Biogeosciences Discuss., 6, C466–C469, 2009
www.biogeosciences-discuss.net/6/C466/2009/

Biogeosciences Discussions

BGD

6, C466–C469, 2009

Interactive Comment

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

D. Hydes (Referee)

david.hydes@noc.soton.ac.uk

Received and published: 19 May 2009

The topic is within the scope of Biogeosciences, and both the approach and results will be of interest to a wide range of readers. The paper begins the process synthesising the very successful data gathering campaigns in the North Atlantic particularly in 2005 funded by CarboOcean and other projects. The work is timely and will have an immediate impact in biogeosciences community.

This work uses the self-organising neural network approach to produce maps of the basin-wide monthly pCO2 distribution across the North Atlantic. Data on pCO2 from 2004-2006 collected by ships of opportunity is used to train the neural network and thence generate maps based on data for SST, CHL and MLD. That data is available over the whole grid from data derived satellite observations and numerical models, which assimilate in situ data. The generated maps are then reviewed with respect to the picture they give of change in different biogeochemical provinces and compared to data from the Takahashi et al., (2008) data based model of global variations in ocean surface water pCO2.

The English and clarity could be improved if consistent use were made of the definite and indefinite articles ("the" and "a")

Specific comments

1. The work is a natural progression from that of Lefevre 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.

2. Section 2 could (and should) be reduced considerably in length if more reference was made to the Lefevre et al papers. The explanation of the methodology was more clearly expressed in the Lefevre 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.

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.

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

BGD

6, C466–C469, 2009

Interactive Comment

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

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?

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?

6.1 Page 3387, line 18. 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 ?

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.

7. Page 3390, 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.

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?

9. 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

BGD

6, C466–C469, 2009

reference case and then challenges us to explain the differences we see.

9.1 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?

9.2 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?

10. Section 3.3. “Interannual variability”. This section would be better supported by figures which show the actual difference between years. See comment 6.

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