at 1° resolution of TCO_2, TAlk, pH, Ω_{Ca} and Ω_{Ar}, phosphate, nitrate, silicate, oxygen, salinity and potential temperature on 33 pressure surfaces. These 33 *pressure* levels are numerically identical to the 33 *depth* levels used for the World Ocean Atlas 2009 and its past versions (Locarnini et al., 2010). As therein, the pixel center latitudes and longitudes run from -89.5 to 89.5ºN and -179.5 to 179.5ºE, respectively. Each property map comes with an associated error map, giving the mapping error in the unit of the parameter in question. This mapping error does not include the measurement uncertainty.

For the strongly time-dependent parameters TCO_2, pH, Ω_{Ca} and Ω_{Ar} we distinguished between the WOCE and CLIVAR eras in the upper 1000 dbar. The former was mapped using data from 1986-1999, and the latter using data from 2000-2013. This is not to say that a mere subtraction of these fields gives the changes due to uptake of anthropogenic CO_2 over the ten years; the separation into two periods simply reduces the risk of translating temporal variability in the input data to spatial variations in the mapped output.

In pixels where the relative mapping error exceeds one standard deviation in the input data on a given pressure level, the mapped value was replaced by -999. This essentially equates to masking values in data-poor regions; notable examples include the Caribbean, the WOCE-era Mediterranean, and the CLIVAR-era Northwest Indian Ocean.

Lauvset et al. (2015) provide the full details of the approaches used for the mapping. The mapped product was prepared from the bias-corrected global product file using a three-step procedure. First, the data for each station were interpolated to the 33 pressure surfaces using a piecewise cubic Hermite interpolating polynomial as implemented in the MATLAB function *pchip* and respecting the maximum separation criteria given in Table 4 of Key et al. (2010). Secondly, the data were bin-averaged into a 1º latitude and longitude grid before, finally, each of the 33 surfaces was mapped using the Data-Interpolating Variational Analysis mapping method (DIVA, Beckers et al., 2014; Troupin et al., 2012). The main advantage of DIVA over the Objective Analysis used for the horizontal interpolation in GLODAPv1.1 is that DIVA has improved functionality for taking the presence of topography into account; otherwise the two methods yields very similar results. Using DIVA allowed us to map the global oceans in a single analysis; for the Objective Analysis of GLODAPv1.1, the mapping had to be carried out per basin to avoid interpolation across barriers such as the Panama isthmus, the southern tip of South America or even the mid-Atlantic Ridge for deep surfaces in the northern basin. The correlation length scale used for mapping GLODAPv2 was a uniform 7º, except for the time-dependent parameters (TCO_2, pH, Ω_{Ca} and Ω_{Ar}) for which 10º was used in the upper 1000 dbar and 7º below. Meridionally this is equivalent to that used for the mapping in GLODAPv1.1, but zonally it is much smaller. Instead of increasing the zonal correlation, advection contraint was used to allow for easier mixing in that direction. The signal-to-noise ratio was set to 10 for all parameters at all depths. Again, this matches that used for GLODAPv1.1. The error fields that accompany the mapped data were derived using the "almost exact" error calculation algorithm distributed as part of the DIVA package.

6. USER REQUIREMENTS

GLODAPv2 is released without any restrictions and can be downloaded free of charge or log-in requirements. However, we ask users to generously acknowledge the contribution of GLODAPv2 data providers in the form of invitation to co-authorship, especially in case of regional studies that depend critically on data from a manageable number of principal investigators (PIs). We also ask users to reference relevant scientific articles by data providers. Those that we are aware of are available on a per cruise basis in the CST. Data providers should inform us of any references that they feel are missing.

GLODAPv2 is supplied with a list of known errors in data files and product. Users should consult this when GLODAPv2 is used, so they may correct these. Any errors that are discovered should be reported to the GLODAPv2 group, so these may be included in this document of known errors.

Please cite this document and Olsen et al. (2015) when GLODAPv2 is used. For the mapped climatology, please also cite Lauvset et al. (2015). Any changes or additions to this request will be posted on the data product download page at CDIAC. For individual cruise data files please acknowledge the original data generators.

7. ACKNOWLEDGEMENTS

GLODAPv2 would not have been possible without the effort of hundreds of sea-going scientists who generated funding and dedicated time to collect the data, and their willingness to share their data with the wider scientific community. 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. R. Key was supported by KeyCrafts grant 2012-001, CICS grants NA08OAR4320752 and NA14OAR4320106, NASA grant NNX12AQ22G, NSF grants OCE-0825163 (with a supplement via WHOI P.O. C119245) and PLR-1425989, and DOE UT-Battelle contract #4000133565 to CDIAC. Importantly, the EU-IP CARBOCHANGE (FP7 264878) 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 Horizon 2020 project (grant agreement 633211) and the Norwegian Research Council project SNACS (229752). S.K. Lauvset and E. Jeansson appreciate 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 U.S. National Science Foundation grant (OCE- 1243377) to the Scientific Committee on Oceanic Research. K. Bakker (NIOZ) shared his invaluable insight with us, helping with the secondary quality control of the nutrient data. A. Kozyr was supported by DOE contract DE-AC05-00OR2272 to UT-Battelle, operators of CDIAC under ORNL.

8. HOW TO OBTAIN THE DATA AND DOCUMENTATION

All individual cruise data, latest data product and mapped climatology are available from the GLODAPv2 website at http://cdiac.ornl.gov/oceans/GLODAPv2/

Contact information:
Alex Kozyr
CDIAC, Environmental Sciences Division , Oak Ridge National Laboratory
U.S. Department of Energy , Building 4500N, Mail Stop 6290
Oak Ridge, TN 37831-6290 U.S.A.
Telephone: (865) 576-8449
E-mail: kozyra@ornl.gov
Ocean CO_2 WWW Page: http://cdiac.ornl.gov/oceans/

9. REFERENCES

Beckers, J.-M., A. Barth, C. Troupin, and A. Alvera-Azcárate (2014), Approximate and Efficient Methods to Assess Error Fields in Spatial Gridding with Data Interpolating Variational Analysis (DIVA). *Journal of Atmospheric and Oceanic Technology, 31*(2), 515-530. doi: 10.1175/JTECH-D-13-00130.1

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

Johnson, G., P. E. Robbins and G. E. Hufford (2001), Systematic adjustments of hydrographic sections for internal consistency, *J. Atmos. Oceanic Technol.*, *18*, 1234-1244.

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

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

Lauvset, S. and T. Tanhua (2015), A toolbox for secondary quality control on ocean chemistry and hydrographic data, *Limnol. Oceanogr. Methods*, doi: 10-1002/lom3.10050.

Lauvset, S., R. M. Key, A. Olsen, T. Tanhua, M. Hoppema, S. Jutterström, R. Steinfeldt, E. Jeansson, A. Kozyr, B. Pfeil, T. Suzuki, M. Ishii, S. van Heuven and A. Velo (2015), A new global interior ocean carbon climatology: the 1ºx1º GLODAP version 2, *Earth Sys. Sci. Data. Discuss*, *submitted*.

Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M.M. Zweng and D. R. Johnson (2010), *World Ocean Atlas 2009, Volume 1, Temperature*, in S. Levitus (ed) *NOAA Atlas NESDIS 68*, US Government Printing Office, Washington D. C., 184 pp. http://www.nodc.noaa.gov/OC5/indprod.html

Olsen, A., R. M. Key, S. van Heuven, S. K. Lauvset, A. Velo, X. Lin, C. C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, S. Jütterström, R. Steinfeldt, E. Jeansson, M. Ishii, F. F. Perez and T. Suzuki (2015), An internally consistent data product for the world ocean: the Global Ocean Data Analysis Project, version 2 (GLODAPv2), *Earth Syst. Sci. Data Discussions*, *submitted*.

Sabine, C., R. M. Key, A. Kozyr, R. A. Feely, R. Wanninkhof, F. J. Millero, T.-H. Peng, J. L. Bullister and K. Lee (2005), *Global Ocean Data Analysis Project (GLODAP): Results and Data*, ORNL/CDIAC-145, NDP-083, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U. S. Department of Energy, Oak Ridge, Tennessee, 110 pp.

Suzuki, T., M. Ishii, A. Aoyama, J. R. Christian, K., Enyo, T. Kawano, R. M. Key, N. Kosugi, A. Kozyr, L. A. Miller, A. Murata, T. Nakano, T. Ono, T. Saino, K. Sasaki, D. Sasano, Y. Takatani, M. Wakita, and C. L . Sabine (2013), *PACIFICA Data Synthesis Project*, ORNL/CDIAC-159, NDP-092, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U. S. Department of Energy, Oak Ridge, Tennessee, 5pp.

Swift, J. and S. C. Diggs (2008), Description of WHP exchange format for CTD/Hydrographic data. CLIVAR and Carbon Hydrographic Office, UCSD Scripps Institution of Oceanography, 19 pp.

Tanhua, T., S. van Heuven, R. M. Key, A. Velo, A. Olsen and C. Schirnick (2010), Quality control procedures and methods of the CARINA database, *Earth Syst. Sci. Data*, *2*, 35-49.

Tanhua, T., A. Olsen, M. Hoppema, S. Jutterström, C. Schirnick, S. van Heuven, A. Velo, X. Lin, A. Kozyr, M. Alvarez, D. C. E. Bakker, P. Brown, E. Falck, E. Jeansson, C. Lo Monaco, J. Olafssson, F. F. Perez, D. Pierrot, A. F. Rìos, C. L. Sabine, U. Schuster, R. Steinfeldt, I. Stendardo, L. G. Andeson, N. R. Bates, R. G. J. Bellerby, J. Blindheim, J. L. Bullister, N. Gruber, M. Ishii, T. Johannessen, E. P. Jones, J. Köhler, A. Körtzinger, M. Metzl. A. Murata, S. Musielewicz, A. M. Omar, K. A. Olsson, M. de la Paz, B. Pfeil, R. Rey, M. Rhein, I. Skjelvan, B. Tilbrook, R Wanninkhof, L. Mintrop, D. W. R. Wallace and R. M. Key (2009), *CARINA Data Synthesis Project*, ORNL/CDIAC-157, NDP-091, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U. S. Department of Energy, Oak Ridge, Tennessee, 12pp.

Troupin, C., Barth, A., Sirjacobs, D., Ouberdous, M., Brankart, J. M., Brasseur, P., Rixen, M., Alvera-Azcárate, A., Belounis, M., Capet, A., Lenartz, F., Toussaint, M. E., and Beckers, J. M.: Generation of analysis and consistent error fields using the Data Interpolating Variational Analysis (DIVA), Ocean Model., 52–53, 90-101, 2012.