Biogeosciences Discuss., 12, C7894–C7900, 2015
www.biogeosciences-discuss.net/12/C7894/2015/

Biogeosciences Discussions

Interactive comment on "Data-based estimates of the ocean carbon sink variability – first results of the Surface Ocean pCO$_2$ Mapping intercomparison (SOCOM)" *by* C. Rödenbeck et al.

C. Rödenbeck et al.

christian.roedenbeck@bgc-jena.mpg.de

Received and published: 24 November 2015

Reply to Anonymous Referee 2

We would like to thank referee 2 for her/his detailed and helpful comments.

The manuscript presents a comparison of a number of spatial-temporal interpolation methods used to map surface pCO2 data of the newly established large data sets. This is a timely and very useful effort in itself. The authors then proceed to infer seasonality and interannual variability in regional and global air-sea CO2 fluxes from the mapped pCO2 interpolations.

The manuscript is generally well written, though sometimes the heavy use of footnotes and comments in brackets disrupts the flow and makes some details difficult to understand. Why not give up on footnotes and simply include them in the text (perhaps in brackets like this?). I don't think that there is any reader who will not read the footnote - thus jumping to the bottom of the page and than not finding back into the main text can be avoided when including it in the main text.

The footnotes were meant to weight different pieces of information differently, actually with the intention to increase readability. However, we also see the referee's concern of the need to jump back from a footnote. We inlined all footnotes.

Overall I found the manuscript a very interesting and useful scientific contribution, but have a few general issues and a number of more specific comments listed below.

General comments: (i) Given that intercomparisons tend to be tedious efforts with difficult choices to be made as to how present exciting science without falling into a ranking trap, the authors have done a great job here. However, I still think that the readers would like, and should receive, some more information about what methods are 'best' for regional or global purposes, what aspects of different methods are most/least problematic, and where future research should focus on. It would also be good to get back to the very good scheme of Fig.1 and say in the concluding section of the paper to what extent this scheme has been confirmed by your study.

Thank you for these suggestions. The need to give more explicit advice had also been highlighted by Referee 1. As detailed in our reply there, we enhanced the last part of the conclusion accordingly. We also added more statements on research needs and the importance of sustained measurements.

(ii) the maps are compared against the SOCATv2 gridded product. It is not clear to what extent the SOCAT data are independent and to what extent these have been used by all/some interpolation methods. Are all interpolated maps at the same level when it comes to a fair comparison against SOCAR data?

SOCAT data are not independent, because all methods use either SOCAT itself, or LDEO which shares a large portion of data (Tab. 3). This is why the mismatch criterion only represents a necessary condition, but not an independent validation.

The question whether using SOCATv2 as comparison data may bias the rating, was dealt with by also using LDEO2013 data, as described in the last paragraph of Sect 3.5. We reformulated the introductory sentence into "To verify that the selection criterion is not unduely biased by the fact that some methods use SOCAT data and others use LDEO data (Tab. 3), IAV mismatch diagnostics ...."

(iii) Wanninkhof 1992 is used to compute air-sea CO2 fluxes from pCO2 differences and uncertainties in air-sea gas exchange are not considered. It would still be useful to compare uncertainties in CO2 fluxes due to pCO2 interpolation to uncertainties in the air-sea gas exchange. Which one is larger?

Gas exchange uncertainty has been considered in the literature. For example, Sweeney et al. (2007) find 30% difference in the flux when different formulations are used. Landschützer et al. (2014) used 4 different parametrizations and find 37% uncertainty in the integrated flux, with the largest fraction originating from the flux formulations (though not considering different winds products). A proper direct comparison of mapping and gas exchange uncertainties for the same time and space scales is certainly a valuable target for further studies.

(iv) Seasonality is shown for the relatively well-behaved North Atlantic subtropical permanently stratified biome (sec.4.1.1). Results for the East Pacific Equatorial Biome look much worse (typical deviations of 20-30uatm compared to 10uatm in the Atlantic) and are shown only in the Appendix (Fig.A1). This is relevant as the East Pacific is the region used in the analysis of interannual variability (sec.4.1.2). Apparently the 12month running mean helps to remove some(?)/most(?) errors in the description of the seasonal cycle? This has to be discussed in more detail. The first point of the conclusions, that seasonality is constrained mostly within 10uatm may have to be revised.

To clarify this point, we introduced the optical weighting among the methods by line thickness also for the monthly variations. This shows that the larger and more irregular seasonal cycles in the East Pacific Equatorial Biome fit the data less good than the smaller ones. There is a clear tendency that methods with larger IAV mismatch also have larger monthly mismatch (ie. the referee's conjecture that the 12month running mean removes errors is not confirmed). The only exception in the East Pacific Equatorial Biome is the AOML-EMP method with $R^{\mathrm{iav}} \approx 60\%$ below the chosen threshold but a high $R^{\mathrm{month}} \approx 90\%$.

Disregarding methods with $R^{\mathrm{month}} > 75\%$, the first point of the conclusions stays valid.

(v) Having read all the positive comments about the interpolation methods and interannual CO2 flux variability in the equatorial Pacific, I was quite surprised to see so little agreement in the interannual global sea-air CO2 fluxes (Fig.5c). It would be very useful to understand this better: What are the regions responsible for the very different behavior of different interpolation schemes? Where would more data be most useful? Or do some interpolation routines particularly well/poorly in some regions? This is potentially a very important figure that may be copied and used a lot. Thus it would be reasonable to provide a robust explanation to avoid giving the impression that "data-based estimates of interannual CO2 flux variability are all over the place".

On the one hand, there actually is agreement in the ensemble regarding certain global features, such as the decadal trends. We made this a bit more prominent in Sect. 4.2.2, also by an additional figure in the Appendix. Also the IAV amplitude is estimated rather consistently among the methods.

On the other hand, the contrast in the ensemble spread between equatorial Pacific and the global flux reflects real differences in the ability to estimate these fluxes from the available data. This also was the very reason to set both plots into a common figure, to demonstrate that the data-based estimates can do a good job where the data density is sufficient. We agree that this message may not have been strong enough, and tried to reformulate accordingly, both in Sect 4.2.2 and the enlarged last part of the Conclusions, where we now also highlight the need for sustained and extended measurement programs.

The regions of origin of the spread have been discussed in Sect 4.2.2. We added another general remark about biomes 15 and 16.

Specific comments: p14051,l9 spread p14051,l9 "mapping methods with closer match to the data also tend to be more consistent with each other." Not clear what is meant here by "consistent". It would be redundant information that points closer together (because closer to the data) are somehow more 'consistent' with each other.

We reformulated into "mapping methods that fit the data more closely also tend to agree more closely with each other in regional averages".

p14052, l14: For model tuned against all WOCE and pre-WOCE data (e.g. ECCO), I would expect an enhanced model skill in predicting trends and variability during the tuning period. Perhaps adding 'beyond the tuning period' expresses better what was meant here? Still, I'm not sure whether this is correct. I think I would always prefer a tuned model over an untuned one, and I would also think that a tuned model could be better outside the tuning period. So I think "cannot be expected" is not always right.

We realize that our formulations were not entirly clear. We rewrote and enlarged this piece, but then moved it to Sect 2.3 in order not to interrupt the flow of thought in the Introduction.

p14054,l24. How can process model simulations provide information on correlation scales? How do you ensure that the process models resolve the right scales and feedbacks/correlations? Perhaps I didn't understand what was meant by 'process' model?

We concretise into "derived from EOF analysis of an ensemble of process model simulations". Of course, this trusts that the models capture the essential modes of variability sufficiently well.

p.14055 "for a quantity determining pCO2" sounds very cryptic and the reader doesn't have much chance to understand it. Can you explain this better?

We concretise into "for the field of ocean-internal carbon sources and sinks which determine the pCO2 field".

p14056,l21 I don't understand why SOM have this advantage and FFNs don't have it. Why should FFNs need a-priori knowledge? Is this really the case? Why can't you feed the same information into a SOM and FFN?

We agree that this formulation was misunderstandable, and rewrote the paragraph.

p14058, l1 In high latitudes (where, presumably, most of the data gaps occur) this procedure is different from that used by Takahashi. Why do you assume that pCO2 of high-latitude surface waters increases at the same rate as atmospheric pCO2? Isn't the upwelling of deep old waters keeping a tendency towards pre-industrial surface-water pCO2 despite increasing atmospheric pCO2?

It may well be that the pCO2 field in the filled-in areas behaves in the way described by the referee, but in the absence of observational knowledge, data-driven schemes have to resort to some simple assumption. Assuming a rise parallel to the atmospheric rise means that the fluxes calculated from these pCO2 fields will not have a trend in that region. Luckily, in any case, the contribution of the filled-in areas is not dominating any of our conclusions.

p14059,l10ff: "we use the amplitude (temporal standard deviation) of the average difference between map and data". This is not completely clear: what average difference do you mean? Spatial and annual difference for each biome?

The subsequent remark in paratheses was meant to clarify this point. We reformulated the whole paragraph into a point-by-point list to make it more accessible.

p14062,l11 "where data exist" Does this refer to locations in the maps where data used by the interpolation exist, or does this refer to locations of the SOCAT gridded product *with data in this particular month? Please clarify.*

It refers to the SOCATv2 data used for comparsion. We add "where SOCATv2 comparsion data exist".

p14064,l20ff and Fig. 4c. Not clear how you compare biome/yearly pCO2 values with monthly SOCAT data. Shouldn't one see the seasonal cycle?

The plot shows the yearly average of the monthly difference, thus removing the seasonal cycle. We added "biome/yearly average", and a reference to "(Sect. 3.5)".

p14066,l26. Not clear what is meant by "the higher IAV amplitudes are likely. Higher compared to what? Likely = plausible? ?

We reformulated to "... UEA-SI likely underestimates the IAV amplitude."

Fig.1 Why does a linear regression require more model assumptions than a nonlinear one?

A linear regression restricts the functional relationship to a linear function, while the nonlinear regressions allow a much wider class of funtions to be chosen from by the data.

Fig. A6: What is the meriodional empty stripe in the Pacific Ocean UEx-MLR? What are "other reasons" in the last sentence of the figure caption?

The empty stripe is a row of invalid pixels, that are filled here by the standard map as described in Sect. 3.2.

"other reasons" are missing SST input values or numerical reasons, respectively (see Sect. 3.2). We added this concretization into the caption.

Interactive comment on Biogeosciences Discuss., 12, 14049, 2015.