

## ***Interactive comment on “An update to the Surface Ocean CO<sub>2</sub> Atlas (SOCAT version 2)” by D. C. E. Bakker et al.***

**D.K. Woolf (Referee)**

d.k.woolf@hw.ac.uk

Received and published: 26 November 2013

This paper provides a description of a major update to the “SOCAT” CO<sub>2</sub> Atlas. This atlas represents a major community effort of broad and deep utility and it is wholly appropriate that a transparent, detailed and definitive description of this product (and the data itself) should be in the public domain and ESSD provides a very suitable portal. Therefore, I have no doubt that in final form the description should be published in ESSD and I can hope only to assist slightly in ensuring the clearest and most meticulous version is completed.

Generally the structure of the manuscript is well judged. All material (text, figures and tables) are at least useful. If necessary some of the material could be removed

C203

or condensed, but I can think of no good reason to do that. My main issue with the manuscript is that in a few instances insufficient detail is provided on the data itself and especially on the processing that leads to three SOCAT products. The main function of the paper should presumably be to present and explain the data, thus while the other material is interesting enough an emphasis must be on substantiating the statement that “The data . . . have been subject to fully documented quality control”. In this respect the current manuscript is inferior to two previous papers in ESSD on SOCAT (Pfeil et al., 2013; Sabine et al., 2013), which did provide a large amount of detail and explained the philosophy behind the products. To some extent, the current manuscript can “lean on” the previous papers, but where it does, it should refer to the detail in those papers and state clearly whether there has been any revision. Where repetition is unnecessary, I suggest referring to specific sub-sections of Pfeil et al or Sabine et al, since this would make the reader’s task much easier. Where the new products differ from the old, the differences should be more carefully explained. Specific instances where the description could be more detailed (or alternatively, a precise reference to a previous description could be included) are scattered among the detailed notes below.

I cannot argue with the order of sections, but I found often that the text raised questions in my mind that were only answered much later (this may be apparent in some of my specific comments below). A few more ‘signposts’ such as “this is described later in Section X.Y” might help the reader.

The abstract told me what I needed to know, but it is in my view written rather eccentrically: P. 469, line 2. I’d prefer “is a product of” to “an effort by” p469 l3. Simplify to “It improves access to . . .” p469 l5 onwards could be written more plainly and informatively e.g. “Version 2 of SOCAT is an evolution of the previous release (version 1) with more data (increased from 6.3 million to 10.1 million values), extended data coverage (from 1968-2007 to 1968-2011) and a slight revision of processing.” The sentence “Version 2 enables . . . until 2011” adds nothing except a little ambiguity.

P470 l4. Remove “has”. P470 l27. Perhaps change “has been demonstrated . . . in

C204

the” to “of CO<sub>2</sub> uptake is apparent in the” P471 l8-9. I suggest ‘unpacking’ the sentence “Underway . . . commercial routes” to “The fugacity can be measured underway in the surface water supply of ships. This method enables the use of a variety of ships including ‘ships of opportunity’ on commercial routes”. P471 l19. To be very pedantic, the use of the term ‘voluntary observing ships’ is potentially confusing since it risks confusion with the Voluntary Observing Ship (VOS) Program (WMO etc.), which is a very distinct and long-established activity. IOCCP recognises a Global Volunteer Observing Ship (VOS) program for CO<sub>2</sub>, which also risks confusion with the original VOS program. I think sticking to ‘ships of opportunity’ is safer. Most of the references at l19-23 could be eliminated. P471 l29 – P472 l6. This sentence actually refers to one part of the process of calculating fluxes, but as written sounds like ‘interpolation’ is the only thing that is needed apart from the measurements of fugacity. This manuscript and the associated data stop short of flux calculations except peripherally in Section 5, but one of the products is a gridded data product of surface water fCO<sub>2</sub>. I suggest “Various methods have been used to infer basin wide distributions from limited observations of surface water fCO<sub>2</sub>. . . .” Also, I am not convinced the long list of methods and references that follows is useful in the context of this paper.

P472-P473. No detail is given about the recalculation of fugacity from the original values reported by the observers to the ‘\_rec’ definition. It is true that this is covered in reasonable detail by an accessible reference (Pfeil et al., 2013), but since this step is key to the exact meaning of “fCO<sub>2</sub>\_rec”, I think that a brief explanation of two points should be given here. Firstly, what is meant by sea surface temperature (SST) in this context (you could refer to previously published explanations of intake temperature and TSGs)? Secondly, what is the premise for calculating fCO<sub>2</sub>\_rec from the actual measurements (in most cases at a different temperature)? Similarly, there is a case for noting the most significant features of the gridding at this point and then referring to 2.4.4 and Sabine et al. for more detail. Particularly, since you have mentioned various ‘interpolation methods’ on the preceding page, readers cannot be expected to guess the actual process if you don’t explain the philosophy behind the ‘Sabine et

C205

al.’ approach. It is especially important to know the criteria for eliminating data. It is clearly important to remove data that is known to be technically defective, but I would argue that it is disastrous to risk introducing “publication bias” by eliminating data only because it appears as an “outlier” in some sense to the “quality controller”. Pfeil et al. (2013) explain quality control for version 1, therefore my main concern is the statement (P473 l17-19) “About 70 cruises contained in version 1 were removed from version 2 upon identification of data quality concerns”. If these data were considered suitable for version 1 (presumably passing the QC described by Pfeil et al.), what precisely led to their exclusion from version 2? Is part or all of the explanation related to “WOCE flags 3 and 4 were reset unintentionally . . .” {quoting from Table 1}? This should be explained immediately following P473 l19, or the relevant sub-section ‘signposted’.

P473 l19-20. It may be useful to describe how much data is contributed by “4 time series, moorings and drifters”. Was this “4 time series from either a mooring or a drifter”? What is the instrumentation? In these cases, how (if at all) does the ‘recalculation’ differ from underway ship data? Is all the ship data from a pumped underway supply or are there other collection methods? If there are other methods, how does the ‘recalculation’ differ? Are there identifiers in the databases and switches in the data viewer that enable isolation of particular methods (I have noted Section 2.2.2 on Expocodes that clearly identifies the platform, so my query is on the collection method and type of instrumentation)?

P474 l5. Change “of “ to “to” P475 l17. Add “for” after “available”.

P476-477, Section 2.2.4. This section is very informative, but could be even more informative. The preferred procedure can be debated (for example, is it appropriate to exclude data from the synthesis files only because the seawater temperature is atypical for the season and region?), but this manuscript is not the place for that debate. In my opinion, it is worth finding space in this manuscript for one or two extra tables, firstly to define as precisely as possible what criteria were imposed for the WOCE flags (e.g. a specific temperature anomaly criterion?) and secondly to give statistics of the number

C206

of data excluded from the synthesis, as far as possible sub-divided according to the active criterion. In the case of criteria, much of the explanation missing in this sub-section is finally given in Section 2.3.2 (secondary quality control), therefore it is worth mentioning 2.3.2 here. Also, Section 4 of Pfeil et al is more informative and a sufficient answer may be “as Pfeil et al.”, but it would be better to state that in 2.2.4.

P477 I20 – p478 I4. These sentences don't really belong under “data entry” and supply a little more detail to the introductory material on P471. This material may be better gathered together (in Section 2.1?). As already commented for P473 I19-20, it would be useful to tabulate the contributions of different sensors and platforms to the total database.

P479 I22. Spelling of “resemble”.

Section 2.4.3. There seems to be a contradiction between P480 I20, “The synthesis files only contain . . . with WOCE flag 2”, and P481 I8-9 “The user can include data with a WOCE flag of 3 . . .”. Perhaps this arises since the Cruise Data Viewer is not accessing just the global synthesis data? But in that case, doesn't this sub-section belong in Section 2.4.2? Please clarify.

Section 2.4.4 is - I think - an example where an appropriate level of detail is supplied.

Section 4.2. The title identifies ‘instrumentation’ and ‘sensors’; it may be better to discuss ‘sensors’ and ‘platforms’. Both the sensor and the platform can affect what is measured and how, while the current text only mentions the platform in the context of the choice of sensor. A talking point rather than something to be addressed in a revision of this manuscript: Placing sensors on platforms such as drifters and gliders undoubtedly raises issues (e.g. very limited calibration), but issues of disturbance are much less than an underway ship (where the moving ship disturbs the water column and the sample may be altered during its journey to the sensor). Thus for sea surface temperature, a temperature measured from a drifter or glider may be superior to intake temperature.

C207

P484 I8. Replace “operation” with “operating” P484 I13-14. I suggest “The working group has recommended an appropriate set of quality control criteria for each sensor”.

Section 4. I note from within Section 4.1.3 of Pfeil et al. (2013). “. . . Apart from these, no strict criteria for QC were defined . . . This will be improved in future versions of SOCAT.” That ambition is laudable (without clear criteria, the definition of the SOCAT products is less clear than it could be). What are the plans for Versions 3 and 4?

Section 4. As described in Section 2.4.3 the original data and metadata is accessible from the global synthesis product via a doi number for the cruise file in PANGAEA. Has there been any consideration to giving a limited amount of metadata (e.g. a sensor-type code) directly within the synthesis product?

---

Interactive comment on Earth Syst. Sci. Data Discuss., 6, 465, 2013.

C208