

Support for this workshop was provided by:



# Coastal and Arctic Surface Ocean CO<sub>2</sub> Atlas (SOCAT) Quality Control Workshop

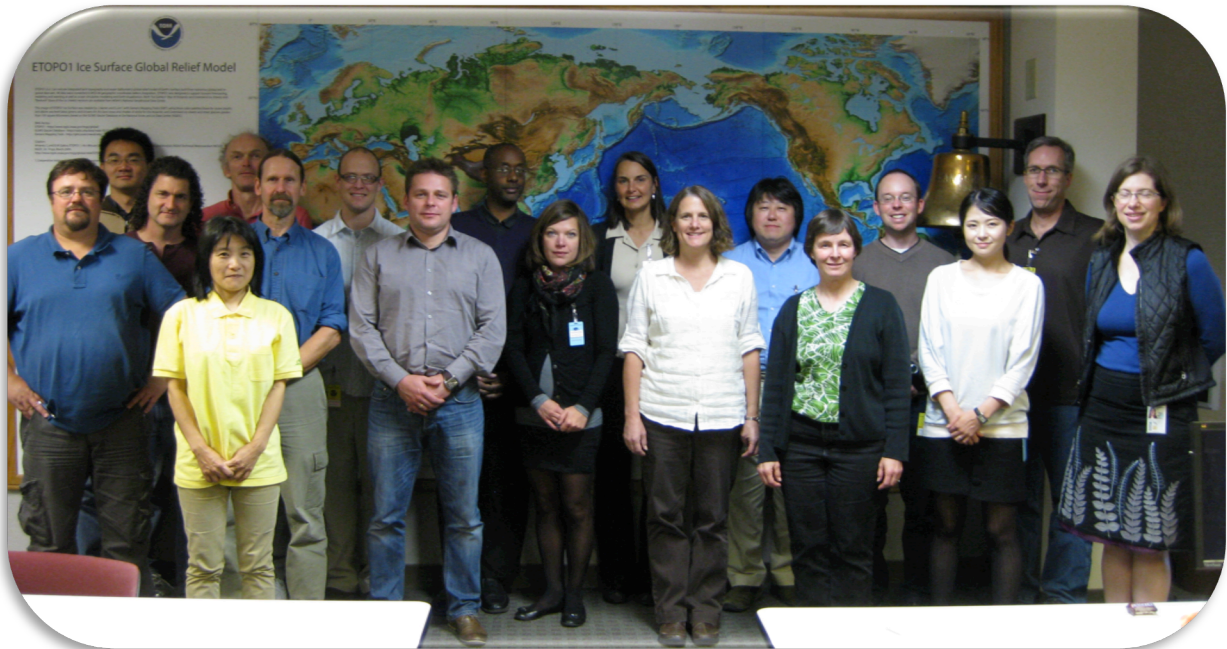
NOAA-PMEL, Seattle, USA, October 2-4, 2012



IOCCP Report 4/2012

## Abstract

Select members of the SOCAT Coastal Regional Group met at the NOAA Pacific Marine Environmental Laboratory from 2-4 October 2012. The workshop was chaired by Maciej Telszewski (IOCCP) and Simone Alin (NOAA-PMEL) and attended by eighteen scientists from four countries. The Coastal group met with the developers of the Live-Access Server (LAS) tools to learn how LAS can be used in the quality control (QC) effort for SOCAT. The participants accessed the tools and software on the LAS via their own computers, downloaded data files for their regions, set up the shared QC environment, and worked through several exercises to demonstrate the system. The groups began working through the data sets for their region (flagging, determining which 2nd level QC tests may be applied, testing those, etc.). Significant progress was made in the QC of new submissions to the SOCAT database for the coastal region.



**Workshop participants**

For bibliographic purposes this document should be cited as:  
IOCCP Report 4/2012, *Coastal and Arctic Surface Ocean CO<sub>2</sub> Atlas (SOCAT) Quality Control Workshop, NOAA-PMEL, Seattle, USA, October 2-4, 2012.*

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the workshop sponsors or of the U.S. National Science Foundation (NSF).

## Table of Contents

1. Introduction.....	4
2. Meeting Objectives.....	4
3. Interactive tutorial to the SOCAT Quality Control (QC) procedures with Questions and Answers session. ....	5
4. Distribution of QC work for the workshop and future SOCAT releases. ....	1
5. Use of the LAS for online SOCAT data analysis and hands-on QC on the Live Access Server (LAS). ....	1
6. Suggested improvements to the SOCAT QC procedures as well as online SOCAT data analysis to be implemented following the approval of the SOCAT Global and Technical Groups. ....	1
7. Discussion on the Coastal Ocean Synthesis Approaches .....	1
8. Appendices	
I. Agenda.....	8
II. Participant List.....	9

## 1. Introduction

The idea for SOCAT ([www.socat.info](http://www.socat.info)) was developed during the *Surface Ocean CO<sub>2</sub> Variability and Vulnerability* workshop held at IOC-UNESCO in Paris in April. Mostly independent from funding agencies, SOCAT is powered primarily by volunteer efforts by the scientific community. SOCAT was divided into region-specific working groups based on major oceans (North and Tropical Atlantic, North and Tropical Pacific, Indian, Arctic and Southern Oceans) and with a coastal working group active in each region, except for the Southern Ocean. SOCAT released the second-level quality-controlled (QC) data set (SOCAT v1.5) in September 2011. The SOCAT Live Action Server (LAS) was activated in 2010, allowing for community input on data QC, and was subsequently made publicly available for data viewing and download with the release of SOCAT v 1.5.

The need for a dedicated coastal group was recognized early in the formation of SOCAT, and a meeting was held in Kiel, Germany, in February of 2009 (prior to the release of the LAS) to engage the coastal CO<sub>2</sub> community. Results of that meeting are summarized in the [report](#) available from the SOCAT website. Since that time, coastal data QC was erratic, with some coastal sub-regions receiving more thorough efforts than others. Prior to the October 2012 meeting in Seattle, organizers identified nearly 900 cruises as at least partially coastal in the SOCAT database in need of QC. This was attributed to the geographically scattered distribution of coastal carbon cycle scientists, and the SOCAT community felt that a meeting of the coastal working group, with active training in the SOCAT LAS and data QC procedures, would be a useful framework for engaging new participants with the SOCAT Coastal Group. This meeting was held at the NOAA-PMEL facility in Seattle, Washington from 2-4 October 2012, and included coastal carbon cycle scientists and data managers from the U.S., Poland, Japan, and Germany. This was the first meeting of the Coastal Group since 2009, and the first since Simone Alin, Wei-Jun Cai, and Burke Hales were introduced as new co-leaders of the SOCAT Coastal Group.

Early plans for this meeting called for joint gathering of Arctic and Coastal groups; however, the small number of datasets and unavailability of SOCAT Arctic researchers at that particular time led organizers to focus on the Coastal Group. All new Arctic data sets (~20) were considered in the QC sessions.

The meeting followed a pattern of plenary sessions interspersed with QC performed by meeting participants. Plenary sessions are listed below in sequential order. The list of participants and the actual agenda are included as Appendices I and II at the end of the report.

## 2. Meeting Objectives

The meeting was intended to: 1) familiarize the coastal QC group with the QC criteria; 2) provide hands-on training with the SOCAT LAS for contributors performing data QC; 3) define QC responsibilities; and 4) to complete as much of the unfinished coastal data QC as possible during the meeting.

### **3. Interactive tutorial to the SOCAT Quality Control (QC) procedures with Questions and Answers session.**

Steve Hankin presented the online tools of SOCAT and described how to use them. Attendees learned that QC Level 1 at the whole-dataset level is the responsibility of the submitting PIs, and should happen usually before submitting. Individual data points can be flagged (e.g. as in WOCE). Level 2 QC is provided by the SOCAT QC community, and consists of a total cruise flag, which, as described in the SOCAT 'cookbook' can have value of A-D, ranking accepted data sets; F, for failure; or S, for data suspended for future implementation. Additional flags include U, for data that have been updated, and X, for data that have been excluded (mostly because of duplication). To receive a flag of 'A,' crossovers are needed; if there are more than 50 points flagged the cruise should be flagged with F. Improvements planned for SOCAT v3 include greater automation for data ingestion, and implementation of new tools and procedures. This will necessitate clear standards for submitting data.

Cathy Cosca presented some examples of previous QC efforts, both good and bad. The agreed-upon procedure was to evaluate metadata, then data, then crossovers. Incomplete metadata is usually the reason for lower-than-optimal quality flags, and the QC volunteer should strive to put specific information in the comments that spell out the reason for low quality flags based on incomplete metadata. It was also agreed that the highest QC rankings would require good data for all parameters seen as central to determination of fCO<sub>2</sub>, specifically, water temperature, equilibrator headspace pressure, and water salinity.

The group identified several issues that will need more attention before future SOCAT releases. SOCAT QC volunteers would like to have input on PI-provided Level 1 QC, and there is currently no way to do this. The 50-flag threshold on individual bad data points for a data set 'F' award was seen as overly restrictive, particularly given the wide range of dataset resolution and magnitude. Attendee Rik Wanninkhof suggested that a percentage of bad-data flags would be more appropriate than an absolute number. The crossover criterion for awarding a dataset an 'A' flag was seen as lacking specificity, especially in the context of coastal data. Time and space 'closeness' is likely to have significantly different meaning in coastal settings than in the open ocean. Crossovers acceptable in the open ocean might not be meaningful in the more heterogeneous ocean margins; conversely, the close time and space overlap needed for a true coastal 'crossover' might make it nearly impossible for coastal data to ever receive an 'A' ranking. The lack of a cookbook for metadata was seen as an important deficiency for submitting PIs, given the importance of metadata in the QC process.

### **4. Distribution of QC work for the workshop and future SOCAT releases**

While the organizers considered simply randomly assigning QC participants cruises, they chose to assign cruises based on QC regionality and interest. For example, Wanninkhof was assigned all cruises performed by NOAA-AOML personnel, while Suzuki, Harasawa, and

Wasada were assigned all Western North Pacific data. This approach was taken to the extent possible, but there was a mismatch between the geographic representation of the participants and that of the data.

## **5. Use of the Live Access Server (LAS) for online SOCAT data analysis and hands-on QC on the LAS.**

Hankin presented a tutorial on the SOCAT Live Access Server, an interactive user-controlled web tool for evaluating, displaying, and downloading data. Data can be selected based on multiple criteria, including time, space, sampling region, QC level, etc. A number of fine points on the operation of the LAS were discussed and demonstrated, largely having to do with subsampling on the graphical displays, and the numerous ways that the user could enter mutually exclusive data constraints. He introduced the audience to the LAS and its options for obtaining information and how to use the QC trainer and the QC editor.

Throughout the meeting individual participants had several opportunities to perform QC on actual data sets with the assistance of SOCAT and LAS personnel. As of the submission of this report, all but 9 (of 894 prior to the meetings) cruises had been QC'ed by the coastal group volunteers.

## **6. Suggested improvements to the SOCAT QC procedures as well as online SOCAT data analysis to be implemented following the approval of the SOCAT Global and Technical Groups.**

A number of issues and suggestions were raised over the course of the meeting. Some of these were:

- Can SOCAT provide greater on-line training tools, for example YouTube-style training videos?
- Is it possible to make a comment on a cruise without adding a flag?
- Can all QC comments be compiled into a single file?
- There were concerns about ingestion of position and time data, given multiple possible user formats (decimal degrees versus minute second for lat and lon; reversed month/day in date formats). Some of these should be obvious, but less obvious are uncertainties with regard to time zone (UTC vs local, DST, e.g.). Are there ways to identify these kinds of issues?
- Is there a way to reconcile disagreements with previous flagging?
- Can the 'masks' that define which region SOCAT assigns to data be made into a servable product? For example, can the coastal mask be provided to LAS users?
- There appears to be a limit (~10) to how many data can be flagged at once. Can this be changed? Data are often bad in blocks that exceed this number.
- It currently appears that multiple URLs are allowed for metadata. The URL of the metadata should never be changed; the procedure should be appending new information.

- Is it possible to incorporate automated notifications to originating PIs when developments regarding their data occur?
- SOCAT should enforce a policy for missing data, preferably NaN.
- There appear to be mismatches between metadata requirements for SOCAT and CDIAC metadata; for example, standard gas concentrations are required in one and not the other. Can this be standardized?
- Metadata is requested for ancillary sensors, but the online metadata forms don't allow enough information to be reported.
- There is a need for a cookbook for metadata.
- Use of third-party sensors (e.g. ship-provided) needs to be noted in metadata, along with estimated uncertainty.
- Versioning control needs to be improved. For example, many files from SOCAT version 1.X were listed as U (updated), but had no obvious changes. QC and update history needs to be retained, and linked to a single originating dataset.

These suggestions were considered and many will be implemented in SOCAT v3. Others have already been implemented. For example, the coastal region perimeter and mask have been made available as a downloadable product on the SOCAT LAS.

## **7. Discussion on the Coastal Ocean Synthesis Approaches**

The group discussed several synthesis products, including simple bin-averaging of the coastal data, following the Chavez et al. (2007) effort for North American coastal waters. Hales presented results of the recent Hales et al. (2012) satellite map/meta-model approach applied to North American Pacific coastal waters. Participants suggested that it made sense to proceed using the above approaches to estimating air-sea CO<sub>2</sub> flux in coastal oceans first on a regional basis, to be followed by a subsequent global synthesis. There were informal agreements made to move forward with these efforts after the release of SOCAT v2, with the SOCAT Coastal Group leads (Hales, Alin, Cai) providing leadership in soliciting participation from PIs within sub-regions of the global coastal oceans and suggesting consistent spatial and temporal resolution to be applied across all regions to facilitate a later global synthesis. The NOAA-PMEL TMAP group (Hankin and colleagues) can provide support for this effort by testing different gridding approaches within the LAS system.

## **Appendix I: Participants list**

### **Present:**

1. Simone Alin, NOAA-PMEL, USA
2. Cathy Cosca, NOAA-PMEL, USA
3. Kevin O'Brien, NOAA-PMEL, USA
4. Burke Hales, CEOAS, Oregon State University, USA
5. Steve Hankin, NOAA-PMEL, USA
6. Wei-Jen Huang, University of Georgia, USA
7. Wiley Evans, University of Alaska Fairbanks, USA
8. Toru Suzuki, Marine Information Research Center, Japan
9. Sumiko Harasawa, NIES, Japan
10. Chisato Wada, NIES, Japan
11. Rik Wanninkhof, NOAA-AOML, USA
12. Chris Hunt, University of New Hampshire, USA
13. Penny Vlahos, University of Connecticut, USA
14. Maciej Telszewski, IOCCP of SCOR and IOC-UNESCO
15. Kirsten Isensee, IOC-UNESCO, Ocean Carbon Sources and Sinks Program
16. Karl Smith, NOAA-PMEL, USA
17. Baba Kofi Weusijana, NOAA-PMEL, TMAP LAS UI Developer
18. Ansley Manke, NOAA-PMEL, TMAP LAS UI Developer

### **Remote:**

1. Ute Schuster
2. Benjamin Pfeil
3. Wei Jun Cai



## Appendix II: Agenda

### Tuesday October 2, 2012

- 09:00-09:20 **Welcome, workshop objectives, round of introductions and housekeeping**
- 09:20-10:30 **Interactive tutorial to the SOCAT QC with Q&A**
- 10:30-11:00 Coffee break
- 11:00-11:30 **Distribution of QC work for the workshop**
- 11:30-12:45 **Hands on QC on the Live Access Server (LAS) – first steps .**
- 12:45-14:00 Lunch
- 14:00-16:45 **QC by individual participants.**
- 16:45-17:00 **Q&A and issues to be raised with Pfeil and Schuster**
- 19:00 Group Dinner

### Wednesday October 3, 2012

- 09:00-9:15 **Wrap-up from day 1**
- 09:15 **Skype conversation(s) with Pfeil and Schuster (postponed)**
- 9:15-10:30 **QC by individual participants.**
- 10:30-11:00 Coffee break
- 11:00-12:45 **QC by individual participants.**
- 12:45-14:00 Lunch
- 14:00-14:45 **Use of LAS for SOCAT data analysis.**
- 14:45-16:45 **QC by individual participants.**
- 16:45-17:00 **Q&A and issues to be raised with Benjamin and or Ute**

### Thursday October 4, 2012

- 09:00-9:15 **Wrap-up from day 2**
- 09:15-09:45 **Skype conversation(s) with participants attending remotely.**
- 9:15-10:30 **QC by individual participants**
- 10:30-11:00 Coffee break

11:00-12:45 **QC by individual participants**

12:45-14:00 Lunch

14:00-16:00 **QC by individual participants**

16:00-17:00 **Workshop wrap-up, action items, reporting etc.**