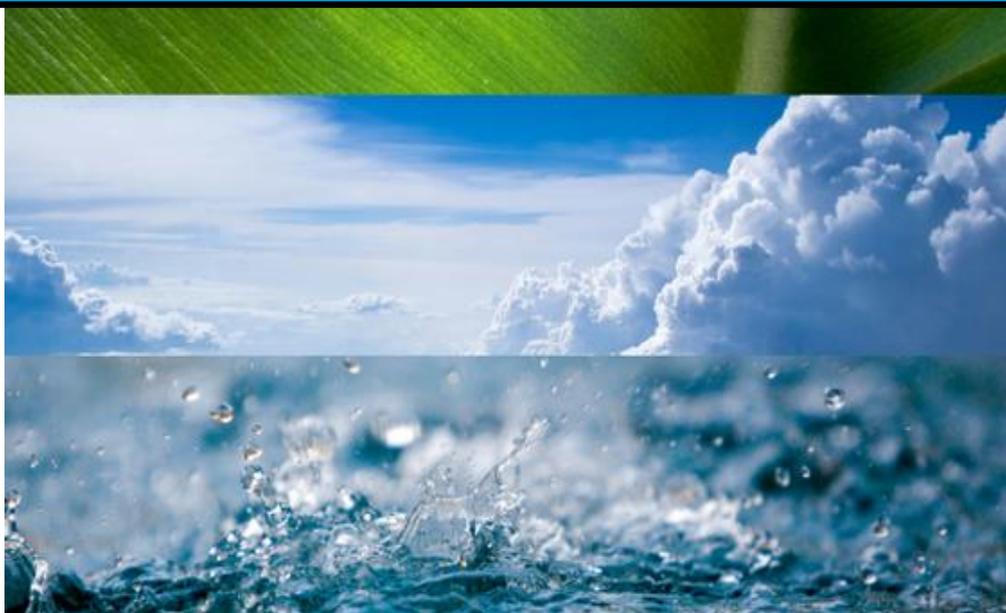




D5.2

GLODAP and SOCAT services fully implemented at ICOS OTC



Deliverable: 5.2. GLODAP and SOCAT services fully implemented at ICOS OTC

Author(s): Benjamin Pfeil, Jonas Henriksen

Date: 30.12.2019

Activity: Data management

Lead Partner: University of Bergen (UiB)

Document Issue:

RINGO (GA no 730944)

Public Document/define level

Dissemination Level:

Contact: Benjamin Pfeil

benjamin.pfeil@uib.no

	Name	Partner	Date
From	Benjamin Pfeil	UiB	Jan. 2 nd 2020
Reviewed by	Janne-Markus Rintala	ICOS ERIC HO	Jan. 21 st 2020
Approved by	Evi-Carita Riikonen	ICOS ERIC HO	17.2.2020
Version	Date	Comments/Changes	Author/Partner
Final Draft	Jan 2 nd	JMR to review	UiB
Review comments	Jan 21 st 2020	request for minor changes	ICOS ERIC
Final corrected version with track changes	Feb. 17 th 2020	Improved version received	UiB
Final Cleaned version approved	Feb. 18 th 2020		UiB

Deliverable Review Checklist

A list of checkpoints has been created to be ticked off by the Task Leader before finalizing the deliverable. These checkpoints are incorporated into the deliverable template where the Task Leader must tick off the list.

- Appearance is generally appealing and according to the RINGO template. Cover page has been updated according to the Deliverable details.
- The executive summary is provided giving a short and to the point description of the deliverable.
- All abbreviations are explained in a separate list.
- All references are listed in a concise list.
- The deliverable clearly identifies all contributions from partners and justifies the resources used.
- A full spell check has been executed and is completed.

DISCLAIMER

This document has been produced in the context of the *project* Readiness of ICOS for Necessities of integrated Global Observations (RINGO)

The Research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730944. All Information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors view.

Amendments, comments and suggestions should be sent to the authors.

ABSTRACT

Climate Change and the related cycles of carbon and GHG are global challenges and cannot be solved with a regional European approach only. ICOS is therefore strongly involved in the development of a global observation system on carbon and GHG. ICOS is in close cooperation with respective data infrastructures in the field and the main objectives within this deliverable were to integrate and develop ICOS OTC towards the European pillar of Global Data Analysis Project (GLODAP) and the Surface Ocean Carbon Dioxide Atlas (SOCAT). Both data products handle data from the Essential Ocean Variable Inorganic Carbon while GLODAP handles discrete (bottle measurements) EOVI Inorganic Carbon from the ocean interior plus other biogeochemistry data from the Global Ocean Observing System (GOOS) GO-SHIP focusing on repeat hydrographic surveys and some data from the GOOS OceanSites focusing on off-shore time series station data. SOCAT focusses on continuous EOVI Inorganic Carbon observations mainly carbon dioxide ($p\text{CO}_2$, $f\text{CO}_2$ and $x\text{CO}_2$) in the surface water (around 5 m depth). While data relevant for SOCAT are the primary focus of the RI ICOS – it was planned to integrate repeat hydrography data relevant for GLODAP into the ICOS portfolio. This has not been achieved yet since those research cruise based surveys are usually performed every few years by (sometimes different) countries, the entire data flow of EOVI Inorganic Carbon data is in delayed mode and it was challenging to introduce the idea of annual station contribution fees for sections that are surveyed not on an annual basis. Nevertheless, those data are of great importance for understanding the marine carbon cycles and the main task of partner UiB in RINGO was to develop the automation of GLODAP by streamlining the entire workflow and making them ready for ICOS. In addition, ICOS OTC partner UiB assisted in data management activities for producing two releases of GLODAP plus three releases of SOCAT. Currently all of the above data from SOCAT and GLODAP is handled at one point by OTC partner UiB making ICOS OTC the European key player and elementary part of these data products. In addition the respective websites for GLODAP (www.glodap.info) and SOCAT (www.socat.info) were established, updated, are being maintained and hosted by ICOS OTC partner UiB.

Contents

1 Introduction	4
2 Surface Ocean Carbon Dioxide Atlas -SOCAT	4
2.1 Global Ocean Data Analysis Project	5
3 Automation of GLODAP	6
3.1 Objectives	7
3.2 System setup	7
3.3 Data upload	7
3.4 Crossovers	8
4 International collaboration	11
5 Websites for GLODAP and SOCAT	11
6 Conclusions	12
7 Acronyms	12
8 References	12

1 Introduction

Sustainability of an RI strongly depends on the integration into existing global data networks in its field. ICOS aims to go far beyond integrating its data into global data streams: ICOS wants to be the European operational pillars of running these integrations. Building on strong involvement in the past all three Thematic Centers will develop or further implement services for the respective global data integration efforts. Since these have different states the resources required in this task vary between the Thematic Centers.

The data products Global Data Analysis Project (GLODAP) and the Surface Ocean Carbon Dioxide Atlas (SOCAT) provide public access to quality controlled, harmonised and integrated global collections of EOV Inorganic Carbon data from the interior and surface ocean. While SOCAT has evolved into a highly automated community effort based on Live Access Server (LAS) on-line documentation, versioning, and QC tools, GLODAP, whose version 2 was released in January 2016, lacks a centralised automated system for data ingestion, documentation and QC. All of these activities are to a large extent prepared and handled on a per scientist basis. This threatens reproducibility and homogeneity of QC results and makes information handling tedious, and non sustainable. Within RINGO we proposed to prepare an online system for data ingestion, versioning, quality control and the merging of individual data files and QC results to actual products, similar to the SOCAT LAS and automation system. This will enable GLODAP to evolve into a routine interior ocean synthesis, with new versions released every year. The automated system can also handle data (including generic quality control) that is currently not handled systematically within GLODAP or SOCAT e.g. data from discrete sampled coastal data from research vessels and fixed ocean time series stations. As mentioned above partner UiB has also supported data management activities for past versions of SOCAT. The developments for GLODAP and for SOCAT made ICOS visible to the international marine biogeochemistry and there is no doubt that the data management infrastructure of ICOS OTC is a global and the main European key player for these. There is no doubt that the objective of the deliverable has been achieved to make ICOS the European pillar of GLODAP and SOCAT by providing core services for both data networks by providing core services, web presentation and global coordination.

2 Surface Ocean Carbon Dioxide Atlas -SOCAT

The Surface Ocean CO₂ Atlas (SOCAT, www.socat.info) is a synthesis activity by the international marine carbon scientists (>100 contributors) with annual public releases. The observational data product of SOCATv2019 (SOCAT version 7) has 25.7 million quality-controlled surface ocean fCO₂ (fugacity of carbon dioxide) observations from 1957 to 2018 for the global oceans and coastal seas, as well as an additional 1.7 million fCO₂ observations of calibrated sensor data. SOCAT enables quantification of the ocean carbon sink, ocean acidification, and evaluation of ocean biogeochemical models in a changing world and represents a milestone in biogeochemical and climate research and in informing policy.

Within SOCAT regional groups perform secondary quality control. This quality control includes the evaluation and setup of the used system and if Standard Operating Procedures (SoPs) as set by the biogeochemistry community are followed (Dickson et al., 2007). A set of property-property plots are used to evaluate the data itself but those give important information about the system as well. The aim of SOCAT is to provide high quality data product and data with a higher uncertainty than 5 μatm is excluded from the basic product but certain data (e.g. fCO₂ from sensor data can be retrieved using the SOCAT Live Access Server). More detailed information can be found in the SOCAT documentation by Bakker et al., 2016.

As introduced above, there are various groups responsible for the global coordination, regional quality control and data management activities – more information on the various tasks can be found in the SOCAT related scientific publications (Pfeil et al., 2013, Bakker et al., 2016). The SOCAT data management is a shared effort of the data integration group of National Oceanographic Atmospheric Administration's Pacific Marine Environmental Laboratory (NOAA PMEL) in Seattle, USA, National

Oceanographic Atmospheric Administration National Centre for Environmental Information (NOAA NCEI) in Silver Spring, USA and the Bjerknes Climate Data Centre at the University of Bergen in Bergen, Norway. RINGO funds were used to support European data management activities such as data ingestion for ICOS data, metadata handling, long-term archival according to the FAIR data management principles (Wilkinson et al., 2016) – this ensured the continuity of SOCAT, international visibility for ICOS and ensured that SOCAT procedures based upon community agreed SoPs have been implemented for marine ICOS stations.

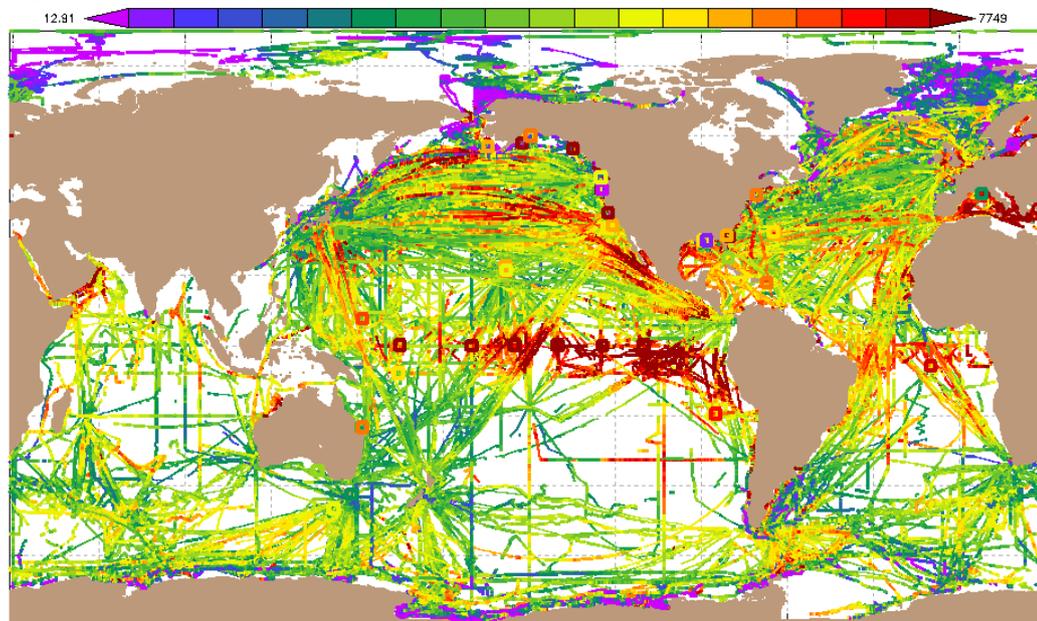


Figure 1: Spatial distribution of surface measurements of the SOCATv6 data product for the period 1957-2017

2.1 Global Ocean Data Analysis Project

The Global Ocean Data Analysis Project (GLODAP, www.glodap.info) is a synthesis project bringing together oceanographic biogeochemistry and tracer data from the open ocean mainly from the GO-SHIP Repeat Hydrography Program and other related cruises. Data from a few coastal surveys and time series (OceanSites) have been included on an opportunistic basis.

It also includes data from some marine time-series stations from the OceanSites program. The central goal of GLODAP is to generate a global climatology of the World Ocean's carbon cycle for use in studies of both its natural and anthropogenically-forced states.

The first version of GLODAP was published in 2004 and consisted of data collected during the 1990s by research cruises on the World Ocean Circulation Experiment, Joint Global Ocean Flux Study and Ocean-Atmosphere Exchange Study programmes. Regional efforts for updating GLODAP were performed in the Atlantic and Southern Ocean with the data product CARINA and PACIFICA for the Pacific Ocean.

The second version of GLODAPv2 published in 2016 includes all data in GLODAPv1.1, CARINA, and PACIFICA, as well as data from 168 new cruises. The new data originate both from recent cruises, completed after production of the previous data syntheses, and from less recent cruises for which the data have only become available recently. The new data were obtained by directly contacting principal investigators known to have carried out relevant cruises and by circulating a request letter to the ocean carbon science community through the IOCCP (International Ocean Carbon Coordination Project), as well as the SOLAS and IMBER core projects of the IGBP. Altogether, GLODAPv2 includes data from 724 cruises. Data from the surveys of WOCE/JGOFS, CLIVAR, and GO-SHIP form the backbone. In addition, data from large-scale surveys of the 1970s and 1980s – and from a multitude of national and regional programs have been included.

All individual cruise data files used for GLODAPv1.1, CARINA, and PACIFICA existed in the required WOCE exchange format and had been subjected to primary QC during the preparation of these products. All

of the new data were merged as necessary, converted to WOCE exchange format, and also subjected to primary QC. The primary QC was carried out following routines outlined in Sabine et al. (2005) and Tanhua et al. (2010), primarily by inspecting property– property plots. Outliers showing up in two or more different property–property plots were generally flagged as such.

GLODAPv2 was updated in 2019 (GLODAPv2.2019), it adds data from 116 cruises to the previous version, extending its coverage in time from 2013 to 2017, while also adding some data from prior years. GLODAPv2.2019 includes measurements from more than 1.1 million water samples from the global oceans collected on 840 cruises (Olsen et al., 2019). Figure 2 shows the regional distribution of the latest update.

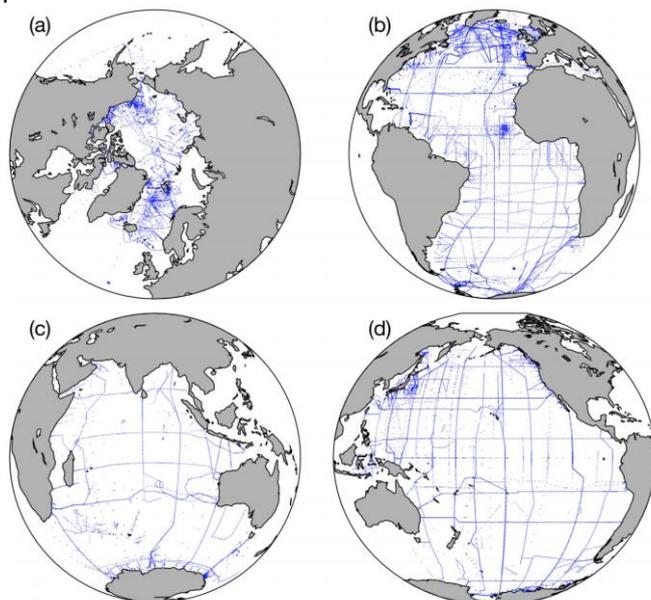


Figure 2: Geographical distribution of stations included in GLODAPv2.2019

3 Automation of GLODAP

As mentioned in the introduction one the main task for UiB in task 5.2 within RINGO was to develop an online system for data ingestion, versioning, quality control and the merging of individual data files and QC results to actual products for GLODAP comparable to what has been achieved to SOCAT in 2014. There are several advantages of automation since it ensures basic quality for incoming data, better manages data submission, version control, standardization of metadata and data formats, and non the less it eases the burden of individuals. Especially the later has been an issue partly for CARINA and PACIFICA but even more for GLODAPv2 and GLODAPv2.2019 since more data have been added, frequently updated and the previous workflow through a single scientist, Dr. Robert Key (Princeton University, Princeton, NJ, USA) that gathered, initially quality controlled, documented and reformatted incoming data in previous versions became a bottleneck.

The developed online automated data ingestion by ICOS OTC allows scientists to upload data and features all secondary quality control procedures used by the GLODAP group needed to do the manual cruise adjustments performed by the regional and variable specific experts.

The secondary quality control of the GLODAP data set can be done by finding crossovers where the track of the cruise you want to investigate, crosses the tracks of other previous quality controlled cruises. The following sub-sections give an overview about the features of the developed systems.

3.1 Objectives

The routines for doing the analysis existed in Matlab scripts, but one of the objectives was a need for avoiding proprietary software to make it more accessible for users using an online platform. The other objectives were:

- Creating a structure for handling reference data
- Allow users to register, and upload their own data
- Implement functionality to mimic the behaviour of the Matlab scripts
- Implement functionality to merge CTD-parameters with bottle samples
- Allow export of maps, plots and results after analysis.

3.2 System setup

The system is based on free, open source software. It is setup to run on an Ubuntu 18.04 Linux server, but should work on any server system without large modifications. Vagrant is used as a virtual server environment during development, ensuring similar conditions as in the live server environment.

Main tools, libraries and frameworks:

- Ubuntu 18.04
- Apache 2 web server
- Vagrant with Virtualbox for virtualization
- PostgreSQL 10 database
 - PostGIS spatial database extension
- Python 3
 - Django, python web framework
 - Pandas, python data handling
- Javascript ES6
 - Plotly, javascript plotting library
 - Openlayers, javascript mapping library
 - Datables, javascript data table library
 - jQuery javascript swiss army knife
- Git - version control

3.3 Data upload

User have to register and verify by email (see Figure 3).

GLODAP Crossover QC tools

Deep water cross-over analysis

- [Home](#)
- [Login](#)

Login

Username:

Password:

Forgot your password? [Click here to reset it!](#)
Or you can [Sign up for an account!](#)

Figure 3: Screenshot of the import mask of the GLODAP automation

Key features of the file upload, review and import function:

- Uploaded files are not automatically imported into the database - files can be reviewed after upload
- Clicking "Import file" starts the import job that runs in the background
- Any import errors are reported

- Measured non-standard variables in the file are ignored, but reported
- Currently the system is accepting data in the WOCE exchange data format - defined data types can be found at the [GLODAP github library](https://github.com/BjerknesClimateDataCentre/glodap/blob/master/util/data_type_dict.py):
https://github.com/BjerknesClimateDataCentre/glodap/blob/master/util/data_type_dict.py

3.4 Crossovers

Setting of the plotting properties for crossovers

The user can choose preconditions for crossover calculations and plots. Those are based upon the minimum depth, crossover radius in meters, plot using Sigma 4 or pressure and the inclusion of only quality controlled data. The preconditions are used for both merging and crossover calculations.

Mapping crossovers

Figure 4 shows the initial maps: all stations from the current dataset in blue, matching crossover stations from selected crossover datasets are red and stations without measurements shown as empty circles. The buffer-rings around filled stations indicates crossover radius settings. All plots and current map views are downloadable as png image files.

Stations with CTDOXY sampling

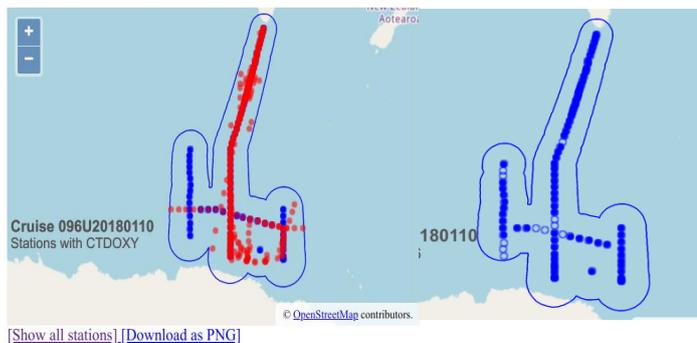


Figure 4: Example of the mapping crossover

Reviewing variables for a dataset (Figure 5)

- For data sets you can review variables for the whole data set
- Available crossover data sets with matching stations, are listed
- Variables used for calculating sigma4 are starred in the variable list
- If you need to merge some variables that are continuous observed by a sensor and discrete measured (eg salinity and oxygen) - you can click on merge data
- Cruise profiles from all crossover data sets are displayed for enabling visual evaluation of the respective data

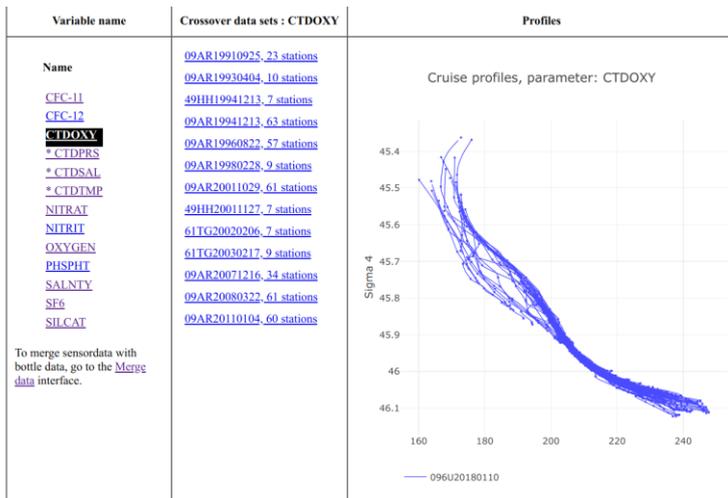


Figure 5: Reviewing variables for crossovers

Crossover profiles and weighted means (Figure 6)

- Profiles of selected parameters
- Area used for weighted mean calculations indicated in light blue
- Weighted mean, standard deviation and difference from total mean in a separate plot
- Zooming and panning interactivity
- Hover-functionality displays actual data values
- All plots can be exported to png images of the current view

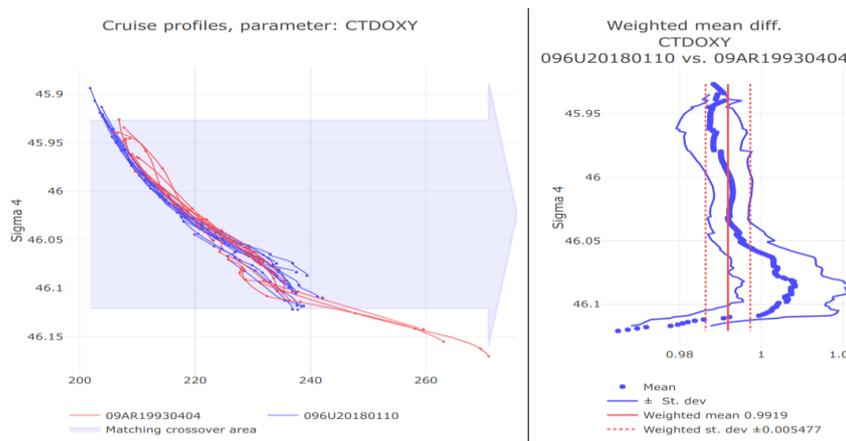


Figure 6: Crossover profiles and weighted means

Statistic summary plots (Figure 7)

- Displays matching crossover datasets
- Total weighted mean and offsets

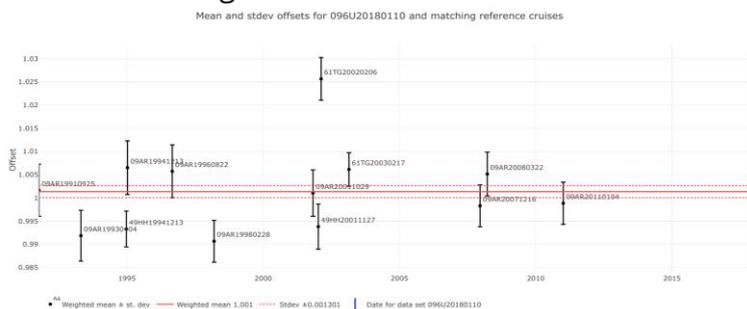


Figure 7: Statistical offsets plots for various cruises

Crossover results and statistical summary (Figure 8)

- Summary of crossover calculations
- Option to download table as excel
- Download includes all calculation preconditions:
 - Crossover range
 - Minimum depth
 - Whether only qc controlled data are used
 - Whether sigma 4 or pressure is used as depth metric

Export data to Excel Search:

Crossover results for 096U20180110: CTDOXY

Expocode	Startdate	Weighted mean offset	Weighted offset stdev (s)
09AR19910925	1991-10-08	1.0016468839277148	0.005618112802418104
09AR19930404	1993-04-24	0.9918552579476579	0.005476593905855006
49HH19941213	1994-12-24	0.9932985367517465	0.0038604698925105724
09AR19941213	1995-01-08	1.0065210853677724	0.005785034719919178
09AR19960822	1996-08-31	1.0057354388594197	0.005689623811188456
09AR19980228	1998-03-14	0.990647954142065	0.004508578617537601
09AR20011029	2001-10-30	1.0010479828456398	0.004996201833142784
49HH20011127	2002-01-09	0.9938039019335261	0.00485553703349696
61TG20020206	2002-02-16	1.0256667413398608	0.004604625393297004
61TG20030217	2003-02-24	1.006141736159501	0.003603344475697906
096U20180110	2018-01-11	Total weighted mean offset: 1.0013502358663489	Total weighted offset stdev: 0.0013008400602186867

Showing 1 to 10 of 13 entries Previous 2 Next

Figure 8: Crossover results

Parameter merging functionality (Figure 9)

- Select parameters to merge data from sensor data and discrete measurements e.g. oxygen or salinity
- Plot to compare / evaluate
- Select from the available merge strategies
 - Only secondary, no primary parameter
 - Only primary, no secondary parameter
 - Primary (few or no secondary or primary > 80%)
 - Mean of primary and UNCHANGED secondary
 - Mean of primary and fitted secondary
 - Primary because bad fit of secondary
- A merge adds a new parameter to the dataset, that can be used for other calculations

ID: 725 - 096U20180110

Primary **Secondary Name**

- CFC-11
- CFC-12
- CTDOXY
- * CTDPBS
- * CTDSAL
- * CTDTMP
- NITRAT
- NITRIT
- OXYGEN
- PHSPHI
- SALNTY
- SE6
- SILCAT

Minimum depth used to calculate linear fit:

Plot selected parameters

Merging

To merge the selected parameters, choose the appropriate action from the select box below, and click Save merge parameter. This will create a new variable in the database

Diff OXYGEN minus CTDOXY

OXYGEN vs CTDOXY; $y = 0.99771 * x + 0.66748$

Linear fit

Figure 9: Merging functionality

4 International collaboration

Similar to SOCAT data and information management activities are a shared effort between the data integration group of National Oceanographic Atmospheric Administration’s Pacific Marine Environmental Laboratory (NOAA PMEL) in Seattle, USA, National Oceanographic Atmospheric Administration National Centre for Environmental Information (NOAA NCEI) in Silver Spring, USA and the Bjerknes Climate Data Centre at the University of Bergen in Bergen, Norway.

The current system is a stand alone version accepting WOCE exchange format files. It is planned that the data ingestion group of NOAA PMEL which has been responsible for the automation of SOCAT will extend the current SOCAT ingestion to a shared SOCAT and GLODAP ingestion system. This will allow users to upload files using various file formats and benefit from already existing services within SOCAT. It will also integrate metadata handling and a basic automated first level quality control.

The previously described secondary quality control procedures using crossover analysis was designed that it can be integrated into the to-be-developed framework avoiding any duplicative efforts.

In addition, RINGO and ICOS OTC funds were used to organize an international workshop for Quality Control for Biogeochemical Data procedures in September 2018 in Seattle, WA, USA where scientists and data managers of the SOCAT, GLODAP, ICOS, NOAA Ocean Acidification Program, US Biogeochemistry ARGO networks participated. The objectives of the workshop were to provide the scientific guidance that can be applied by the ICOS OTC and NOAA Ocean Acidification Program developer groups towards an implementable solution, to determine what degree of level 1 and level 2 quality control is acceptable for which variables, and what can be implemented by developer groups, to determine what degree of QC automation is possible for both level 1 and level 2.

5 Websites for GLODAP and SOCAT

As introduced previously were the respective websites for GLODAP (www.glodap.info) and SOCAT (www.socat.info) were established, updated, are being maintained and hosted by ICOS OTC partner UiB since 2017. The content management system WordPress (www.WordPress.org) based on PHP and MySQL are used for both websites and credit is giving to the ICOS RI and the Bjerknes Climate Data Centre creating and hosting the websites.

Figure 10 shows the respective opening pages of the websites.



Figure 10: Screenshots of the SOCAT and GLODAP websites

6 Conclusions

The developments for GLODAP and for SOCAT made ICOS RI visible to the international scientific marine biogeochemistry and marine data management community. There is no doubt that the data management infrastructure of ICOS OTC is a global and the main European key player for these and there is no doubt that the objective of the deliverable has been achieved to make ICOS the European pillar of GLODAP and SOCAT by providing core services for both data networks. Even though repeat hydrographical data relevant for GLODAP is not yet a part of ICOS RI – the road and procedures are in place to integrate it at a later stage. In addition the new online GLODAP automation platform eases the workload of the global GLODAP group and provides the technical platform to produce annual efforts. This will have a positive feedback for the objectives of ICOS RI and the global scientific community.

7 Acronyms

CARINA – Carbon Dioxide in the Atlantic Ocean

CLIVAR - Climate Variability and Predictability

GO-SHIP – Global Ocean Ship-based Hydrographic Investigations Program

GLODAP – Global Data Analysis Product

IMBER - Integrated Marine Biosphere Research project

JGOFS – Joint Global Ocean Flux Study

IGBP - International Geosphere-Biosphere Programme

OceanSites - Worldwide system of long-term, deepwater reference stations measuring dozens of variables and monitoring the full depth of the ocean, from air-sea interactions down to 5,000 meters.

PACIFICA - PACIFIC ocean Interior CARbon)

SOCAT – Surface Ocean Carbon Dioxide Atlas

SOLAS- Surface Ocean - Lower Atmosphere Study

SOP – Standard Operating Procedures

WOCE- World Ocean Circulation Experiment

8 References

Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S.-I., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibáñez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., and Xu, S.: A multi-decade record of high-quality fCO₂ data in version 3 of the Surface Ocean CO₂ Atlas (SOCAT), *Earth Syst. Sci. Data*, 8, 383-413, <https://doi.org/10.5194/essd-8-383-2016>, 2016.

Dickson, A. G., Sabine, C. L., and Christian, J. R. (Eds.): Guide to best practices for ocean CO₂ measurements, PICES Special Publication 3, IOCCP Report 8, 191 pp., 2007.

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

Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnack, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Pérez, F. F., and Suzuki, T.: The Global Ocean Data Analysis Project version 2 (GLODAPv2) – an internally consistent data product for the world ocean, *Earth Syst. Sci. Data*, 8, 297-323, <https://doi.org/10.5194/essd-8-297-2016>, 2016.

Olsen, A., Lange, N., Key, R. M., Tanhua, T., Álvarez, M., Becker, S., Bittig, H. C., Carter, B. R., Cotrim da Cunha, L., Feely, R. A., van Heuven, S., Hoppema, M., Ishii, M., Jeansson, E., Jones, S. D., Jutterström, S., Karlsen, M. K., Kozyr, A., Lauvset, S. K., Lo Monaco, C., Murata, A., Pérez, F. F., Pfeil, B., Schirnack, C., Steinfeldt, R., Suzuki, T., Telszewski, M., Tilbrook, B., Velo, A., and Wanninkhof, R.: GLODAPv2.2019 – an update of GLODAPv2, *Earth Syst. Sci. Data*, 11, 1437–1461, <https://doi.org/10.5194/essd-11-1437-2019>, 2019.

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 CO₂ Atlas (SOCAT), *Earth Syst. Sci. Data*, 5, 125-143, <https://doi.org/10.5194/essd-5-125-2013>, 2013.

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

Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C. E., Pfeil, B., Olsen, A., Metzl, N., Kozyr, A., Fassbender, A., Manke, A., Malczyk, J., Akl, J., Alin, S. R., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W.-J., Chavez, F. P., Chen, A., Cosca, C., Feely, R. A., González-Dávila, M., Goyet, C., Hardman-Mountford, N., Heinze, C., Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., 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., Salisbury, J., Santana-Casiano, J. M., 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., Vandemark, D., Veness, T., Watson, A. J., Weiss, R., Wong, C. S., and Yoshikawa-Inoue, H.: Surface Ocean CO₂ Atlas (SOCAT) gridded data products, *Earth Syst. Sci. Data*, 5, 145-153, <https://doi.org/10.5194/essd-5-145-2013>, 2013.

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

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016) [doi:10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18), 2016