Geoscientific

Model Development
Discussions

[GMDD]

Interactive
comment

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.

Rödenbeck (Referee)

christian.roedenbeck@bgc-jena.mpg.de

Received and published: 30 November 2018

The authors present a method to interpolate temporally and spatially discrete surface-ocean pCO2 observations into a gridded field, using non-linear regression against various 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

[Printer-friendly version]

[Discussion paper]

methods from the SOCOM ensemble. I clearly recommend to publish this manuscript. Below some suggestions mainly to further improve clarity at a few places.

Comments:

The method is very similar to CARBONES-NN (not published in the peer-reviewed literature but described in the SOCOM paper Rödenbeck 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).

On reading the manuscript, a few open questions arose that I'd find interesting to consider, either for this paper (in case you already made some corresponding tests) or in a follow-up study. (1) You mention the inclusion of SSH, but do not discuss it. How does it influence, and possibly improve, the results? (2) How much do the data outside the period improve the estimates of seasonality from the 1st step? (3) 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.

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.

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

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

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.

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

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.)?

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

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

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

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

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

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

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

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.

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

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.

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.

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

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

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

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").

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".

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

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

Fig S7 caption: "according" probably means "agreement"?

Typos:

Line 69: "Feed Forward"

Line 135: the overbar of mean SSS should stretch over all three letters

Line 270: "bias" (2 times)

Line 287: "CO2 flux"

Line 297: "latter"

Line 391: Add comma after "Despite this"

Fig 2 caption: "(a)" seems misplaced. It could be after "2016", in accordance with placement of "(b)" after the respective desction.

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