Review of

"Estimating temporal and spatial variation of ocean surface pCO2 in the North Pacific using a Self Organizing Map neural network technique"

by Nakaoka et al. [2013]

The authors are using Self-Organizing Maps to produce basinwide surface pCO2 maps for the North Pacific from VOS-line pCO2 measurements and remote sensing data of SST, SSS, Chl and MLD. The reconstructed pCO2 values are compared to pCO2 data of time-series stations and independent observations.

Overall evaluation:
The present manuscript is a very useful study which will be relevant for a broad scope of readers. It is well written and all details are explained thoroughly. I have only one major point of criticism that deals with the estimate of the overall RMS-error of the method. I believe that the presented estimate of 17.6 µatm is misleading and most likely too low. I am confident that the manuscript can be published after this issue is addressed.

Major point:
As far as I understand the authors the RMS-error of 17.6 µatm is based on a comparison between the in-situ pCO2 measurements used for the labeling process of the SOM and the pCO2 estimates of the SOM. If this is true, the validation was not done against independent data as the SOM contains (a lower dimensional representation of) this training data. The study by *Friedrich and Oschlies* [2009a, *JGR*] cited by the authors clearly showed that the true RMS-error must be expected to be much higher if a validation against training data is used. Why are the data shown in Figure 8 only used to "f*acilitate a discussion about the temporal variations of pCO2*"? These measurements could at least provide an idea of how well the method works for extrapolation to areas not covered by the training data measurements.

Furthermore an estimate of the overall RMS-error needs to include all possible sources of pCO2-mismapping of the method. For example, the remote sensing data are subject to uncertainties which can be quite large and which will affect the SOM-estimate (and the SOM formation process as well). For a first order evaluation of this effect the authors could add noise representative of the remote sensing uncertainties to the data used for the mapping process and compare the resulting pCO2 estimates to the untainted reconstructions. It could be done in a similar way for the training process.

I don't think it would compromise the value of the method if the study came up with a higher overall RMS-error but it would certainly add to its credibility if this RMS-error estimate would be based on a more realistic validation. As much as an RMS-error of around 20 µatm may sound small compared to the overall mean pCO2, *Watson et al.* [1991, *Nature*] stated that a bias of 1 µatm in the global ΔpCO2 would already result in about 0.2 Pg/yr uncertainty in the estimated ocean carbon uptake.

Minor points:
Is the presented method more skillful than simply using the Takahashi climatology? The first guess of pCO2(x,y) would be to refer to Takahashi(x,y). This first guess can be refined -as the authors do it- by

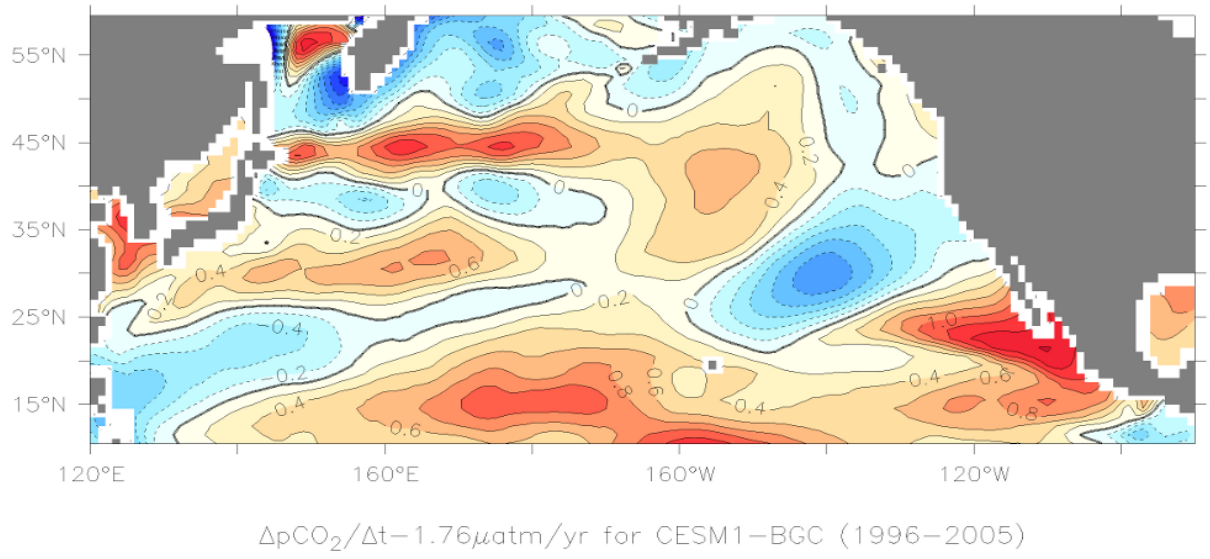
adding a ΔpCO2/Δt * (t-t(ref.)). Does the SOM method result in smaller RMS-errors compared to these two ”cheaper” methods?

page 4578 / line 18:
The authors might want to include a reference for ESTOC (e.g. *González-Dávila et al.* [2010, *Biogeosciences*]) and use a more recent reference for BATS (e.g. *Bates* [2012, *Biogeosciences*]).

Discussion on the use of SSS
page 4579 / line 18
page 4587 / line 2 + 23
page 4592 / line 10
SSS has already been successfully used by *Friedrich et al.* [2009b, *JGR*] to map basinwide pCO2 in the North Atlantic. They have also provided an explanation why it is such a skillful predictor for pCO2: *“Surface water pCO2 is, besides its dependence on sea level pressure, a function of DIC, total alkalinity, SST and SSS. Because for any individual ocean basin total alkalinity can, to good accuracy, be estimated from SSS using a nonlinear empirical fit [e.g., Eden and Oschlies , 2006], ARGO SST and SSS data already provide substantial (though local) information about parameters that determine pCO2”*

page 4585 / section 2.6
I agree with the authors on their treatment of Chl. However, it is a little awkward to say that the difference is “negligibly small” when there is a lack of coverage. At least we need to know what percentage of data coverage this statement is based on.

page 4587 / line 27...
The authors might want to consider refining their method by using ΔpCO2(x,y)/Δt which could be obtained from the CMIP5 data. I included a figure of the deviation from the value of 1.76 µatm/yr used by the authors derived from the CESM1-BGC model.



page 4593 / line 13
I do not understand what the sentence starting with *“Even if ...”* is supposed to tell us.

Figure 2b
Just curios: How are unlabeled neurons treated?