An update to the Surface Ocean CO2 Atlas (SOCAT version 2)


To cite this version:


HAL Id: hal-01130481
https://hal.archives-ouvertes.fr/hal-01130481
Submitted on 7 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
An update to the Surface Ocean CO₂ Atlas (SOCAT version 2)


¹Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, UK
²Geophysical Institute, University of Bergen, Bergen, Norway
³Bjerknes Centre for Climate Research, Bergen, Norway
⁴Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration, Seattle, Washington, USA
⁵Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, Washington, USA
⁶Uni Climate, Uni Research, Bergen, Norway
⁷National Institute for Environmental Studies, Tsukuba, Japan
⁸Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
⁹College of Life and Environmental Sciences, University of Exeter, Exeter, UK
¹⁰International Ocean Carbon Coordination Project, Institute of Oceanology of the Polish Academy of Sciences, Sopot, Poland
¹¹CSIRO Marine and Atmospheric Research, Hobart, Australia
¹²Antarctic Climate and Ecosystems Cooperative Research Centre, Hobart, Australia
¹³Cooperative Institute for Marine and Atmospheric Studies, Rosenstiel School for Marine and Atmospheric Science, University of Miami, Miami, Florida, USA
¹⁴Bermuda Institute of Ocean Sciences, Ferry Reach, Bermuda
¹⁵Sorbonne Universités (UPMC, Univ Paris 06), CNRS, IRD, MNHN, LOCEAN Laboratory, Paris, France
¹⁶CNRS, UMR 7144, Equipe Chimie Marine, Station Biologique de Roscoff, Roscoff, France
¹⁷Sorbonne Universités, UPMC Université Paris 06, UMR7144, Adaptation et Diversité en Milieu Marin, SBR, Roscoff, France
¹⁸School of Marine Science and Policy, University of Delaware, Newark, Delaware, USA
¹⁹Atlantic Oceanographic and Meteorological Laboratory, National Oceanic and Atmospheric Administration, Miami, Florida, USA
²⁰Monterey Bay Aquarium Research Institute, Moss Landing, California, USA
²¹Key Laboratory of Global Change and Marine-Atmospheric Chemistry, Third Institute of Oceanography, State Oceanic Administration, Xiamen, P. R. China
²²Chinese Arctic and Antarctic Administration, Beijing, P. R. China

Published by Copernicus Publications.
23Institute of Marine Research, Tromsø, Norway
24Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden
25National Institute of Water and Atmospheric Research, Dunedin, New Zealand
26Royal Netherlands Institute for Sea Research, Texel, the Netherlands
27Ocean Acidification Research Center, University of Alaska Fairbanks, Fairbanks, Alaska, USA
28Norwegian Polar Institute, Fram Centre, Tromsø, Norway
29College of Earth, Ocean and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA
30CSIRO Marine and Atmospheric Research, Floreat, WA, Australia
31Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany
32Ocean Process Analysis Laboratory, University of New Hampshire, Durham, New Hampshire, USA
33National Research Institute for Fisheries Science, Fisheries Research Agency, Yokohama, Japan
34Tyndall Centre for Climate Change Research, University of East Anglia, Norwich Research Park, Norwich, UK
35IVL Swedish Environmental Research Institute, Gothenburg, Sweden
36Plymouth Marine Laboratory, Plymouth, UK
37GEOMAR, Helmholtz Centre for Ocean Research, Kiel, Germany
38IRD, Sorbonne Universités (UPMC, Univ Paris 06), CNRS, MNHN, LOCEAN Laboratory, Paris, France
39LaboMar, Universidade Federal do Ceará, Fortaleza-Ceará, Brazil
40Japan Agency for Marine-Earth Science and Technology, Yokosuda, Japan
41Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA
42East Sea Research Institute, Korea Institute of Ocean Science and Technology, Ujinn, Korea
43Instituto de Investigaciones Marinas de Vigo, Consejo Superior de Investigaciones Científicas, Vigo, Spain
44Marine Division, Global Environment and Marine Department, Japan Meteorological Agency, Tokyo, Japan
45National Institute of Oceanography, Regional Centre, Visakhapatnam, India
46Marine Information Research Center, Japan Hydrographic Association, Tokyo, Japan
47Lamont Doherty Earth Observatory, Columbia University, Palisades, New York, USA
48National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan
49Centre for Isotope Research, University of Groningen, Groningen, the Netherlands
50Department of Marine Sciences, University of Connecticut, Groton, Connecticut, USA
51Department of Oceanography, Dalhousie University, Halifax, Canada

*formerly at: School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, UK

Correspondence to: D. C. E. Bakker (d.bakker@uea.ac.uk)

Received: 25 June 2013 – Published in Earth Syst. Sci. Data Discuss.: 16 August 2013
Revised: 12 January 2014 – Accepted: 24 January 2014 – Published:

Abstract. The Surface Ocean CO₂ Atlas (SOCAT), an activity of the international marine carbon research community, provides access to synthesis and gridded fCO₂ (fugacity of carbon dioxide) products for the surface oceans. Version 2 of SOCAT is an update of the previous release (version 1) with more data (increased from 6.3 million to 10.1 million surface water fCO₂ values) and extended data coverage (from 1968–2007 to 1968–2011). The quality control criteria, while identical in both versions, have been applied more strictly in version 2 than in version 1. The SOCAT website (http://www.socat.info/) has links to quality control comments, metadata, individual data set files, and synthesis and gridded data products. Interactive online tools allow visitors to explore the richness of the data. Applications of SOCAT include process studies, quantification of the ocean carbon sink and its spatial, seasonal, year-to-year and longer-term variation, as well as initialisation or validation of ocean carbon models and coupled climate-carbon models.

Data coverage

Gridded products:
doi:10.3334/CDIAC/OTG.SOCAT_V2_GRID
Available at: http://www.socat.info/
Coverage: 79° S to 90° N; 180° W to 180° E
Location Name: Global Oceans and Coastal Seas
Date/Time Start: 16 November 1968
Date/Time End: 26 December 2011

1 Introduction

Human activity is releasing large quantities of the greenhouse gas carbon dioxide (CO₂) into the atmosphere. As a result, the atmospheric CO₂ mole fraction has increased from 280 µmol mol⁻¹ in pre-industrial times (Janssen et al., 2007) to 397 µmol mol⁻¹ in April 2013 (Tans and Keeling, 2014).
The rapid, ongoing change in the atmospheric composition by greenhouse gas emissions has been predicted to increase global mean temperature by 1.5 °C to 5.0 °C by the end of the century (Peters et al., 2013). Such warming would be accompanied by sea level rise, increased storm frequency, melting of ice caps and sea ice, changes in precipitation patterns and ocean acidification (Solomon et al., 2007), to name only the most prominent examples. Already many changes in the Earth’s climate are apparent, such as the decline in Arctic sea ice extent (Stroeve et al., 2007), and warming in Alaska, near the Antarctic Peninsula (Vaughan et al., 2003; Mulvaney et al., 2012) and of the upper ocean (Levitus et al., 2005).

The oceans absorb a substantial part of the CO$_2$ emissions by human activity, thereby mitigating climate change. From pre-industrial times to 1994 the oceans have taken up 118 ± 19 Pg C from the atmosphere (Sabine et al., 2004). This is equivalent to roughly 50 % of CO$_2$ emissions from fossil fuel burning and cement production or 30 % of the total anthropogenic emissions, if CO$_2$ emissions from land use change are included. Recent estimates indicate that the oceans are a contemporary sink for roughly 27 % of the annual CO$_2$ emissions by fossil fuel combustion, cement production and land use change (Le Quéré et al., 2013). Uncertainty in the land use change emissions leads to a larger error estimate for the proportion of the anthropogenic emissions taken up by the oceans.

There is uncertainty on how much CO$_2$ the oceans will absorb in a warming climate of the future (e.g. Jones et al., 2013). Considerable year-to-year, decadal and longer-term variation of CO$_2$ uptake is apparent in the North Atlantic Ocean (Corbierre et al., 2007; Schuster and Watson, 2007; Thomas et al., 2008; Schuster et al., 2009; McKinley et al., 2011), the North Sea (Thomas et al., 2007), the North Pacific Ocean (Takamura et al., 2010), the equatorial Pacific Ocean (Feely et al., 2002, 2006; Ishii et al., 2004, 2009; Park et al., 2006, 2012) and the Southern Ocean (Le Quéré et al., 2007; Metzl, 2009), with large differences between ocean regions (Le Quéré et al., 2010; Lenton et al., 2012).

Measurements of CO$_2$ in the surface oceans (generally expressed as the mole fraction of CO$_2$ (xCO$_2$), partial pressure (pCO$_2$), or fugacity (fCO$_2$)) enable estimation of CO$_2$ air–sea fluxes and their variability. The fugacity can be measured underway on the surface water supply of ships. This method is used on a variety of ships, including ships of opportunity on commercial routes. The number of CO$_2$ measurements has greatly increased over the past four decades (Fig. 1) (Sabine et al., 2010). Data collection started in the late 1960s and 1970s, increased in the 1980s and intensified from the 1990s onwards. Roughly four times more data have been collected during the 2000s than in the 1990s. The growth in data collection has partly resulted from large international research programmes, for example JGOFS (Joint Global Ocean Flux Study) and WOCE (World Ocean Circulation Experiment), and regional funding initiatives. The development of autonomous instrumentation for the continuous measurement of surface water fCO$_2$ (e.g. Körtzinger et al., 1996; Cooper et al., 1998; Pierrot et al., 2009), the intercomparison of such instrumentation at sea (Körtzinger et al., 1996, 2000) and its installation on ships of opportunity (e.g. Cooper et al., 1998; Lüger et al., 2004; Schuster and Watson, 2007; Watson et al., 2009; Takamura et al., 2010; Lefèvre et al., 2013) and on moorings and drifters (e.g. Hoed et al., 1999; Emerson et al., 2011) have played an important role in the increase in data collection.

Quantification of global and regional, annual mean ocean CO$_2$ uptake requires observations of surface water fCO$_2$ with adequate spatial and temporal coverage (Sweeney et al., 2000; Lenton et al., 2006). Studies of year-to-year, decadal and longer-term trends in air–sea CO$_2$ uptake necessitate consistent, multi-decade data records of surface ocean fCO$_2$ (e.g. Schuster and Watson, 2007; Park et al., 2012). Statistical techniques and modelling have been developed to infer basin-wide distributions of surface water fCO$_2$ from limited observations, for example a diffusion–advection based interpolation scheme (Takahashi et al., 1997, 2009), (multiple) linear regression (e.g. Boutin et al., 1999; Sarma et al., 2006), neural network approaches (e.g. Lefèvre et al., 2005) and a diagnostic ocean mixed layer model (Rödenbeck et al., 2013).

Uniform procedures for the collection, reporting, processing and archiving of CO$_2$ data, as well as public release of data, are essential for creating global and regional, long-term, consistent surface ocean fCO$_2$ synthesis products. Takahashi
and co-workers have constructed an impressive series of surface ocean CO$_2$ climatologies, the most recent one for the climatological year 2000 (Takahashi et al., 2009), and now provide annual updates to their global surface ocean CO$_2$ data set (Takahashi et al., 2013). The Surface Ocean CO$_2$ Atlas (SOCAT) (Bakker et al., 2012; Pfeil et al., 2013; Sabine et al., 2013) complements this work. The SOCAT and Takahashi data sets benefit from standardisation and intercomparison of measurement and reporting protocols, as well as discussions between data providers and quality controllers on reporting standards and data quality (Dickson et al., 2007; IOCCP, 2008; SOCAT, 2011; Wanninkhof et al., 2013a). Both data sets contribute towards more rapid availability of ocean carbon data for synthesis products and policy-related assessments.

SOCAT is an international activity of ocean carbon scientists. It aims to create, make publicly available and archive the following (IOCCP, 2007):

- A 2nd level quality-controlled, global surface ocean fCO$_2$ data set following internationally agreed-upon procedures and regional review;

- A gridded data product of mean monthly surface water fCO$_2$ on a 1° latitude by 1° longitude grid with minimal temporal or spatial interpolation using the 2nd level quality-controlled, global surface ocean fCO$_2$ data set.

The first SOCAT release was made public as versions 1.4 and 1.5, here jointly referred to as version 1, in September 2011 (Bakker et al., 2012). SOCAT version 1 contains 6.3 million surface fCO$_2$ data points from 1851 data sets in the global oceans and coastal seas between 1968 and 2007 (Fig. 1, Table 1) (Pfeil et al., 2013; Sabine et al., 2013). Version 2 is presented here.

## 2 SOCAT version 2

### 2.1 An update of version 1

Version 2 is an update of version 1 with 60% more data and 4 years extra data coverage. SOCAT version 2 contains 10.1 million surface fCO$_2$ values from 2660 data sets for the global oceans and coastal seas between November 1968 and December 2011 (Figs. 1 and 2). Version 2 was made public on 4 June 2013 at the 9th International Carbon Dioxide Conference in Beijing, China (SOCAT, 2013b).

SOCAT data products provide surface water fCO$_2$ values at sea surface temperature (fCO$_2$rec, with “rec” indicating recommended fCO$_2$), which have been (re-)calculated from the original CO$_2$ values reported by the data provider, following a strict calculation protocol. Sea surface temperature refers to the temperature at the seawater intake, often at about 5 m depth on ships. The procedures for data retrieval, for data entry, for the (re-)calculation of surface water fCO$_2$, for quality control, and for the creation of data products in version 2 are analogous to those used in version 1 (Pfeil et al., 2013; Sabine et al., 2013) and are described in Sects. 2.2, 2.3 and 2.4. The sections also highlight where version 2 differs from version 1 (Table 1).

Version 2 has three data products (Tables 2 and 3):

1. Individual data set files of surface water fCO$_2$ in a uniform format which have been subject to 2nd level quality control;

2. A synthesis data set of surface water fCO$_2$ for the global oceans and coastal seas;


These data products are much the same as those for version 1 (Sect. 2.4) (Pfeil et al., 2013; Sabine et al., 2013). The SOCAT website (http://www.socat.info/) provides access to the data products together with online visualisation tools, data documentation, quality control comments, meeting reports, publications and a list of contributors (Tables 4, 5 and 6).

### 2.2 Data assembly and (re-)calculation of fCO$_2$ in version 2

#### 2.2.1 Data origin

New data sets were either submitted directly to SOCAT or were retrieved from public websites hosted by the
Table 1. Key differences between SOCAT versions 1 (released as versions 1.4 and 1.5) and 2. Further details are in the text.

<table>
<thead>
<tr>
<th>Description</th>
<th>Version 1</th>
<th>Version 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Pfeil et al. (2013); Sabine et al. (2013).</td>
<td>This study.</td>
</tr>
<tr>
<td>Data coverage</td>
<td>1968 to 2007, 6.3 million surface water $f$CO$_2$ values, 1853 data sets.</td>
<td>1968 to 2011, 10.1 million surface water $f$CO$_2$ values, 2660 data sets.</td>
</tr>
<tr>
<td>Time stamp</td>
<td>The time stamp did not contain seconds. Multiple entries for the same time stamp were reported in individual data set files (version 1.4), but were averaged in the synthesis files (version 1.5).</td>
<td>The time stamp includes seconds for all new and updated data sets. Seconds were added to time stamps for version 1 data sets to avoid concurrent entries. Artificial times with tenths and hundreds of a second have been generated for a dozen historical data sets reported at midnight or with few decimals in the time stamp.</td>
</tr>
<tr>
<td>Version numbers</td>
<td>Version numbers 1.4 and 1.5 highlighted the different treatment of multiple entries for the same time stamp.</td>
<td>Version 2 only.</td>
</tr>
<tr>
<td>Expocode</td>
<td>Expocodes were not used for moored and drifting buoys.</td>
<td>Expocodes are used for moored and drifting buoys.</td>
</tr>
<tr>
<td>Arctic region</td>
<td>Arctic data were included under the North Atlantic, North Pacific and coastal regions.</td>
<td>An Arctic region has been defined as all open ocean and coastal waters north of 70° N for 100° W to 43° E and north of 66° N elsewhere.</td>
</tr>
<tr>
<td>Identification of outliers</td>
<td>No systematic search was carried out for outliers and unrealistic values.</td>
<td>A systematic search for outliers and unrealistic values has been carried out. In total 154 data sets have been suspended.</td>
</tr>
<tr>
<td>Suspension of data sets</td>
<td>Data sets part of version 1.</td>
<td>Suspension of 70 data sets included in version 1 upon identification of unrealistic values.</td>
</tr>
<tr>
<td>WOCE flags of 2 (good), 3 (questionable), 4 (bad)</td>
<td>Virtually all $f$CO$_2$ values were reported with a WOCE flag of 2.</td>
<td>WOCE flags of 2, 3 and 4 have been assigned to $f$CO$_2$ values. Flags of 3 and 4 given during version 1 quality control (0.2 % of data) have been reinstated. A total of 20 850 $f$CO$_2$ values (0.2 %) has been given a flag of 3 or 4 in version 2.</td>
</tr>
<tr>
<td>Parameters in the individual and synthesis files</td>
<td>Atmospheric CO$_2$ mole fractions were from GLOBALVIEW-CO2 (2008). The files downloadable from the Cruise Data Viewer contained more parameters than the synthesis files.</td>
<td>Atmospheric CO$_2$ mole fractions are from GLOBALVIEW-CO2 (2012). New parameters are the data set quality control flags of A to D and distance to a major land mass (Table 3). The parameters in files downloadable via the Cruise Data Viewer as “All Variables” and “Current Variable” match those in the synthesis files at CDIAC.</td>
</tr>
<tr>
<td>Gridded Data Viewer</td>
<td>Available</td>
<td>The capabilities of the Gridded Data Viewer have been expanded.</td>
</tr>
<tr>
<td>Release notes</td>
<td>None</td>
<td>Release notes document problems in version 2 data sets and data products.</td>
</tr>
</tbody>
</table>

Carbon Dioxide Information Analysis Center (CDIAC), PANGAEA®, institutions and research projects. Version 2 has an additional 3.8 million surface $f$CO$_2$ values and 807 data sets relative to version 1, mostly from 2006 to 2011 (Fig. 1, Table 1). Figure 3 shows the number of $f$CO$_2$ values from the 30 ships, including 1 ship-based time series, with the most intense data collection effort. The data sets in version 2 originate from 107 different ships, plus 3 ship-based time series, 13 mooring-based time series and 3 drifters (Table 7). This study will adopt the term “data set” rather than “cruise” for individual data sets to reflect the presence of mooring and drifter data (0.7 % of $f$CO$_2$ values in version 2). Tools and parameters available online will be referred to by their name, e.g. “Cruise Data Viewer” (Sect. 2.4.5) and “cruise-weighted means” (Sect. 2.4.6).
Table 2. Key characteristics of the three SOCAT data products for surface ocean $\text{f/CO}_2$ values in version 2 (Sect. 2.4). The synthesis product is available as synthesis files and as subsets of the global synthesis data set. The table lists whether the data products include only $\text{f/CO}_2$ data with a WOCE flag of 2 (good) or also with flags of 3 (questionable) and 4 (bad). Information on access to metadata and quality control comments is provided. All data products can be accessed via the SOCAT website (http://www.socat.info) and via the links in the table.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>WOCE flag</th>
<th>Metadata QC entries</th>
<th>Access and format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual data set files</td>
<td>2, 3, 4</td>
<td>Yes</td>
<td>Text files at Pangaea(^1)</td>
</tr>
<tr>
<td>Synthesis data set</td>
<td>2 only</td>
<td>No</td>
<td>Zip text files(^2) and in Ocean Data View format(^5)</td>
</tr>
<tr>
<td>Subset of synthesis data set (i)</td>
<td>2 as default; 2, 3, 4 upon request</td>
<td>No</td>
<td>Text files via Cruise Data Viewer(^4)</td>
</tr>
<tr>
<td>Subset of synthesis data set (ii)</td>
<td>2, 3, 4</td>
<td>Yes</td>
<td>Text files via Table of Cruises(^4)</td>
</tr>
<tr>
<td>Gridded files</td>
<td>2 only</td>
<td>No</td>
<td>NetCDF files(^5), in Ocean Data View format(^1), and via Gridded Data Viewer(^6)</td>
</tr>
</tbody>
</table>

\(^1\) doi:10.1594/PANGAEA.811776, \(^2\) http://cdiac.ornl.gov/ftp/oceans/SOCATv2/;
\(^3\) http://odv.awi.de/en/data/ocean/socat/_fCO2_data, \(^4\) http://ferret.pmel.noaa.gov/SOCAT2_Cruise_Viewer/;
\(^6\) http://ferret.pmel.noaa.gov/SOCAT_gridded_viewer/

As in version 1 (Sect. 3.1 in Pfeil et al., 2013), most surface water $\text{CO}_2$ values have been measured by equilibration of a headspace with seawater and subsequent analysis of the $\text{CO}_2$ content of the headspace. Historical measurements generally used gas chromatographic analysis, while more recent measurements are based on infrared detection. SOCAT versions 1 and 2 include a small number of historical, discrete surface water $\text{f/CO}_2$ measurements. SOCAT products do not include $\text{f/CO}_2$ calculated from other carbon parameters, such as pH, alkalinity or dissolved inorganic carbon. A small percentage of the $\text{f/CO}_2$ values (0.2% in version 2) is from measurements by a spectrophotometric method using a pH-sensitive dye (Table 7) (e.g. Hood et al., 1999).

2.2.2 Data entry

The data were assembled in a uniform file format, as in version 1 (Sect. 3.2 in Pfeil et al., 2013). Key differences in data entry between versions 1 and 2 relate to the time stamp, version numbering and an expocode for moorings and drifters, as described in Sect. 2.2.3.

Primary quality control was carried out at this stage. Primary quality control included identification of basic problems in the data, for example unrealistic positions, times and orders of magnitude. Additional basic problems were identified during secondary quality control (Sect. 2.3).

2.2.3 Key differences with version 1 in data entry

**Time stamp and version numbering:** The time stamp for SOCAT version 1 products did not contain seconds (Table 1) (Pfeil et al., 2013). In some cases this resulted in multiple entries for a given time stamp. Such multiple entries were averaged in the synthesis files (version 1.5), but not in the individual data set files (version 1.4). Two version numbers (version 1.4 and 1.5) highlight the different treatment of multiple entries for the same time stamp in the version 1 data products (Table 1).

SOCAT version 2 products include seconds, as reported by the data contributor, in the time stamp for all new and updated data sets (Table 1). However, a time stamp including seconds is not available for version 1 data sets. For these
Table 3. Content of the individual data set files (IF) and the synthesis files in SOCAT version 2. The global synthesis product is available as zip text files at CDIAC (ZIP) and in Ocean Data View (ODV) format. Subsets of the global synthesis data set can be created via the Cruise Data Viewer for All Variables (CDV_AV), Current Variable (CDV_CV) and via the Table of Cruises (CDV_TC). The first column (“Notation”) lists column headers for the parameters in the files.

<table>
<thead>
<tr>
<th>Notation</th>
<th>IF</th>
<th>ZIP</th>
<th>CDV_AV</th>
<th>CDV_CV</th>
<th>ODV</th>
<th>CDV_TC</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expocode</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>12-character expocode</td>
<td></td>
</tr>
<tr>
<td>SOCAT_Doi</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>Digital object identifier for the individual data set and metadata</td>
<td></td>
</tr>
<tr>
<td>QC_ID</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>2</td>
<td>Data set quality control flag with value 1 11 for A, 12 for B, 13 for C and 14 for D</td>
<td></td>
</tr>
<tr>
<td>Date/Time</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>Data set quality control flag with value 1 11 for A, 12 for B, 13 for C and 14 for D</td>
<td></td>
</tr>
<tr>
<td>yr</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Year (UTC)</td>
<td></td>
</tr>
<tr>
<td>mon</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Month (UTC)</td>
<td></td>
</tr>
<tr>
<td>day</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Day (UTC)</td>
<td></td>
</tr>
<tr>
<td>hh</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Hour (UTC)</td>
<td></td>
</tr>
<tr>
<td>mm</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Minute (UTC)</td>
<td></td>
</tr>
<tr>
<td>ss</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Seconds (may include decimals)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>−</td>
<td>−</td>
<td>2</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>Hours since 1970</td>
<td></td>
</tr>
<tr>
<td>Day of Year</td>
<td>−</td>
<td>−</td>
<td>2</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>Day of Year (UTC) with 1 January 00:00 as 1.0</td>
<td></td>
</tr>
<tr>
<td>Longitude</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>Longitude (0 to 360)</td>
<td></td>
</tr>
<tr>
<td>Latitude</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Latitude (−90 to 90)</td>
<td></td>
</tr>
<tr>
<td>Sample_depth/Depth water</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Water sampling depth</td>
<td></td>
</tr>
<tr>
<td>Sal</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>Water sampling depth</td>
<td></td>
</tr>
<tr>
<td>Temp/SST</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Sea surface temperature</td>
<td></td>
</tr>
<tr>
<td>Tequ</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Equilibrator chamber temperature</td>
<td></td>
</tr>
<tr>
<td>PPPP</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>Atmospheric pressure</td>
<td></td>
</tr>
<tr>
<td>Pequ</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>Equilibrator chamber pressure</td>
<td></td>
</tr>
<tr>
<td>WOA_SSS/Sal interp</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>Salinity from WOA (2005)</td>
<td></td>
</tr>
<tr>
<td>NCEP_SLP/PPPP interp</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>−</td>
<td>Salinity from WOA (2005)</td>
<td></td>
</tr>
<tr>
<td>ETOPO2_depth/Bathy depth interp</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>m</td>
<td>Bathymetry interpolated from ETOPO2 (2006)</td>
<td></td>
</tr>
<tr>
<td>Distance/d2l</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>Distance to major land mass</td>
<td></td>
</tr>
<tr>
<td>GVCO2/xCO2_air_interpol</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>µmol mol⁻¹</td>
<td>µmol mol⁻¹</td>
<td>Atmospheric xCO2 from GLOBALVIEW-CO2 (2012)</td>
<td></td>
</tr>
<tr>
<td>xCO2/water_equ_dry</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>µatm xCO2 (water) at equilibrator temperature (dry air)</td>
<td></td>
</tr>
<tr>
<td>fCO2/water_SST_wet</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>µatm</td>
<td>µatm</td>
<td>µatm xCO2 (water) at sea surface temperature (air at 100 % humidity)</td>
<td></td>
</tr>
<tr>
<td>pCO2/water_SST_wet</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>µatm</td>
<td>µatm</td>
<td>µatm pCO2 (water) at sea surface temperature (air at 100 % humidity)</td>
<td></td>
</tr>
<tr>
<td>xCO2/water_SST_dry</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>µatm</td>
<td>µatm</td>
<td>µatm xCO2 (water) at sea surface temperature (dry air)</td>
<td></td>
</tr>
<tr>
<td>fCO2/water_equ_wet</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>µatm</td>
<td>µatm</td>
<td>µatm xCO2 (water) at equilibrator temperature (air at 100 % humidity)</td>
<td></td>
</tr>
<tr>
<td>µCO2/water_equ_wet</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>µatm</td>
<td>µatm</td>
<td>µatm pCO2 (water) at equilibrator temperature (air at 100 % humidity)</td>
<td></td>
</tr>
<tr>
<td>µCO2/water_rec_wet</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>µatm</td>
<td>µatm</td>
<td>µatm µCO2 (water) at equilibrator temperature (air at 100 % humidity)</td>
<td></td>
</tr>
<tr>
<td>µCO2rec/µCO2/water_SST_wet</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>µatm</td>
<td>µatm</td>
<td>µatm µCO2 rec calculated following the SOCAT protocol (0: not generated; algorithm 1–14 in Table 8)</td>
<td></td>
</tr>
<tr>
<td>µCO2rec/srch/Algorithm</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>✓</td>
<td>µatm</td>
<td>µatm</td>
<td>µatm µCO2 rec calculated following the SOCAT protocol (0: not generated; algorithm 1–14 in Table 8)</td>
<td></td>
</tr>
<tr>
<td>µCO2rec/flag</td>
<td>✓</td>
<td>✓</td>
<td>2</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>µCO2 rec calculated following the SOCAT protocol (0: not generated; algorithm 1–14 in Table 8)</td>
<td></td>
</tr>
</tbody>
</table>

1 Data reported by the data originator
2 Available upon selection of parameter.
3 If the intake depth has not been reported by the data originator, an intake depth of 5 m has been assumed.
5 Atmospheric pressure interpolated from the NCEP/NCAR (National Centers for Environmental Prediction/National Center for Atmospheric Research) 40-Year Reanalysis Project on a 6-hourly, global, 2.5° latitude by 2.5° longitude grid (Kalnay et al., 1996), available at: http://www.esrl.noaa.gov/psd/gridded/data.ncep.reanalysis.surface.html (last access: 1 May 2013).
7 Individual data set files contain all µCO2 rec data. Synthesis files at CDIAC and via ODV only contain µCO2 rec data with a WOCE flag of 2 (Table 2).
Table 4. Activities and key participants in SOCAT versions 2 and 3 to date. Regional group leads are in Table 5.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Key Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global group for coordination</td>
<td>Bakker (chair), Hankin, Kozyr, Metzl, Olsen, Pfeil, Telszewski</td>
</tr>
<tr>
<td>Data retrieval, data entry, (re-)calculation of fCO₂</td>
<td>Pfeil, Olsen</td>
</tr>
<tr>
<td>Live Access Server</td>
<td>Hankin, O’Brien, Smith</td>
</tr>
<tr>
<td>Individual data set files, synthesis products and gridded products</td>
<td>Pfeil, Smith, Manke, Hankin</td>
</tr>
<tr>
<td>Ocean Data View</td>
<td>Schlitzer</td>
</tr>
<tr>
<td>Matlab files</td>
<td>Pierrot, Landschützer</td>
</tr>
<tr>
<td>SOCAT website</td>
<td>Pfeil</td>
</tr>
<tr>
<td>Data archiving and online access</td>
<td>Pfeil, Sieger, Kozyr, Smith, Manke, Hankin</td>
</tr>
<tr>
<td>Meetings</td>
<td>Alin, Bakker, Hales, Hankin, Nojiri, Telszewski</td>
</tr>
<tr>
<td>Alternative sensors (version 3)</td>
<td>Wanninkhof, Steinhoff, Bakker, Bates, Boutin, Olsen, Sutton</td>
</tr>
<tr>
<td>Automation (versions 3 to 4)</td>
<td>Hankin, S. Jones, Kozyr, O’Brien, Pfeil, Smith, Bakker, Olsen, Schweitzer</td>
</tr>
</tbody>
</table>

data sets, seconds were added artificially to the time stamp to avoid the problem of multiple entries. The next version of SOCAT will include seconds, as reported by the data contributor, for all data sets.

The CO₂ measurements for a dozen historical data sets are listed at midnight or their time stamp in fractional days contains insufficient decimals for retrieving minutes and seconds. Artificial seconds, in some cases including tenths or hundreds of a second, were generated for these valuable data, such that they can remain in SOCAT version 2. Every effort will be made to retrieve a more adequate time stamp for these data sets for future versions. Unlike version 1, which has version 1.4 and 1.5 data products, version 2 data products have a single version number (Table 1).

Excode for moorings and drifters: SOCAT uses twelve character expcodes (Swift, 2008) as stable and unique data set identifiers. For example, 49P120101218 indicates a cruise on the Japanese (49) ship of opportunity Pyxis (P1) with the first day of the cruise on 18 December 2010. In contrast to version 1, expcodes have been assigned for moorings and drifters in version 2, by registering a “vessel code” at International Council for the Exploration of the Sea (ICES) in collaboration with the National Oceanographic Data Center (NODC) and the British Oceanographic Data Centre (BODC) (Table 1).

2.2.4 (Re-)calculation of recommended fCO₂

Surface water fCO₂ values at sea surface temperature (or intake temperature), also known as recommended fCO₂ (fCO₂rec), have been recalculated, analogous to version 1 (Sect. 3.3 in Pfeil et al., 2013). A single set of equations and a strict order of preference for the CO₂ input parameter has been used (Table 8) (Pfeil et al., 2013). Six different CO₂ parameters were reported in the original data files, notably xCO₂, pCO₂ and fCO₂, either at the equilibration temperature (Tequ) or at the sea surface temperature (SST).

The (re-)calculation procedure of fCO₂ has the following philosophy (Pfeil et al., 2013):

1. Whenever possible, (re-)calculate fCO₂;
2. The favourite starting point for the calculations is xCO₂, next pCO₂, followed by fCO₂;
3. Minimise the amount of external data required for the calculations.

Table 8 lists surface water CO₂ parameters and ancillary parameters used for calculation of recommended fCO₂ in version 2 in order of preference with algorithm 1 as the favourite (analogous to Table 4 in Pfeil et al., 2013). The algorithm is provided in the output files (Table 3). Equations recommended by Dickson et al. (2007) have been used for the conversion of the dry CO₂ mole fraction to pCO₂, for the calculation of the water vapour pressure and for the conversion of pCO₂ to fCO₂, similar to version 1 (Sect. 3.3 of Pfeil et al., 2013). As in version 1, the correction of Takahashi et al. (1993) has been applied for temperature change between the seawater intake and the site of equilibration:

\[ f\text{CO}_2^{\text{SST}} = f\text{CO}_2^{\text{equiv}} \exp(0.0423(\text{SST} - \text{Tequ})). \] (1)

Climatological values of salinity and atmospheric pressure from reanalysis have been used in the calculation of recommended fCO₂ (Table 8), if the data contributor did not report in situ salinity and atmospheric pressure, following Pfeil et al. (2013).
Table 5. Regions and regional group leads in SOCAT version 2 (Fig. 5).

<table>
<thead>
<tr>
<th>Region</th>
<th>Definition</th>
<th>Lead(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal Seas</td>
<td>Less than 400 km from land; between 30° S and 70° N for 100° W to 43° E;</td>
<td>Alin, Cai, Hales</td>
</tr>
<tr>
<td></td>
<td>between 30° S and 66° N elsewhere, including coastal waters</td>
<td></td>
</tr>
<tr>
<td>Arctic Ocean</td>
<td>North of 70° N for 100° W to 43° E; north of 66° N elsewhere</td>
<td>Mathis</td>
</tr>
<tr>
<td>North Atlantic</td>
<td>30° N to 70° N</td>
<td>Schuster</td>
</tr>
<tr>
<td>Tropical Atlantic</td>
<td>30° N to 30° S</td>
<td>Lefèvre</td>
</tr>
<tr>
<td>North Pacific</td>
<td>30° N to 66° N</td>
<td>Nojiri</td>
</tr>
<tr>
<td>Tropical Pacific</td>
<td>30° N to 30° S</td>
<td>Cosca</td>
</tr>
<tr>
<td>Indian Ocean</td>
<td>North of 30° S</td>
<td>Sarma</td>
</tr>
<tr>
<td>Southern Ocean</td>
<td>South of 30° S, including coastal waters</td>
<td>Tilbrook, Metzl</td>
</tr>
</tbody>
</table>

Table 6. Meetings for SOCAT versions 2 and 3 to date.

<table>
<thead>
<tr>
<th>Timing</th>
<th>Meeting description</th>
<th>Location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>05/2012</td>
<td>Automation planning meeting</td>
<td>NOAA-PMEL, Seattle, USA</td>
<td>(SOCAT, 2012a)</td>
</tr>
<tr>
<td>07/2012</td>
<td>SOCAT progress meeting</td>
<td>Epochal Centre, Tsukuba, Japan</td>
<td>(SOCAT, 2012b)</td>
</tr>
<tr>
<td>10/2012</td>
<td>Coastal and Arctic SOCAT quality control workshop</td>
<td>NOAA-PMEL, Seattle, USA</td>
<td>(IOCCP, 2012)</td>
</tr>
<tr>
<td>06/2013</td>
<td>SOCAT side event at the 9th International Carbon Dioxide Conference, Public release</td>
<td>Beijing International Convention Center, Beijing, China</td>
<td>(SOCAT, 2013b)</td>
</tr>
<tr>
<td></td>
<td>of version 2.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.3 Secondary quality control in version 2

2.3.1 Secondary quality control criteria

Criteria for 2nd level quality control have been defined in a series of workshops (IOCCP, 2008, 2009, 2010; Pfeil et al., 2013). Second level quality control consists of assigning a quality control flag to each data set and a WOCE flag to individual surface water $fCO_2$ values. The criteria for quality control are identical in versions 1 (Sect. 4.1 in Pfeil et al., 2013) and 2.

Only data sets with a quality control flag of A, B, C and D are included in SOCAT version 1 and 2 data products (Table 9) (Pfeil et al., 2013). The data set quality control flags (formerly known as “cruise flags”) in versions 1 and 2 have been developed for automated shipboard measurement of surface water $fCO_2$, mainly by infrared detection and frequent at sea standardisation using calibration gases with a range of $CO_2$ concentrations (IOCCP, 2008; Pfeil et al., 2013). Much weight is put on whether approved methods or standard operating procedures (SOP) (AOML, 2002; Dickson et al., 2007; Pierrot et al., 2009) were followed by making this a prerequisite for flags of A and B. Citing Pfeil et al. (2013):

“Seven SOP criteria need to be fulfilled for a cruise (or data set) flag of A or B in SOCAT:

1. The data are based on $xCO_2$ analysis, not $fCO_2$ calculated from other carbon parameters, such as pH, alkalinity or dissolved inorganic carbon;
2. Continuous $CO_2$ measurements have been made, not discrete $CO_2$ measurements;
3. The detection is based on an equilibrator system and is measured by infrared analysis or gas chromatography;
4. The calibration has included at least 2 non-zero gas standards, traceable to World Meteorological Organization (WMO) standards;
5. The equilibrator temperature has been measured to within 0.05 °C accuracy;
6. The intake seawater temperature has been measured to within 0.05 °C accuracy;
7. The equilibrator pressure has been measured to within 0.5 hPa accuracy.”

The $fCO_2$ values from data sets with flags of A and B are judged to have an accuracy of ±2 μatm or better. Criterion 1
also needs to be met for flags of C and D, similar to version 1 (Sect. 4.1 in Pfeil et al., 2013). Complete metadata documentation is required for data set quality control flags of A, B and C. Comparison to other data is carried out, if possible. The overall quality of the data needs to be deemed acceptable for flags of A, B, C and D (Table 9) (Pfeil et al., 2013).

The Southern and Indian Ocean groups (Table 5) have applied three additional quality control criteria for the temperature change between the seawater intake and the equilibrator in versions 1 and 2 (IOCCP, 2010), citing Pfeil et al. (2013):

- “Warming should be less than 3 °C;
- Warming rate should be less than 1 °C h⁻¹, unless a rapid temperature front is apparent;
- Warming outliers should be less than 0.3 °C, compared to background data.”

In addition:

- Cooling between the seawater intake and the equilibrator is unlikely in high-latitude oceans for an indoor measurement system;
- Zero or constant temperature change may indicate absence of SST values.

The above five guidelines have been applied widely in version 2 for open ocean data away from sea ice, as part of a systematic search for unrealistic data and outliers (Sect. 2.3.3). Such a systematic search has not been carried out for version 1 (Table 1).

These quality control criteria (Table 9) have also been applied for quality control of surface water CO₂ measurements from moorings and drifters in versions 1 and 2 (Table 7).

Individual fCO₂ values are assignedWOCE flags: 2 (good), 3 (questionable) or 4 (bad) with 2 being the default setting (Sect. 4.1.3 in Pfeil et al., 2013). Outliers in parameters required for the timing, position and (re-)calculation of fCO₂ values are given flags of 3 and 4. Thus, flags of 3 and 4 might indicate an erroneous time or position stamp, an unrealistic seawater temperature, strong warming between the seawater intake and the equilibrator or a large pressure difference between the equilibrator and the atmosphere. Data sets with a large number of flags of 3 and 4 (> 50, as a guide line) are suspended, as was also the case for version 1 (Pfeil et al., 2013).

2.3.2 Secondary quality control in practice

Secondary quality control for version 2 has been carried out by 24 marine carbon scientists from eight countries (Fig. 4, Table 4). Quality control in SOCAT is carried out by groups organised according to region. These regions have been operationally defined and do not necessarily follow common oceanographic definitions. Regions for version 2 are the Coastal Seas, the North Atlantic, Tropical Atlantic, North Pacific, Tropical Pacific, Indian Ocean and Southern Oceans and a newly defined Arctic region (Fig. 5, Table 5). The Arctic region includes both coastal seas and the deep ocean. It encompasses all waters north of 70° N for 100° W to 43° E (Atlantic sector) and north of 66° N elsewhere (Table 1) (Sect. 2.3.3) (SOCAT, 2012b).

The regional group members assign data set quality control flags, WOCE flags and enter quality control comments. The Live Access Server (http://ferret.pmel.noaa.gov/LAS) is used for quality control, as in version 1 (Sect. 4 in Pfeil et al., 2013). Quality control comments are available via the Table of Cruises on the Cruise Data Viewer (Table 2) (Sect. 2.4.5).

All new and updated data sets are subject to quality control. Each data set is assigned a flag of A, B, C, D, S (Suspend) or X (Exclude) for each region it crosses. As a final step the quality controllers need to resolve any “conflicting”
data set flags between regions and decide on the “agreed” flag for a data set. The data set quality control flag has been added as a parameter in the synthesis files (Tables 1 and 3) (Sects. 2.4.4 and 2.4.7).

Overall data quality and reporting of metadata has improved from version 1 to version 2, which we attribute to the SOCAT effort. In version 1, 41% of the data sets were assigned a flag of A or B, 22% obtained a flag of C and, 37% received a flag of D. Version 2 has a larger proportion of data sets with flags of A or B (48%) and smaller proportions of data sets with a flag of C (18%) and D (33%).

2.3.3 Key differences with version 1 in secondary quality control

This section identifies key differences in secondary quality control between versions 1 and 2.

Creation of an Arctic regional designation: in version 1 Arctic data were part of the North Pacific and North Atlantic Oceans and the Coastal Region (Sect. 2.2 in Pfeif et al., 2013). Given the importance of Arctic research and the rapid increase in the quantity of Arctic fCO2 values, an Arctic region has been defined for version 2 (Figs. 5 and 6; Table 1) (Sect. 2.3.2) (SOCAT, 2012b).

Identification of unrealistic values: in version 2 a systematic search has been carried out for unrealistic values or patterns in all data relevant for the timing, position or (re-)calculation of fCO2 (Sect. 2.3.1). This activity considered all the data sets submitted to version 2, including data sets previously included in the version 1 data release. The search applied to the ship’s cruise track, position, time, atmospheric pressure, equilibrator pressure, salinity, sea surface temperature, equilibrator temperature, and temperature change between the seawater intake and the equilibrator. This helped locate problems with data entry, e.g. overlap between data sets, reversal of hours and minutes, reversal of SST and salinity, presence of undefined values (e.g. −999, −99, −9.999, −9.99, −9.9, −9.0), identification of unrealistic values (e.g. an atmospheric pressure of 780 mbar) and problems with water flow (absolute temperature change between intake and equilibrator exceeding 3 °C). Depending on the nature of

Table 7. Drifters and time series in SOCAT version 2 with the location, year(s) of operation, platform type, CO2 instrument type, algorithm used for (re-)calculation of fCO2 (Table 8), number of data sets, number of fCO2 values with a WOCE flag of 2, and reference.

<table>
<thead>
<tr>
<th>Drifters and time series</th>
<th>Location</th>
<th>Year(s)</th>
<th>Platform type</th>
<th>CO2 instrument type</th>
<th>Algorithm</th>
<th>Number of data sets</th>
<th>Number of fCO2 values</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CARIOTA</td>
<td>75.0° N 3.0° W</td>
<td>1996–1997</td>
<td>Drifter</td>
<td>Membrane spectrophotometer</td>
<td>6</td>
<td>1</td>
<td>2668</td>
<td>H1999</td>
</tr>
<tr>
<td>CARIOTA</td>
<td>40.1° S 15.8° E</td>
<td>2005</td>
<td>Drifter</td>
<td>Membrane spectrophotometer</td>
<td>6</td>
<td>1</td>
<td>1451</td>
<td>BM2009</td>
</tr>
</tbody>
</table>

Figure 5. Quality control regions for SOCAT version 2 (Table 5). White shading corresponds to the coastal region. The regions have been defined for operational reasons and do not necessarily reflect common oceanographic definitions.
Table 8. Surface water CO₂ parameters used for the calculation of recommended \( f_{\text{CO}_2} \) (\( f_{\text{CO}_2}\text{rec} \)) at sea surface temperature in version 2 (after Table 4 in Pfeil et al., 2013). The parameters are listed in order of preference (with algorithm 1 as the favourite). The algorithm is provided in the output files (Table 3). In cases of incomplete data reporting, these ancillary parameters have been used for atmospheric pressure and salinity: NCEP (National Centers for Environmental Prediction) atmospheric pressure (Kalnay et al., 1996) and WOA (World Ocean Atlas) salinity (Antonov et al., 2006) (Sect. 2.2.4).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Reported CO₂ parameter</th>
<th>Unit</th>
<th>Data Percentage (%)</th>
<th>Extra variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>xCO₂\text{water_equi_dry}</td>
<td>( \mu \text{mol} \text{ mol}^{-1} )</td>
<td>66.7</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>xCO₂\text{water_SST_dry}</td>
<td>( \mu \text{mol} \text{ mol}^{-1} )</td>
<td>4.5</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>pCO₂\text{water_equi_wet}</td>
<td>( \mu \text{atm} )</td>
<td>4.5</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>pCO₂\text{water_SST_wet}</td>
<td>( \mu \text{atm} )</td>
<td>2.6</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>fCO₂\text{water_equi}</td>
<td>( \mu \text{atm} )</td>
<td>0.2</td>
<td>–</td>
</tr>
<tr>
<td>6</td>
<td>fCO₂\text{water_SST_wet}</td>
<td>( \mu \text{atm} )</td>
<td>10.8</td>
<td>–</td>
</tr>
<tr>
<td>7</td>
<td>pCO₂\text{water_equi_wet} (^1)</td>
<td>( \mu \text{atm} )</td>
<td>0.3</td>
<td>NCEP Pressure</td>
</tr>
<tr>
<td>8</td>
<td>pCO₂\text{water_SST_wet} (^1)</td>
<td>( \mu \text{atm} )</td>
<td>8.3</td>
<td>NCEP Pressure</td>
</tr>
<tr>
<td>9</td>
<td>xCO₂\text{water_equi_dry} (^2)</td>
<td>( \mu \text{mol} \text{ mol}^{-1} )</td>
<td>0.2</td>
<td>WOA Salinity</td>
</tr>
<tr>
<td>10</td>
<td>xCO₂\text{water_SST_dry} (^2)</td>
<td>( \mu \text{mol} \text{ mol}^{-1} )</td>
<td>0.7</td>
<td>WOA Salinity</td>
</tr>
<tr>
<td>11</td>
<td>xCO₂\text{water_equi_dry} (^1)</td>
<td>( \mu \text{mol} \text{ mol}^{-1} )</td>
<td>0.01</td>
<td>NCEP Pressure</td>
</tr>
<tr>
<td>12</td>
<td>xCO₂\text{water_SST_dry} (^1)</td>
<td>( \mu \text{mol} \text{ mol}^{-1} )</td>
<td>1.0</td>
<td>NCEP Pressure</td>
</tr>
<tr>
<td>13</td>
<td>xCO₂\text{water_equi_dry} (^{1,2})</td>
<td>( \mu \text{mol} \text{ mol}^{-1} )</td>
<td>0.01</td>
<td>NCEP Pressure, WOA Salinity</td>
</tr>
<tr>
<td>14</td>
<td>xCO₂\text{water_SST_dry} (^{1,2})</td>
<td>( \mu \text{mol} \text{ mol}^{-1} )</td>
<td>0.1</td>
<td>NCEP Pressure, WOA Salinity</td>
</tr>
</tbody>
</table>

\(^1\) Atmospheric pressure was not reported in the original data file.

\(^2\) Salinity was not reported in the original data file.

In total, 154 data sets have been suspended, of which 70 had previously been included in the version 1 release (Table 1). Table 10 lists grounds for suspension of data sets. These include absence of CO₂ values (14 %), a data entry problem (10 %), use of a constant atmospheric pressure or salinity in the calculation of \( f_{\text{CO}_2} \) (45 %), absence of SST (8 %), and concerns on the quality of \( f_{\text{CO}_2} \) (3 %), temperature (14 %), or atmospheric pressure (2 %). In case of a data entry problem, data sets will be re-entered into the SOCAT quality control system for version 3. In other cases, data sets may need revision before resubmission to SOCAT. Finally, six data sets (4 %) made by a spectrophotometric method were suspended, as the data set flags of A to D were not deemed appropriate by the quality controller. In response, a new data set flag of E has been defined for use in version 3 (Sect. 4.2) (Wanninkhof et al., 2013a).

Suspension of 70 data sets included in version 1: 70 data sets previously included in version 1 were suspended from version 2 upon identification of data quality concerns (Tables 1 and 10), as discussed above. Most of these (59) were suspended as a constant atmospheric pressure had been used in the calculation of \( f_{\text{CO}_2} \). These 59 data sets have since been revised and resubmitted to SOCAT for version 3. Six data sets were suspended for a data entry problem, while three data sets lacked surface water CO₂ values. Concerns on the quality of a temperature or atmospheric pressure reading were grounds for suspension of a further two data sets.

WOCE flags of 3 and 4: WOCE flags of 2 (good), 3 (questionable) and 4 (bad) have been assigned to all \( f_{\text{CO}_2} \) values in version 2, including for data sets part of the version 1 release. During version 1 quality control, 0.2 % of the \( f_{\text{CO}_2} \) values had been assigned a flag of 3 or 4. However, these flags were accidentally reset to a flag of 2 prior to the version 1 release, such that most \( f_{\text{CO}_2} \) values in version 1 were reported with a flag of 2. The initial flags of 3 and 4 set during version 1 quality control have been reinstated in version 2. In version 2, a total of 20 850 \( f_{\text{CO}_2} \) values (0.2 %) has been given a flag of 3 or 4.

2.4 Version 2 data policy and data products

2.4.1 Data policy

Users of the SOCAT data products are requested to do the following (SOCAT, 2013a, b):

1. Recognise the contribution of SOCAT data contributors and quality controllers in the form of invitation to co-authorship or citation of relevant scientific articles by data contributors;

2. Cite all SOCAT data products by reference to publications documenting SOCAT.
Table 9. Criteria for assigning data set quality control flags based on the expected quality of the recommended $f$CO$_2$ values (per Table 6 in Pfeil et al., 2013). All criteria need to be met for assigning a data set flag. Only data sets with a quality control flag of A, B, C and D are included in version 1 and 2 data products. SOP is Standard Operating Procedures (Dickson et al., 2007); QC is quality control.

<table>
<thead>
<tr>
<th>Data set flag (ID)</th>
<th>Criteria</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A (11)</td>
<td>1. Followed approved methods or SOP criteria and 2. Metadata documentation complete and 3. Extended QC was deemed acceptable and 4. A comparison with other data was deemed acceptable.</td>
<td></td>
</tr>
<tr>
<td>B (12)</td>
<td>1. Followed approved methods or SOP criteria and 2. Metadata documentation complete and 3. Extended QC was deemed acceptable.</td>
<td></td>
</tr>
<tr>
<td>C (13)</td>
<td>1. Did not follow approved methods or SOP criteria but 2. Metadata documentation complete and 3. Extended QC was deemed acceptable (including comparison with other data if possible).</td>
<td></td>
</tr>
<tr>
<td>D (14)</td>
<td>1. Did or did not follow approved methods or SOP criteria and 2. Metadata documentation incomplete but 3. Extended QC was deemed acceptable (including comparison with other data if possible).</td>
<td></td>
</tr>
<tr>
<td>S (suspend)</td>
<td>1. Did or did not follow methods or SOP criteria and 2. Metadata documentation complete or incomplete and 3. Extended QC revealed non-acceptable data but 4. Data are being updated.</td>
<td></td>
</tr>
<tr>
<td>X (15) (Exclude)</td>
<td>The data set duplicates another data set in SOCAT.</td>
<td></td>
</tr>
<tr>
<td>N (No flag)</td>
<td>No data set flag has yet been given to this data set.</td>
<td></td>
</tr>
<tr>
<td>U (Update)</td>
<td>The data set has been updated. No data set flag has yet been given to the revised data.</td>
<td></td>
</tr>
</tbody>
</table>

3. Send references of publications using SOCAT products to submit@socat.info.

2.4.2 SOCAT data products

The SOCAT data products provide access to recommended surface ocean $f$CO$_2$ values in a uniform format for the global oceans and coastal seas. Three different SOCAT data products are available: individual data set files, synthesis files and gridded files. User-defined subsets of the synthesis files are available via the Cruise Data Viewer. The Gridded Data Viewer facilitates querying of the gridded data products. All data products can be accessed via the SOCAT website (http://www.socat.info/) or via the web-links provided below. Table 2 identifies the key characteristics of the SOCAT data products, while Table 3 lists the contents of downloadable files. The version 2 data products resemble those for version 1 (Pfeil et al., 2013; Sabine et al., 2013), apart from further standardisation and extra parameters. The data products and tools are discussed below, followed by a description of key differences between version 1 and 2 (Table 1) (Sect. 2.4.7).

2.4.3 Individual data set files

Individual data set files provide surface water $f$CO$_2$, the parameters used to (re-)calculate $f$CO$_2$ and the original CO$_2$ parameter(s) reported by the data contributor for data sets with a flag of A, B, C or D (Table 2). The files include all surface water $f$CO$_2$ values with WOCE flags of 2, 3 and 4. Individual data set files are archived at PANGAEA® (doi:10.1594/PANGAEA.811776). Each data set has a digital object identifier (doi) (Table 3). Metadata provided by the data contributor accompany the data set files. As in version 1, the individual data set and synthesis files include the climatological values of salinity and atmospheric pressure from reanalysis. The files also contain values for the water depth, the distance to a major land mass and the atmospheric CO$_2$ mole fraction interpolated from GLOBALVIEW-CO2 (2012). Via PANGAEA®, version 2 is made available to the World Data System (WDS) of the International Council for Science (ICSU), to the Group of Earth Observations (GEO) Portal and to the Global Earth Observation System of Systems (GEOSS).
Table 10. Grounds for suspension of data sets from SOCAT version 2. A distinction is made between data sets previously included in version 1 and new data sets in version 2. Abbreviations are SST for sea surface temperature, Tequ for equilibrator temperature and dT for the difference between the equilibrator temperature and the sea surface temperature.

<table>
<thead>
<tr>
<th>Ground for suspension</th>
<th>Number data sets version 1</th>
<th>Number data sets version 2</th>
<th>Percentage of total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap with other data set</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Data entry problem (incomplete data set, time, position, SST, salinity)</td>
<td>6</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Constant atmospheric pressure in calculation of fCO2rec</td>
<td>59</td>
<td>5</td>
<td>42</td>
</tr>
<tr>
<td>Constant salinity (0 or −999) in calculation of fCO2rec</td>
<td>0</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>No xCO2, pCO2 or fCO2 reported</td>
<td>3</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>No SST reported</td>
<td>0</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Concerns on quality of fCO2rec</td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Concerns on quality of SST, Tequ or dT</td>
<td>1</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Concerns on quality of atmospheric pressure</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>No appropriate sensor flag</td>
<td>0</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>84</td>
<td>100</td>
</tr>
</tbody>
</table>

2.4.4 Global synthesis product

The global synthesis data set consists of individual data sets with flags of A, B, C and D and contains fCO2 values with a WOCE flag of 2 (Table 2). The synthesis files do not contain the original CO2 values (Table 3). Each line in the files lists the doi-number of the corresponding individual data set file at PANGAEA® (Sect. 2.4.3). The synthesis data set is available as global and regional files for the SOCAT regions (Fig. 5). The regional files only contain data from within that region, so that data from most cruises are split between several regional files. The global and regional files are publicly available as compressed zip text files via CDIAC (ftp://cdiac.ornl.gov/ftp/oceans/SOCATv2/). Matlab code is available for reading these synthesis files. The global synthesis product is also available in Ocean Data View format (ftp://odv.awi.de/en/data/ocean/socat_fCO2_data).

2.4.5 Subsetting the global synthesis product

The Cruise Data Viewer (http://ferret.pmel.noaa.gov/SOCAT2_Cruise_Viewer/), an interactive tool on the Live Access Server, enables searching and subsetting the global synthesis data set by year, month, day, region, parameter, expocode, cruise name, vessel, and data set quality control flag. One may define search limits, for example salinity below 32. The user can create property-property plots and download data. The default setting is access to fCO2 values with a WOCE flag of 2 (Table 2). However, the user can include data with flags of 3 (questionable) and 4 (bad) by selecting “Include SOCAT invalids”. Figures 2 and 7 have been made with the Cruise Data Viewer.

The Table of Cruises, available via the pull-down menu “Tables” on the Cruise Data Viewer, enables the user to find metadata and read quality control comments for specific data sets (Table 2). Files downloadable via the Table of Cruises contain fCO2 values with WOCE flags of 2, 3 and 4 and the original CO2 data (Table 3).

2.4.6 Gridded products

Sabine et al. (2013) detail the gridding of the fCO2 values on a 1° latitude by 1° longitude grid with a higher 0.25° latitude by 0.25° longitude resolution product for the coastal seas in version 1. The procedures for gridding the data are identical in versions 1 and 2.

Several gridded products of surface ocean fCO2 means with minimal interpolation are available (doi:10.3334/CDIAC/OTGSOCAT_V2_GRID). Surface water fCO2 values with a flag of 2 have been put on a 1° latitude by 1° longitude grid in four ways: per year, monthly per year, monthly per decade, and per climatological month from 1970 to 2011 (Table 2). A higher resolution of 0.25° latitude by 0.25° longitude is available as monthly means per year for the coastal region (Fig. 5).

Gridded fCO2 values are reported as unweighted means and as cruise-weighted (or data set-weighted) means (Sabine et al., 2013). In an unweighted mean all the recommended fCO2 values in a grid cell have been given equal weight in calculating the mean. In a cruise-weighted mean, first averages of the recommended fCO2 values per cruise (or data set) have been calculated within a grid cell, before averages of the cruise means have been determined. Grid cells without fCO2 values are empty. No correction has been made for the expected long-term increase in surface water fCO2.
something users of the monthly climatological and decadal gridded products should keep in mind. Furthermore, the gridded products may have a temporal bias in grid cells with uneven temporal data coverage. For example, an annual gridded product may have a strong seasonal bias, if only summertime $f$CO$_2$ values are available.

Gridded $f$CO$_2$ products can be accessed as NetCDF files from CDIAC (http://cdiac.ornl.gov/ftp/oceans/SOCATv2/SOCATv2_Gridded_Dat/), in Ocean Data View format (http://odv.awi.de/en/data/ocean/socat_fCO2_data) and via the Gridded Data Viewer (http://ferret.pmel.noaa.gov/SOCAT_gridded_viewer/) (Table 2). Matlab code is available for reading the NetCDF files.

The capabilities of the Gridded Data Viewer have been expanded in version 2. The number of different years is a new variable in the monthly climatological gridded data set. The Gridded Data Viewer now shows the 400 km continental margin mask at 1 min resolution used for defining the Coastal Region (Table 5) and the distance to the nearest major land mass from 0 to 1000 km at 20 min resolution. The Gridded Data Viewer has an option for animation of gridded products. The interface has a new comparison capability for up to four gridded data sets. This enables the user to visualise, for example, gridded data products in SOCAT versions 1 and 2 in a multiple-plot view.

### 2.4.7 Key differences with version 1 in the data products

This section identifies key differences between the data products for versions 1 and 2.

**Parameters in the individual and synthesis data set files:** The data set quality control flags of A to D have been added as numerical values 11 to 14 to the synthesis files in version 2 (Tables 1 and 3). The distance to a major land mass is a new parameter in the files. Atmospheric CO$_2$ mole fractions from the 2012 GLOBALVIEW-CO2 are reported in version 2 output files; this represents an update from the 2008 GLOBALVIEW-CO2 values which were reported for version 1. The number of parameters in the downloadable files available via the Cruise Data Viewer as “All Variables” and “Current Variable” has been strongly reduced to match those in the synthesis files at CDIAC (Tables 1 and 3).

**Gridded Data Viewer:** the number of different years has been added as a variable to the monthly climatological gridded data set (Sect. 2.4.6). Data sets for the 400 km continental margin mask and the distance to the nearest major land mass are now available. The visualisation tools of the Gridded Data Viewer have been expanded.

**Release notes:** release notes document issues identified with individual data sets or data products in version 2. The notes are available on the CDIAC (http://cdiac.ornl.gov/ftp/oceans/SOCATv2/) and SOCAT (http://www.socat.info/access.html) websites (Table 1).

### 3 Spatial and temporal data coverage

SOCAT version 2 includes surface ocean $f$CO$_2$ values collected between 1968 and 2011 for the global oceans and coastal seas (Figs. 1 and 2). Data availability has increased over time for most ocean regions (Fig. 1b). A notable exception is the Indian Ocean, for which data are available from the 1990s, but where few subsequent observations have been made. Marked increases in data collection are apparent in the Gulf of Mexico (not shown) and the Arctic Ocean (Fig. 1b).
The number of data sets (colour bar on top of subplots) with surface water \( fCO_2 \) values in each 1° latitude by 1° longitude grid cell from 1970 to 2011 in SOCAT version 2. Figure 5 updates a similar figure for version 1 in Sabine et al. (2013, Fig. 5).

Figure 8. The number of (a) months of the year and (b) total months with surface water \( fCO_2 \) values in each 1° latitude by 1° longitude grid cell from 1970 to 2011 in SOCAT version 2. Figure 8a updates a similar figure for version 1 in Sabine et al. (2013, Fig. 5).

For example, version 2 has a total of 40 data sets in the Arctic Ocean, of which 10 data sets were collected in 2011 alone. Data coverage remains sparse in large parts of the Southern Hemisphere oceans (Fig. 2).

On average 3.4 surface water \( fCO_2 \) values have been collected per 100 km\(^2\) between 1968 and 2011 in the global oceans and coastal seas. Data density ranges widely from 0.8 \( fCO_2 \) values per 100 km\(^2\) in the Indian Ocean to 6.7 values per 100 km\(^2\) in the North Atlantic Ocean (Fig. 6). Data density in the Southern Ocean appears somewhat high with 2.6 values per 100 km\(^2\) relative to the North Pacific, the Tropical Pacific and the Tropical Atlantic Oceans. However, the Southern Ocean includes coastal waters with higher than average data density (Fig. 5), while the other three open ocean regions do not. Five of the ten most “productive” ships in terms of data collection are active south of 30°S, notably the Nathaniel B. Palmer, the Laurence M. Gould, the Southern Surveyor, the Polarstern and the Aurora Australis (Fig. 3).

The seasonal distribution of surface water \( fCO_2 \) values is shown in Fig. 7 for the period 2000 to 2009. The maps demonstrate the near absence of wintertime data in the high-latitude regions. The Ross Sea (Southern Ocean) has about 20 months of observations spanning five months from autumn to spring (Fig. 8).

The installation of automated \( fCO_2 \) systems on ships of opportunity and Antarctic supply ships has greatly improved the data availability along shipping routes and including for coastal seas near major ports (Fig. 9). For example, between 2000 and 2009 more than 40 individual ship visits have been made to the 1° latitude by 1° longitude grid boxes in the Florida Straits, the English Channel, off the coast of Japan and near the Antarctic Peninsula.

The numbers of unique months and total months with \( fCO_2 \) values per 1° latitude by 1° longitude grid cell shed light on data collection activities for 1970 to 2011 (Fig. 8). High data density along shipping routes highlights the repeated \( fCO_2 \) observations. For example, numerous grid boxes east of Japan have observations in all months of the year for about 50 months in total, reflecting an intense \( CO_2 \) observational effort over a large number of years. This ongoing data collection effort is critical for the quantification of the variability and trends in \( CO_2 \) air–sea exchange.

4 Future plans

4.1 Progress towards version 3

Surface water \( CO_2 \) values and accompanying metadata can be submitted to CDIAC in the IOCCP-recommended formats (http://cdiac.ornl.gov/oceans/submit.html) at all times. Ideally data are submitted as they become available. The SOCAT global group sets deadlines for consideration of data in specific SOCAT versions; for example, the deadline for submission to SOCAT version 3 was 28 February 2014. Version 3 quality control is scheduled to take place during the...
summer and autumn of 2014 with the release of version 3 planned for mid-2015.

4.2 Quality control flags for alternative sensors on a range of platforms

The SOCAT data quality control flags have been primarily designed for shipboard, continuous surface water CO$_2$ measurements by gas chromatography or infrared detection (Pfeil et al., 2013). Since the definition of these flags, high-precision and stable cavity ring-down spectroscopy (CRDS) has become available for surface water CO$_2$ measurements (Friedrichs et al., 2010; Becker et al., 2012). The quality control criteria in SOCAT are deemed adequate for the measurements by CRDS. Measurements made by CRDS will be included in future SOCAT versions, provided calibrations have been carried out at least daily (SOCAT, 2012b).

The quality control criteria, as used in versions 1 and 2, need revision for fCO$_2$ values from sensors on surface moorings, surface drifters and self-propelled surface vehicles (SOCAT, 2012b). These measurements do not follow all the standard operating procedures and at-sea calibration of such fCO$_2$ measurements is often infrequent or non-existent. Also, the sensors tend to use fewer gas standards than on ships due to logistical and power constraints. A working group on alternative sensors (Table 4) has developed a vision on how to include such fCO$_2$ values, as measured for example by infrared analysis and spectrophotometry, in future SOCAT versions (Wanninkhof et al., 2013a). The working group has determined which quality control criteria should apply to surface water CO$_2$ data from these new sensors and platforms. The term “data set quality control flag” replaces “cruise quality control flag”. The accuracy of data with data set flags of C and D has been specified as 5 µatm. A new data set quality control flag of “E” with an accuracy better than 10 µatm has been defined. The platform and the CO$_2$ instrument type will be introduced as parameters. These quality control criteria and other recommendations of the working group will be adopted for SOCAT version 3.

4.3 Automation

The large effort for data entry and quality control is a major obstacle for regular and prompt SOCAT updates, especially with more data becoming available each year. The need for automating SOCAT was formally recognised in September 2011 (SOCAT, 2011) and an automation team was created (Table 4). The automation vision proposed by the team was accepted by regional and global group leads (SOCAT, 2012a, b). The automation system will allow the data provider to upload, review and submit data and metadata. It will calculate surface water fCO$_2$. The automation system will provide a single portal for data providers, data managers and quality controllers. Manual data entry by the SOCAT data managers will be strongly reduced. Regular, prompt releases of SOCAT will be more straightforward once the automation system is fully operational. The automation system is expected to become the primary mode of data submission from version 4 onwards.

5 Scientific applications of SOCAT

Several scientific studies have used SOCAT data products. The global synthesis product is the most popular SOCAT product in scientific publications. Both files in zip text format (Lourantou and Metzl, 2011; Tjiputra et al., 2012; Nakaoka et al., 2013; Rödenbeck et al., 2013; Wanninkhof et al., 2013b) and the Ocean Data View collection (Chierici et al., 2012) are used for data access. Two studies utilise a global gridded product (Landschützer et al., 2013; Schuster et al., 2013).

Scientific applications of SOCAT include:

- Visualisation of surface ocean fCO$_2$ data coverage (Chierici et al., 2012) and data requirements (Wanninkhof et al., 2013b);
- Use of the SOCAT continental margin mask (Evans and Mathis, 2013);
- Process studies (Lourantou and Metzl, 2011);
- Creation and validation of surface water fCO$_2$ and CO$_2$ air–sea flux maps by a variety of techniques, including multiple linear regression (Schuster et al., 2013), neural network approaches (Landschützer et al., 2013; Nakaoka et al., 2013) and an ocean mixed layer model (Rödenbeck et al., 2013);
- Quantification of the annual mean ocean carbon sink (Schuster et al., 2013);
- Studies of variation in the ocean carbon sink on seasonal (Rödenbeck et al., 2013), year-to-year (Landschützer et al., 2013) and decadal timescales (Lourantou and Metzl, 2011);
- Initialisation and validation fields for ocean carbon cycle models (Tjiputra et al., 2012).

These applications highlight the importance of SOCAT for regional and global air–sea CO$_2$ flux assessments, process studies and ocean carbon modelling.

6 Conclusions

SOCAT version 2 represents a 44 yr record of surface water fCO$_2$ values from 1968 to 2011 for the global oceans and coastal seas (Figs. 1 and 2). Version 2 extends version 1 by four years, while also adding more fCO$_2$ values for the years 2006 and 2007. The data are in a uniform format and have been subject to documented quality control. The quality
of data and of data reporting has improved in version 2 relative to version 1. The temporal data distribution partly reflects activities in large international research programmes. Over time, data coverage in all ocean regions has increased, with the exception of the Indian Ocean. Data coverage has increased four-fold from the 1990s to the 2000s, thus providing much better seasonal and spatial coverage for large parts of the Northern Hemisphere oceans and coastal seas. Data coverage remains sparse in large parts of the Southern Hemisphere and the Indian Ocean.

The international importance of SOCAT is evident from recent scientific articles using SOCAT data products for quantification of the ocean carbon sink, process studies and ocean carbon modelling. Regular updates to SOCAT will extend the SOCAT data record and ensure that new data are promptly made available for flux assessments and modelling. Future plans include use of the revised quality control criteria for $f$CO$_2$ values from alternative sensors and platforms, as well as automation.

**Acknowledgements.** The International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS), and the Integrated Marine Biogeochemistry and Ecosystem Research program (IMBER) promote SOCAT. We thank both referees, anonymous referee 1 and D. Woolf, for their constructive and detailed comments, which have considerably improved the Cruise Data Viewer and this manuscript. Support for SOCAT has been received from the University of East Anglia (UK), the Bjørknes Centre for Climate Research, the Geophysical Institute at the University of Bergen (Norway), the Climate Observation Division of the Climate Program Office of the US National Oceanic and Atmospheric Administration, the University of Washington, Oak Ridge National Laboratory (USA), PANGAEA® Data Publisher for Earth and Environmental Data (Germany), the Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (Germany), the National Institute for Environmental Studies (Japan), the Korean Institute of Ocean Science and Technology, the European Union projects CarboChange (FP7 264879) and GEOCARBON (FP7 283080), the UK Ocean Acidification Research Programme (NE/H017046/1; funded by the Natural Environment Research Council, the Department for Energy and Climate Change and the Department for Environment, Food and Rural Affairs), the Research Council of Norway (CARBON-HEAT), the Scientific Committee on Oceanic Research (SCOR, US, OCE-0938349), the US National Science Foundation (OCE-1068958), Swedish Research Council projects (2009–2994; 2008–6228; 2009–2994) and the Swedish National Space Board (RESCUE project; REMote Sensing ocean Carbon UptakE).

Edited by: F. Schmitt

**References**


