Biogeosciences Discuss., 10, C4816–C4834, 2013
www.biogeosciences-discuss.net/10/C4816/2013/

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.

P. Landschützer et al.

p.landschutzer@uea.ac.uk

Received and published: 6 September 2013

We would like to thank reviewer#1 for the detailed review of our manuscript and the constructive and thoughtful suggestions and comments which will be answered point by point below.

Specific comments:

Reviewers comment: The paper gives contradictory statements on whether the variability of the Atlantic CO_2 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.

Authors response: We agree that the statements, both in the abstract and in the text, were not clear. In the abstract the statement p8800 line 3 "... but this sink is known to vary substantially in time" refers to findings from previous studies in the North Atlantic (also investigating different time-scales), whereas p8801 line 1 "The Atlantic carbon sink varies relatively little on inter-annual time-scales" refers to our basin-wide findings from this study. We suggest re-phrasing both statements to:

(p8800 line 3) "... but previous studies suggest that the sink strength varies substantially in time"

(p8801 line 1) "Within our study period from 1998 to 2007, the Atlantic Ocean carbon sink varies relatively little on inter-annual time-scales"

On p8817 line 17 the statement "The sea surface $p\mathrm{CO}_2$ exhibits substantial year-to-year variability within the North Atlantic north of 40 N and the eastern Equatorial and South Atlantic" was meant to indicate where we find the strongest year-to-year variability of the sea surface $p\mathrm{CO}_2$ by region, while statement p8817 line 21 "Integrating our monthly air-sea CO_2 flux estimates for each year over the Atlantic ocean reveals small but significant inter-annual variability" refers to annual average fluxes over the entire basin, but falsely uses the term inter-annual variability instead of year-to-year variability. In general the term year-to-year variability is used when comparing annual means, whereas inter-annual variability is calculated as a 12 month running average of the integrated CO_2 flux.

We therefore suggest re-phrasing both statements and avoid the use of the terms "substantial" and "significant" (as both annual minimum and maximum are within each other's uncertainty range):

(8817 line 17): "The largest year to-year variability's of the sea surface pCO$_2$ are found within the North Atlantic north of 40°N and in the eastern Equatorial and South Atlantic."

(8817 line 21 to 23): "Integrating our monthly air-sea CO$_2$ flux estimates for each year over the Atlantic Ocean reveals the largest annual mean flux differences during the second half of our study period (Fig. 11a), where annual mean fluxes range from -0.39±0.13 Pg C · yr^{-1} in 2001 up to -0.56±0.18 Pg C · yr^{-1} in 2006."

Reviewers comment: 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?).

Authors response: The phrase "overcome most of these limitations" refers to the methods described in the previous paragraph. While we use a relatively fine spatial resolution, the neural network method is further capable of reconstructing small variabilities, due to the non-linear input-output relationship. The different methods described in the introduction are "designed" to answer specific questions, like e.g. long term trends. Our method aims to investigate the inter-annual variability of the CO$_2$ flux. Therefore, we intended to use this paragraph to briefly explain the features of the method, rather than making a statement of which method performs better or worse. This would be more suitable for an in-depth inter-comparison study.

For more clarification, we suggest to re-phrase p8804 line 8 to:

"In order to investigate the variability of the sea surface pCO$_2$ and the resulting carbon flux, we overcome most of these limitations by presenting a new neural network based approach, which is capable of capturing a large amount of variability due to the

non-linear predictor-observation relationship on a fine 1°×1° spatial grid. The method determines the non-linear..."

Reviewers comment: p8805 line 7: "same" -> "analogous"?

Authors response: We don't think that a change of wording is appropriate here. The provinces are not bound to a particular spatial region; hence if e.g. province number 7 occurs in both the subtropical Atlantic and the Pacific Ocean, the feed-forward network will use all data within province 7, i.e. the same province, to estimate the input-output relationship.

Reviewers comment: 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?

Authors response: This paragraph aims to explain how the data sets (FINP, FINP2, etc.) are arranged. This includes data sets that do and don't have corresponding targets (see Table 1). This is a necessary step for the method and we felt this should be stated explicitly. We do agree that this paragraph is not fully clear and re-phrasing also provides more clarity regarding the importance of all data sets, as their meaning was not always clear (comment below regarding data set FINP2)

We therefore suggest to re-phrase the text to:

"In the next step the monthly 1°×1° input data are rearranged into 3 major data sets. Each of these data sets consists of input vectors ($\boldsymbol{p}_n$) where the input data are organized as row vector elements, for example SST, log(MLD), SSS, and pCO$_{2,Takahashi}$ for the self-organizing map input (SINP) dataset, sampled at the same space-time point

(Table 1). Two of these sets, SINP and the feed-forward network input 2 set (FINP2) are global sets and do not have a corresponding target dataset (Table 1). Input vectors with empty vector elements, e.g. where no salinity data are available, were removed from these data sets. The third major set, the feed-forward network input set (FINP), consists only of input vectors where corresponding SOCAT v1.5 observations, or targets (t), are available, i.e. they are subsampled in time at the locations where observations are available. In order to train the feed-forward network, two sub sets of the FINP set are created, namely the actual training (FITR) set and a validation (FIVAL) set (Table 1, Appendix A2)."

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

Authors response: We expanded the paragraph for more clarification to:

"We use a self-organizing map (SOM) method (Kohonen, 1987, 2001) to partition the global ocean into 16 biogeochemical provinces, characterized by all data observations having a similar relationship among all input variables of the SINP data set, i.e., climatological $p\mathrm{CO}_2$ as well as the independent variables SST, log(MLD) and SSS. The provinces change in shape from one month to the next and further change slightly between years. A SOM is a neural network based cluster algorithm that can detect regularities within the provided input data and then learns to group them together. Similar input data, arranged as input vectors, are identified via their Euclidean distance towards the nodes (or neurons) of the network."

Reviewers comment: p8807 line 20: "and" -> "but" (to clarify sentence)

Authors response: We have replace "and" with "but" for clarification.

Reviewers comment: 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.

Authors response: Static provinces do have the advantage that one can easily analyse the results within the province. One will encounter, however, strong $p\mathrm{CO}_2$ variability (seasonal, inter-annual, etc.) within one province. For most parts of the ocean the seasonal signal can be much larger than the inter-annual signal, and therefore regression methods might burry this small inter-annual "noise" in order to fit the seasonal signal. Our approach is to keep the seasonal $p\mathrm{CO}_2$ variability as small as possible within each province, so that (in an ideal case) the only "noise" left within a province stems from the inter-annual variability.

We have now added the following sentence prior to line p8807 line 26:

"We do not provide any additional time or space information to the SOM, hence the regions are strongly influenced by the temporal variability of the input data, in particular the seasonal variability within the climatological $p\mathrm{CO}_2$, and are therefore not static, unlike conventional provinces or biomes. Despite ..."

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

Authors response: We have included a brief explanation earlier (p8807 lines 6-9) in the data section. This paragraph states: "To highlight anomalies and year-to-year trends within our data sets, we further produced de-seasonalized sets of our input variables by removing their long-term mean seasonal cycle from 1990 to 2010 (1998 to 2010 in the case of chlorophyll a and 1992 to 2010 in case of MLD and SSS) from the original dataset." We think it is more convenient to leave this paragraph within the data section.

Reviewers comment: 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.)

Authors response: We have now included a better description of the data sets for more clarification (see comment above regarding the re-phrasing of paragraph p8807 lines 1-5)

Reviewers comment: 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.

Authors response: Available observations are not equally sampled in time and space and the $p\mathrm{CO}_2$ distribution as well as the distributions of the other input parameters co-located to the observations are skewed. Therefore, it was important to us to test if the residuals reveal a larger misfit between estimated $p\mathrm{CO}_2$ and observed $p\mathrm{CO}_2$ in e.g. cold waters, or high chlorophyll regions, where we do not have a large amount of data. Figure 3 shows that this is not the case. The average of the residuals stays close to 0, independent of the data density (shown as bar plots below each sub figure). We can therefore say that "Figure 3 shows that there is no indication of a substantial degeneration of the fits (upper panels) as a function of data density (lower panels)."

Reviewers comment: 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.

Authors response: We have re-phrased this statement to:

"While the phase of the seasonal cycle is captured fairly well, Fig.4 shows that the neural network estimates in general underestimate the winter minima at Bermuda from January to April and the autumn maxima at ESTOC from August to November."

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

Authors response: We agree that this statement and the reference are misleading. It is not an artefact of neural networks in general, but rather results from the generalization approach we chose. This sentence has now been changed to:

"This underestimation of the seasonal amplitude is likely linked to the early stopping approach (see appendix) which prevents our neural network from overtraining."

p8811 (error calculation):

Reviewers comment: * Shouldn't there be a component from $p\mathrm{CO}_2$ measurement error?

Authors response: We agree that there are sources of uncertainty, such as the $p\mathrm{CO}_2$ measurement error that we did not account for. However, there are 2 main reasons we decided to not include the measurement error into our analysis. The first one is that measurement systems differ depending on the vessel or timeseries station which collected the data. One would have to evaluate this measurement uncertainty for each system individually to accurately quantify the overall error. The second reason is that in general measurement uncertainties are in the order of a few percent, which makes them negligible compared to the other sources of uncertainty we account for (especially when considering autocorrelation).

Reviewers comment: * It is not clear to me how the "discretization error" arises - does it refer to pCO$_2$ variability within the 1x1 degree grid boxes? But if so, is the 400 km decorrelation length scale approprate to be used in this context?

Authors response: As we use the gridded observations for our analysis, we simply use the standard deviation of the pCO$_2$ as reported by Sabine et al. 2013. We then use the 400 km decorrelation length scale to identify how many of the $1^\circ \times 1^\circ$ boxes are decorrelated. This is indeed a simplification of the problem, but considering that the uncertainty derived by this estimate is minor compared to the other sources of uncertainty we account for, we think this is an appropriate way to reduce the degrees of freedom derived from this source of uncertainty.

Reviewers comment: * 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.

Authors response: We do agree that this could be the case. For example, Fig. 2 shows that the neural network misfit is not homogeneous in space. However, in this study we only report fluxes and their uncertainty for large and basin scale regions. Therefore we think that our estimate is appropriate. If we were to report uncertainties of the air-sea fluxes integrated over much smaller regions, then it would be appropriate

to use the RMSE of the individual region. Underlying all of this is the power of large numbers, i.e. that the error of the mean decreases with the square root of the number of independent observations.

Reviewers comment: * The error of 0.2 uatm for atmospheric pCO$_2$ 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 measurement stations to whole latitude belts. (Of course, this error component is a minor one anyway.)

Authors response: The 0.2 μatm estimate stems from Takahashi et al. 2009. The authors also use the GLOBALVIEW product and further account for the effect of barometric pressure. As this error component is a minor one we decided to keep this estimate unchanged. For clarification where this estimate actually comes from, we now included the reference to Takahashi et al. 2009 in the manuscript p8811 line 22.

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

Authors response: We have now changed this in the manuscript.

Reviewers comment: p8814 line 15: Give equation number from Takahashi et al (1993) for unique documentation.

Authors response: We have now included the equation number (Equation 2) with the reference as well as stated the actually used partial derivative.

Reviewers comment: p8814 line 17-29: As the Neural Network implicitly establishes a relation between $p\mathrm{CO_2}$ 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)?

Authors response: By nature, variabilities as they occur in the measured $p\mathrm{CO_2}$ data are not only driven by SST changes. Using only SST as regression parameter could lead to false relationships where the sea surface $p\mathrm{CO_2}$ is driven by another input parameter. An example would be the subpolar North Atlantic north of 40°N where the seasonal $p\mathrm{CO_2}$ cycle is antiphased to the SST cycle. Here, the neural network could only use SST to explain $p\mathrm{CO_2}$ variabilities and would therefore suggest a negative relationship between SST and $p\mathrm{CO_2}$ which is clearly not the case as the drawdown of $\mathrm{CO_2}$ is driven by biology.

Reviewers comment: Sect 3.6, discussion of trends: If I understand correctly, then trends in estimated $p\mathrm{CO_2}$ can only arise from trends in one of the driving variables. Which of the variables is there 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.

Authors response: Correct. The atmospheric $\mathrm{CO_2}$ is the main driving variable for the trend. We are confident that we do give a realistic trend estimate based on the results presented in Table 2 (p8835). Here we show that there is no temporal bias between our estimates and the available observations considering the entire basin. However, we cannot make a statement where we do not have observations or any other validation data. We added the following short paragraph on page 8815 at the beginning of section 3.6:

"The main driving variable for trends in the sea surface $p\mathrm{CO_2}$ is the atmospheric $\mathrm{CO_2}$, but within our study period the neural network estimates show that these trends are not in parallel. Across ..."

We further added at the end of the first paragraph in section 3.6:

"Table 2 shows that the bias between estimates and observations is fairly constant at each year individually, suggesting that trends are captured well where observations exist."

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

Authors response: "differential increase" means atmospheric trend minus ocean trend. We suggest re-phrasing this sentence for more clarity:

"The small difference in the increase in surface ocean $p\mathrm{CO_2}$ relative to the atmosphere results in an almost steady strength of the Atlantic carbon sink in time ..."

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

Authors response: We have changed "carbon flux increase" to "carbon sink increase" on page 8816 line 24

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

Authors response: Correct. "Northern Hemisphere" and "Southern Hemisphere" in this context refer to the Atlantic Ocean. For more clarity we changed p8818 line 1 to:

"... for the Atlantic Ocean north of the Equator, south of the Equator and the entire basin ..."

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

Authors response: We have re-phrased the sentence on page 8818 line 4 to avoid the term often:

"Inter-annual variability of the sea-surface pCO$_2$ in the North Atlantic has previously been linked to variations in the North Atlantic Oscillation (NAO) (e.g. Gruber et al., 2002; Schuster and Watson, 2007; Thomas et al., 2008)."

Reviewers comment: p8818 lines 4-25: If the drivers of the Neural Network correlate to NAO, the resulting pCO$_2$ estimates may correlate to NAO as well, even if the pCO$_2$ 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 pCO$_2$ itself? If not, then this caveat should be stated.

Authors response: We agree that signals within our estimates stem from the input data, however the transformations of these input data into pCO$_2$ are non-linear, so a linear correlation in the input data with NAO does not automatically mean that pCO$_2$ correlates with NAO. Therefore, we regard this as an emergent property. We agree that this needs to be mentioned in the manuscript and we added following to page 8818 line 22:

"The correlation pattern are derived from the neural network estimates, hence the NAO signal stems from the signal of the input data. Clearly, an important ..."

Reviewers comment: 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.

Authors response: We agree that this part is very technical and hard to follow. We have now re-phrased parts of appendix A1 (below) to give the reader an easier introduction to the details regarding the SOM algorithm used. The closest cluster algorithm to our knowledge is k-means clustering. However there are substantial differences, like the use of neurons, their initialization and their distance relation which are used in the SOM algorithm but not in k-means. The regimes are derived purely mathematical, i.e. we do not add any subjective criterion to it, e.g. time or space information. Therefore the provinces evolve freely and are not fixed in time, unlike conventional steady biomes. We therefor tried to avoid using the term "biome", as we thought this could cause confusion. We adjusted page 8820 paragraph 1 and it reads now:

"A map with 16 neurons was chosen, organized on a 2 dimensional 4$\times$4 point hexagonal grid. This means that the input data are clustered into 16 neurons, which represent the 16 biogeochemical provinces. The term neuron refers to a processing unit, which consist of a weight vector, where each element of the weight vector corresponds to one input parameter. In our case each weight vector consists of 4 elements (SST, log(MLD), SSS, $p\text{CO}_{2,Takahashi}$), representing its co-ordinates and the distance between 2 neurons is calculated via a distance function. These processing units are initially spread over a 2 dimensional field, in our case in a hexagonal formation, forming a single layer of neurons. Our experience has shown, however, that the choice of neuron topology does not have a significant effect on the final province distribution. The use of neurons, their initialization and their distance relation describes the biggest difference towards other clustering algorithms, e.g. k-means clustering. For our study, the Euclidean distance between a neurons weight vector and the input vectors of the SINP dataset was used for the distance function. The weight matrix ($\mathbf{W}_{m=16,n=4}$), which is formed by the 16 neurons with their 4 vector elements, was randomly initialized."

We added following sentence at page 8820 line 15:

"... of the input vector $\boldsymbol{p}_n^j$. The smallest element of the distance vector, i.e. the shortest distance element, marks the distance towards the closest neuron, called the winning neuron. The neuron i, gets updated ..."

We edited page 8821 line 13 onwards:

"We forced the relative weights of the input data toward the climatological $p\text{CO}_2$ data, in order to minimize the variance of $p\text{CO}_2$ within each biogeochemical province. To do so, we did not normalize our input data, with the exception of MLD, which we log-transformed (Table 1). As a consequence, the range between the lowest and highest value of $p\text{CO}_2$ is one order of magnitude larger than that for SST, and about another order of magnitude larger than that for the remaining input parameters (log(MLD), SSS)."

Reviewers comment: 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)?

Authors response: This is possible. As we do derive our provinces from our input data, steep gradients between neighbouring pixels in the input dataset, e.g. the $p\text{CO}_2$ climatology, can lead to small "island" provinces in certain regions. We do however leave these regions unaltered.

Reviewers comment: p8822 line 18: Spell out "infinity"

Authors response: Infinity is now spelled out

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

Authors response: The correct wording is "parabolically". This stems from the open plane curve "parabola". This has now been corrected.

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

Authors response: r_{opt} is a size ratio and depends on the number m of modifiable parameters (weights [**W**] and bias [***b***] - see equation A3). Therefor the split is not random, but changes depending on the number of neurons (which we increase parabolically during the training) We agree that this is not clear on page 8823 line 12 and we suggest re-phrasing to: "Amari et al. (1997) suggested an optimal split (r_{opt}) between training and validation data as a function of the modifiable parameters m:

$$r_{opt} = \frac{1}{\sqrt{2 \cdot m}} \quad (1)$$

where m consists of weights (**W**) and biases (***b***) of the network."

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

Authors response: We calculate the standard deviation of the residuals within each pixel only when there are at least 2 observations available in time. Hence, if there is only 1 observation available in time, this pixel will have a mean error (a) but no standard deviation (b) and the pixel is empty.

To clarify this issue we added the following to the figure caption:

"Pixels that have a value in (a) but not in (b) indicate where only 1 observation in time is available."

Reviewers comment: Fig 3: Explanation of the box symbols is missing.

Authors response: We have added a more detailed explanation to the figure caption. It states now:

"The upper plot in each panel shows the residuals, shown as a box-and-whiskers plot. The red line in the box show the median, the blue box indicates the 25 and 75 percentiles and red plusses mark residuals outside this interval. The lower plot shows the relative number of observations within each bin"

Reviewers comment: 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.

Authors response: We now changed the lower extreme colour from purple to a darker shade of blue.

Typos:

Reviewers comment: p8802 line 1 "amplitude dominated"

Authors response: done

Reviewers comment: p8813 line 17: "network"

Authors response: done

Reviewers comment: p8816 line 17: "thermally driven"

Authors response: done

Reviewers comment: p8824 line 17: spurious "in"

Authors response: "in" has now been removed

Reviewers comment: p8825 line 11 (and later): "borders"

Authors response: done

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