Biogeosciences
Discussions

Open Access

Interactive comment on "A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink" *by* P. Landschützer et al.

Anonymous Referee #1

Received and published: 18 June 2013

The authors estimate the oceanic CO2 exchange in the Atlantic using the SOCAT set of pCO2 observations and a Neural Network technique to map them in space and time. They validate their results based on fit residuals and using independent data, assess uncertainty, and discuss decadal means, seasonality, interannual variations, and trends. As data-based estimates are essential for a quantitative understanding of the global carbon cycle, this is an interesting and important contribution. The method and findings are thoroughly and convincingly described. I would definitely like to recommend this study for publication in Biogeosciences. The comments below are intended to help making this paper even more clear.

Specific comments:

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

The paper gives contradictory statements on whether the variability of the Atlantic CO2 sink is "substantial" or "small", both in the Abstract (p8800 line 3 versus p8801 line 1) and in the Discussion (p8817 line 17 "substantial" versus line 21 "small"). Please clarify.

p8804 line 8: "overcome most of these limitations": This also seems to refer to the statement 3 lines earlier that MLR methods only explain little variance. I missed some proof that the Neural Network indeed does better, and why (more explanatory variables used, or non-linearity, or something else?).

p8805 line 7: "same" –> "analogous"?

p8807 lines 1-5: I felt this was not very clear. Couldn't you say that there is one vector/target pair for each SOCAT data point, and that the input fields are sampled at the location and time of the SOCAT data?

p8807 line 18: It should be said more specifically what "similar patterns" refers to.

p8807 line 20: "and" –> "but" (to clarify sentence)

p8807 line 26: "dynamics in time" - does that mean that the provinces are time-dependent? If so, this needs to be mentioned explicitly, because the conventional understanding of provinces seems to be static.

p8808 line 11: A brief explanation would be helpful why "deseasonalized representations" have been included.

p8808 line 12: The meaning of "FINP2" is not clear - either explain here or defer mentioning it. (By the way, it would give the reader a better feeling to know where the abbreviations "FINP" etc. actually come from.)

p8809 line 23: Not clear what "degeneration of the fits as a function of data density" means, and how you saw that this is not a problem.

p8810 line 13: Given the substantial differences in amplitude, I would not share the statement "captured fairly well", unless you only refer to the phase of the seasonal

cycle.

p8810 line 16-17: This seems to suggest that the underestimation was a feature of Neural Networks. Why should that be the case?

p8811 (error calculation):

* Shouldn't there be a component from pCO2 measurement error?

* It is not clear to me how the "discretization error" arises - does it refer to pCO2 variability within the 1x1 degree grid boxes? But if so, is the 400km decorrelation length scale appropriate to be used in this context?

* If the error of the Neural Network is estimated from the statistics of the residuals, isn't there the possibility of potentially larger errors in data-sparse regions that would remain undetected, resulting in an underestimation of regional errors? This could e.g. be the case in those areas where you find large differences in Fig 6.

* The error of 0.2 uatm for atmospheric pCO2 seems low: Even though individual atmospheric measurements will indeed have errors on this order, you are using the GLOBALVIEW product which involves heavy spatial extrapolation from the actual mea- surement stations to whole latitude belts. (Of course, this error component is a minor one anyway.)

p8811 line 12: "effective number of degrees of freedom"

p8814 line 15: Give equation number from Takahashi et al (1993) for unique documen- tation.

p8814 line 17-29: As the Neural Network implicitly establishes a relation between pCO2 and SST, couldn't this be used in the discussion of thermal and non-thermal components (e.g. by driving the network with only SST variability)?

Sect 3.6, discussion of trends: If I understand correctly, then trends in estimated pCO2 can only arise from trends in one of the driving variables. Which of the variables is there

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

to provide a trend? Do you expect that there may be trends in reality that are not in parallel to trends in the inputs? A short discussion of these issues may be appropriate here.

p8816 line 18: Not clear what "differential" means here.

p8816 line 24: "carbon sink increase" (because it becomes more negative).

p8818 line 1: "Northern Hemisphere" - this probably only means the Atlantic?

p8818 line 4: Avoid "often" as it is not clear what this means.

p8818 lines 4-25: If the drivers of the Neural Network correlate to NAO, the resulting pCO2 estimates may correlate to NAO as well, even if the pCO2 data do not contain such a signal. The presented material does not seem to exclude this possibility of artificial correlation. Do you have further indication that the NAO correlation indeed come from pCO2 itself? If not, then this caveat should be stated.

Appendix A1: The authors did a good job in describing only main method features in Sect. 2.2, and giving details in the appendices. Nevertheless, for readers interested in more details but lacking a deeper Neural Network background (like myself) the appendix is difficult to follow, as the context of terms like "distance" (p8820 line 8), "winner" (line 16), "log-normalizing" etc. is not explained. It also remained unclear to me what "neurons" means in the SOM context, and what the meaning of a "hexagonal grid" is and how it may affect the outcoming provinces. Is the algorithm related to cluster algorithms? Are the provinces purely mathetical, or can you interpret them in terms of more conventional biome classifications? I think a bit more explanation (and motivation) is needed here.

I was also wondering whether the SOM algorithm (to the extend I understood it right) couldn't lead to fuzzy provinces (if vectors in close-by locations would be classified into different provinces)?

p8822 line 18: Spell out "infinity"

p8823 line 6: What is "paraboloidal" here? And should it be "paraboloidally"?

p8823 line 12: Say what "r_opt" is (size ratio?). Is the split into training and validation done randomly?

Fig 2: Why do (a) and (b) not cover the same pixels?

Fig 3: Explanation of the box symbols is missing.

Fig 6 and 10: The positive and negative extremes of the color scale are disadvantageous, as they are similar, and in particular as negative extremes (reddish color) appear as positive values.

Typos:

p8802 line 1 "amplitude dominated"

p8813 line 17: "network"

p8816 line 17: "thermally driven"

p8824 line 17: spurious "in"

p8825 line 11 (and later): "borders"

———————————————————

Interactive comment on Biogeosciences Discuss., 10, 8799, 2013.