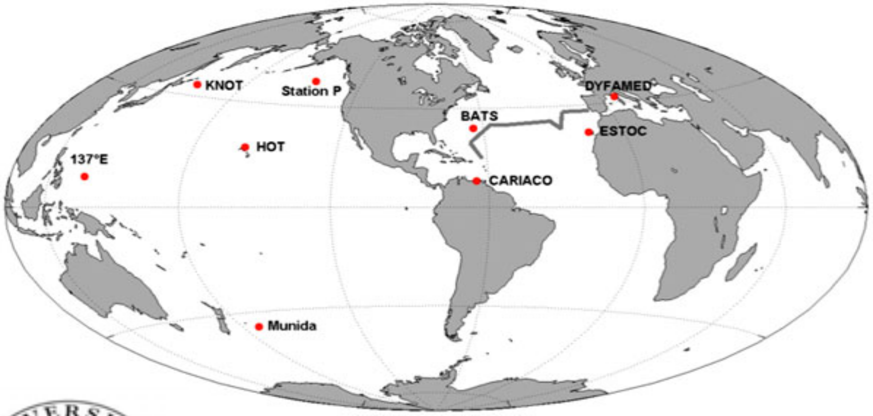
Carbon relevant ship-based time series station synthesis – A data product pilot



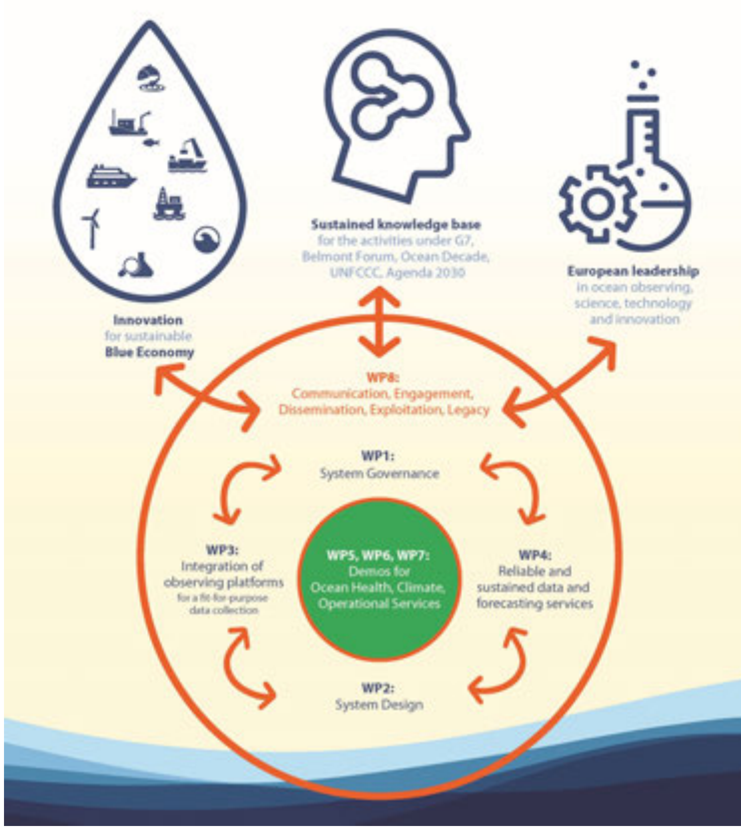
Nico Lange
University of Bergen
Nico.Lange@uib.no

With contributions from:
Björn Fiedler, Benjamin Pfeil, Toste Tanhua & Are Olsen

EuroSea

Improving and Integrating European Ocean Observing and Forecasting Systems for Sustainable use of the Oceans

- **Strengthen** the European and Global Ocean Observing System (EOOS and GOOS) and support its partners
- Increase the technology readiness levels (**TRL**) of the ocean observations systems and tools,
- **Improve** the: coordination, design, networks, data delivery, integration, and forecasting capability
- Work towards an **integrated** observing system

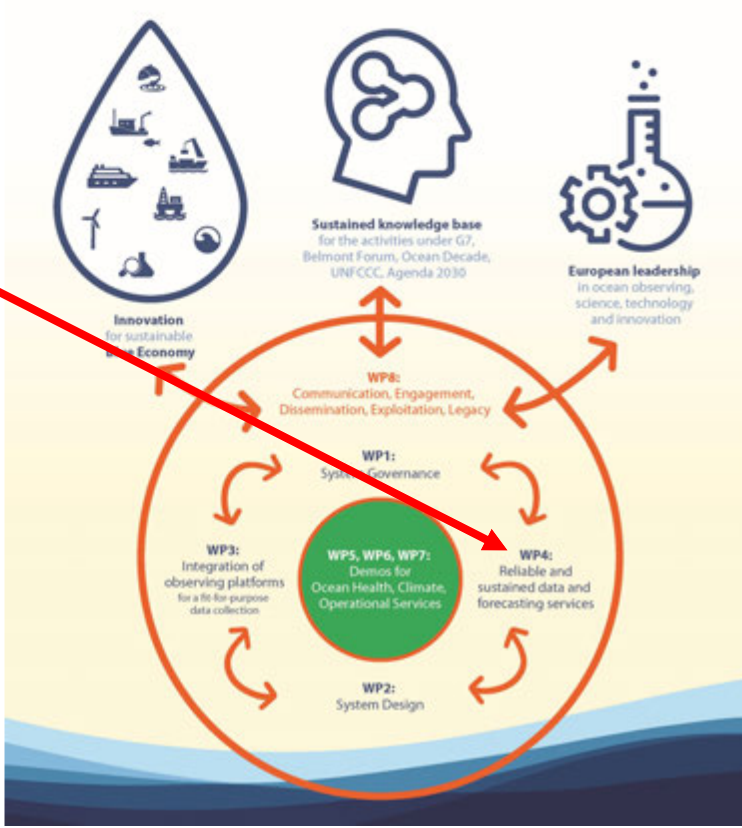


EuroSea

Improving and Integrating European Ocean Observing and Forecasting Systems for Sustainable use of the Oceans

WP4: Data integration, Assimilation, and Forecasting:

- Ensure that new or consolidated in-situ observation **data sets are integrated** in the European modelling and forecasting systems.
- Ensemble **forecasting at regional level** to extract Extreme Forecast Indices (EFI)
- Produce new **carbon EOV synthesis** products
- **Assess the skill** of ocean variables from the Copernicus Climate Change seasonal forecasting systems
- All new products will be **integrated in Copernicus**, reaching TRL7 and 8



A carbon relevant ship-based time series stations data product is in particular relevant for:

- Assessments of ocean acidification/carbon budgets on a global scale
- Cross-validation of evolving autonomous observational networks
- Biogeochemical modelling studies
- Simplification of joint time series studies

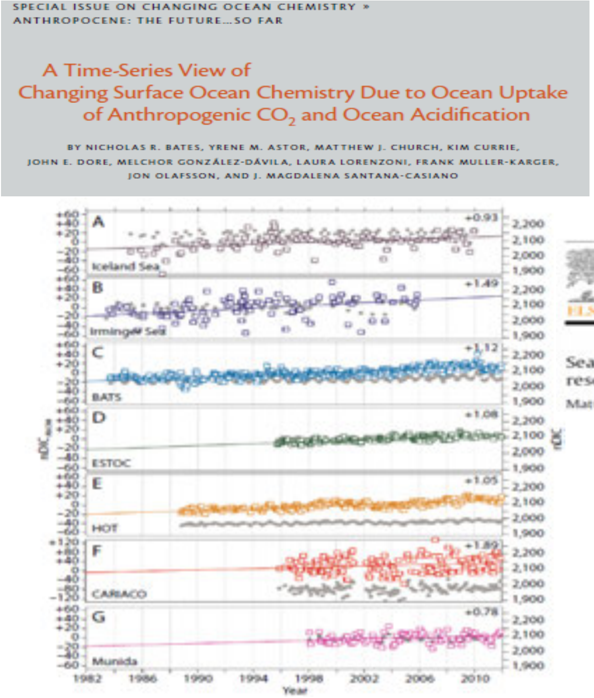
SINGLE TIME SERIES ANALYSIS
- Identification of temporal patterns
- Understanding of local processes
- Establishment of local baselines

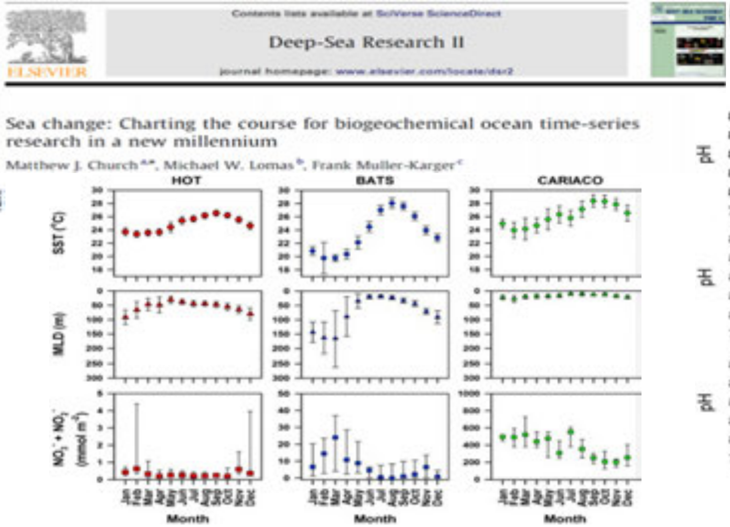
JOINT TIME SERIES ANALYSIS
- Identification of temporal and spatial patterns
- Establishment of regional baselines
- Separation of stressors
- Understanding of regional and global processes
- Projection and forecasting

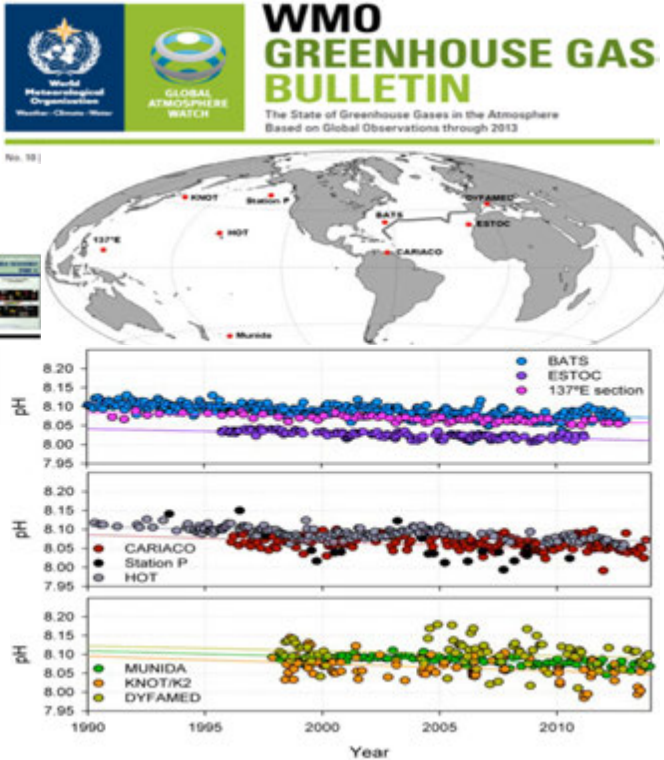
IGMETS

A carbon relevant ship-based time series stations data product is in particular relevant for:

➡ **Increase efficiency of ocean observations & efforts**


SPECIAL ISSUE ON CHANGING OCEAN CHEMISTRY » ANTHROPOCENE: THE FUTURE…SO FAR
A Time-Series View of Changing Surface Ocean Chemistry Due to Ocean Uptake of Anthropogenic CO_2 and Ocean Acidification
BY NICHOLAS R. BATES, YRENE M. ASTOR, MATTHEW J. CHURCH, KIM CURRIE, JOHN E. DORE, MELCHOR GONZÁLEZ-DÁVILA, LAURA LORENZONI, FRANK MULLER-KARGER, JON OLAFSSON, AND J. MAGDALENA SANTANA-CASIANO
A Iceland Sea +0.93
B Irminger Sea +1.49
C BATS +1.12
D ESTOC +1.08
E HOT +1.05
F CARIACO +1.89
G Munida +0.78
Year


Contents lists available at SciVerse ScienceDirect
Deep-Sea Research II
journal homepage: www.elsevier.com/locate/dsr2
ELSEVIER
Sea change: Charting the course for biogeochemical ocean time-series research in a new millennium
Matthew J. Church, Michael W. Lomas, Frank Muller-Karger
HOT
BATS
CARIACO
SST (°C)
MLD (m)
Month


WMO
GREENHOUSE GAS BULLETIN
The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2013
GLOBAL ATMOSPHERE WATCH
World Meteorological Organization
KNOT
Station P
HOT
BATS
DYFAMED
ESTOC
CARIACO
137°E
Munida
BATS
ESTOC
137°E section
CARIACO
Station P
HOT
MUNIDA
KNOT/K2
DYFAMED
pH
Year

Munida

Papa

Vizag

JGOFS

IMARPE

KERPIX

ESTOC

COPAS

PAP

CARIACO

?

DYFAME
D

KNOT

HOT

CVOO

CATS

BATS

SEATS

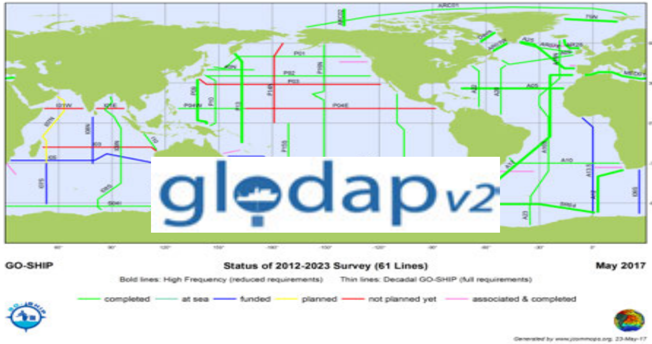
KOPR

Existing Efforts: Time-Series-stations

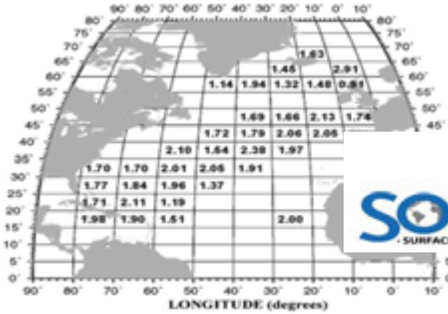
Ocean Sites	Mooring focus
IGMETS	No carbon focus No 2nd QC
BCO-DMO	US-Limitation
GOA_ON IOCCP/OCB	Focus on consistent metadata & measuring guidelines

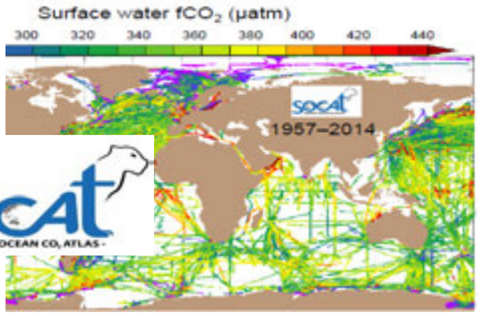
Existing Efforts: Carbon relevant ocean data products

Repeat Hydrography

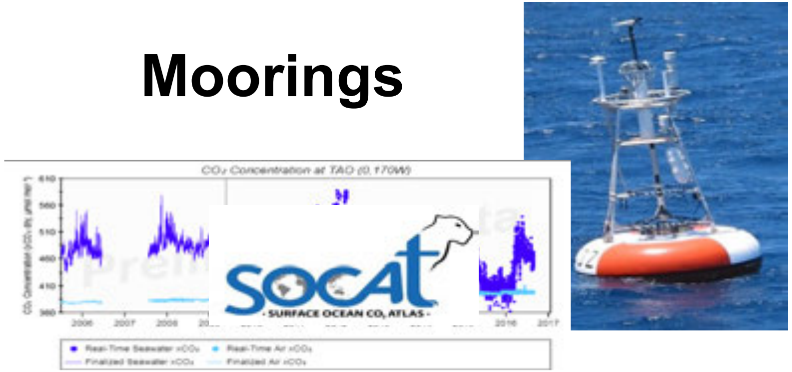


SOOP





Moorings



GLODAP: *A uniformly calibrated open ocean data product of inorganic and carbon-relevant variables*

glodap v2 Global Ocean Data Analysis Project

- A yearly updated synthesis of ship-based hydrographic survey data (1972 – today)
- Contains data of more than 840 cruises (*GLODAPv2.2019 – an update of GLODAPv2*, Earth Syst. Sci.)
- Involves a consistent QC routine and bias correction (precision & accuracy) of 8 core variables: S, O2, Nutrients, DIC, TA, pH
- Examples of success include the calculation of the ocean anthropogenic CO_2 uptake (Gruber et al. 2018)

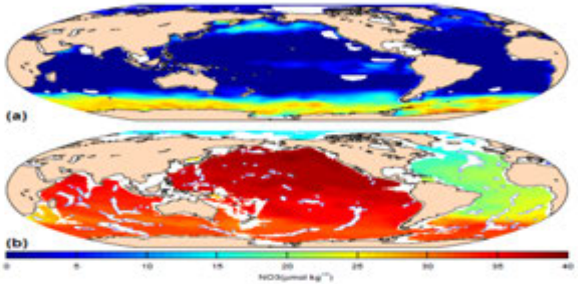


Figure 5. Mapped climatology of nitrate at 10 m (a) and 3000 m (b).

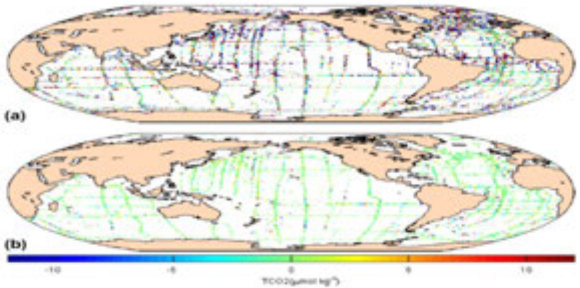


Figure 6. Difference between the gridded TCO_2 input data and the mapped climatologies at 10 m (a) and 3000 m (b).

Sanity Check	1st QC (Precision)	2nd QC (Accuracy)
Position check	Unit check	Crossover (w/-, w/o inversion)
Impossible values	Property-Property plots	Carbon interconsistency
Density instabilities	(Saturation ratios)	MLR
Stuck value test		Depth Averages
Inverted profile check		(Relation among tracers)

Secondary QC

Secondary quality control is a process in which the data are objectively studied in order to quantify systematic bias in the reported values. The identified data biases are then subjectively compared to predetermined accuracy limits.

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

https://glodap.geomar.de/adjustments/list?page=4

Meistbesucht | Erste Schritte

GLODAP :: GLODAPv2 :: ALL regions :: Adjustments

Toste Tanhua | Edit profile | Admin

GLODAPv2 @ CDIAC | Comments | Uploads, Downloads, Documentation — Filter adjustments: New and added | GLODAPv2 | DISMISSED

Adjustments Table | Cruises | Ships — Filter adjustments by ...

Recommended adjustment values (+ = add; x = multiply)

Cruise: Expocode (Regions)	Salinity [+]	CTD-Sal. [+] adj.	slope	intercept	action [ID]	TCO_2 [+]	Alkalinity [+]	pH [+]	pCO_2 [x]
09AR19941213 (IO)	0.0				0	0.0	0.0	-999	-999
09AR19960822 (IO)	0.0				0	0.0	0.0	-999	-999
09AR19980228 (IO)	0.0	0.0	1.0	0.0	5	0.0	0.0	-999,	
09AR20011029 (IO)	0.0	0.0	1.0	0.0	5	0.0	0.0	-999,	-999
09AR20030103 (IO)	0.0	0.0	1.0	0.0	5	-999	-999	-999	
09AR20041223 (IO)	0.0	0.0	1.0	0.0	5	0.0	0.0	-999,	
09AR20060102 (IO)	0.0	0.0	1.0	0.0	5	0.0	0.0	-999,	
09AR20071216 (IO)	0.0	0.0	1.0	0.0	5	-999	-999	-999	
09AR20080322 (IO)	0.0	0.0	1.0	0.0	5	-4.0	0.0	-999,	
09AR20110104 (IO)	0.0	0.0	1.0	0.0	5	-999	-999	-999	
09AR20120105 (IO)	0.0	0.0	1.0	0.0	5	-999	-999	-999	
09FA19930624 (PO)	0.0	0.0	1.0	0.0	5	-999	-999	-999	
09FA19930624 (PO)	0.0				0	-999	-999	-999	-999
09FA19941112 (IO)	-888				0	-999	-999	-999	-999
09FA20000926 (IO)	0.0	0.0	1.0	0.0	5	0.0	0.0	-999,	
09FA20010524 (PO)	0.0	-999	-999	-999	3	0.0	0.0	-999,	
09FA20010524 (PO)	0.0				0	0.0	0.0	-999	-999
09SS20090203 (PO)	0.0	0.0	1.0	0.0	5	0.0	0.0	-999,	
11BG20060425 (AO)	-888	-999	-999	-999	3	-999,	-888	-888	
11BG20060531 (AO)	-888	-999	-999	-999	3	-999,	-888	-888	

Cruise: Expocode (Regions)	Nitrate [x]	Phosphate [x]	Silicate [x]	Oxygen [x]	CTD-Oxyg. [x] adj.	slope	intercept	action [ID]	CFC12 [x]	CFC11 [x]	CFC113 [x]	CCl4 [x]
09AR19941213 (IO)	1.0	1.0	1.03	1.0				0	-888	-888	-888	-888
09AR19960822 (IO)	1.0	1.0	1.0	0.96				0	-999	-999	-999	-999
09AR19980228 (IO)	1.0	1.0	1.03	1.0	-999	-999	-999	3	-999	-999	-999	-999
09AR20011029 (IO)	1.0	1.05	0.98	0.985	0.985	1.0	0.0	5	1.0	1.0	1.0	-999
09AR20030103 (IO)	0.99	1.05	0.98	1.015	1.015	1.0	0.0	5	-999	-999	-999	-999
09AR20041223 (IO)	1.0	1.0	1.0	1.0	1.0	1.0	0.0	5	1.0	1.0	1.0	-999
09AR20060102 (IO)	1.0	-999	1.0	1.0	1.0	1.0	0.0	5	-999	-999	-999	-999
09AR20071216 (IO)	1.0	1.0	0.975	1.0	1.0	1.0	0.0	5	1.0	1.0	1.0	-999
09AR20080322 (IO)	1.0	0.97	0.97	1.0	1.0	1.0	0.0	5	1.0	1.0	1.0	-999
09AR20110104 (IO)	1.02	1.0	1.0	1.0	1.0	1.0	0.0	5	-999	-999	-999	-999
09AR20120105 (IO)	1.0	1.0	1.0	1.0	1.0	1.0	0.0	5	-999	-999	-999	-999
09FA19930624 (PO)	0.91	0.97	1.0	0.99	-999	-999	-999	3	-999	-999	-999	-999
09FA19930624 (PO)	0.92	0.95	0.98	0.99				0	-999	-999	-999	-999
09FA19941112 (IO)	-888	-888	-888	-888				0	-999	-999	-999	-999
09FA20000926 (IO)	1.0	1.0	1.0	1.0	1.0	0.9888	0.0	6	-999	-999	-999	-999
09FA20010524 (PO)	1.0	1.0	1.0	1.0	-999	-999	-999	3	1.0	1.0	-777	1.0
09FA20010524 (PO)	1.0	1.0	1.0	1.0				0	-888	-888	-888	-888
09SS20090203 (PO)	1.0	1.0	1.0	1.0	1.0	1.0157	-2.7746	6	1.0	1.0	-999	-999
11BG20060425 (AO)	-999	-999	-999	-888	-999	-999	-999	3	-999	-999	-999	-999
11BG20060531 (AO)	-999	-888	-888	-888	-999	-999	-999	3	-999	-999	-999	-999

« Previous 1 2 3 4 5 6 7 8 9 … 35 36 Next » | Adjustments Table

Footnotes for adjustments:

- *-999=Not available | -888=Not considered | -777=Poor data, no adjustment suggested*
- *DISMISSED!: If one of the three carbon parameters TCO_2, alkalinity or pH is calculated, this one is marked red and annotated as e.g. -0.123456,!*
- *Actions (act.) taken for (CTD) **salinity** calibration: 7="using BOTsal because bad fit of CTDsal."; 6="using mean of BOTsal and fitted CTDsal."; 5="using mean of BOTsal and UNCHANGED CTDsal."; 4="using BOTsal (few or no CTDsal or BOTsal>80%)."; 3="using BOTsal (no CTDsal)"; 2="using CTDsal (no BOTsal)"; 1="using none"*
- *Actions (act.) taken for (CTD) **oxygen** calibration: 7="using BOToxy because bad fit of CTDoxy."; 6="using mean of BOToxy and fitted CTDoxy."; 5="using mean of BOToxy and UNCHANGED CTDoxy."; 4="using BOToxy (few or no CTDoxy or BOToxy>80%)."; 3="using BOToxy (no CTDoxy)"; 2="using CTDoxy (no BOToxy)"; 1="using none"*
- *To calibrate CTDsal use: CTDsal_calibrated = (CTDsal - CTDsal_intercept) / CTDsal_slope*
- *To calibrate CTDoxy use: CTDoxy_calibrated = (CTDoxy - CTDoxy_intercept) / CTDoxy_slope*

Carbon relevant ship-based time series station synthesis

Create a living data product on an overarching framework with existing (SOCAT & GLODAP) carbon data products

Leverage from existing tools/routines, e.g. AutoQC4Env, Crossover toolbox, Merging-Routines

A Novel Concept for Automated Quality Control of Atmospheric Time Series

Najmeh Kaffashzadeh, Sabine Schröder, and Martin G. Schultz
Jülich Supercomputing Center, Forschungszentrum Jülich, Germany

Objective

Atmospheric chemistry and physics time series measurements are used for various purposes such as model evaluation, trend analysis, prediction, etc. The quality of these data can be affected by various error types. Moreover, when the time series from various measurement instruments or data providers are merged together, assessing the quality of the data presents a major challenge and often relies on subjective screening. Here, we propose a new tool (auto-qc) based on a novel concept for the automated quality control of atmospheric time series data.

Aspects of Concern

- The presence of poor quality values in measured data influences the results of statistical analysis.
- The quality assessment of large data volume based on visual screening is not possible.
- Inconsistent standard criteria for automated data flagging at several major environmental agencies.

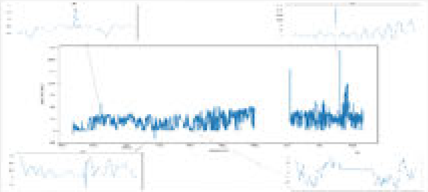


Figure 1: A manipulated ground level ozone time series at an arbitrary station

Fig.1 shows an ozone time series where various poor quality values have been added for demonstration, such as:

(a) a very high value (extreme outlier),

(b) negative values,

(c) constant values over a period of time,

(d) a large individual value over a short time interval.

Properties of the auto-qc Tool

1) Data classification based on user-defined thresholds

One of the uniqe features of the auto-qc tool is the "probability" concept by which a probability (p) is assigned to each value. The p shows the likelihood that a reported measurement value belongs to the real distribution of the atmospheric quantity under consideration.

Fig.2 shows a framework that has been implemented to assign the p in the auto-qc tool. This framework consists of several groups:

G0: exclude large gross errors from the data for the further analysis,

G1: check the quality of a single value test,

G2: check the quality of a single value comparing with neighboring data points,

G3-Gn: check the internal and spatial consistency.

Advantages:

- runs in parallel (e.g. several variables or stations)
- users defined classification and flagging

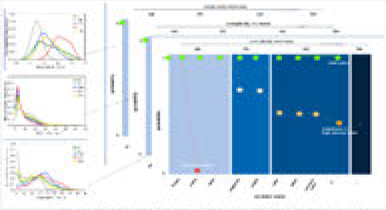


Figure 2: The probability concept framework of the auto-qc tool

2) Flexible software framework

The auto-qc tool is hosted in the GitLab repository at the Super Computing Center in the Forschungszentrum Jülich, and follows accepted coding standards (Fig.3).

Advantages:

- easy integration into an automated workflow,
- application to both archived and near-real time data processing.

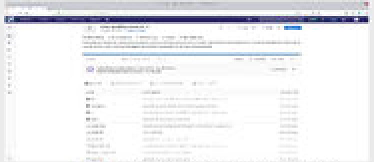


Figure 3: A snapshot of the auto-qc tool in the Gitlab

The set of statistical qc-tests and any user arguments can easily be configured with variable-specific control files in the JSON format (Fig.4).

Advantages:

- easily expandable (any number of tests can be defined and run sequentially),
- allows data users to adapt the tool for their needs (fit for purpose).

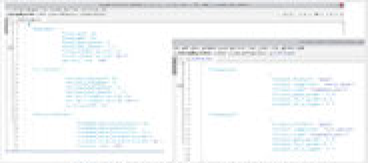


Figure 4: An example of the settings file for the qc-tests

Results

The auto-qc tool was used for the quality check of a multi-annual hourly ozone time series from the database of the Tropospheric Ozone Assessment Report (TOAR).

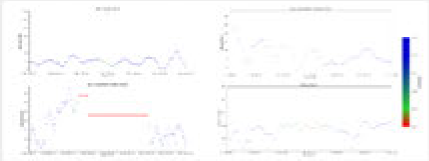


Figure 5: Quality control results of the ozone time series shown in Fig.1

Fig.5 shows that the auto-qc tool is able to detect the poor quality values:

(a) the extreme outlier has been excluded by the *range test* in G0,

(b) the probability (p) of the negative value has been lowered by the *negative value test* in G1,

(c) p of the constant values has been lowered by the *constant value test* in G2,

(d) p of the local outlier has been reduced by the *q test* in G2.

Conclusions

Preliminary results indicate that the probability concept works well and is able to deal with a large and heterogeneous dataset such as the global collection of ozone data in the TOAR database.

Recent Achievements and Future Work

The tool has been released to the Helmholtz Digital Earth community for further adaptations, tests, and experiments.

Our future work focuses on the consolidation of the statistics in G0-G2 and the development of further qc-tests for making use of the full potential of the comprehensive TOAR metadata information (G3-Gn).

Contact: n.kaffashzadeh@fz-juelich.de

2nd QC adaptions to be made:

- Add seasonality filter to crossover routine (sinusoidal regression curves/running averages)
- Add long-term trend filter (?)

Symbiosis with GLODAP through case study:

- Determine zero-trend baselines (lowest depth) of no detectable linear trends based on GLODAP crossovers
- Ground truthing through new time-series data product using „footprint areas“ (Henson et al. 2016) of time-series-stations

→ Exemplarily demonstrate benefits of the common framework supporting automatization of 2nd QC (*automated minimum depth selection*)

- Set of Core Stations : JGOFS?

- Set of (Core) Variables: GLODAP vs. TS-Workshop?

- Sanity and 1st QC points: GLODAP?
  Flags: GLODAP/WOCE?

- Bias correction for entire station?
  Set of accuracy limits: GLODAP?

- Carbon calculations: GLODAP/CO2SYS?
  PH-Scale: GLODAP/Total at 25°C and 0dbar?

- Merging of bottle- and CTD data?

- (Inversion – yes/no?)

Current Situation:

- Meta data and bgc bottle data “well” distributed
- implementation of data intercomparability recommendations unknown (→2012 workshop)
- need for harmonized and quality controlled data and meta data
- several efforts underway with individual priorities

Challenges:

- large variety of data formats & documentation
- accessibility of data (data policy & restrictions of different sites)
- intercomparability of data
- developing appropriate QC routines

Benefits:

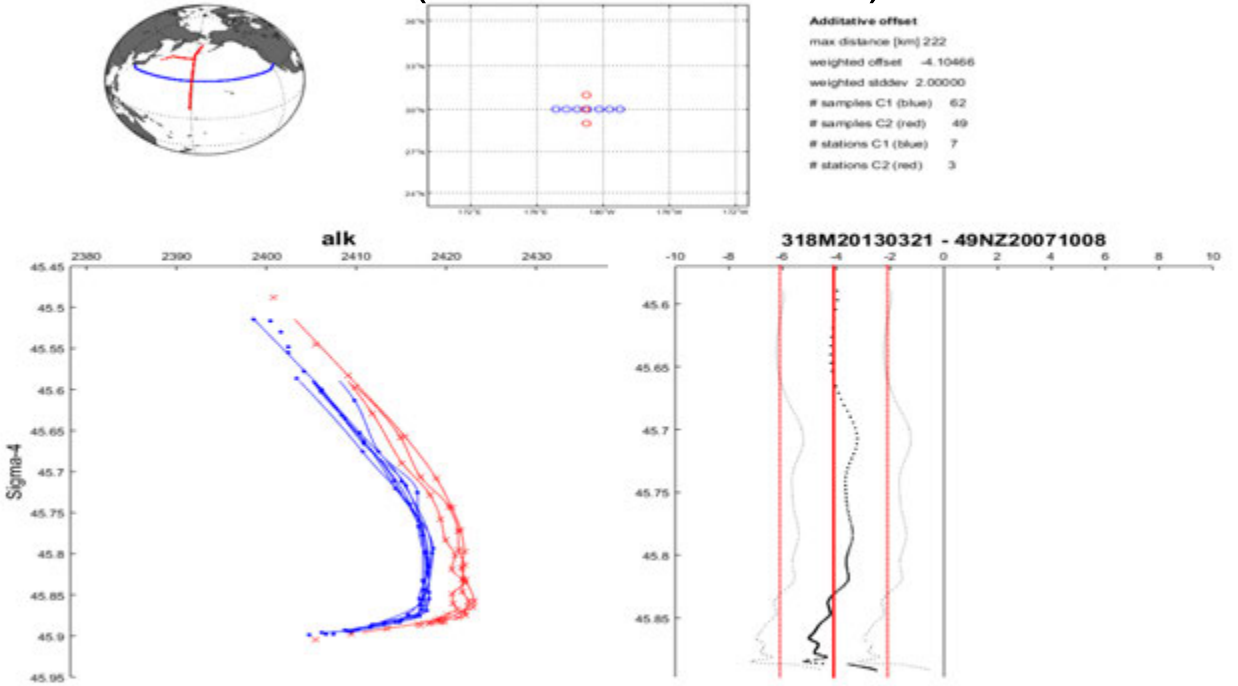
- forming a TS community with a defined target
- pushing forward implementation of quality control measures (→2012 workshop)
- making QCed TS data easy accessible for the scientific community and other synthesis efforts

Opportunities:

- leveraging on established data products (blueprints)
- benefitting from ongoing developments in data infrastructures

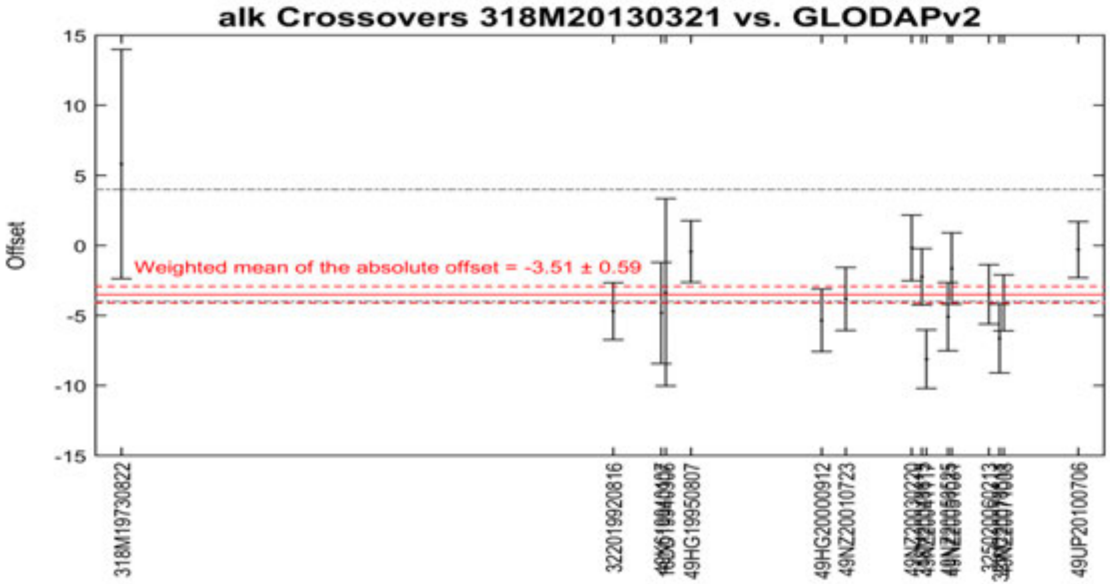
Crossover Analysis

An objective comparison of deep water data from one cruise with data from other cruises in the same area (Tanhua et al. 2010).

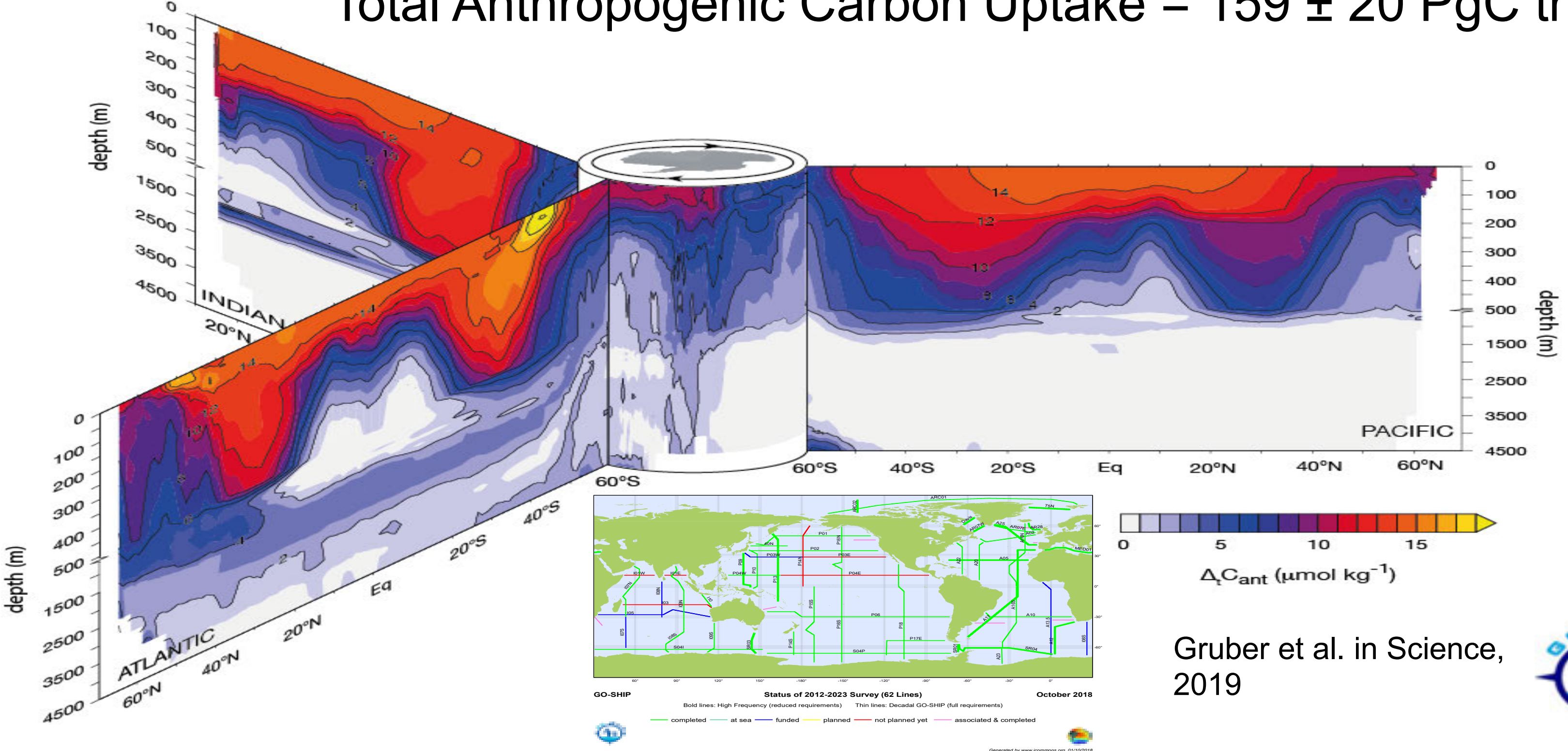


Crossover Analysis

An objective comparison of deep water data from one cruise with data from other cruises in the same area (Tanhua et al. 2010).



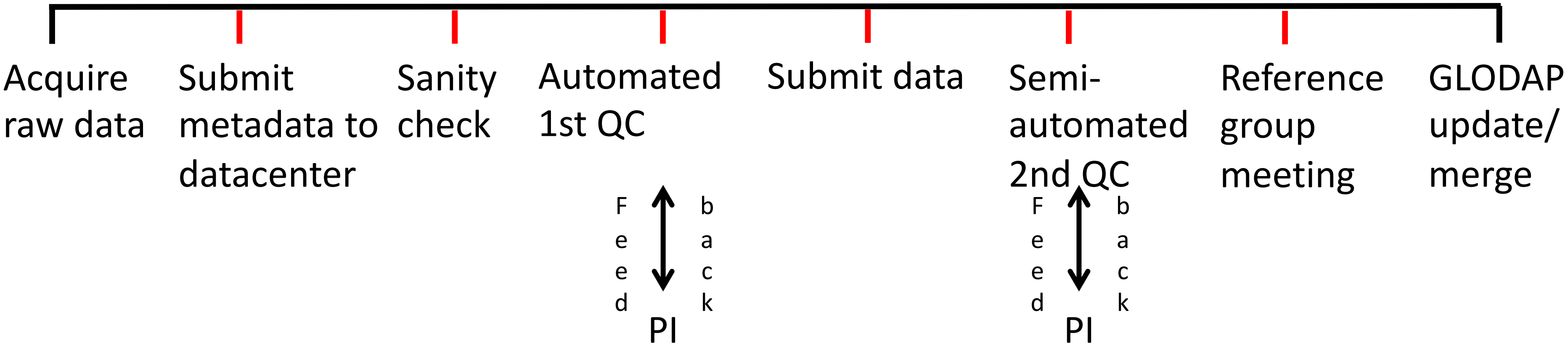
Anthropogenic Carbon Uptake = 32 ± 4 PgC from 1994 to 2007
Total Anthropogenic Carbon Uptake = 159 ± 20 PgC thru 2010



Gruber et al. in Science, 2019

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

Internal consistency = Weighted mean of absolute weighted mean offsets


Acquire raw data
Submit metadata to datacenter
Sanity check
Automated 1st QC
Submit data
Semi-automated 2nd QC
Reference group meeting
GLODAP update/ merge
Feed back
PI
Feed back
PI