Biogeosciences Discuss., 6, C900–C917, 2009
www.biogeosciences-discuss.net/6/C900/2009/

Biogeosciences Discussions

Interactive comment on "Estimating the monthly pCO$_2$ distribution in the North Atlantic using a self-organizing neural network" *by* M. Telszewski et al.

M. Telszewski

m.telszewski@uea.ac.uk

Received and published: 23 June 2009

Dear Editor,

Below follow comments on the manuscript: "Estimating the monthly pCO2 distribution in the North Atlantic using a self-organizing neural network" by Telszewski et al.

We thank all the referees for their thorough and thoughtful reviews, and constructive comments on our manuscript. We agree with most of the suggested edits and we think that the corrections applied have clarified and improved the text. In our response below the issues, which have been pointed out by more than one referee, are addressed collectively. Otherwise, we first respond to comments made by Referee 1 (Yonggang

Liu), then to comments made by Referee 2 (David Hydes) followed by comments made by Referee 3 (anonymous). We hope that the manuscript in its revised form will be acceptable for publication in Biogeosciences.

—Yonggang Liu (General comment 1): As a novel and powerful data analysis tool, the SOM is becoming popular in various disciplines. I believe this timely work will have an immediate impact in biogeosciences community. The description of the method is clear. The assumptions and the implementation of the experiments (SOM analysis) are acceptable, and the results are better than those obtained by others.

ˆˆˆOUR REPLY: Thank you.

—Yonggang Liu (General comment 2): The paper can be more concise, and English writing also needs to be improved in general.

—David Hydes (General comment 1): The English and clarity could be improved if consistent use were made of the definite and indefinite articles ("the" and "a").

ˆˆˆOUR REPLY: The article will be shortened whenever possible. The co-authors will check the English of the final version of the manuscript.

—Yonggang Liu (Specific comment 1): P3385 L20 and P3396 L8-9; Another explanation of the smoothed patterns is the shape of the neighbourhood function. The Gaussian neighbourhood function was chosen in the SOM training. According to a sensitivity study on the choice of neighbourhood function in extracting the known patterns of progressive sine waves, the Gaussian neighbourhood function returns the smoothest SOM patterns, while the Epanechikov (ep) neighbourhood function gives the most accurate patterns (see Figure 5 of Liu et al., 2006b). This difference is also seen from the real geophysical data (moored time series of coastal ocean currents). The patterns of strong currents due to the hurricane and tropical storm activities were extracted using the "ep" neighbourhood function (Liu et al., 2006b), however, these extreme current patterns were not seen from the SOM patterns using the Gaussian neighbourhood function (Liu and Weisberg, 2005). So, if the "ep" function is used instead, the SOM patterns might be less smoothed. A repetition of the analysis with a different neighbourhood function is not suggested here, however, a few sentences need to be added in the text to clarify the reason why some strong blooms are smoothed by the SOM. This would be indicative to other SOM users as well. Both papers mentioned above (Liu and Weisberg, 2005; Liu et al. 2006b) have already been cited in the manuscript.

^^^OUR REPLY: We agree with the referee and have added the proposed adjustment to the SOM mapping procedure, as a potential improvement of SOM estimates. However, the real benefit of using "ep" neighbourhood function in this particular case is minor. Our experiments (not implemented in the manuscript) suggest that the monthly values for July and August in the eastern subtropics increase by 1-2 μatm when an "ep" neighbourhood function is used instead of the Gaussian neighbourhood function. Although not negligible, such a change is relatively minor. The SOM estimates are still $\sim$ 10 μatm lower than other reports suggest for July through September in the NAST(E), and other causes for the SOM to underestimate (overestimate) the highest (the lowest) values are being investigated. We are currently exploring suggestions from this manuscript such as the introduction of additional training parameters (e.g. salinity) and an increase in the spatial and temporal resolution of the training data to reduce the potential effect of data averaging. We are also investigating the sensitivity of the method to the input data source (satellite data as opposed to the reanalysis data). Finally, we are looking to improve the SOM estimates by using the maximum chlorophyll a concentration (instead of the average concentration) for each pixel during the training and labelling as suggested by Ono et al. (2004).

—Yonggang Liu (Specific comment 2): Authors state in P3380 L14-15: "The SST analyses were done weekly and interpolated linearly to daily values", and later in P3381, L4: "All parameters were re-gridded onto weekly frequency". It seems that the interpolation to daily time series is not necessary, because the SST product is already weekly. An interpolation in time may be necessary if the two weekly time stamps are different.

Even though the extra linear interpolation may not substantially modify the SST, the less manipulation of the original data, the better.

^^^OUR REPLY: The sentence in P3380, L14-15 was supposed to inform the reader that the product we are using, despite being called "daily" is in reality linearly interpolated Reynold - or NOAA Oiv2 weekly SST data. Note at http://www.cdc.noaa.gov/data/gridded/data.ncep.reanalysis.surfaceflux.html under Caveats: "The skt.sfc files contain skin temperature as described in the March, 1996 BAMS article. As such, over land and sea ice, the temperature is a prognostic variable. Over open water, the skin temperature is fixed at its initial value; i.e., the Reynolds SST as seen by the model. The Reynolds' SST analyses were done weekly and the reconstructed SST done monthly. The analyses were linearly interpolated to daily values which were used for all four analyses". The skt.sfc files are used in this study, and the four analyses mentioned at the end are 6-hourly, daily, weekly and monthly. We acknowledge the confusion pointed out by the referee, and the modified sentence reads: "The NCEP SST data (used in this study) contain the temperature data as described in Kalnay et al. (1996). As such, over open water the temperature is fixed at its initial weekly value and linearly interpolated to daily frequency in the NCEP data product". As the referee rightly noticed, an additional interpolation in time is necessary because the two weekly time stamps are different. Our eight-daily week is determined by the minimum CHL data frequency (daily CHL data proved too patchy in a basin-wide context) and therefore we could not use weekly (seven-daily) temperature data.

—Yonggang Liu (Specific comment 3): In a recent paper (Friedrich and Oschlies, 2009), basin-wide monthly maps of pCO2 in 2005 were derived from model results and satellite SST and CHL using the SOM (they called it "KFM" instead). I suggest the authors to cite that paper and compare this work with theirs. As I see, the data sets are different (there are some overlaps, though), and the results are much better in the present analysis according to the reduced RMS values (12 versus 20 μatm).

—David Hydes (Specific comment 6.1): You report an RMS of 11.55 uatm. This is smaller than the RMS reported by Freidrich and Oschlies (2009). Can you comment on this?

—Anonymous Referee #3 (General comment 1-partial): Coming mainly from the same project (CarboOcean) this paper can be regarded as the practical part following the methodological study published 3 month ago by Friedrich and Oschlies in JGR (F&O in the following). I assume that the presented manuscript was already with the numerous co-authors for approval when F&O was published. However, in the revised version of the manuscript the findings of F&O with respect to the basin-wide uncertainties of the pCO2 estimates need to be discussed. The text describing the method's uncertainty in estimating pCO2 is very confusing and the RMS-error of 11.55 uatm given in the abstract is really misleading. The pCO2 values memorized by the SOM are averages of the VOS-line pCO2 data. Thus, the given overall RMS-error represents a validation against a data-set that is at best semi- independent. F&O pointed out that this way of validation is not representative of the basin-wide error (see their Figure 9). In fact they found the basin-wide error to be about 3 times higher (including water depths < 500m and the Mediterranean and Labrador Sea where there were no data available). For a more indicative uncertainty estimate I highly recommend to take 2 of the 3 years for labeling the SOM and to validate it against the remaining year and repeating this for all 3 per- mutations. One main focus of the presented study should be to present an uncertainty estimate as reliable and as representative as possible rather than attracting attention with a low RMS-error that may not be realistic.

ˆˆˆOUR REPLY: The suggestions by the referees to add a discussion on Friedrich and Oschlies (2009) to the manuscript reflect our plans. It is worth pointing out that we did not discuss our findings with F&O in our initial submission, as this manuscript was accepted for discussion in BGD before the F&O paper was published. However, there are several fundamental differences between F&O and this study in terms of the SOM utilization. These three are the most relevant to the above comments.

Firstly, F&O's monthly maps are not for 2005. They are for model year 11, which aims

to mimic the preindustrial seasonal pCO2 cycle in the North Atlantic. The model is integrated over a 10 year spin-up period during which atmospheric pCO2 remains on a preindustrial level, and DIC is taken from the preindustrial estimate. Therefore, any comparison between the two approaches needs to be performed very carefully.

Secondly, their "satellite" SST and CHL are also taken from model year 11. They are synthetic "measurements" modelled with the preindustrial set up. This again, makes comparing the results somewhat tricky, since the environment in which the SOM (their KFM) is set up is far from similar.

Thirdly, the most fundamental difference lies in the application of the SOM training procedure. As described in section 2.3 of this manuscript we use 3 years' worth of the whole grid data (SST, MLD and CHL) to train the SOM. This way the SOM "sees" the relationships between the training parameters in every grid point in the North Atlantic, with weekly frequency for the three years. This enables maximum SOM efficiency, regardless of the spatio-temporal VOS cover, and ensures that the SOM has been preconditioned with comprehensive, basin-wide training knowledge with regards to the relevant biogeochemical processes. F&O decided to train the SOM (KFM) only with values (SST, CHL) "collected" along the VOS lines, using 2005's coverage metadata. Such a small data set carries very limited training knowledge, despite the very successful data gathering campaign in 2005. Some processes occurring in the vast extent of the basin are never sampled (and therefore not included in the training), and when they are sampled, it very often happens only a few times during the year (Friedrich and Oschlies, 2009; Fig. 2 for monthly cover and Fig. 6 for seasonal cover). It is hardly surprising that their KFM produces poor estimates for regions outside the sampling network (their Fig. 6). The SOM, by definition cannot reliably estimate output values for input values from outside the training data space, and that is what F&O are essentially trying to achieve.

Finally, the reason for the two RMS values being so different, or one being better than the other, needs an explanation. F&O had the luxury of knowing the "real" modelled pCO2 distribution in the North Atlantic to which they can compare the KFM estimates. The $\sim$20 μatm RMS is the basin-wide KFM error estimate based on their mapping procedure. Our RMS by default can only relate to data points along the sampling network, which we state in the text. We understand the need for a wider error estimate and therefore in an underway study we combine the training procedure described in this manuscript with the basin-wide modelled fields to show how an RMS along the VOS lines relates to the basin-wide RMS. The preliminary results of this study show that an RMS along the VOS lines of 9 μatm relates to the basin-wide RMS of 9 to 16 μatm, depending on the season. This suggests that the two RMS estimates (along the VOS lines and basin-wide) are much more closely related in the training scheme employed in this study than in that employed by F&O.

—Yonggang Liu (Technical corrections): (1) P3375 L16, "cover" should be "coverage". This change should also apply to many other places in the text, e.g., P3376 L15&27, P3381 L1&3, P3390 L25, P3393 L2, Table 1. (2) P3384 L2, "Many more" should be "More". (3) P3384 L7, "few" should be "less". (4) P3386 L12-13, the sentence "Most neurons. . .of the data" needs rewriting. (5) P3388 L10, "daily" should be "day". (6) P3399 L9, "Weisberg R." should be "Weisberg R. H." (7) P3399 L11, "Liu, Y. and Weisberg R." should be "Liu, Y., Weisberg, R. H., and He, R." (8) P3399 L13, "Weisberg R." should be "Weisberg R. H."

ˆˆˆOUR REPLY: Done.

—David Hydes (Specific comment 1): The work is a natural progression from that of Lefèvre et al., (2005). I would have like to have seen more acknowledgement of this in the introduction, explaining the relative advances made in this paper.

—Anonymous Referee #3 (General comment 2-partial): It should be mentioned what the additional benefit of this study is compared to e.g. Lefevre et al. [2005] and Jamet et al. [2007]. At first sight their approaches result in similar uncertainties, although Lefevre et al. [2005] were not able to use Chl or MLD.

ˆˆˆOUR REPLY: We will extend the discussion of advances of the recent work in relation to Lefèvre et al. (2005) and will add a few lines describing the differences between this work and that of Jamet et al. (2007).

—David Hydes (Specific comment 2): Section 2 could (and should) be reduced considerably in length if more reference was made to the Lefèvre et al papers. The explanation of the methodology was more clearly expressed in the Lefèvre paper. In this paper, which is aimed at biogeochemists the aim should be to try and explain the method in words they can easily grasp rather than repeating text, which reads like the software manual probably did.

—Anonymous Referee #3 (Specific comment 3): The description of the methodology is very long. Maybe it would be enough to refer to Kohonen and Lefevre et al. [2005] and focus on the different labeling scheme used here.

—Anonymous Referee #3 (General comment): ...the description of the method is thorough and clear.

—Yonggang Liu (General comment): The description of the method is clear.

ˆˆˆOUR REPLY: While we understand the position of David Hydes and the Anonymous Referee, we feel that it is of interest to our peers in the field to have an improved insight into application of this novel and powerful data analysis tool. Based on questions and discussions during meetings and conferences where we presented our work, there is a fair amount of confusion amongst marine biogeochemists as to what the SOM exactly is and how it "works". Therefore we decided to extend the description of the method beyond that in Lefèvre et al. (2005). An expert in the application of neural networks to environmental data, Yonggang Liu stated that our explanation of methodology is clear. In times when ever increasing number of variables is measured from space, offering the full grid coverage, the SOM could become a standard tool for approximating several marine variables. For that to have happened, a successful, field-specific applications must be reported together with in-depth method's description. Additionally, as discussed above, the SOM can be applied to the data in at least two very different ways and we believe that it is important for readers to know what they are. Therefore we argue that section 2 of this manuscript should not be reduced in length. Nevertheless, should the referees or the editor feel strongly about this, then we will reduce its length by referencing particular elements of the SOM mapping procedure to previously published applications.

—David Hydes (Specific comment 3): Page 3380, line 25. Why was the criterion of a change of 0.05 kg m-3 used to determine the MLD? No reference is given to validate this choice and the criterion chosen can have significant effect on the MLD found. I bring this point up because in Figures 11 and 12, the deeper MLDs seen in 2006 do not correspond to lower surface temperatures and in the NADR production (indicated by the change in chlorophyll) took place in April and May with apparent MLDs of 350 and 200 m respectively. Did you check these MLDs against Argo data? Please comment.

ˆˆˆOUR REPLY: We concur with the referee that the criterion chosen to compute the MLD should be referenced. The criterion was chosen by the Met Office scientists running the FOAM model. It is beyond the scope of this manuscript to argue the choice of the particular model parameters in FOAM. We address referee's concerns by rephrasing the text which now reads:

"The mixed layer depth (used in this study) is determined in the FOAM model with a density based criterion, as the depth where the density increase of 0.05 kg m-3 from the surface value occurs (Chunlei Liu, Environmental Systems Science Centre of the U.K. National Environmental Research Council, personal communication, 2007)."

The MLDs were checked against Argo data for the period of January to May 2006. No systematic bias was found. The FOAM model assimilates Argo data together with other observations, therefore one would expect that the model output is accurate for positions and times when Argo data is available. Indeed we note that the deeper 2006 MLDs in Figures 11 and 12 do not correspond to lower SSTs and our assumption is

that perhaps changes in the MLDs between 150m and 350 m produce such a weak response in SSTs for this region.

—David Hydes (Specific comment 4): Page 3383, line 4. "Figure(s) 2a–c show the distribution of neurons within the input data space, visualizing how the SOM accounts for the non-linear relations between the components. The SOM is well equipped for such a complicated setup, e.g. the distribution of the neurons closely follows the data distribution, even in such an extreme case as MLD versus SST (Fig. 2b)." Can you quantify the correlations you are trying to show in Figure 2. I can't see that any "close following" is happening in 2a and 2b.

ˆˆˆOUR REPLY: Every neuron (three-dimensional vector) represents a number of input data points (three-dimensional vectors). Due to the incorporation of the neighbourhood function during training, it is impossible to assign specific input vectors to specific neurons. Figure 2a-c show the stretch of the so called map space (neurons) within the input data space, which is a qualitative measure of SOM's ability to represent the input data set. For better visualization three 2D plots rather than one 3D plot are used. We now added a histogram showing the frequency distribution of SST, MLD and CHL for the input data and SOM neurons. In the text we also compare the mean values (SST_SOM = 18.1, SST_INPUT = 19.4; MLD_SOM = 63.6, MLD_INPUT = 66.3; CHL_SOM = 0.37, CHL_INPUT = 0.27) and ranges of input data and SOM neurons with regards to all parameters. We argue that correlations are not a good measure for quantifying the non-linear and complex relationships between the parameters. Therefore, despite the correlations being very similar for each relationship shown in Figure 2a-c, (Figure 2a: r2_SOM = -0.04, r2_INPUT = -0.02; Figure 2b: r2_SOM = -0.23, r2_INPUT = -0.27; Figure 2c: r2_SOM = -0.25, r2_INPUT = -0.31) we do not discuss them in the text.

SOM neurons and input data points have similar frequency distributions, ranges and mean values for all parameters. We hope that these together with the histograms sufficiently address the concern expressed by the referee.

—David Hydes (Specific comment 5): Page 3384, line 18. A similar point is that a concentration of chlorophyll of 65 mg m-3 is probably an order of magnitude higher than is a any likely real value. How carefully was this data set reviewed before use? Do you believe the value of 65 mg m-3?

ˆˆˆOUR REPLY: We note that the sentence is misleading and could confuse the reader about the quality of the data set used. We left those very high (probably unrealistic) values in the data set for two reasons. Their influence on the mapping is negligible due to their small number. There are 142 (out of 389,336) data points with CHL values higher than 10 mg m-3 (0.04%), and there are 603 (out of 389,336) data points with MLD values higher than 1000 m (0.15%) in the data set. The second reason was to make use (during training) of the remaining two values in each data point. We have now added the above information to the text which now reads: "SST varies between -1.8°C and 30°C, the depth of the mixed layer ranges from ∼10 m to more than 1000 m (0.15% of data has MLD values greater than 1000 m) and chlorophyll a concentrations vary from 0 to ∼10 mg/m3 (0.04% of data has CHL values greater than 10 mg/m3). Table 1 is updated accordingly.

—David Hydes (Specific comment 6): Pages 3387 - 3390. "Monthly pCO2 maps": I would like to see a full set of monthly map presented. I would also like to see 2004 and 2006 presented as the difference to 2005, otherwise seeing the differences referred to later in the text is difficult. In addition the difference between the monthly maps for 2005 and equivalent maps based on the Takahashi et al (2008) data, adjusted to 2005, should be shown. Given the effort that has gone into producing the maps I think it would be good idea to show them. What do you think?

—David Hydes (Specific comment 10): Section 3.3. "Interannual variability". This section would be better supported by figures which show the actual difference between years. See comment 6.

ˆˆˆOUR REPLY: Figures 11 and 12 show interannual variability for specific regions and

Section 3.3. uses these graphs. To accommodate the referee's request we will add a figure (similar to Figure 6) showing 2004 and 2006 presented as the difference to 2005. Four different months (one from each season) will be used. We also agree that given the fact that the purpose of this study was to introduce the mapping technique, showing all the monthly maps computed is worth considering. However, we feel that adding those (7 pages with 12 panels each) to the manuscript will interrupt the flow of the argument and will not significantly improve the manuscript. Furthermore we quantify variability between the three years in Figures 11 and 12. Therefore we are exploring the possibility of placing the maps on the CarboOcean website, from where they will be freely available. The link to such a long term storage depository will be referenced in the appropriate sections of the manuscript. This way those interested can spend time analyzing them, while others will still have sufficient overview by looking at Figures 6, 11 and 12.

—David Hydes (Specific comment 6.2): Page 3388, line 3 to line 28. This is an information filled and important paragraph as far as an overview of the data fields is concerned. It would be good if it were expanded slightly and some sub-headings inserted.

ˆˆˆOUR REPLY: The text will be amended to incorporate a more detailed analysis of the maps. We will also divide the text by adding two sub-headings.

—David Hydes (Specific comment 7): Page 3389, line 17. "In both cases (the) SOM reproduces the label(l)ing data set well." As you point out later page 3390 line 6, this not the case. Please reword.

ˆˆˆOUR REPLY: We have removed the first sentence.

—David Hydes (Specific comment 8): Page 3390, line 21. "For each province we show the number of data points available for training and label(l)ing of the SOM." Could you take this a stage further a comment on the relative goodness of fit for different numbers of data points?

ˆˆˆOUR REPLY: We have added RMS and r2 values for all provinces and comment on the relative goodness of fit with relation to different numbers of data points.

—David Hydes (Specific comment 9, 9.1 and 9.2): Page 3390, beginning line 27 and Figure 9. I was pleased that a comparison was made to the Takahashi data set. I think one of its important uses is, that it provides a reference case and then challenges us to explain the differences we see. Page 3390, line 28 and Figure 9. I am puzzled as to what is being compared here. What do you mean when you say "we compare SOM estimates for a reference year 2005 (mean of the monthly SOM estimates for 2004 to 2006)" - what SOM output are you using? Similarly Page 3391 line 4 and Figure 9. Do you need to show the Takahashi data before and after adding 1.8 uatm per year?

ˆˆˆOUR REPLY: The only SOM output used in this study are 36 monthly maps between January 2004 and December 2006. For comparison with Takahashi's climatology we eliminated the interannual variability by averaging the 3 monthly values estimated for each grid point. For example, the grid point centred at 40°N and 40°W has one value estimated for January 2004, one for January 2005 and one for January 2006. The mean of the three values is our "SOM estimate for a reference year 2005". In theory we should add 1.8 μatm to all 2004 values and deduct 1.8 μatm from all 2006 values to match modifications performed on the Takahashi data, but because there is exactly one year on both "sides" of the reference year this modifications cancel out.

Showing Takahashi data before and after adding 1.8 μatm per year has its roots in an inconsistency emerging from Takahashi et al. (2009). These authors estimate the mean rate of annual pCO2 increase for the North Atlantic at 1.8 $\pm$ 0.4 μatm. However for their calculations they use the global mean value of 1.5 μatm. Therefore we decided to show the original Takahashi data as well as that after adding 1.8 μatm per year to give readers a chance to draw their own curve should they disagree with our reasoning.

—Anonymous Referee #3 (General comment 1-partial): The comparison with the MV Santa Maria data remains unclear to me. What is meant by "absolute value of mean

monthly residuals". Why not calculate the error as in equation 4 for having a validation against a truly independent data set that can be used for comparison with the results of F&O in order to get at least some first order estimate for the basin-wide accuracy?

ˆˆˆOUR REPLY: The reviewer is correct in pointing out the misuse of the metrics. We have removed the use of the absolute value of mean monthly residuals in section 3.1. Additionally, this measure is no more used in section 3.2 where a small independent data set (1600 data points) is introduced in an attempt to show how the SOM estimates compare with in situ data which has large sub-grid variability. We have changed the text on P3389 L19-20 and P 3390 L4-5 so it discusses an RMS (6.3 μatm and 19.3 μatm respectively) rather than an absolute value of mean monthly residuals. We also relate it to RMS values calculated for all the labelling data in 5 provinces requested by David Hydes in his Specific Comment 8.

—Anonymous Referee #3 (General comment 2-partial): I believe that this manuscript will be of great benefit for the biogeoscience community. Only, this benefit should be clarified to a broader scope of readers. Probably most readers are familiar with the necessity to better constrain the marine carbon uptake. So, what are the metrics of success for a basin-wide pCO2 mapping in the North Atlantic? How large is the uncertainty of the presented method with respect to CO2 uptake and it's interannual variability? Can we detect the anthropogenic impact on oceanic pCO2 with this method and this VOS-line coverage?

ˆˆˆOUR REPLY: Addressing these questions is beyond the scope of our investigation. This manuscript aims at providing the insight into the application of this novel and powerful data analysis tool. We have decided to concentrate on the pCO2 distribution as a scientific problem from which the progress will be steadily made towards the SOM-based air-sea flux estimates. We have recently submitted a manuscript dealing with air-sea flux estimates (Watson et al., 2009, submitted to Science), but quantitative SOM-based estimates of the anthropogenic impact on oceanic pCO2 are unlikely to be ever possible by the very nature of the method. However, we can detect long term regional and basi-wide trends in the pCO2 distribution and relate them to accompanying changes in other environmental parameters. Such an indirect analysis might provide useful qualitative information for more specific studies. In the Summary we state that our maps should serve as an alternative (to climatological) input for studies aiming at constraining marine carbon uptake. We will take it a step further and hint at our plans to verify the mapping quality with regards to air-sea flux estimates.

—Anonymous Referee #3 (Specific comment 1): Figure 3. The figure is somehow deceptive as it shows the cumulative coverage instead of what is available monthly/seasonally. The great challenge the authors are confronted with is (besides the large pCO2 variability) the lack of coverage for pCO2 observations. This should be illustrated by the figure. e.g. Similar to Figure 6: (4 Seasons) x (3 years) (Also the black lines on the blue background are hardly recognizable.)

ˆˆˆOUR REPLY: We will fully accommodate this request.

—Anonymous Referee #3 (Specific comment 2) Figure 5. The density of the scattered points is not clear. The way it is shown as a contour plot in Figure 2 is much better.

ˆˆˆOUR REPLY: Agreed. The figure will be changed.

—Anonymous Referee #3 (Specific comment 4) Friedrich and Oschlies [2009] pointed out that depending on the mapping procedure (daily, monthly) there might by a considerable impact of remote sensing errors on the pCO2 estimates. MLD products are still subject to unknown (and probably high) uncertainties. How would a MLD-error of 5%, 10%, 25%... affect the pCO2-error?

ˆˆˆOUR REPLY: The referee raises an interesting aspect of F&O's analysis. Showing the impact of artificially biased input data on the mapping performance gives the reader important sensitivity information. Knowing the "real" distribution of all parameters is advantageous for such analyses and therefore it is a real pity that F&O haven't extended theirs to include the impact that MLD related error has on the SOM estimates. With

regards to this manuscript, we do not have the means to perform such an analysis robustly. We hope that as more similar studies are performed using alternative model outputs, we will be in a better position to estimate such impact in the near future.

—Anonymous Referee #3 (Specific comment 5) Page 3398, line 21: Reference for Jamet et al. [2007] gives the wrong pages.

ˆˆˆOUR REPLY: Corrected.

—Anonymous Referee #3 (Specific comment 6) Page 3395, lines 4-19: I am not sure I understand the argument presented considering the impact of MLD on pCO2 in the Subtropics. How is the entrainment of DIC-rich water by a deeper MLD in the year 2006 balanced by a lower SST if the SST is virtually the same for all 3 years? In Olsen et al. [2008] (their Figure 9), I see a large impact of changes in MLD on pCO2 for the considered depth range right at the bottom of the euphotic zone. Also Jamet et al. [2007] find a positive coefficient for MLD for Winter in their multiple linear regression (their Table 2, last row)

ˆˆˆOUR REPLY: Indeed we note that the deeper 2006 MLDs in Figure 12 do not correspond to lower SSTs and our assumption would be that perhaps for this region changes in MLD between 100m and 300 m produce such a weak response in SST that monthly averages do not show it. As mentioned in this manuscript, both Olsen et al. (2008) and this study find non-linear relationship between sea surface pCO2 and MLD. Also, similarly to Jamet et al. (2007), we find the strongest correlation between the two in late winter-early spring. We will add this quantitative information to the reviewed version of the manuscript.

—Anonymous Referee #3 (Specific comment 7) Page 3378, equation 1: Lefevre et al. [2005] and Friedrich & Oschlies [2008] successfully used Latitude, Longitude and Time as additional input parameters for their SOM-based mapping. The latter ones reported that neglecting position leads to larger (about 5-10 uatm) RMS-errors (their Figure 8 + paragraph 32). Since Latitude, Longitude and Time are available 'for free' and have been shown to improve mapping accuracy why doesn't this study utilize them? Is it the different labeling scheme applied in this study that impedes the use of Latitude, Longitude and Time?

ˆˆˆOUR REPLY: Our different (than those used in these earlier studies) training scheme allows for much better determination of the statistical structure of the basin-wide data but also the patterns found are more strongly implemented in the output. Using latitude or longitude causes a concentration of similar values along east-west or north-south (respectively) lines. Using both parameters causes clustering of similar pCO2 values in patches with surprisingly equal distances between one another. Using time increases the influence of seasonality on the pCO2 maps. Thus, use of latitude (longitude) and time does not improve the RMS much in our scheme. These issues are the topic of another manuscript (in preparation). Out of the three, the use of longitude seems to cause the least "artificial" distribution and it also slightly improves the RMS. That is probably because it helps the SOM to differentiate between regions at the same latitude with similar SST, MLD and CHL but which have a different pCO2 values because of the different origin of the water. However, we decided against using longitude because we are uncertain whether the similar values concentrated along certain north-south lines are artefacts related to the use of longitude.

References:

Friedrich, T., and Oschlies, A.: Neural network-based estimates of North Atlantic surface pCO2 from satellite data: A methodological study, J. Geophys. Res., 114, C03020, doi:10.1029/2007JC004646, 2009.

Ono, T., Saino, T., Kurita, N. and Sasaki, K.: Basin-scale extrapolation of shipboard pCO2 data by using satellite SST and Chla, Int. J. Remote Sensing, 25(19), 3803-3815, 2004.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R.A., Chipman, D.W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A.J., Bakker,

D.C., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T.S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C.S., Delille, B., Bates, N.R., and de Baar, H.J.W.: Climatological mean and decadal change in surface ocean pCO2, and net sea-air CO2 flux over the global oceans, Deep-Sea Res. II, 56, 554-577, 2009.

Watson, A. J., Schuster, U., Bakker, D .C. E., Bates, N., Corbière, A., Friedrich, T., González-Dávila, M., Hauck, J., Heinze, C., Johannessen, T., Körtzinger, A., Metzl, N., Olafsson, J., Olsen, A., Oschlies, A., Padin, X.A., Pfeil, B., Santana-Casiano, M., Steinhoff, T., Telszewski, M., Ríos, A., Wallace, D.W.R., Wanninkhof, R.: Accurately tracking the variation in the North Atlantic sink for atmospheric CO2, submitted to Science, 2009.

Interactive comment on Biogeosciences Discuss., 6, 3373, 2009.