Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-247-AC2, 2019

Geoscientific Model Development Discussions

Interactive comment on "FFNN-LSCE: A two-step neural network model for the reconstruction of surface ocean pCO_2 over the Global Ocean." *by* Anna Denvil-Sommer et al.

Anna Denvil-Sommer et al.

anna.sommer.lab@gmail.com

Received and published: 13 March 2019

We thank Dr Christian Rödenbeck for his comments and suggestions. The manuscript was revised to address each point.

Christian Rödenbeck: The authors present a method to interpolate temporally and spatially discrete surface- ocean pCO2 observations into a gridded field, using non-linear regression against var- ious environmental drivers. Compared to similar methods, two specific features are (a) estimating seasonality separately based on data also outside the calculation period, thereby improving its data constraint, and (b) the addition of SSH to the set of predictors. This is an interesting contribution to the ensemble of existing interpolation methods, enlarging the range of plausible outcomes. The manuscript nicely describes the details of the method, and presents evaluation metrics, also in the context of other methods from the SOCOM ensemble. I clearly recommend to publish this manuscript. Below some suggestions mainly to further improve clarity at a few places.

Authors: Thank you very much for your positive evaluation. We provide hereafter a detailed point-by-point reply.

Christian Rödenbeck: The method is very similar to CARBONES-NN (not published in the peer-reviewed literature but described in the SOCOM paper RoÌĹdenbeck et al., 2015), developed at LSCE as well. I therefore assume that the presented method builds on and supersedes CARBONES-NN. Is this correct? If so, this is an interesting piece of information to users of the SOCOM ensemble and should be mentioned in this paper. Further, I would find it fair to give credit to the authors of CARBONES-NN (to my knowledge, Abdou Kane and Philippe Peylin).

Authors: The present study (by Denvil-Sommer et al.) reflects our current vision of this particular estimation problem. It is part of a long-standing activity at LSCE with focus on the surface ocean carbon system. Abdou Kane and Philippe Peylin's work, the development of the CARBONES-NN model, is an earlier example. The model has been largely redesigned to turn into the one presented in the current submission. Hence, the new model presented in this paper replaces the former one. Philippe Peylin is of course fully aware of our latest efforts, but his work has now other priorities.

Christian Rödenbeck: You mention the inclusion of SSH, but do not discuss it. How does it influence, and possibly improve, the results?

Authors: First tests suggested that the inclusion of SSH does not significantly improve the accuracy of reconstructed pCO2 at global scale. We added a corresponding sentence in the manuscript. At basin and regional scale, however, adding SSH improves the spatial pattern of reconstructed pCO2 and the accuracy of our method. The full assessment of SSH will be part of a follow-up study.

C.R.: How much do the data outside the period improve the estimates of seasonality from the 1st step?

A.: The climatology by Takahashi et al. (2009) is based on data obtained from the early 1970s to 2007. Coastal pCO2 measurements and values obtained in the Eastern Equatorial Pacific are excluded. It represents average open ocean pCO2 conditions and is centred on the year 2000. A way to estimate the contribution of data outside the reconstruction period to estimates of seasonality would consist in computing the climatology over the period covered by the reconstruction only. A sensitivity analysis could also be carried out subsampling a numerical model. However, since Takahashi et al. excluded years with strong anomalies, we expect the influence of data outside of the reconstruction period to be small.

C.R.: Couldn't step 1 also be fitted against the SOCAT data (folded into a climatological year as done by Takahashi et al., 2009) as well, rather than against the already-interpolated climatology? This would also bring in the data from the most recent years not yet in the Takahaski et al. (2009) climatology.

A.: This could indeed by an alternative approach to step 1, which could be tested in the future. The use of an interpolated climatology allows however to benefit from a full data coverage (a larger data set, as stated in line 78), which would not be the case with SOCAT.

C.R.: Title: While there is no question that the authors are fully free to choose a name for their method, it seems that they intended to use a "SOCOM-style" name. Therefore I'd like to point out that all names in the SOCOM paper are of the form INSTITUTION-METHOD, not the other way round.

A.: We followed the reviewer's suggestion and changed the name of the model to LSCE-FFNN.

C.R.: Line 78: "On a larger data set" - please clarify what this means

A.: The use of a gridded climatology during the first step allows benefiting from a global and continuous data coverage. It provides more data for training and validation. It was clarified in the text.

C.R.: Lines 79-81 and 446: Unclear statement - isn't it the 2nd step (not the 1st) which takes care of changes?

A.: The 1st step reconstructs the climatological seasonal variability. The 2d step adds interannual variability to the signal if present. It contributes information on anomalies with respect to the climatological state.

C.R.: Line 114: If you took surface pressure from the s76_v4.1 run, it is in fact from NCEP (Kalnay et al.). This should be mentioned."

A.: It was modified and mentioned in the text.

C.R.: Lines 141, 142, 191: The "1" and "2" are awkward, and should rather be part of the subscript.

A.: It was modified.

C.R.: Line 174: Isn't the 1st step using climatologies of the driver variables? Please clarify. Also, is it correct that the 1st step uses un-normalized drivers (rather than SST_n etc.)?

A.: Yes, driver climatologies are used during step 1. All data have to be normalized for both steps of the model. It was clarified in the text.

C.R.: Line 191: I assume that the "n" in the subscript to pCO2 is not correct, is it?

A.: It is correct. "n" stands for "normalized". All variables are normalized prior to their use in the model. It was clarified.

C.R.: Line 197: Maybe say "for each climatological month".

A.: Corrected.

C.R.: Line 209: Does "chosen automatically" mean randomly? Please clarify.

A.: Yes, it is randomly. It was modified in the text.

C.R.: Line 212: What does "final model" mean? Is it the model of step 2?

A.: Yes it is step 2. We have clarified it.

C.R.: Line 236: "(4)" should probably be "(5)".

A.: Corrected.

C.R.: Line 238: Is it from run s76_v4.1 as well? Please mention.

A.: Yes, it is. We have clarified it.

C.R.: Line 244: Isn't "did not participate" in contradiction to Sect 2.2 b) which says "100% used for training"?

A.: We distinguish between the development of the model presented in section 2.2 a) and its application to pCO2 reconstruction in section 2.2 b). It is stated in the manuscript as follows: "The previous section presented the development of a FFNN model for the reconstruction of global surface ocean pCO2, and the estimation of its accuracy". During development, the accuracy of the method was evaluated based on 25% of data that did not participate in training. We added a sentence for clarification.

C.R.: Line 283: Though the Jena scheme indeed uses SST etc in its parameterizations, the sentence can be misunderstood as meaning regression drivers. Maybe just remove the list of drivers here.

A.: We prefer to provide a list of variables used. We have added that these variables are used in its parameterizations.

C.R.: Line 285: I feel "combined" is confusing and can be removed.

A.: We agree and removed it.

C.R.: Line 286: Please mention whether you used the same versions of these methods as used in the SOCOM paper, or whether you used updated versions. If updated, the respective version IDs should be given.

A.: We have added the number of versions used.

C.R.: Line 286-287: The sentence "Qualification...(2015)." is repeated in the next paragraph and may be deleted here. I feel the next sentence "The time series..." should start a new paragraph.

A.: We agree, the sentence was deleted.

C.R.: Line 392: I guess "small IAV" should be "small relative IAV mismatch", isn't it?

A.: It is. It was corrected.

C.R.: Line 394, 296, 403, 455: The figures after the decimal point are certainly meaningless and should be removed.

A.: We agree, it was corrected.

C.R.: Lines 412, 425: It is unclear to me what "total negative trend" means.

A.: It is a linear trend averaged over the globe. It was clarified in the text.

C.R.: Line 413 and below: "PgC/yr" is the unit of a flux, not of a flux trend. Please clarify the unit (it should be something like "PgC/yr/yr" or "PgC/yr/decade").

A.: Yes, indeed, it is PgC/yr/yr. This has been corrected.

C.R.: Lines 421ff: Are trends like 0.0004PgC/yr/yr statistically significant given the IAV? If not, it is not actually appropriate to call them "positive".

A.: Yes, t-test showed that it is significant (p-val is 0.05).

C.R.: Figures: In my print-out, the annotations and labels of most figures are much too small to be readable.

A.: The figures will be changed in the revised version.

C.R.: Fig 9, color bar: In this figure, the numbers should be within the colors, not between them, otherwise the meaning is ambigious.

A.: Done.

C.R.: Fig S7 caption: "according" probably means "agreement"?

A.: Corrected.

All typos were corrected.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-247, 2018.