

Review of bg-2017-64

Global high-resolution monthly pCO<sub>2</sub> climatology for the coastal ocean derived from neural network interpolation by Laruelle et al.

Reviewer: Rik Wanninkhof, NOAA/AOML

This is largely a descriptive paper of procedures to create monthly estimates of coastal pCO<sub>2</sub> levels. As mentioned in the abstract, Laruelle et al. use a modified version of a two-step artificial neural network method (SOM-FFN) to interpolate the pCO<sub>2</sub> data along the continental margins with a spatial resolution of 0.25 degrees and with monthly resolution from 1998 until 2014.

The effort is clearly an impressive one and an important contribution to coastal ocean science. However there are some shortcomings. Many readers will not fully understand the approach and assumptions in SOM-FNN, and this needs more discussion. The manuscript lacks in context and interpretation. Some of the procedural shortcomings that were in the initial global open ocean effort as described in Landschützer et al., (2013; 2015) prevail.

While there are comparisons and validations of the SOM-FNN approach it mostly in terms of a RMSE. It is unclear what impact the RMSE would have on the phenomena investigated. Other means of comparison of how well the approach works should be performed. Rödenbeck et al (2015) present some nice diagnostics that could be applied. At very least examples of the distribution of errors in pCO<sub>2</sub> should be shown in histograms.

As the authors indicate, their definition of the coastal realm (200 nm or 1000 m depth) covers a much greater region than commonly viewed as coastal. The outer edge of the domain for much of the ocean can be considered "blue water". Therefore it is surprising that the differences between the coastal SOM-FFNN and open ocean SOM-FNN in Landschützer et al. are large. A more comprehensive diagnostic comparison should be made as it could suggest some fundamental issues with the approach.

The validation approach is weak. There is significant (complete?) overlap between the data in SOCAT and that of Takahashi. The biases in datasets are likely due to different data reduction approaches. More comparisons should be made with actual data not used in the training, and more data should be excluded from the training for validation purposes.

It is unclear how the change in surface water over time is dealt with. Are the pCO<sub>2</sub> data normalized like in the Takahashi monthly climatology? SST and SSS from the WOA are used but are these monthly climatologies that do not reflect change over time. This exercise provides monthly maps from 1998-2014 and it is clear how this is done. Also, the product is referred to as a climatology but it sounds like it is a monthly time series. That is, climatology mostly refers to a (multi) decadal average.

The grouping of provinces such that a coastal region can include an inshore and open ocean province is odd. Perhaps limit the coastal area to just one province. It is difficult to assess the data density for the different provinces using as validation or training.

Specific comments often relating to the general observations are below. The referenced text is in italics:

Line 125: "*motivated a number of modifications of the global ocean SOM-FFN method, including a 16 fold increase in spatial resolution from 1 degree to 0.25 degree, the introduction of a second neuron layer in the FFN calculations, the addition of new environmental variables as biogeochemical predictors, and a shortening of the simulation period to the period 1998 through 2014, rate of sea ice SST, SSS, bathymetry, sea-ice concentration and chlorophyll a second artificial neuron layer*". Some more detail on how these modification impact the results would be worthwhile .

Line 175: "*SOM-FFN from generating negative values.*" This suggests that there are issues with the original setup. Adding a second neuron layer to prevent negative values certainly is unorthodox.

Line 193: "*All the datasets used in our calculations were converted from their original spatial resolutions to a regular 0.25 degree resolution grid.*" Specify what the original resolution was for each dataset.

Line 196: "*SST and SSS maps were taken from the World Ocean Atlas (Antonov et al., 2010 for SST and Locarnini et al., 2010 for SSS).*" Are these monthly climatologies or monthly time series? If the former it is unclear how the time element from 1998-2014 is incorporated.

Line 203 and beyond: "*validation are extracted from the LDEOv2014 database The coastal SOM-FFN results are validated through a comparison with the LDEOv2014 data (Takahashi et al., 2016).*" This is not independent data and not a proper validation in statistical sense.

Line 280: "*Considering these complexities, the achieved RMSE is quite good.*" Two issues here. How are the complexities determined? That is, we know the coastal region is complex but it is unclear if the complexity is incorporated into the analysis using T, S, chl-a and sea ice. And, based on what criteria is the RMSE quite good.

Line 306: "*which compares with the most robust pCO<sub>2</sub> regional coastal estimates from the literature (Chen et al., 2016)*". Chen et al. 2016 use a crude remote sensing approach. These are by no means "most robust".

Line 349: "*highlight the current knowledge gap regarding the mean state and variability of the transition zone.*" It is unclear if this highlights a knowledge gap or highlights issues with the SOM\_FNN approach. This warrants some discussion  
Line 358: "*Our results indicate that the very nearshore processes controlling the CO<sub>2</sub> dynamics likely*" Again the SOM-FNN is a mathematical construct. So I guess what the authors are stating is that the SOM-FNN cannot address adequately nearshore dynamics.

Line 429 " $\square\square\square\square^2$ ". The "n" generally refers to salinity normalization. Perhaps use  $pCO_2(SST_{mean})$ .

Line 470: "*cells at a 0.25° spatial resolution for each of the 204 months of the simulation period (from January 1998 to December 2014). Climatologically averaged pCO<sub>2</sub> maps for each month are*". The use of the term climatology is ambiguous here.

Line 471: The province names are peculiar "Deep Polar, Polar Very deep Polar"  
Table 1 suggests that Ice is a predictor in the tropics? Also P3 and P4 appear to have the same "distribution".

Figure 1 shows a peculiar extension off of New Zealand. Is this the Chatham Rise and is this considered coastal?

Figure 2: Perhaps comment on the absence of high pCO<sub>2</sub> in the SOM-FNN for the summer monsoon upwelling region in the Arabian Sea. Data of the Takahashi climatology clearly show this. Figure 2 does not show the high pCO<sub>2</sub> Arabian Sea seasonal (JAS) upwelling off the coast of the Arabian Peninsula.