

Interactive comment on “A multi-decade record of high-quality fCO₂ data in version 3 of the Surface Ocean CO₂ Atlas (SOCAT)” by Dorothee C. E. Bakker et al.

Dorothee C. E. Bakker et al.

d.bakker@uea.ac.uk

Received and published: 21 July 2016

Final authors' response to reviewers

Authors: We thank the referees for their thoughtful and constructive comments which have substantially improved the manuscript and will improve the ongoing SOCAT effort. Below we copy the comments by the reviewers (R1 and R2), provide the author response (A) and indicate changes in the text (T).

Interactive comment on “A multi-decade record of high-quality fCO₂ data in version 3 of the Surface Ocean CO₂ Atlas (SOCAT)” by Dorothee C. E. Bakker et al. Anonymous Referee #1

C1

R1: General comments: R1: The manuscript by Bakker et al. describes SOCATv3 and its improvements relative to version 2. Inclusion of new flags in this version is discussed in details, in addition to plans for future SOCAT automated system. Impact of SOCAT database for the wider scientific community highlights the critical role of this database.

R1: The manuscript is useful for SOCAT users, providing a clear technical documentations as well as guidelines on how to acknowledge the data generators. As the authors state, the SOCAT data is an important contribution to the carbon cycle community from large number of scientists and technicians, whose continuing fundings depend on proper acknowledgements of their works.

R1: As an end-user of SOCAT who is not all familiar with the data collection and quality control processes involved, there are sections of the manuscripts that I found very informative. Nevertheless, note that I may not be qualified to provide extensive reviews on some sections such as evaluating the correctness or the technicality/criteria described in the quality control or flagging, etc. The paper is well structured and written and I fully support its publication in ESSD. Below, I have several (mostly minor) comments that, if addressed, could further improve the manuscript.

A: We thank the referee for his/her thoughtful and constructive comments which have substantially improved the manuscript and will improve the ongoing SOCAT effort. Below we copy the comments by the reviewer (R1), provide the author response (A) and indicate changes in the text (T).

R1: On Page 5, the paper mentions other pCO₂ data set, namely the LDEO surface pCO₂. I think it could be useful to elaborate what are the main differences between this and SOCAT database. Is there any plan to merge them together? If not, how one dataset complements the other?

A: Text has been added shortly before the end of Section 1. In short, there is no plan to merge the two data products. The data treatment and documentation differ between the two data products. Merging SOCAT with the LDEO data set would go against one

C2

of the key features of SOCAT, notably that it only accepts original data, as reported by the data originator. This enables SOCAT to apply a coherent data treatment and quality control.

T: This text has been added to the manuscript (towards the end of Section 1): 'Two large surface ocean CO₂ data synthesis products, the LDEO and SOCAT synthesis products, are now available (Takahashi et al., 2009, 2015; Pfeil et al., 2013; Sabine et al., 2013; Bakker et al., 2014; this study). While there is substantial overlap in the data sets they contain, the LDEO and SOCAT synthesis products are independent and differ in their data treatment and quality control. There is no intention to merge the LDEO and SOCAT synthesis products, which from a SOCAT perspective would not meet its aim of full documentation and coherence of data treatment and quality control. This said, the SOCAT data managers regularly check which data sets are in the LDEO data product, but are not (yet) included in SOCAT, and invite the data providers to submit their original data sets to SOCAT. In reverse, SOCAT expects data providers to make their original data sets public upon submission to SOCAT or upon publication of the SOCAT version of which these data sets are part (Sect. 6.1). The frequent SOCAT releases therefore increase the data availability in general, including for the LDEO data product. Overall, both data products reinforce each other. Furthermore, the existence of the two data products with slightly different time lines enables the use of independent data from the LDEO data set (i.e. data not (yet) included in SOCAT) in testing interpolation methods built using SOCAT (Landschützer et al., 2015) and vice versa.'

R1: It is clear from the manuscript that in the case of regional studies which significantly use SOCAT data in their analysis, the authors should invite data providers, especially those collecting the data into potential publications. It is less clear however in the case of global studies (e.g., large-ensemble model inter comparison studies involving model data evaluation, where SOCAT and other physical and biogeochemical data sets and synthesis are used. Please clarify.

A: SOCAT encourages authors of global studies to include large SOCAT data providers

C3

and contributors as co-authors. However, in view of the large number of SOCAT data providers globally, it would not be realistic for global studies to invite all data providers as co-authors. T: No changes made.

R1: The development of SOCAT automation system is certainly appealing and a progressive feature of SOCAT quality control scheme, allowing for annual SOCAT releases, as the authors state. My question is whether it is justified to aim for 'annual' SOCAT releases. There are a lot of works going into documenting a data release, e.g., publication, public announcement, etc. But is it really necessary for this? A public release every 2-3 years seems to be sufficient and would allow more time for the data providers to get a first hand to study and analyze them. In addition, I am concern about having too many fragmentation (versions) of the SOCAT database without any clear improvements or distinguishable features (other than simply one more year of extra data) from one version to the others.

A: The Global Carbon Project (GCP) puts much weight on surface ocean fCO₂ measurements from the most recent year for its annual Global Carbon Budgets. In the past such recent fCO₂ data have been provided to the GCP via early data access. However, many scientist in the international marine carbon community feel that data should be available to all studies in the same way, not just to the GCP. Annual SOCAT releases will provide open data access to all users and will be fully transparent.

A: We are not worried about fragmentation of the SOCAT data base, as each consecutive version contains all the previous versions in addition to new data added. Quality controllers have access to earlier data sets. The strict SOCAT protocols will ensure a coherent data set.

A: The referee correctly identifies the considerable amount of work involved in the release of each SOCAT version and the publications documenting SOCAT. It is likely that SOCAT will make (some, but not all) future versions public online (by email, blogs and newsletters), without a festive meeting. Future ESSD manuscripts will describe

C4

several SOCAT versions.

T: No changes made.

R1: Section 6.3. Does it make sense or feasible to also include measurements of surface wind speed, which would allow for real-time estimate of CO₂ fluxes?

A: At present, there are no plans for inclusion of wind speed measurements, as provided by data contributors, in future SOCAT versions. Quality control of wind speed measurements is non-trivial. A number of wind products exist and users of SOCAT can combine these with SOCAT fCO₂rec values.

T: No changes made.

R1: Section 8: Poor coverage in the Indian Ocean and Southern Ocean: Is there any plans or possibilities to address this in the future (e.g., integration with other pCO₂ measurements from other observing platforms not necessarily underway)?

A: In the Indian Ocean data coverage is poor in recent years, but it was good during the WOCE/JGOFS era. It is likely that new data will be added for the Indian Ocean as part of the Second International Indian Ocean Expedition (IIOE-2). In the Southern Ocean, the coverage is especially poor in austral winter.

A: A new data set flag of E has been added to SOCAT version 3 to enable inclusion of good quality fCO₂ measurements (with an accuracy of better than 10 μatm) made on alternative platforms, such as moorings and drifters. Version 3 now includes fCO₂ measurements made by 15 drifting CARIOCA buoys in the Southern Ocean (Boutin and Merlivat, 2009). This is a start.

T: Text has been added to Section 8: 'The new data set flag of E in version 3 now enables inclusion of good quality surface ocean fCO₂ measurements (with an accuracy of better than 10 μatm) made on alternative platforms, such as moorings and drifters, in remote and less remote ocean regions.'

C5

T: Text on autonomous platforms has been added to this sentence: 'The above reinforces the need for the continuing collection and synthesis of accurate, well-calibrated and well-documented observations and investment in high-quality surface ocean CO₂ measurements on autonomous platforms.'

R1: P4, L23-29: For unfamiliar reader, I suggest adding a sentence or two on the mechanisms contributing to these observed large spatial and temporal variabilities. This comment also applies to the ensuing paragraph on pH variability.

A: A clarification has been added. T: The text now reads: 'Seasonal and spatial variation in surface water fCO₂ and pH tend to be larger in coastal waters than in the open ocean, as a result of relatively strong tidal forces, large temperature changes, fresh-water river inputs and strong primary production in coastal waters (e.g. Simpson and Sharples, 2012).'

R1: P5, L21-23: I think this is better clarified towards the end of the manuscript, but I suggest adding few sentences specifically describing SOCAT's role for the Global Carbon Project.

A: The use of SOCAT-based estimates of the ocean carbon sink in the annual Global Carbon Budgets is discussed in Section 7.3.

T: This text has been added to Section 1: 'Recent Global Carbon Budget studies include ocean carbon sink estimates based on the LDEO and SOCAT synthesis products (Sect. 7.3) (Le Quéré et al., 2014, 2015a, 2015b).'

R1: P7, L16-20: please clarify if the increase in flag A/B data (41% to 49%) is simply due to the higher quality of the newly added data or as a result of improved quality control method applied to the earlier SOCAT version dataset.

A: The percentages suggested for version 2 were incorrect. Actually, the percentage of data sets with a flag of A or B is very similar in versions 2 and 3 (48% and 49%, respectively). One might speculate that a small decrease in the percentage of data sets

C6

with a flag of A results from the introduction of the high-quality cross-over criterium.

T: The text now reads: 'Overall, the quality of the data is comparable to that of version 2, with a small improvement in the documentation of the data. In version 3, 14% of the data sets (509 data sets) received a quality control flag of A, 35% (1260 data sets) a flag of B, 23% (840) a flag of C and 27% (990) a flag of D. This compares to 17% (454), 31% (834), 18% (491) and 33% (881), respectively, in version 2. The percentage of data sets receiving a flag of A or B is remarkably similar between both versions (49% in version 3 and 48% in version 2). The small reduction in the percentage of data sets with a flag of D (27% in version 3, 33% in version 2), which implies incomplete metadata, highlights an improvement in the documentation accompanying data sets.'

R1: P10, L14: It is not clear to me why datasets with less than 50 duplicate get a flag 4 (bad)?

A: Not sure we fully understand the question. A flag of 4 was ONLY given to fCO₂ values with duplicate time stamps, NOT to the whole data set!

A: Duplicate time stamps for different fCO₂ values are problematic. Adding artificial seconds for duplicate times is time consuming. There was insufficient time to add artificial seconds to duplicate time stamps in all the data sets. The SOCAT data managers have a limited amount of time to spend on SOCAT. The fCO₂ values in data sets with less than 50 duplicate time stamps have been given a WOCE flag of 4.

T: The text has been rephrased as follows: 'However, if there were less than 50 duplicate times in a data set, a WOCE flag of 4 was generated for the fCO₂rec values with duplicate time stamps during the automated data checks (Sect. 4.3). Adding artificial seconds is time consuming and there was insufficient time available for adding artificial seconds to all duplicate times in all data sets.'

R1: Page 4, Line 21: 'surface' pH in the subtropical . . .

A: Thank you.

C7

T: 'surface' added.

R1: P7, L1: WOCE 'flag' of 3 . . .

A: Thank you.

T: Correction made.

R1: P8, L26: 'a' good scientific practice.

A: Thank you.

T: Correction made.

R1: P12, L28: 0.3C and 5 micro-atm: What are these criteria based on?

A: These criteria reflect the test for a high-quality cross-over between two data sets with a flag of A or B, i.e. each with an accuracy for fCO₂rec of better than 2 μatm or a joint accuracy of better than 4 μatm with 1 μatm added to account for differences in time and space. A temperature difference of 0.3°C roughly corresponds to an fCO₂ difference of 5 μatm.

T: The above text has been added to the manuscript.

R1: P15, L14-15: I am not familiar with the term 'property-property' plots. Can the authors give examples?

A: 'Property-property' plots are scatter plots. They display values for two variables of a data set or of multiple data sets. Examples are scatter plots of fCO₂ as a function of time or latitude. This has been clarified in the text.

T: The text has been revised as follows: 'Scatter plots or property-property plots, available via the Correlation Viewer, can be used to depict two variables of a data set or data sets, enabling further investigation. Examples are figures of fCO₂ or sea surface temperature as a function of time, salinity or latitude.'

R1: Throughout the manuscript, some of the text sections are italicized. I assume this

C8

is just a technical error and needs to be fixed.

A: The manuscript has been submitted as a 'Living Data'. ESSD asks authors to clearly identify changes relative to the former manuscript. Text in italic is significantly different from Bakker et al. (2014). ESSD preferred italic over a blue (SOCAT) colour.

T: All italic has been removed (except in names of research ships).

References (not previously cited in the ESSDD manuscript) Simpson, J. H., and Sharples, J.: Introduction to the physical and biological oceanography of shelf seas. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 424 pp., 2012.

Interactive comment on "A multi-decade record of high-quality fCO₂ data in version 3 of the Surface Ocean CO₂ Atlas (SOCAT)" by Dorothee C. E. Bakker et al. Anonymous Referee #2

R2: The paper by Bakker et al. presents the latest version (version 3) of the Surface Ocean CO₂ Atlas (SOCAT) and describes its improvements relative to SOCAT version 2. The new database includes a large number of additional data and extends the record from 2011 to 2014. Moreover, data quality check and documentation are improved, as well as its graphical interface. Authors also highlight the importance of such database for the carbon cycle scientific community by providing an interesting description of previous applications of SOCAT data. Additionally, they provide useful guidelines to acknowledge the work of SOCAT data providers and community. The manuscript is well written and detailed. As a consequence, I believe that, after modest revisions that should not be hard to address, the paper should be published.

R2: SOCAT database, including its future versions as described in this manuscript, is and will be a very useful tool for a wide scientific community. Keep up the good progress.

A: We thank referee 2 for his/her kind words and thoughtful comments on our

C9

manuscript. We much appreciate his/her constructive review and aim to continue the SOCAT synthesis activity by the international marine carbon community. Below we copy the suggestions by referee 2 (R2), discuss them (A) and indicate how we have addressed them in the manuscript (T).

R2: Page 5, line 6: I am not familiar with LDEO database, but I am wondering why it is not merged with the SOCAT database yet and if there is any plan to do that in the future. I suggest clarifying this aspect in the manuscript.

A: Referee 1 asks a similar question. See the response to referee 1.

T: Text has been added as detailed in the response to referee 1.

R2: Page 6, line 32: Please specify what WOCE stands for.

A: WOCE is World Ocean Circulation Experiment.

T: The abbreviation has been clarified in Section 2.

R2: Page 7, line 16: Please clarify how the number of data with quality control flags A, B and C have been increased relative to SOCAT version 2.

A: Referee 1 asks a similar question. See the response to referee 1.

T: Text has been amended as detailed in the response to referee 1.

R2: Page 10, lines 17-25: It is unclear to me what the difference between the 'recommended' fCO₂, calculated by using the 14 algorithms listed in Table 6, and the (re-)calculated fCO₂, calculated by the SOCAT algorithms, is. I suggest reformulating the text to clarify better this crucial aspect.

A: Recommended fCO₂ and (re-)calculated fCO₂ are identical. The 14 algorithms are identical to the SOCAT algorithms. This has been clarified in the text by adding 'or recalculated' and '14' in the text.

T: The text has been changed to: 'In total, 14 algorithms were used for calculating

C10

these 'recommended' or 'recalculated' fCO₂ (fCO₂rec) values (Table 6). The term 'recommended' fCO₂ values is used here to distinguish the fCO₂ values (re-)calculated by the 14 SOCAT algorithms from the xCO₂, pCO₂ and/or fCO₂ values reported by the data providers.'

R2: Page 10, lines 29-30: Does the use of external data (e.g. atmospheric pressure from NCEP or salinity from WOA) affect the data quality flags?

A: Good point. We might consider if an automated check can be added. The text has been amended.

T: A sentence has been added to Section 4.2: 'The use of external of atmospheric pressure data would rule out data set quality control flags of A and B during subsequent quality control, while use of external salinity values would not affect the data set quality control flag.'

R2: Page 10, lines 30-31: Please specify which version of WOA is used to obtain salinity when in situ data are not available, as done for the NCEP version.

A: WOA 2005 has been used (as specified in Table 9).

T: These sentences have been added to Section 4.2: 'Sea surface salinity was from the World Ocean Atlas (WOA) 2005 (Antonov et al., 2006). Full details on the external pressure and salinity products is in the footnotes of Table 9.'

R2: Page 10, line 32: Please clarify why the new Ferret script implemented in SOCAT version 3 is "an important change" relative to previous versions using a Matlab code.

A: The implementation of the Ferret scripts enables full integration of SOCAT data submission, (re-)calculation of fCO₂ and quality control on a single software platform. This allows parallel data submission on one data set and quality control on another data set. It streamlines and simplifies the SOCAT data flow.

T: The above sentences have been added to the text in Section 4.2.

C11

R2: Page 12, lines 30-31: I believe it is worth providing a quantification of how much "rarely" is. I would suggest providing percentages of high-quality cross-overs found in coastal waters, iced-covered regions and ROFIs.

A: Interesting point. Provision of such percentages is beyond the scope of this manuscript, and is likely to require considerable programming effort. However, we might consider a review of cross-overs in coastal waters, ice-covered regions and ROFIs for future SOCAT versions.

T: No changes made.

R2: Page 13, line 10: Please clarify why the additional guidelines in SOCAT version 3 are not systematically applied to open ocean fCO₂rec data away from sea ice.

A: At present SOCAT does not have an automated check for these five additional guidelines. Only with an automated check could one unequivocally state that guidelines have been systematically applied.

T: 'but not systematically applied' has been removed from the text. The indication that these guidelines were considered is sufficient.

R2: Page 15, line 14: Please provide an example of what a property-property plot is.

A: Referee 1 asks a similar question. See the response to referee 1.

T: The text has been amended as detailed in the response to referee 1.

R2: Page 22, line 22: I cannot find this information in Regnier et al., 2013. I believe they refer to 0.1 Pg C yr⁻¹ as the increase of the lateral export of land-derived carbon, associated to anthropogenic activities, to the open ocean. Please check.

A: The conversion of contemporary air-sea CO₂ fluxes to anthropogenic air-sea CO₂ fluxes, as done for e.g. the Global Carbon Budget, takes the land-derived carbon term reaching the open oceans as an extra ocean outgassing term. The extra 0.1 Pg C yr⁻¹ of outgassing by the oceans is included in Figure 3 of Regnier et al.

C12

T: No changes made.

R2: General: Some sections of the manuscript are in italic and, as mentioned by authors, indicate significant differences from the paper Bakker et al., 2014 describing SOCAT version 2. I am not sure it is necessary.

A: Referee 1 asks a similar question. See the response to referee 1.

T: All italic has been removed (except where needed in vessel names etc.).

Please also note the supplement to this comment:

<http://www.earth-syst-sci-data-discuss.net/essd-2016-15/essd-2016-15-AC1-supplement.pdf>

Interactive comment on Earth Syst. Sci. Data Discuss., doi:10.5194/essd-2016-15, 2016.