

## ***Interactive comment on “Estimating temporal and spatial variation of ocean surface $p\text{CO}_2$ in the North Pacific using a Self Organizing Map neural network technique” by S. Nakaoka et al.***

**R. Wanninkhof (Referee)**

rik.wanninkhof@noaa.gov

Received and published: 15 April 2013

Review of estimating temporal and spatial variation of ocean surface  $p\text{CO}_2$  in the North Pacific using a Self Organizing Map neural network technique By S. Nakaoka, M. Tel-  
szewski, Y. Nojiri, S. Yasunaka, C. Miyazaki, H. Mukai, and N. Usui

Reviewer: R. Wanninkhof, NOAA/AOML, Miami USA

Nakaoka and co-authors present a well-written and thorough overview of surface water  $p\text{CO}_2$  patterns in the North Pacific using a self-organizing map (SOM) neural network technique. This is the first time this approach is used for the North Pacific and it takes advantage of remotely sensed (SST & Chl) and modeled (SSS & MLD) parameters  
C976

for the training step. They then use the extensive and high-quality dataset  $p\text{CO}_2$  from NIES to perform the labeling phase. Finally they use the SST, SSS, MLD and Chl to determine the temporal and spatial changes of surface water  $p\text{CO}_2$  over 6 years. The results are compared with the Takahashi climatology and SOCAT database. The SOM shows very good correspondence but with higher fidelity and the approach has the ability to determine finer scale structures than seen in the climatology.

The paper is clear and well-structured. It does a good job describing the general principles of the technique and it does a very thorough comparison of the results with other approaches. It ends with some comments on future efforts and possible improvements. It would be of interest if the authors provided some suggestions on the amount of  $p\text{CO}_2$  data that is necessary for labeling and how the uncertainty is related to the amount of labeling data. The quantification of the sources of uncertainty would be of interest although this might be a paper in itself.

My other comments are minor.

Page 4577 line 24: replace “current systems” with “current regimes”. Page 4578 line 14: place space between “high” and “latitude”. Page 4580 line 7 and elsewhere: in Takahashi et al. 2006 they point out regions in the N Pac. where the ocean  $\text{CO}_2$  increases are slower than atmospheric increases. These regions appear to be the same areas where the inclusion of the rate term in eq. 1 does not improve the fit between the SOM and observed data. This should be mentioned. Page 4582 line 1: There needs to be some discussion on how well the assimilated MLD and SSS used for the training data match observations. I expect that SSS is measured along with all the  $p\text{CO}_2$  observations such that a large comparison set is available. Assuring that there are no biases in the training data is important. Page 4584 line 1: the difference between  $f\text{CO}_2$  and  $p\text{CO}_2$  needs to be described as this will lead to confusion for some readers. Perhaps the values should be converted to a single unit (either  $p\text{CO}_2$  or  $f\text{CO}_2$ ) (although the difference is very small). Page 4585 line 9: Emphasize here that the MLD is expressed in log form in training and that log 200 and log 400 m are

close in magnitude. Page 4585 line 26: give the magnitude of “negligibly small” Page 4586 line 12: it is of note that the RMS of 18 uatm is fairly large for some applications (e.g. determining regional sea-air CO<sub>2</sub> fluxes). Page 4587 line 19: Is 37 N, 140 W a region or location- If region, specify approximate size Page 4588, line 7: see comment above: Page 4580 line 7. Page 4589, line 5 and beyond: spell out acronyms in text the first time they are used e.g. KE, WST, VOS Page 4590, line 19: “This confirms the notion that the SOM technique overcomes problems associated with temporal and spatial data scarcity.” Is this a universal conclusion or does it apply to certain regions only? It would appear that a mechanistic understanding of controlling processes would be important to back this up.

References appear comprehensive. Table 1. Perhaps mention in caption that these data are used for labeling. Figures are appropriate. Their size and resolution should be checked in final version. Figure 4. Perhaps list the total number of data points in this figure. Figure 6 and 8 (and also in text): Why is 2-sigma used as an indication of spatial variability? I've generally seen 1 or 3 sigma.

I saw no technical/typographical errors. The paper appears to have been produced and checked by the authors with great care.

---

Interactive comment on Biogeosciences Discuss., 10, 4575, 2013.