Dear Anonymous Referee #4,

We very much appreciate your constructive comments, which will allow us to improve the overall quality of our manuscript.

General comments:

Section 1 (Introduction):

Anonymous Referee #4 Comment (AC): *Please include a few sentences outlining any previous data-based work that has focused on advancing our understanding of ocean carbon variability in the NWES. Also include a sentence on how this study builds on previous work in the NWES, or if this is the first, clearly state it.*

Please include a sentence outlining the rest of the paper.

Author Reply (AR): We included all your suggestions, see Section 1.

Section 3.1. (FerryBox dataset):

AC: *Page 5646, line 14: Please provide more details on your stated uncertainty in calculated pCO_2 via DIC/TA.*

AR: We added the following explanation in the revised manuscript and refer to Marrec et al. (2014) for more details on the analytical methods used to measure DIC/TA: "The methods used for the analytical determinations of DIC and TA are described in details in Marrec et al. (2014) and gave accuracies of ± 2 and 3 µmol kg^{-1}, respectively. Thus, the computed values of pCO_2 from DIC and TA have uncertainties at the lower end of ± 6 µatm (Zeebe and Wolf-Galdrow, 2001)."

Section 3.3. (Development of pCO_2 algorithms):

AC: *In section 2 (study region), you state the frontal zone separating the sWEC/nWEC oscillates and could be precisely locate from data. I suggest you use this ability to create a dynamically shifting front (via SST) from which measurements are partitioned prior to the MLR. This might also reduce your current pCO_2 discontinuity at the region boundaries.*

AR: We based our separation of the different provinces on a 10 year dataset of SST covering the entire shelf (Fig. 2), which provides robust estimates of the mean location of thermal fronts. We feel that the use of fixed boundaries allow a clear discussion of our datasets and

direct comparison between the representative provinces. The sharp boundaries between permanently well-mixed and seasonally stratified systems can appear as surprising, especially between August and October. However, these sharp boundaries are a fact that we observed every year between sWEC and nWEC waters and that can occur elsewhere on the shelf. To support this we made 2 new figures (Fig. S3 in the supplement material) showing a comparison between in-situ pCO_2 data acquired during 2 crossing performed in August and September 2014 between Roscoff and Cork (Ireland) (from a newly exploited Voluntary Observing Ship, the ferry Pont-Aven) and mean pCO_2 data along the ferry tracks calculated from our MLR from 2003 to 2013. We did not have access yet to the requested satellite and modeled products in 2014, which explained the choice of using monthly mean pCO_2 estimates instead of newly computed pCO_2 estimates from remotely sensed and modeled data. These two figures and the new in-situ data between Roscoff and Cork clearly show the presence of these sharp boundaries. Moreover, region boundaries represent the shifting area of thermal fronts. As we could not estimate pCO_2 using our algorithms in frontal zones, such an approach appeared suitable. We hope that we now provide enough evidence for the choice of our different provinces. We added a short statement at the end of Section 2 for the choice of fix boundaries and added the discussion above with reference to the new figures of the supplement material at the end of Section 4.3.2. in the revised manuscript.

AC: *Page 5652, line 4: Please state whether this increase in ocean surface pCO_2 (1.7 uatm/yr) is a global or regional estimate. Also please include a sentence discussing if this estimate is representative of your study region.*

AR: Following reviewer #3 suggestion, we computed an atmospheric pCO_2 increase of 1.8 µatm/year based on our Mace Head (Ireland) xCO_2 dataset. Based on the paper by Signorini et al. (2013), we assumed that ocean surface pCO_2 increase is trending at the same pace as the atmosphere pCO_2. Thus, we consider an increase of seawater pCO_2 of 1.8 µatm/year representative of our study region. This is now explained in the revised manuscript at the end of the Section 3.4.

Section 4.1. (Performance of MLR):

AC: *I'm concerned about the significance of time (TI) in your regression model. This is a non-physical parameter which captures up to 50% of the observed variability, thus indicating*

key physical information is missing from your MLR (be it salinity, nutrients, ect). Could you please include in table 1 a regression model where TI was not included as a predictor parameter, and discuss what it means if temperature and biological indicators can only capture ~50% of the pCO_2 variability.

AR: We included the variable TI at the end of Table 2 in order to observe the relative importance of each physical and biological variable in the MLR without TI. We acknowledge that it allowed the reader to better understand the respective role of each of these variables. We now explain why the variable TI captures up to 50% of the observed variability in the second paragraph of Section 4.1.

AC: *Please include a sentence discussing why K is used as a predictor for ocean surface pCO_2. Also remember that K is calculated via temperature, so K and T are not independent variables. This may in part explain the observed difference between your Temp co-efficient and the Takahashi SST/pCO_2 relationship.*

AR: We now use wind speeds instead of K in the MLR and only in sWEC following suggestion of other reviewers.

AC: *Page 5656 line 17: I disagree that the distribution of residuals in the nWEC looks more homogenous. To strengthen your claim you could colour the points in Fig 6D to indicate sample year (or Latitude).*

AR: We followed your advice and colored the points in Fig. 6c and 6d to indicate sample years. Our statement is now well supported by the figures since the residuals during year 2011 in the nWEC were clearly more homogenous than in the sWEC. We modified the text specifically for year 2011 in the revised manuscript in the 3rd paragraph of Section 4.1.

Section 4.2.:

AC: *While I believe empirical approaches are extremely valuable in predicting ocean carbon variability in data-limited regions, they do have limitations. From Figure 8, it seems your MLR predictions compare well to the SOCATv2 measurements in the sCS. In the IS and nCS however, you have no (or very few) pCO_2 measurements from which to justify your pCO_2 predictions are accurate. You should state that you have no way of quantifying uncertainties in your pCO_2 predictions beyond the two WEC regions. One possibility to strengthen your approach would be to look at correlation length scales in you predictor variables. If your*

predictor variables are highly correlated between the WEC and the IS and nCS, it suggests pCO_2 concentrations could be predicted from your WEC model in these regions.

AR: This issue was also raised by reviewer #1 and he advised us to use pCO_2 data from the LDEO database. Thanks to his suggestion we now have access to new pCO_2 data, particularly in IS and nCS, which consolidate our comparison between observed and modeled pCO_2 in other regions than the English Channel. These new in-situ pCO_2 data are represented by yellow dots on the updated Fig. 8 of the revised manuscript. These new results greatly enhance our extrapolation in these poorly studied regions and therefore support the main purpose of this study, which is to have access for the first time to pCO_2 estimates in this area where only few pCO_2 data are currently available. Table 1 in the revised manuscript has also been updated to include these new sources of in-situ pCO_2 data. We also included a new figure in supplement material (Fig. S2) for statistical data-model comparison as recommended by reviewer #2.

AC: *Was the anthropogenic increase factor (eq 4) included in these pCO_2 predictions? If so, please state these estimates are representative of the year 2012, and discuss why you observe any trends (as is evident in the nWEC).*

AR: For a comparison of the observed and modeled pCO_2 over three years in the WEC, we used July 2012 as a reference month. This allowed a direct estimation of the performance of our model without any impact of the anthropogenic increase factor. However, for the ten years estimates in other regions of the shelf, the regional pCO_2 anthropogenic increase of 1.8 µatm $year^{-1}$ was included in the data using July 2012 as the reference year. This explains why on Fig. 8 the estimated regional anthropogenic signal might be visible over a decade in several provinces. We clarified the revised manuscript at the end of Section 3.3.

Specific comments :

AC: *Page 5643, line 1: Please rephrase for easier reading. I suggest, 'Continental shelf seas form a complex interplay between the land, ocean and atmosphere, hosting a multitude. . .'*

AR: Done.

AC: *Page 5644, line 6: Neural network is the name given to the family of statistical learning models of which the self-organising map is one of. Please correct this sentence.*

AR: We agree and we corrected the sentence as follow: "More complex neural networks techniques using self-organizing map have also given promising results."

AC: *Page 5646, line 1: I suggest rephrasing this sentence for easier reading. Perhaps 'The WEC forms part of the North-West European continental shelf - one of the world's largest margins.'*

AR: Done

AC: *Page 5646, line 7: I suggest removing 'by' from 'depths and by intense'*

AR: Done

AC: *Page 5652, line 13/14: Please remove 'the' from 'the SD', and include 'the' in 'of p over study period'.*

AR: Done

AC: *Page 5656, line 9/10: Suggest changing 'on' to 'in'*

AR: Done

AC: *Page 5663, line 5: Please change 'confirm' to 'further supports'*

AR: Done

AC: *Page 5664, line 20: Please include time (TI) in the list of predictor variables.*

AR: Done

AC: *Fig 10,11,12: What happened in December?*

AR: Because no monthly Chl-a data available in December. This is now mentioned in the caption of Figure 10.