Searth System Discussion Science Solutions Data



# A new global interior ocean mapped climatology: the 1°x1°

# 2 GLODAP version 2

- 3 Siv K. Lauvset<sup>1,2</sup>\*, Robert M. Key<sup>3</sup>, Are Olsen<sup>1,2</sup>, Steven van Heuven<sup>4</sup>, Anton Velo<sup>5</sup>, Xiaohua
- 4 Lin<sup>3</sup>, Carsten Schirnick<sup>6</sup>, Alex Kozyr<sup>7</sup>, Toste Tanhua<sup>6</sup>, Mario Hoppema<sup>8</sup>, Sara Jutterström<sup>9</sup>,
- 5 Reiner Steinfeldt<sup>10</sup>, Emil Jeansson<sup>2</sup>, Masao Ishii<sup>11</sup>, Fiz F. Perez<sup>5</sup>, Toru Suzuki<sup>12</sup>, Sylvain
- 6 Watelet<sup>13</sup>.
- <sup>1</sup>Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research,
   Allègaten 70, 5007 Bergen, Norway
- <sup>2</sup>Uni Research Climate, Bjerknes Centre for Climate Research, Allegt. 55, 5007 Bergen,
   Norway
- <sup>3</sup>Atmospheric and Oceanic Sciences, Princeton University, 300 Forrestal Road, Sayre Hall,
   Princeton, NJ 08544, USA
- <sup>4</sup>Royal Netherlands Institute for Sea Research (NIOZ), Marine Geology and Chemical
- 14 Oceanography, P.O. Box 59, 1790 AB Den Burg, The Netherlands
- 15 <sup>5</sup>Instituto de Investigaciones Marinas CSIC, Eduardo Cabello 6, 36208 Vigo, Spain
- <sup>6</sup>GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel,
- 17 Germany
- <sup>7</sup>Carbon Dioxide Information Analysis Center, Environmental Sciences Division, Oak Ridge
- 19 National Laboratory, U.S. Department of Energy, Building 4500N, Mail Stop 6290, Oak
- 20 Ridge, TN 37831-6290, U.S.A.
- <sup>8</sup>Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bussestrasse 24,
   27570 Bremerhaven, Germany
- <sup>9</sup>IVL Swedish Environmental Research Institute, Ascheberggatan 44, 411 33 Göteborg,
   Sweden
- <sup>10</sup>University of Bremen, Institute of Environmental Physics, Otto-Hahn-Allee, 28359 Bremen,
   Germany
- <sup>11</sup>Oceanography and Geochemistry Research Department, Meteorological Research Institute,
- 28 Japan Meteorological Agency, 1-1 Nagamine, Tsukuba, 305-0052 Japan
- 29 <sup>12</sup>Marine Information Research Center, Japan Hydrographic Association, 1-6-6-6F,
- 30 Hanedakuko, Otaku, Tokyo, 144-0041 Japan
- 31 <sup>13</sup>Department of Astrophysics, Geophysics and Oceanography, University of Liège, Liège,
- 32 Belgium
- 33 \*Communication to Siv K. Lauvset (siv.lauvset@uib.no)





## 34 Abstract

35 We here present the new GLODAP version 2 (GLODAPv2) mapped climatology, which is 36 based on data from all ocean basins up to and including 2013. In contrast to its predecessor, GLODAPv1.1, this climatology also covers the Arctic Ocean and Mediterranean Sea. The 37 38 quality controlled and internally consistent data product files of GLODAPv2 (Olsen et al., 2015; Key et al., 2015) were used to create global 1°x1° mapped climatologies of total 39 40 dissolved inorganic carbon, total alkalinity, and pH using the Data-Interpolating Variational 41 Analysis (DIVA) mapping method. Climatologies were created for 33 standard pressure 42 surfaces. To minimize the risk of translating temporal variability in the input data to spatial 43 variations in the mapped climatologies, layers with pressures of 1000 dbar, or less, were mapped for two different time periods: 1986-1999 and 2000-2013, roughly corresponding to 44 the "WOCE" and "CLIVAR" eras of global ocean surveys. All data from the 1972-2013 45 period were used in the mapping of pressures higher than 1000 dbar. In addition to the marine 46 47 CO<sub>2</sub> chemistry parameters listed above, nitrate, phosphate, silicate, oxygen, salinity and theta 48 were also mapped using DIVA. For these parameters all data from the full 1972-2013 period 49 were used on all 33 surfaces. The GLODAPv2 global 1°x1° mapped climatologies, including 50 error fields and ancillary information have been made available at the GLODAPv2 web page 51 at the Carbon Dioxide Information Analysis Center (CDIAC, 52 http://cdiac.ornl.gov/oceans/GLODAPv2/).

53

#### 54 1 Introduction

55 Accurate estimates of recent changes in the ocean carbon cycle, including how these changes will influence climate, requires high quality data. The fully quality controlled and 56 57 internally consistent Global Ocean Data Analysis Project (GLODAPv1.1, Key et al., 2004) 58 has for the past decade been the only global interior ocean carbon data product available. 59 GLODAPv1.1 has been and continues to be of immense value to the ocean scientific 60 community, which is reflected in the almost 500 scientific studies that have used and cited 61 GLODAPv1.1 so far. GLODAPv1.1 has been used most prominently for calculation of the 62 global ocean inventory for anthropogenic  $CO_2$  by *e.g.* Sabine et al. (2004) and for validation of global biogeochemical or earth system models by e.g. Bopp et al. (2013). 63 64 The GLODAPv1.1 data product is dominated by data from the World Ocean Circulation Experiment (WOCE) survey of the 1990s, but contains data from the entire period 65





66 1972-1999, though very few data north of 60°N in the Atlantic and no data in the Arctic 67 Ocean or Mediterranean Sea. Many more seawater CO<sub>2</sub> chemistry data have been collected at research cruises after 1999, particularly within the framework of the global repeat 68 69 hydrography program CLIVAR/GO-SHIP (Feely et al., 2014; Talley et al. 2016) so that 70 significantly more interior ocean carbon data exists today than was available in 2004. 71 In response to the shortcomings of GLODAPv1.1 and to include more recent data, the 72 updated and expanded version GLODAPv2 (Key et al., 2015; Olsen et al., 2015), has been 73 developed. This new data product combines GLODAPv1.1 with data from the two recent 74 regional synthesis products: Carbon in Atlantic Ocean (CARINA, Key et al., 2010); and 75 Pacific Ocean Interior Carbon (PACIFICA, Suzuki et al., 2013). In addition, data from 168 76 cruises not previously included in any of these data products-both new and historical-have 77 been included. Notably, 116 cruises in GLODAPv2 cover the Arctic Mediterranean Seas, i.e., 78 the Arctic Ocean and the Nordic Seas (>65°N). GLODAPv2 data are available in three forms: 79 as original, unadjusted data from each cruise in WOCE exchange format files; as a merged 80 and calibrated data product, where adjustments have been applied to minimize measurement 81 biases and several calculated data have been added to complete the data coverage; and as a 82 mapped climatology. This paper presents the methods used for creating the mapped 83 climatology and its main features, while the assembly of the data and construction of the 84 product, including the broad features and output of the secondary quality control are described 85 by Olsen et al. (2015).

86 As opposed to a gridded data product, which e.g. the Surface Ocean CO<sub>2</sub> Atlas (Pfeil 87 et al., 2013; Bakker et al., 2014) provides (Sabine et al., 2013), we have created mapped 88 climatologies. The difference is that gridded data are observations projected onto a grid, using 89 some form of binning and averaging, but no interpolation or other form of calculation is used 90 to fill gaps in the observational record. In mapped data the gaps have been filled, in the case 91 of GLODAPv2 using an objective mathematical method. The method used to create the 92 mapped climatologies from the merged and calibrated data product is presented in Section 93 2.2. Some of the resulting data fields and their associated error estimates are shown in Section 94 3 to highlight important features in the data product; and finally some recommendations for 95 use and interpretation are given in Section 4.

96





# 97 2 Methods

# 98 2.1 Input data

99		The input data for the GLODAPv2 mapped climatology consisted of the bias corrected
100	and m	erged data product from GLODAPv2 (Olsen et al., 2015). Whereas the complete data
101	produ	ct contain many variables, we mapped the GLODAPv2 primary biogeochemical
102	variab	les (Olsen et al., 2015): total dissolved inorganic carbon (TCO <sub>2</sub> ), total alkalinity (TAlk),
103	pH, sa	turation state of calcite and aragonite ( $\Omega_{\rm C}$ and $\Omega_{\rm A}$ ), nitrate (NO <sub>3</sub> <sup>-</sup> ), phosphate (PO <sub>4</sub> <sup>3-</sup> ),
104	silicat	e (Si), dissolved oxygen (O <sub>2</sub> ), salinity, and potential temperature, where the latter two
105	variab	les are to be used as a reference for the biogeochemical variables. The GLODAPv2 data
106	produ	ct includes vertically interpolated data for the nutrients, oxygen and salinity if any of
107	those	were missing from a bottle data-point, and calculated seawater CO <sub>2</sub> chemistry data
108	whene	ever pairs of measured CO <sub>2</sub> chemistry parameters were available (Olsen et al., 2015).
109	These	were all included in the mapping. The following pre-mapping data treatments were
110	carrie	d out:
111	1.	$\Omega_C$ and $\Omega_A$ were calculated from the TCO <sub>2</sub> and TAlk pair at <i>in situ</i> temperature and
112		pressure using the MATLAB version (van Heuven et al., 2009) of CO2SYS (Lewis
113		and Wallace, 1998). We used pressure, temperature, salinity, phosphate, and silicate
114		from the GLODAPv2 data product, the dissociation constants of Lueker et al. (2000)
115		for carbonate, Dickson (1990) for sulphate, and the total borate concentration of
116		Uppstrom (1974).
117	2.	All data were vertically interpolated onto 33 surfaces: 0, 10, 20, 30, 50, 75, 100, 125,
118		150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500,
119		1750, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500 dbar. The interpolation was
120		done station by station, using a cubic hermite spline function. This interpolation
121		method is quite robust, but can give unreliable results in a few unusual circumstances.
122		Consequently, if this interpolation gave values more than 1% different from those
123		produced using a simple linear vertical interpolation the linear results were used. We
124		used the maximum distance criteria specified in Table 1 to avoid interpolation over
125		excessive vertical distances between data points. These maximum distance criteria are
126		similar to those used by Key et al. (2004) for GLODAPv1.1. Note that the
127		GLODAPv2 climatologies cover 33 pressure surfaces, which is slightly different from





128		the depth surfaces, originally chosen by Levitus and Boyer (1994), used in
129		GLODAPv1.1.
130	3.	The vertically interpolated data for each pressure surface were then gridded by bin-
131		averaging all data in each 1°x1° grid cell. The mapped climatologies are thus based on
132		gridded data. We do this because the repeat hydrography program means that there are
133		several transects in the ocean that have observations at the same points in space at
134		different points in time.
135	4.	GLODAPv2 includes data for the 42 year period 1972-2013 and in this time frame
136		there have been significant changes in pH, TCO <sub>2</sub> , and the saturation states, due to
137		increasing atmospheric levels of CO <sub>2</sub> (e.g. Orr et al., 2001; Lauvset et al., 2015;
138		Sabine and Tanhua, 2010). Therefore the upper ocean data, here defined as pressure
139		less than, or equal to, 1000 dbar, were separated into two time periods for TCO <sub>2</sub> , pH,
140		$\Omega_C$ , and $\Omega_A$ : 1986-1999 and 2000-2013, roughly corresponding to the WOCE and
141		CLIVAR eras of global hydrography programs. These were then mapped separately to
142		reduce the risk of transforming time trends into spatial variations in the mapped
143		climatologies. No additional corrections were used to account for the seasonal cycle or
144		potential bias due to uneven temporal sampling. Below 1000 dbar, and for all other
145		mapped parameters, all available data from 1972-2013 were used in the mapping. The
146		inherent assumption is that the change in dissolved inorganic carbon and pH below
147		1000 dbar is negligible. This assumption is further discussed in Section 4.

148 **2.2 Mapping method** 

149 The Data-Interpolating Variational Analysis (DIVA) mapping method (Beckers et al., 150 2014; Troupin et al., 2012) was used to create the mapped climatologies. DIVA is the 151 implementation of the Variational Inverse Method (VIM) of mapping discrete, spatially 152 varying data. A major difference between this and the Optimal Interpolation (OI) method used 153 in GLODAPv1.1 is how topography is handled. DIVA takes the presence of the seabed and 154 land into account during the mapping and gives better results in coastal areas and around islands. In addition, the entire global ocean can be mapped at once, e.g, DIVA does not 155 156 propagate information across narrow land barriers such as the Panama isthmus so there is no 157 need to split the data into ocean regions which are then stitched together to form a global map. 158 Hence, each climatology is a global analysis for the range 180°W to 180°E with a 1°x1° 159 resolution. To ensure that the analysis converges on the boundaries, *i.e.* the dateline and the





160 North Pole, the input data were duplicated for 10° on either side of the dateline and for every 161 30° in a circle around the North Pole (i.e., north of 82.5°N data in the longitude band 0-30° 162 were duplicated into the 30-60° and so on), before mapping. This removes most 163 discontinuities along these boundaries, but in some cases discontinuities still appear along the 164 North Pole boundary, most notably in the top 125 dbar in the 2000-2013 climatology for 165 TCO<sub>2</sub>. Note that this approach creates an artificially large amount of data in the Arctic Ocean, 166 but the spatial patterns in the observations are retained. 167 Apart from the data, the most important DIVA input parameters are the spatial 168 correlation length scale (CL) and the data signal-to-noise ratio (SNR). The CL defines the 169 characteristic distance over which a data point influences its neighbors. For GLODAPv2 this 170 was defined a priori as  $7^{\circ}$  for all parameters, except for TCO<sub>2</sub> and pH. In latitude this approximately matches the 750 km north-south CL used for the GLODAPv1.1 mapped 171 172 climatologies, but is in longitude much smaller than the 1500 km east-west CL used for 173 GLODAPv1.1. For TCO<sub>2</sub>, pH,  $\Omega_{\rm C}$ , and  $\Omega_{\rm A}$  a CL of 10° was used for the two time periods in 174 the top 1000 dbar, and 7° in the deep ocean. This was required because separating these 175 parameters into two time periods significantly reduced the data density (Figure 1), and the 176 smaller CL led to gaps in the climatology. Since the oceans generally tend to mix easier 177 zonally than meridionally, a pseudo-velocity field and advection constraint was used such that 178 the correlation becomes stronger in the east-west direction even though the input CL is the 179 same for both directions (see the DIVA user manual available at 180 http://modb.oce.ulg.ac.be/mediawiki/index.php/Diva\_documents for details). Setting the CL 181 is partly a subjective effort, aiming to strike the optimal balance between large values, which 182 tends to smooth the data fields and reduce mapping errors, and small values, which leads to 183 more correct rendering of fronts and other features. We also want to stay within the physical 184 constraints set by ocean dynamics and natural spatial variability. It is possible to optimize CL 185 in DIVA, but this works well only when the data density is reasonably high. The sparse global data distribution in GLODAPv2 gives optimized CL in the order of 25°. Doing a cruise-by-186 187 cruise analysis following Jones et al. (2012) to get spatially varying CL is possible, but would leave large gaps. For GLODAPv2 it was therefore decided to use a globally uniform a priori 188 189 choice since this is the most transparent and easily reproducible. 190 The SNR defines how representative the observations are for the climatological state.

- 191 For spatially varying data sets like GLODAPv2, it is the assumed ratio of climatological
- 192 spatial variability ("signal") to the short term variability ("noise") in the data. For the





193 GLODAPv2 mapped climatologies the SNR was defined a priori to be 10 (i.e. the noise is 194 10% of the signal), following Key et al. (2004). To understand the importance of SNR, and 195 the reason for a subjective a priori choice, a brief discussion of the differences between 196 interpolation and approximation/analysis is necessary. When interpolating between points in a 197 data set, gaps between data points are filled but the existing points are not replaced. When 198 approximating, a function, e.g. a regression line, is applied that describes the original data 199 points to some degree. The resulting approximated data set has new values at every point and 200 is smoother than the interpolated data set. None of the approximated data points exactly match 201 the original ones, and that assumes more uncertainty - or non-climatological variability - in 202 the data. In the case of very high SNR, the observed values are retained in the mapped 203 climatology and DIVA interpolates between them, while smaller values allow for larger 204 deviations from these and an increasingly smooth climatology.

205 Working with real world observations, we know that the observations are indeed 206 affected by shorter term variations and in addition have uncertainties associated with them. 207 They do not represent the "true" climatological value. For this reason the SNR should always 208 be kept quite small when making mapped climatologies, but this needs to be balanced by the 209 need to keep the error estimates reasonable. The lower the SNR the further the approximation 210 is allowed to deviate to be from the original data and the higher the error associated with the 211 approximation becomes. The SNR can be calculated from observations using generalized 212 cross validation, but for GLODAPv2 such calculations give very high SNR (in the order of 213 100). This is maybe not completely unreasonable, since GLODAPv2 has been carefully 214 quality controlled and we have high confidence that the measurement uncertainties are small. 215 Using gridded data covering 14 years as input also means that the input is reasonably 216 representative relative to the climatology. However, both increasing the SNR and increasing 217 the CL will decrease the error estimates, because this assumes small representativity errors 218 (*i.e.* that what is observed is the true climatology) and a large circle of influence. If those 219 assumptions are wrong the errors will be significantly underestimated. Therefore, even with a 220 high confidence that the input data are climatologically representative the mapping errors are 221 likely to be underestimated if we use the SNR calculated from general cross validation. 222 A DIVA analysis is created by minimizing a cost function which is defined by the 223 difference between observations and analysis; the smoothness of the analysis; and the

- 224 physical laws of the ocean (Troupin et al., 2012). The result is thus the analysis with the
- smallest global mean error, but determining the spatial distribution of errors is important. In

7





226	DIVA this is non-trivial as, in contrast to OI, the real covariance function, which is necessary
227	to obtain spatial error fields, is not formulated explicitly, but is instead the result of a
228	numerical determination (Troupin et al., 2012). Determining the real covariance to get error
229	estimates is the most exact method, but is computationally expensive. There are several error
230	estimation methods implemented in DIVA, from the very simple to the very exact (Troupin et
231	al., 2012, Beckers et al., 2014) and for GLODAPv2 the error fields are based on the "almost
232	exact error calculation". This method calculates the exact error, using real covariances, in a
233	few locations and then uses DIVA to interpolate between them. While not tested for all
234	climatologies due to the significant computational cost, the almost exact errors generally only
235	differ significantly from the exact error in regions where the data fields have very high errors,
236	which only happens in coastal areas and in areas with no data. Since a mask removing the
237	result in all grid cells where the mapping error exceeds one standard deviation in the input
238	data (on a given pressure surface) have been applied to the mapped climatologies, the almost
239	exact errors are considered equivalent to the real covariance errors.

240

## 241 3 Results

#### 242 3.1 Data fields

The mapped climatologies are available as one netcdf files per parameter from CDIAC 243 (http://cdiac.ornl.gov/oceans/GLODAPv2/). Each of these contain the global 1°x1° 244 245 climatology, the associated error fields, and the gridded input data (Table 2) for the parameter 246 in question. The files containing TCO<sub>2</sub>, pH,  $\Omega_C$ , and  $\Omega_A$  are four-dimensional due to the two 247 different time periods (1986-1999 and 2000-2013) for the top 1000 dbar. For all surfaces below 1000 dbar (*i.e.* surfaces 20-33) the TCO<sub>2</sub>, pH,  $\Omega_C$ , and  $\Omega_A$  climatologies are identical 248 249 for both time periods. The fields for the other parameters are three-dimensional, since the full 250 time period 1972-2013 was used in the mapping.

Figures 2-4 show the mapped climatologies for TCO<sub>2</sub>, TAlk, and nitrate, respectively, at two different pressure surfaces. These all show the spatial patterns expected from biological dynamics, global ocean salinity, and large-scale circulation. Figures 5-7 show, for the same parameters and pressure surfaces, the difference between the gridded input and the mapped climatologies, which is relatively large and variable near the surface and generally within the data uncertainties in the deep ocean. Figures 8-10 show the error fields associated with the climatologies shown in Figures 2-4. There are large differences in spatial data coverage





258 between the 1986-1999 and the 2000-2013 periods (Figure 1), which affect the error estimates for the top 1000 dbar (Figure 8). The 1986-1999 period (Figure 1a) has no data in the 259 Caribbean and Mediterranean Seas, and whereas the 2000-2013 period (Figure 1b) has data in 260 261 those regions it lacks coverage in the Indian Ocean and has lower data density in the Pacific 262 Ocean. The spatial variability in mapping errors is a function of the observational network, 263 and further study of this variability would be of great use in optimizing the existing and future 264 observational networks. The biggest improvement in GLODAPv2 compared to GLODAPv1.1 is that the former includes the Arctic Ocean and Mediterranean Sea which was missing in 265 266 GLODAPv1.1 due to a near total absence of data from these regions.

# 267 3.2 Error fields

While the mapping error reflects the data distribution and the choice of input variables 268 269 (*i.e.* CL and SNR), it represents only the errors due to the mathematical mapping of the input 270 data, and does not take into account all the uncertainty in the input data (although some of this 271 uncertainty is assumed in the choice of SNR). For details regarding the accuracy and precision 272 of the GLODAPv2 discrete data the reader is referred to Olsen et al., (2015), but note that 273 these uncertainties in the input data overall are smaller than the mapping errors. Overall the 274 spatial error distribution is as expected: relatively small where there are observations and 275 larger elsewhere. In regions without data the mapping errors may approach, and sometimes 276 exceed, the climatological value. In these regions the climatology cannot be trusted, and therefore all grid cells where the mapping error exceeds one standard deviation of the input 277 278 data on a given pressure surface have been masked (*i.e.* set to -999). This still leaves regions 279 with high mapping errors, so the relative error fields (i.e. error scaled to the standard deviation 280 in the data) are provided in the netcdf files making it possible for the user to create alternative 281 masks if needed. For  $TCO_2$  the average error in the masked data across all surfaces is in the range 14-27  $\mu$ mol kg<sup>-1</sup>; for TAlk in the range 8-29  $\mu$ mol kg<sup>-1</sup>; for pH at standard temperature 282 and pressure in the range 0.016-0.056; for pH at *in situ* temperature and pressure in the range 283 0.011-0.042; for  $\Omega_{\rm C}$  at *in situ* temperature and pressure in the range 0.029-0.46; for  $\Omega_{\rm A}$  at *in* 284 285 situ temperature and pressure in the range 0.020-0.31; for nitrate in the range 1.5-2.4 µmol kg<sup>-</sup> <sup>1</sup>; for phosphate in the range 0.10-0.17  $\mu$ mol kg<sup>-1</sup>; for silicate in the range 4-13  $\mu$ mol kg<sup>-1</sup>; for 286 oxygen in the range 10-19  $\mu$ mol kg<sup>-1</sup>; for salinity in the range 0.02-0.48; and for potential 287 temperature in the range 0.19-2.5 °C. The ranges reflect the variability in data density on 288





different surfaces and also the larger variability in surface ocean observations, because thisleads to larger background variance in the input data.

291 The TCO<sub>2</sub> and TAlk mapping errors for our GLODAPv2 climatology have been 292 compared with those of GLODAPv1.1 (Figures 11-12). Note, however, that here we compare 293 pressure surfaces with depth surfaces. The most obvious result is the large spatial variability 294 in the differences, which seem to correlate with the data distribution. Very generally, there are large differences (>10 µmol kg<sup>-1</sup>) in error estimates between the GLODAPv2 2000-2013 295 TCO<sub>2</sub> climatology and the GLODAPv1.1 climatology (Figure 11b), and between the 296 297 GLODAPv2 1986-1999 and GLODAPv1.1 TCO2 climatologies in the top 200 dbar of the Southern Ocean (exemplified by the 10 dbar surface in Figure 11a and 11b). In both cases the 298 error estimate in GLODAPv2 is frequently more than 15 µmol kg<sup>-1</sup> higher than in 299 GLODAPv1.1. For TCO<sub>2</sub> the 1986-1999 climatology has comparable error to GLODAPv1.1 300 in the Atlantic, but is smaller by 10-15 µmol kg<sup>-1</sup> in the Pacific (Figure 11a). For the 2000-301 2013 TCO<sub>2</sub> climatology the GLODAPv2 mapping errors are frequently larger than those in 302 303 GLODAPv1.1, but here also smaller in the Pacific Ocean (Figure 11b). Below 1000 dbar 304 (exemplified by the 3000 dbar surface in Figure 11c) the mapping errors are overall larger by 5-10 µmol kg<sup>-1</sup> in GLODAPv2 than in GLODAPv1.1. For TAlk the GLODAPv2 mapping 305 306 errors exceed those of GLODAPv1.1 in the Southern Ocean in the top 200 dbar (exemplified 307 by the 10 dbar surface in Figure 12a), but otherwise for the top 1000 dbar errors in the two 308 products are comparable (not shown). Below 1000 dbar the GLODAPv2 TAlk errors are typically 5-10 µmol kg<sup>-1</sup> larger than those of GLODAPv1.1 (exemplified by the 3000 dbar 309 310 surface in Figure 12b).

311 A scientific study of the differences in mapping error between GLODAPv1.1 and 312 GLODAPv2 and the mechanisms behind these would be worthwhile, and could perhaps 313 improve future climatologies of the marine  $CO_2$  chemistry, but is beyond the scope of this 314 paper. The reasons for the differences seen in Figures 11-12 are currently not clear, but 315 several things are likely to contribute: (i) differences in the methods used; (ii) in general 316 GLODAPv2 uses a smaller CL and this results in larger errors, but for TCO<sub>2</sub> and pH in the 317 top 1000 dbar the CL is larger and this explains the smaller errors on these surfaces; (iii) there 318 are differences in data density and data distribution between the two versions, and the 319 improved distribution in GLODAPv2 leads to larger natural variability and thus larger, more 320 realistic, errors.





321	For the macronutrients (nitrate, phosphate, silicate) the GLODAPv2 climatologies can
322	be compared to the World Ocean Atlas (WOA) nutrient climatologies, but before doing so
323	several things need to be considered: (i) methods used for mapping WOA (Garcia et al., 2010)
324	are very different from those used in GLODAPv2; (ii) WOA also does not provide mapped
325	error estimates for their climatologies; (iii) WOA reports nutrients with units $\mu$ mol L <sup>-1</sup> while
326	GLODAPv2 uses units $\mu$ mol kg <sup>-1</sup> . Given the lack of error fields in WOA a direct comparison
327	of errors like we did for TCO <sub>2</sub> and TAlk cannot be performed for the nutrients. Instead we
328	have looked at the differences between the nitrate climatology in GLODAPv2 and in WOA09
329	(for the purpose of this comparison roughly converted the $\mu mol \; L^{\text{-1}}$ data in WOA to $\mu mol \; kg^{\text{-1}}$
330	by dividing by 1.024). These differences will be due to a combination of the differences in
331	input data and the differences in mapping methods, and also here we compare pressure
332	surfaces with depth surfaces. When comparing the GLODAPv2 gridded observations with the
333	WOA09 climatology we see certain patterns (Figure 13). Near the surface the GLODAPv2
334	observations are overall smaller than the WOA09 climatology in high latitudes (Figure 13a),
335	and this is most likely a manifestation of the seasonal bias in GLODAPv2 which in these
336	regions contains almost only summertime data. In the tropics and subtropics the differences
337	are within the data uncertainties. In the deep ocean (Figure 13b), however, the differences
338	between the GLODAPv2 observations and the WOA09 climatology are very similar to the
339	differences between the GLODAPv2 observations and climatology (Figure 7b). This suggests
340	that below the seasonally influenced surfaces the differences between GLODAPv2 and
341	WOA09 stem mainly from differences in mapping method, but that the climatologies
342	otherwise are comparable. The biggest difference between GLODAPv2 and WOA09 is that
343	the latter has considerably more input data, and is thus able to provide monthly climatologies
344	which GLODAPv2 cannot. Note that we have compared the GLODAPv2 nitrate climatology
345	to WOA09 since the WOA13 has a very different vertical resolution.

346

# 347 4 Best practices for using the GLODAPv2 1°x1° data fields

For the marine CO<sub>2</sub> chemistry parameters with known, large, temporal trends two climatologies are provided for the surface ocean (<1000 dbar). This division was implemented to reduce the risk of converting time trends into spatial variations in the climatology. Alternative and more sophisticated approaches certainly exist, and these will be considered for future versions; here, however, we choose as simple and transparent an

approach as possible. The 1986-1999 climatology is centered on the early 1990s and the

11





- 354 2000-2013 climatology on the mid-2000s. The difference between the climatologies for the 355 two time periods in no way represents an estimate of decadal change in global ocean CO2. The errors in each TCO<sub>2</sub> climatology, mainly a consequence of limited spatial coverage of the 356 input data, approach 100 µmol kg<sup>-1</sup> in some regions, which is much larger than any expected 357 358 trends. Additionally, each climatology was created using data from more than ten years, hence 359 some fraction of our spatial features is a consequence of time trends in each of these periods. 360 Users interested in time trends are better served by evaluating differences between repeat 361 sections in the data product.
- 362 Planned future work includes creating mapped climatologies of several additional 363 parameters available in the GLODAPv2 data product: water ages based on the halogenated 364 transient tracer data and the <sup>14</sup>C data. As estimates of the anthropogenic CO<sub>2</sub> content based on 365 the GLODAPv2 product become available we will consider creating new climatologies for 366 TCO<sub>2</sub> and pH, where the anthropogenic trend in the data has been removed.
- 367

#### 368 Acknowledgements

369 The work of S. K. Lauvset was funded by the Norwegian Research Council through the 370 projects DECApH (214513/F20). The EU-IP CARBOCHANGE (FP7 264878) provided 371 funding for A. Olsen, S. van Heuven, T. Tanhua, R. Steinfeldt, as well as the project 372 framework that instigated GLODAPv2. A. Olsen additionally acknowledges generous support 373 from the FRAM - High North Research Centre for Climate and the Environment, the Centre 374 for Climate Dynamics at the Bjerknes Centre for Climate Research, the EU AtlantOS (grant 375 agreement No 633211) project and the Norwegian Research Council project SNACS 376 (229752). A. Kozyr acknowledges funding from the US Department of Energy. M. Ishii 377 acknowledges the project MEXT 24121003. A. Velo and F. F. Pérez were supported by 378 BOCATS (CTM20134410484P) project cofounded by the Spanish Government and the 379 Fondo Europeo de Desarrollo Regional (FEDER). The International Ocean Carbon 380 Coordination Project (IOCCP) partially supported this activity through the U.S. National 381 Science Foundation grant (OCE-1243377) to the Scientific Committee on Oceanic Research. 382 "The research leading to the last developments of DIVA has received funding from the 383 European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 283607, SeaDataNet 2, and from the project EMODNET (MARE/2012/10 - Lot 4 384 Chemistry - SI2.656742) from the Directorate-General for Maritime Affairs and Fisheries." 385

386





# 387 References

388	Bakker, D. C. E., Pfeil, B., Smith, K., Hankin, S., Olsen, A., Alin, S. R., Cosca, C., Harasawa,
389	S., Kozyr, A., Nojiri, Y., O'Brien, K. M., Schuster, U., Telszewski, M., Tilbrook, B.,
390	Wada, C., Akl, J., Barbero, L., Bates, N. R., Boutin, J., Bozec, Y., Cai, W. J., Castle,
391	R. D., Chavez, F. P., Chen, L., Chierici, M., Currie, K., de Baar, H. J. W., Evans, W.,
392	Feely, R. A., Fransson, A., Gao, Z., Hales, B., Hardman-Mountford, N. J., Hoppema,
393	M., Huang, W. J., Hunt, C. W., Huss, B., Ichikawa, T., Johannessen, T., Jones, E. M.,
394	Jones, S. D., Jutterström, S., Kitidis, V., Körtzinger, A., Landschützer, P., Lauvset, S.
395	K., Lefèvre, N., Manke, A. B., Mathis, J. T., Merlivat, L., Metzl, N., Murata, A.,
396	Newberger, T., Omar, A. M., Ono, T., Park, G. H., Paterson, K., Pierrot, D., Ríos, A.
397	F., Sabine, C. L., Saito, S., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Sieger, R.,
398	Skielvan, I., Steinhoff, T., Sullivan, K. F., Sun, H., Sutton, A. J., Suzuki, T., Sweenev,
399	C., Takahashi, T., Tiiputra, J., Tsurushima, N., van Heuven, S. M. A. C., Vandemark,
400	D Vlahos P Wallace D W R Wanninkhof R and Watson A I: An undate to
401	the Surface Ocean CO <sub>2</sub> Atlas (SOCAT version 2) Earth Syst. Sci. Data 6 69-90
402	2014
403	Beckers I-M Barth A Troupin C and Alvera-Azcárate A Approximate and Efficient
404	Methods to Assess Error Fields in Spatial Gridding with Data Interpolating Variational
405	Analysis (DIVA) Journal of Atmospheric and Oceanic Technology 31, 515-530
406	2014
407	Bonn I. Resplandy I. Orr I.C. Doney S.C. Dunne I.P. Gehlen M. Halloran P.
408	Heinze C Ilvina T Seferian R Tiiputra I and Vichi M Multiple stressors of
409	ocean ecosystems in the 21st century: projections with CMIP5 models
410	Biogensciences 10, 6225-6245, 2013
411	Dickson A : Standard potential of the reaction: $AgCl(s) + 1/2H_0(g) = Ag(s) + HCl(ag)$ and
412	the standard acidity constant of the ion HSO <sub>4</sub> in synthetic sea water from 273 15 to
413	318 15 K I Chem Thermodyn 22 113–127 1990
414	Feely R Talley I Bullister I I Carlson C A Doney S Fine R A Firing F Gruber
415	N Hansell D A Johnson G C Key R Langdon C Macdonald A Mathis I
416	Mecking S Millero F I Mordy C Sabine C Smethie W M Swift I H
417	Thurnherr A M Wanninkhof R and Warner M : The US Repeat Hydrography
418	CO2/Tracer Program (GO-SHIP): Accomplishments from the first decadal survey US
419	CLIVAR and OCB Report 2014-5 US CLIVAR Project Office 47 pp 2014
420	Garcia H E Locarnini R A Bover T P Antonov L I Zweng M M Baranova O K
421	and Johnson D R : World Ocean Atlas 2009 Volume 4: Nutrients (phosphate nitrate
121	silicate) S. Levitus Ed. NOAA Atlas NESDIS 71. U.S. Government Printing Office
423	Washington D C 398 np np 2010
424	Gouretski V V and Jancke K : Systematic errors as the cause for an annarent deen water
425	property variability: global analysis of the WOCE and historical hydrographic data
425	Progress in Oceanography 48, 337-402, 2001
420	Johnson G. C. Robbins P. E. and Hufford G. E. Systematic adjustments of hydrographic
127	sections for internal consistency. Journal of Atmospheric and Oceanic Technology, 18
420	1234 1244 2001
429	1254-1244, 2001. Jonas S. D. La Quéré, C. and Bödenback, C.: Autocorrelation characteristics of surface
430	Jones, S. D., Le Quere, C., and Rouenbeck, C.: Autocorrelation characteristics of surface
437	Kay P. M. Kozyr A. Sabina C. I. Las K. Wanninkhof D. Dullister, I. L. Essly, D. A.
432	Millero E. I. Mordy C. and Dong T. H.: A global according to primetal any Desults
433	from Clobal Data Analysis Project (CLODAD). Clobal Discrease homical Cycles 19
434	CP4021 doi:10.1020/2004CP002247.2004
433	004051, 001:10.1029/200400002247, 2004.

Earth System Discussion Science usions



436	Key, R. M., Tanhua, T., Olsen, A., Hoppema, M., Jutterström, S., Schirnick, C., Van Heuven,
437	S., Kozyr, A., Lin, X., Velo, A., Wallace, D. W. R., and Mintrop, L.: The CARINA
438	data synthesis project: introduction and overview, Earth Syst. Sci. Data, 2, 105-121,
439	2010.
440	Lauvset, S. K., Gruber, N., Landschützer, P., Olsen, A., and Tjiputra, J.: Trends and drivers in
441	global surface ocean pH over the past 3 decades, Biogeosciences, 12, 1285-1298,
442	2015.
443	Levitus, S., and Bover, T.P.: World Ocean Atlas, vol. 4, Temperature, NOAA Atlas NESDIS
444	4. National Oceanic and Atmospheric Administration, Silver Spring, Md., 1994.
445	Lewis, E. and Wallace, D. W. R.: Program developed for CO2 system calculations. In:
446	ORNL/CDIAC-105. Carbon Dioxide Information Analysis Center. Oak Ridge
447	National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, 1998.
448	Lueker, T. L. Dickson, A. G. and Keeling, C. D.: Ocean pCO(2) calculated from dissolved
449	inorganic carbon, alkalinity, and equations for K-1 and K-2: validation based on
450	laboratory measurements of CO <sub>2</sub> in gas and seawater at equilibrium Marine
451	Chemistry 70 105-119 2000
452	Olsen A Key R M Lauyset S K Van Heuven S Lin X Schirnick C Tanhua T
453	Honnema M Jutterström S Steinfeldt R Jeansson F Jshij M Suzuki T Velo
454	A Pfeil B and Kozyr A : Building a fully consistent carbon-relevant data product
455	for the world ocean: the Global Ocean Data Project version 2 (GLODAPv2)
456	submitted to: Earth Syst. Sci. Data Discuss. 2015
457	Orr I C Maier-Reimer F Mikolaiewicz II Monfray P Sarmiento I I. Togoweiler I
458	R Taylor N K Palmer I Gruber N Sabine C I Le Quéré C Key R M and
459	Boutin L: Estimates of anthronogenic carbon untake from four three-dimensional
460	global ocean models. Global Biogeochemical Cycles 15, 43-60, 2001
461	Pfeil B Olsen A Bakker D C F Hankin S Koyuk H Kozyr A Malczyk I Manke
462	A Metzl N Sahine C L Akl I Alin S R Bates N Bellerby R G I Borges
463	A Boutin I Brown P I Cai W I Chavez F P Chen A Cosca C Fassbender
464	A I Feelv R A González-Dávila M Govet C Hales B Hardman-Mountford
465	N Heinze C Hood M Hoppema M Hunt C W Hydes D Ishii M
466	Johannessen T. Jones S. D. Key R. M. Körtzinger A. Landschützer P. Lauvset
467	S. K., Lefèvre, N., Lenton, A., Lourantou, A., Merlivat, L., Midorikawa, T., Mintrop.
468	L., Miyazaki, C., Murata, A., Nakadate, A., Nakano, Y., Nakaoka, S., Nojiri, Y.,
469	Omar, A. M., Padin, X. A., Park, G. H., Paterson, K., Perez, F. F., Pierrot, D., Poisson,
470	A Ríos A F Santana-Casiano I M Salisbury I Sarma V V S S Schlitzer R
471	Schneider, B., Schuster, U., Sieger, R., Skielvan, L. Steinhoff, T., Suzuki, T.,
472	Takahashi, T., Tedesco, K., Telszewski, M., Thomas, H., Tilbrook, B., Tiiputra, J.,
473	Vandemark, D., Veness, T., Wanninkhof, R., Watson, A. J., Weiss, R., Wong, C. S.,
474	and Yoshikawa-Inoue, H.: A uniform, quality controlled Surface Ocean CO2 Atlas
475	(SOCAT). Earth Syst. Sci. Data. 5, 125-143, 2013.
476	Rixen, M., Beckers, J. M., Brankart, J. M., and Brasseur, P.: A numerically efficient data
477	analysis method with error map generation. Ocean Model., 2, 45-60, 2000.
478	Sabine, C. L., Feelv, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R.,
479	Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, TH., Kozvr, A.,
480	Ono, T., and Ríos, A. F.: The oceanic sink for anthropogenic CO <sub>2</sub> , Science, 305, 367-
481	371, 2004.
482	Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C. E., Pfeil, B., Olsen, A., Metzl, N., Kozvr,
483	A., Fassbender, A., Manke, A., Malczyk, J., Akl, J., Alin, S. R., Bellerby, R. G. J.,
484	Borges, A., Boutin, J., Brown, P. J., Cai, W. J., Chavez, F. P., Chen, A., Cosca. C.,
485	Feely, R. A., González-Dávila, M., Goyet, C., Hardman-Mountford, N., Heinze, C.,





486	Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., Key, R. M.,
487	Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lourantou,
488	A., Merlivat, L., Midorikawa, T., Mintrop, L., Miyazaki, C., Murata, A., Nakadate, A.,
489	Nakano, Y., Nakaoka, S., Nojiri, Y., Omar, A. M., Padin, X. A., Park, G. H., Paterson,
490	K., Perez, F. F., Pierrot, D., Poisson, A., Ríos, A. F., Salisbury, J., Santana-Casiano, J.
491	M., Sarma, V. V. S. S., Schlitzer, R., Schneider, B., Schuster, U., Sieger, R., Skjelvan,
492	I., Steinhoff, T., Suzuki, T., Takahashi, T., Tedesco, K., Telszewski, M., Thomas, H.,
493	Tilbrook, B., Vandemark, D., Veness, T., Watson, A. J., Weiss, R., Wong, C. S., and
494	Yoshikawa-Inoue, H.: Surface Ocean CO2 Atlas (SOCAT) gridded data products,
495	Earth Syst. Sci. Data, 5, 145-153, 2013.
496	Sabine, C. L. and Tanhua, T.: Estimation of Anthropogenic CO2 Inventories in the Ocean,
497	Annual Review of Marine Science, 2, 175-198, 2010.
498	Suzuki, T., Ishii, M., Aoyama, M., Christian, J. R., Enyo, K., Kawano, T., Key, R. M.,
499	Kosugi, N., Kozvr, A., Miller, L., Murata, A., Nakano, T., Ono, T., Saino, T., Sasaki,
500	KI., Sasano, D., Takatani, Y., Wakita, M., and Sabine, C.: PACIFICA Data Synthesis
501	Project, Carbon Dioxide Information Analysis Center, Oak Ridge National
502	Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, 2013.
503	Talley, L. D., Feely, R. A., Sloyan, B. M., Wanninkhof, R., Baringer, M. O., Bullister, J. L.,
504	Carlson, C. A., Doney, S. C., Fine, R. A., Firing, E., Gruber, N., Hansell, D. A., Ishii,
505	M., Johnson, G. C., Katsumata, K., Key, R. M., Kramp, M., Langdon, C., Macdonald,
506	A. M., Mathis, J. T., McDonagh, E. L., Mecking, S., Millero, F. J., Mordy, C. W.,
507	Nakano, T., Sabine, C. L., Smethie, W. M., Swift, J. H., Tanhua, T., Thurnherr, A. M.,
508	Warner, M. J., and Zhang, JZ.: Changes in Ocean Heat, Carbon Content, and
509	Ventilation: A Review of the First Decade of GO-SHIP Global Repeat Hydrography,
510	Annual Review of Marine Science, 8, null, doi:10.1146/annurev-marine-052915-
511	100829, 2016.
512	Touratier, F., Azouzi, L., and Goyet, C.: CFC-11, Delta C-14 and H-3 tracers as a means to
513	assess anthropogenic CO2 concentrations in the ocean, Tellus Series B-Chemical and
514	Physical Meteorology, 59, 318-325, 2007.
515	Troupin, C., Barth, A., Sirjacobs, D., Ouberdous, M., Brankart, J. M., Brasseur, P., Rixen, M.,
516	Alvera-Azcárate, A., Belounis, M., Capet, A., Lenartz, F., Toussaint, M. E., and
517	Beckers, J. M.: Generation of analysis and consistent error fields using the Data
518	Interpolating Variational Analysis (DIVA), Ocean Model., 52–53, 90-101, 2012.
519	Uppstrom, L. R.: Boron/chlorinity ratio of deep-sea water from the Pacific Ocean, Deep-Sea
520	Research, 21, 161-162, 1974.
521	van Heuven, S., Pierrot, D., Lewis, E., and Wallace, D.: MATLAB Program developed for
522	CO <sub>2</sub> system calculations, ORNL/CDIAC-105b, Carbon Dioxide Information Analysis
523	Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge,
524	Tennessee, 2009.
505	

### 525

# 526 Tables

# 527 Table 1. Maximum distance criteria used when vertically interpolating the input data.

Range (dbar)	Maximum distance allowed
0-200.99	100
201-750.99	200
751-1500.99	250
1501-12000	500





Variable name	Description
lon	Longitude in degrees east, range -180 – 180
lat	Latitude in degrees north, range -90 – 90
tco2, talk, pH, $\Omega_C$ , $\Omega_A$ , nitrate, phosphate, silicate, oxygen, salinity, theta	Mapped climatology with land mask and $3\sigma$ mask applied.
_error	Mapping error associated with the mapped climatology
_relerr	Mapping error scaled with the global standard deviation of the input data
Input_mean	Binned and averaged mean of the observations on the same grid as the climatology
Input_std	Standard deviation of the binned mean.
Input_N	Number of observational data points in the grid cell
Input_lon	Binned and averaged longitude in degrees east
Input_lat	Binned and averaged latitude in degrees north

### 529 Table 2. List of information available in the netcdf data files.

530

#### 531 Figures

- Figure 1. a) Data density of TCO<sub>2</sub> in the years 1986-1999; b) Data density of TCO<sub>2</sub> in the years 2000-2013.
  Both figures show data at the 10 dbar surface.
- 534Figure 2. Mapped climatology of TCO2 at 10 dbar (a, b) and 3000 dbar (c). a) is the climatology for the5351986-1999 period while b) is for the 2000-2013 period.
- 536 Figure 3. Mapped climatology of TAlk at 10 dbar (a) and 3000 dbar (b).
- 537 Figure 4. Mapped climatology of nitrate at 10 dbar (a) and 3000 dbar (b).
- 538 Figure 5. Difference between the gridded TCO<sub>2</sub> input data and the mapped climatologies atv10 dbar (a,b) 539 and 3000 dbar (c). a) is the climatology for the 1986-1999 period while b) is for the 2000-2013 period.
- 540 Figure 6. Difference between the gridded TAlk input data and the mapped climatologies at 10 dbar (a) 541 and 3000 dbar (b).
- 542 Figure 7. Difference between the gridded nitrate input data and the mapped climatologies at 10 dbar (a) 543 and 3000 dbar (b).
- 544 Figure 8. Mapping error for TCO<sub>2</sub> at 10 dbar (a, b) and 3000 dbar (c). a) is the error for the 1986-1999
- 545 climatology while b) is for the 2000-2013 climatology. Notice how the error is large between repeat 546 transect and creates a spatial pattern of square-like features in the Pacific.
- 547 Figure 9. Mapping error for TAlk at 10 dbar (a) and 3000 dbar (b). Notice how the error is large between 548 repeat transect and creates a spatial pattern of square-like features in the Pacific.
- 549 Figure 10. Mapping error for nitrate at 10 dbar (a) and 3000 dbar (b). Notice how the error is large
- 550 between repeat transect and creates a spatial pattern of square-like features in the Pacific.
- 551 Figure 11. Difference in error estimates for TCO<sub>2</sub> between GLODAPv2 and GLODAPv1.1. a) compares
- 552 the 10 dbar surface from the 1986-1999 climatology in v2 with the 10 m surface in v1.1; b) compares the
- 553 10 dbar surface from the 2000-2013 climatology in v2 with the 10 m surface in v1.1; c) compares the 3000 554 dbar surface in v2 with the 3000 m surface in v1.1.





- 555 Figure 12. Difference in error estimates for TAlk between GLODAPv2 and GLODAPv1.1. a) compares
- 556 the 10 dbar surface in v2 with the 10 m surface in v1.1, while b) compares the 3000 dbar surface in v2 with 557 the 3000 m surface in v1.1.
- 558 Figure 13. Differences between the GLODAPv2 nitrate gridded input data and the WOA09 annual
- 559 mapped nitrate climatology. a) compares the 10 dbar surface in GLODAPv2 with the 10 m surface in
- 560 WOA09, while b) compares the 3000 dbar surface in GLODAPv2 with the 3000 m surface in WOA09.







Fig. 1







































































