Summary:

Xu and colleagues investigate the regional sea surface pCO_2 and air-sea flux in the Prydz Bay Antarctica using observations from the CHINARE cruise in February 2015. The authors divide the study regions into 3 sub-regions, based on the physical and biogeochemical controls of these sub-regions. Using a self-organizing map approach, the authors extrapolate the cruise data to the entire study region in order to estimate the carbon exchange of the Prydz Bay.

The Southern Ocean is still among the least observed and certainly least well understood ocean basins, hence I found this process study – investigating carbon variability and air-sea exchange in the Prydz Bay – to be very interesting and certainly relevant for the GB readership. More details on the strengths and weaknesses are listed below.

Strengths:

I found the manuscript and particularly the discussion of the processes comprehensive and logically built-up. The authors further make use of an appropriate and previously applied method based on machine learning (i.e. the SOM method) to extrapolate the cruise information to the full region of interest. They use independent validation data to test how well their approach reproduces observations from the SOCAT dataset and use this information to estimate the uncertainty of their integrated air-sea flux.

Weaknesses:

Up-front, I would like to note that there are several language issues – too many to be all named here (just one example: line 232: "In Pacific Ocean" should be "In the Pacific Ocean") – hence I do recommend English language editing.

During my review, I have encountered a few things that need clarification or some more information from the authors. They are listed from the most to least concerning. Additional comments (not of major concern) with line-numbers can be found at the end of this document:

1) Method section: At the moment, it is impossible for a reader who has not worked with the SOM approach to understand the methods section. Sentences like: "The SOM is trained using unsupervised learning to project the input space of training samples to a feature space (Kohonen, 1984), which is usually represented by grid points in tow-dimension space." Imagine a BG reader who is interested in the carbon exchange of the Prydz Bay but has never worked with a SOM. How is that person supposed to understand wording like"unsupervised, feature space, weight vector, training data, labeling data, etc." without reading several other papers first? As a SOM user I had no issues to follow this section but in my view, it has to be simplified for the more general BG audience. Furthermore, the authors miss to mention what distance function the SOM uses to detect the "winner neuron" (Euclidean distance maybe?). Furthermore, I don't think the phrase "resolve nonlinear relationships" (see abstract) is appropriate, since a SOM is a clustering algorithm that clusters based on similarities, but does not explicitly "resolve" a relationship.

<u>Response:</u> We have revised the introduction part about the SOM method to make it easy to understand what is SOM. And in the 2.2 section we have revised the sentences and adjusted the structure to make it easy for the reader to know how SOM works. In our SOM analysis we used Euclidean distance (the shortest distance) to select winner neurons and we have added this to the manuscript. We agree with the reviewer's suggestion and have changed the phrase 'resolve nonlinear relationships' to be 'to overcome a complex relationship among the biogeochemical and physical conditions in the Prydz Bay region'.

2) Training data: This links a bit to my point above but goes a bit more in-depth: I am not sure how data have been handled. On line 177 the authors state that the data have been "the four proxy parameters were logarithmically normalized" but table 1 suggests otherwise. In table 1 all values are absolute values. Besides that, I am not convinced that it makes sense to logarithmically normalize all 4 proxies. It makes sense for the skewed MLD and CHL-a but not really for salinity and temperature. Besides, I wonder how the normalization effects the distance function (which is not mentioned). Euclidean distances depend on the data-value range of each proxy. Also, what I am missing is a discussion why exactly the 4 proxies have been chosen? Why not sea surface height, wind speed, sea level pressure? What makes the 4 proxies so unique? I know they have been used by other authors, but the reader of THIS study needs this information.

Response: In table 1 all values are absolute values of the four proxies to show the value range. For the skewness and the N coverage percentage, the normalized data are shown in parenthesis. According to the change of skewness and N coverage percentage we found out only MLD and Chla data needed to be normalized for both the training and labeling dataset. Since we used Euclidean distance function to select the winner neuron and it depends on the data-value range of each proxy. The normalization for MLD and Chla dataset is to avoid weighting issue raised from the different magnitude among the variables.

In section 2.1 we have discussed the four proxies which will affect the distribution of pCO_2 in the surface sea water. The dissolution of CO_2 into water is mainly affected by temperature and pressure of water. The variation of salinity has little effect on the dissolution of CO_2. However the sea ice changed quickly in the study region and we chose salinity to be a proxy to simulate pCO_2. Moreover, in the region where local biology activities are active, pCO_2 will be affect strongly by photosynthesis. The mixed layer depth will prevent the upward mixing of nutrients and limits the biological production therefore we chose MLD as another proxy to simulate pCO_2. Sea surface height and sea level pressure are not major factors to the distribution of oceanic pCO_2. Wind speed is vital for the sea-air gas exchange and it is included in the air-sea flux equation.

3) Uncertainty: line389 states: "increased from week-1 (2.13 TgC) to week-2 (2.24TgC) due to increased wind speed." I was a bit disappointed here. First there is the effort to calculate uncertainties, then it is neglected in the text. Given the final uncertainty estimate, it is very unlikely that this regional difference of 0.1 TgC is significant. In general I suggest to add uncertainties wherever possible to avoid such misinterpretations.

Response: We have added uncertainties to the carbon uptake in section 3.4 and we have changed 'increased' to be 'changed mildly'.

4) Validation, comparison: I appreciate that the authors do a comparison with SOCAT data and include this in the overall flux uncertainty. I think that there need to be a bit more info in the text what cruise from SOCAT you are comparing to (this information is available on socat.info), or what the average spatial and temporal distance (which should be possible since a nearest grid method was used) between the cruises is. That certainly contributes to the mismatch as well. Otherwise, I was quite impressed by the relatively small (~22µatm)

difference. It might not sound small at first but your are comparing small special scale and high frequency temporal scale data based on the extrapolation of a single cruise. Therefore, 22μatm is impressive in my view. Furthermore, the RMSE tells the reader about the spread, but it would be valuable to add the mean (or absolute mean) difference between the SOM derived CO_2 and the SOCAT cruise. This would give you an indication of the bias.

Response: We have added the information of the cruise we selected from SOCAT in section 2.3. We have calculated the absolute mean difference between the SOM derived CO_2 and the SOCAT cruise. According to the validation, the SOM derived pCO_2 is generally lower than the SOCAT. Since the dataset from SOCAT does not cover the low- pCO_2 area towards the south, the precision might be of great uncertainty.

Methods section: On many occasions the authors re-grid data to the desired 0.1*0.1 resolution, but a bit more information on all data that were regridded and the algorithm would be appreciated. Ideally in form of a table. Additionally, I am missing the motivation why 0.1*0.1 was chosen. Why not 0.5*0.5 or even 0.05*0.05. Just to be clear, I don't suggest changing the resolution, but the text needs some motivation/technical explanation on why the current resolution was chosen that justifies all the data handling (i.e. regridding of proxy data)

Response: The 0.1*0.1 resolution of our study was desired according to the study area. It is a small area from 63E to 83E and 64S to 70S and the 0.1 resolution is the optimal. In the paper of Telszewshi et al. (2009), it was a basin-wide area from 9.5E to 75.5E and 10.5N to 75.5N, so their resolution was a 1 latitude by 1 longitude resolution. For a global area, Takahashi et al.(2012) chose 4*5 resolution. For our study area, it would be too rough if the resolution of 0.5, and the matrices would be too big if the resolution of 0.05.

The other data including remote sensing data and modeled data of different resolution were regridded to be the same resolution of 0.1 * 0.1 by Kriging method. We have added some explanation in the text. We think it is clear in the text.

Recommendation:

I have found this study to be interesting and to be of value to the BG readership. While I have raised some (partly major) concerns above I think that they can be resolved by the authors. I therefore recommend major revisions of the manuscript.

Specific and minor comments to the text:

1. Abstract line 14: Please also add the temporal resolution to the spatial resolution

   Response: We have added 'weekly' to the spatial resolution in abstract.

2. Abstract lines 27-29: This last sentence is out of context and is not something you can conclude from this study, hence it needs to be removed.

   Response: We have removed the last sentence.

3. Lines 32-33 reads "The role of the ocean south of 60S in the transport of CO_2 to or from the atmosphere is still uncertain despite of its importance of reducing anthropogenic CO_2 in the atmosphere" – that is a conflicting statement as it currently reads. If we know the importance of reducing atmospheric CO_2 how can its role be uncertain?

Response: It was a mistake. Here we mean ‘the amount of carbon uptake in the ocean south of 60’. We have revised it.

4. Lines 76-77: “Therefore, the direction of the sea-air CO_2 transfer is mainly regulated by the oceanic pCO_2” – this statement needs a reference
   Response: We have added the references needed.

5. Line 84:”The SOM analysis, based on neural network (NN), a type of artificial neural network” – the second part (based on neural network) can be removed
   Response: It has been removed.

6. Line 117: “Salinity records the physical processes” – When I read this sentence I also think of larger scale circulation and mixing in the context of physical processes, whereas this statement links to the follow-up discussion about brine rejection. Maybe a different term would be more appropriate.
   Response: It has been revised.

7. Line 130: How was the interpolation done?
   Response: We gridded the chlorophyll-a data from Modis according the cruise track.

8. Lines 133-136: “The mixed layer links the atmosphere to the deep ocean and plays a critical role in climate variability. Very few studies have emphasized the importance of accounting for the vertical mixing through the mixed layer depth” – Firstly, I disagree. Several studies have emphasized the importance of vertical mixing of carbon (but also nutrients, etc) through the mixed layer. Secondly, I caution the authors to mention the role in climate variability here. Their study does not resolve the necessary timescales to discuss either seasonal or interannual or decadal (whatever variability the authors refer to) variability.
   Response: We have made the correction and have removed the mention about the role in climate variability since in our study it didn’t relate to that.

9. Lines 154-155 ’SOM based multiple non-linear regression’ – This must have been a mistake or typo here, since the SOM (unlike e.g. a back propagation network) does not perform a regression (also not a non-linear one). Instead the SOM clusters data based on similar environmental conditions.
   Response: Yes, we agree the reviewer’s suggestion and have removed ‘multiple non-linear regression’.

10. Lines 194-195: “until the neural network sufficiently represents the nonlinear interdependence of proxy parameters used in training.” – how is this judged? When do you know that its sufficient? I suppose this is judged by the number of SOM iterations, but how is set?
    Response: Because SOM analysis is a powerful technique to estimate pCO_2 from among the non-linear relationships of the parameters (Telszewski et al., 2009;), actually, we presumed the nonlinear interdependence of proxy parameters are sufficiently represented after the

training procedure. Also, we used the som_make() function in the SOM toolbox for training data. Thus, we updated the sentence accordingly.

11. Line 215: "I could not figure out where the factor 30.8*10-4 comes from? Please explain in the text
Response: The factor is induced according to the simplification of the equation. We have added the explanation in the text.

12. Line 264: "robustly divided" – I caution the authors here: How can you be sure the division is "robust"? Have you done any test that would proof robustness?
Response: Three regions are divided according to the distribution of oceanic pCO_2. From the distribution of pCO_2 as shown in Fig.2-a there are three ranges. One is from 291.98 μatm to 379.31 μatm, the second is from 200 to 310μatm and the third is below 200μatm. We roughly divided the study region according to the three ranges of pCO_2 and the range of the depth of water in the Prydz Bay region. It was a mistake to use the word 'robustly'.

13. Lines 281-282: "region atmospheric pCO_2 was stable from 374.6μatm to 387.8μatm" That is a difference of 13μatm – I would not call this stable at all! I suppose this difference is largely the result of sea level pressure variability and relative humidity in the surface layer, hence it would be interesting to see the molar fractions (in ppm) for comparison if available.
Response: We don't have sea level pressure data and relative humidity in the surface layer. We have revised this sentence and removed 'stable'.

14. Line 285: "biological consume" – should be "biological uptake"
Response: It has been revised.

15. Line 318-319:"for a same period" – This would be important information. Furthermore, have you considered ARGO biogeochemistry floats from the SOCCOM array? They are deployed since 2013 and may add some additional independent estimate. This might however be beyond this manuscript.
Response: Thanks for letting us know the SOCCOM. We have searched from SOCCOM but we can't find dataset useful for our study. However SOCCOM is a helpful website and we will turn to it when we other analyses in the Southern Ocean next time.

16. Figure 4b: It would be easier visible if x-axis and y
Response: We have changed the x-axis and y to be the same range.