

5th October 2021.

Biogeosciences.

Dear Editor,

Thank you for your and the reviewer's comments on our manuscript entitled 'Derivation of seawater $p\text{CO}_2$ from net community production identifies the South Atlantic Ocean as a CO_2 source' by Ford, Tilstone, Shutler and Kitidis. We have addressed all of the comments raised and implemented the necessary changes to the updated version of the manuscript. We provide detailed responses to the comments below and hope you find these changes satisfactory.

We look forward to hearing from you
Yours sincerely,



Daniel Ford



Registered Office:
Prospect Place
The Hoe, Plymouth
PL1 3DH, UK

T +44 (0)1752 633100
E forinfo@pml.ac.uk
W www.pml.ac.uk
@PlymouthMarine

Patron: James Cameron
Registered charity number 1091222.
PML is a company limited by guarantee,
registered in England & Wales,
company number 4178503

Response to Jonathan Sharp (Reviewer #1)

General Comments:

Daniel Ford and coauthors use a feed-forward neural network (FNN) to estimate surface ocean partial pressure of CO₂ (pCO_{2(sw)}) in the South Atlantic Ocean. The authors test satellite chlorophyll a (Chl a), satellite-derived net primary production (NPP), and satellite-derived net community production (NCP) as biological predictors in the neural network to determine which produces the most accurate pCO_{2(sw)}. They find that using satellite-derived NCP as a predictor in the FNN scheme produces the most reliable pCO_{2(sw)} reconstructions for the Amazon River plume and upwelling regions. They also show that, among the neural networks examined, the NCP-based FNN (SA-FNN_{NCP}) has the highest capacity for improved performance under scenarios of reduced uncertainty. For these reasons, the authors suggest that using satellite-derived NCP as a proxy for biological effects in surface reconstructions of pCO_{2(sw)} may be desirable going forward. Finally, Ford et al. find that SA-FNN_{NCP} indicates that the South Atlantic Ocean is a source of CO₂ to the atmosphere, whereas the FNNs with Chl a or NPP as a biological proxy or with no biological proxy all indicate that the South Atlantic Ocean is a CO₂ sink.

This manuscript fits well within the scope of Biogeosciences: it explores the implications of choosing different biological predictor variables in estimation schemes for sea surface pCO₂ and demonstrates the consequences of those choices for carbon cycling calculations. It is based on the very logical assumption that NCP, which captures all biological processes that modulate CO₂ concentrations in the surface ocean, should serve as a better biological predictor than Chl a for estimates of pCO_{2(sw)}. This work has the potential to shift the way in which studies of this nature are typically performed. That shift could result in better representations of sea surface pCO₂ in regions that are highly influenced by biological processes, regions that may contribute a disproportionately large fraction of global CO₂ flux across the air–sea interface.

In general, the manuscript is well-written and the figures and tables are effective in communicating the results. The manuscript addresses an important aspect of the global carbon cycle and is forward-thinking in its assessment of algorithm performance in response to reduced uncertainties. A couple concerns of mine, however, include the lack of quantitative or graphical support for the conclusion that SA-FNN_{NCP} produces the best representations of pCO_{2(sw)} compared to the other FNNs (section 4.2) and the shortage of further investigation into one of the manuscript's major conclusions: that SA-FNN_{NCP} flips the South Atlantic from a CO₂ sink to a CO₂ source. More discussion of these concerns as well as some minor comments can be found in the following sections.

Response: Thank you for the appraisal of our manuscript and the comments that you have provided which have improved the manuscript. Detailed responses to each comment are given below.

Specific Comments:

1. Performance of SA-FNN_{NCP}:

I mainly would like to see some quantitative or graphical evidence supporting the assertion that SA-FNN_{NCP} outperforms the other FNNs in the Amazon River plume and upwelling regions. Figure 3 shows differences between the mean climatologies given by different FNNs at different stations and Figure 4 shows that some of these differences are statistically significant (in comparison to SA-FNN_{NCP}), but neither says anything about the performance of any one FNN. That is left to the more



qualitative discussion in section 4.2 that compares general patterns in $pCO_{2(sw)}$ from previous studies to those indicated by the FNNs.

The points made in that qualitative discussion are compelling and certainly do appear to indicate superior performance of SA-FNN_{NCP} in the Amazon River plume, Benguela upwelling system, and equatorial regions. However, following along with the discussion takes some effort from the reader, and a lot of flipping back and forth between the text and Figure 3. A new figure comparing FNN results to some $pCO_{2(sw)}$ observations or a brief presentation of some relevant statistics would be more compelling. In particular, for example, $pCO_{2(sw)}$ data from the moorings at 6° S 10° W and/or 8° N 38° W could be plotted along with the SA-FNNs results to demonstrate the superior performance of SA-FNN_{NCP}.

I think this area is especially important to improve upon given that the bulk error statistics (Figures 2, A1, and A2) indicate SA-FNN_{NCP} to be the least accurate of the three FNNs that have biological predictors.

Response: Thank you for your suggestions, which we agree with as they make perfect sense. We have now added graphical evidence to support the discussion, as you suggest. Fig. 3 has been updated to include the PIRATA buoy data available in the SOCATv2020 Flag E dataset for 8°N 38°W and 6°S 10°W. These data were reanalysed to a consistent temperature and depth dataset in the same process as the SOCATv2020 data, so that all data have been treated in a consistent manner. We have also added graphical representations of the literature values described in the text to make the discussion easier to follow. Fig. 3 now also includes two SA-FNN variants (SA-FNN_{NO-BIO-1} and SA-FNN_{NO-BIO-2}) which have no biological parameters as input. The updated Fig. 3 with its caption is given below:

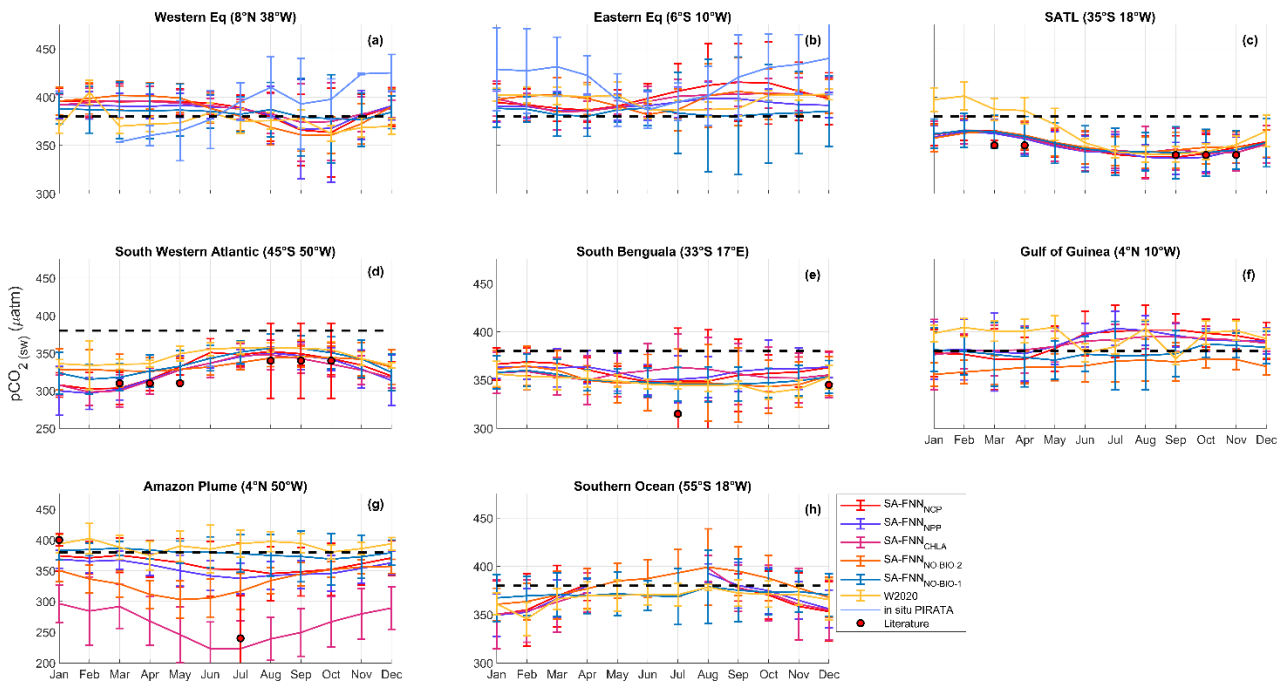


Fig. 3: Monthly climatologies of $pCO_{2(sw)}$ referenced to the year 2010 for the 8 stations marked in Fig. 1 from the SA-FNN_{NCP}, SA-FNN_{NPP}, SA-FNN_{CHLA}, SA-FNN_{NO-BIO} and W2020 (Watson et al., 2020b). Light blue lines in Fig. 3a, b indicate the in situ $pCO_{2(sw)}$ observations from PIRATA buoys. The atmospheric CO_2 increase was set



as $1.5 \mu\text{atm yr}^{-1}$. Black dashed line indicates the atmospheric $p\text{CO}_2$ ($\sim 380 \mu\text{atm}$). Error bars indicate the 2 standard deviation of the climatology ($\sim 95\%$ interval), where larger error bars indicate a larger interannual variability. Red circles indicate the literature values of $p\text{CO}_2(\text{sw})$ described in section 4.2. Note the different y-axis limits in Fig. 3g and Fig. 3d.

The updated Fig. 3 shows estimates of $p\text{CO}_2(\text{sw})$ climatologies from 5 SA-FNN variants and the W2020 at 8 stations in the South Atlantic Ocean. The figure also shows climatologies computed from the in situ PIRATA buoy observations (in Fig. 3 a, b,) and the literature values described in the text.

We have not however, produced statistical comparisons between the neural network estimates and the PIRATA data, because at $8^\circ\text{N } 38^\circ\text{W}$ the seasonal cycle is not fully captured, and the data only covers 4 years with patchy temporal coverage, so the datasets are not statistically comparable. For $6^\circ\text{S } 10^\circ\text{W}$ the data does capture the full seasonal cycle, but this consists of 6 years of data weighted towards June-December. As the neural network estimates cover 16 years, a direct statistical analysis is not possible. The inclusion of these data are beneficial to highlight that the SA-FNN_{NCP} does improve on the $p\text{CO}_2(\text{sw})$ estimates, but there are still some differences between December and April. These have now been discussed directly in section 4.2.

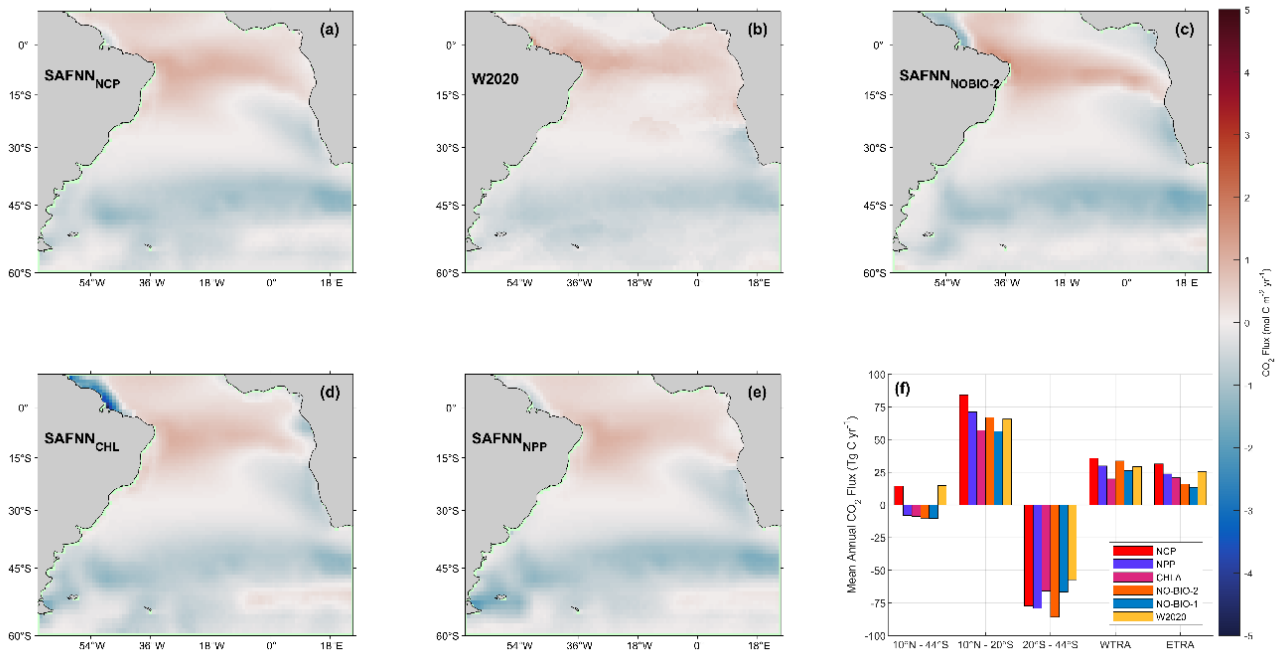
2. Sink to source transition:

The change in the cumulative regional sink from -7 Tg C yr^{-1} with the NPP-based FNN (SA-FNN_{NPP}) to $+14 \text{ Tg C yr}^{-1}$ with SA-FNN_{NCP} seems rather drastic, and I'm curious to know more about why such a significant change occurs. The reason is not obvious from Figure 5 alone. If indeed the transition occurs because high outgassing events in biologically-controlled regions with relatively limited geographic extent are captured by SA-FNN_{NCP} but not the other FNNs, as is suggested in lines 399–412, that point should be demonstrated and emphasized more explicitly.

This could perhaps be explored by breaking down the annual fluxes into different sub-regions (e.g., the biogeochemical provinces from Figure 1) and/or into average monthly fluxes to clearly show the spatial and/or temporal differences that lead to the significant discrepancy between SA-FNN_{NPP} and SA-FNN_{NCP}. This information could be presented in a table, figure, or even just in the body of the manuscript (like the geographic comparison in lines 419–420 between SA-FNN_{NCP} and the Watson et al. [2020] product).

Response: We have now updated the text and relevant figures to strengthen the comparison between the SA-FNN_{NCP} and the other methods through the addition of a bar chart to Fig. 5. The bar chart shows the mean annual CO_2 fluxes (Fig. 5f) for 5 regions; (1) the whole South Atlantic, (2) 10°N to 20°S , (3) 20°S to 44°S , and additional the (4) WTRA and (5) ETRA provinces. The results for the SA-FNN_{NO-BIO-1} and SA-FNN_{NO-BIO-2} appear very similar. For completeness, both have been included in the bar chart, but since they are so similar, only the spatial distribution of the fluxes for SA-FNN_{NO-BIO-2} are included (Fig. 5c). The updated Fig. 5 is given below:





The bar chart shows the increased outgassing in the Equatorial region observed by the SA-FNN_{NCP} compared to the other SA-FNN variants, and the overall CO₂ source for the South Atlantic Ocean estimated by the SA-FNN_{NCP}. We have expanded the discussion of the SA-FNN_{NCP} and SA-FNN_{NPP} differences. This appears in section 4.2 and reads: "The incremental inclusion of parameters to account for the biological signal starting with Chl a (-9 Tg C yr⁻¹) then NPP (-7 Tg C yr⁻¹) then NCP (+14 Tg C yr⁻¹) switched the South Atlantic Ocean from a CO₂ sink to a source, which is driven by differences in the pCO₂(_{sw}) estimates in regions that are biologically controlled. This 21 Tg C yr⁻¹ difference between the SA-FNN_{NCP} and SA-FNN_{NPP} is due to additional outgassing in the Equatorial Atlantic provinces of the WTRA and ETRA (Fig 1a; Fig.5f). Compared to the in situ pCO₂(_{sw}) observations at the Equatorial stations (Fig. 3a, b), it is likely that the outgassing is still underestimated by the SA-FNN_{NCP} but does improve these estimates within the upwelling season (June – September).".

Callbacks in Discussion section:

In general, because the work is presented in separate Results and Discussion sections, I'd make sure to refer specifically to figures, tables, and statements from the Results section when commenting on them in the Discussion section. This will make it easier for the reader to follow what exactly is being discussed without having to determine for themselves where those results are presented. I mention a couple specific instances of this in the following section.

Response: Thank you for this suggestion, which will help the reader to follow the logic. We have now included within the discussion the requested specific references to the relevant tables and figures. For example, please see throughout section 4.1 where we have now referred to Fig. 2, and Table 2 and 3.

Minor Comments and Technical Corrections

Line 17: There shouldn't be a comma after pCO₂(_{sw}).



Registered Office:
Prospect Place
The Hoe, Plymouth
PL1 3DH, UK

T +44 (0)1752 633100
E forinfo@pml.ac.uk
W www.pml.ac.uk
@PlymouthMarine

Patron: James Cameron
Registered charity number 1091222.
PML is a company limited by guarantee,
registered in England & Wales,
company number 4178503

Response: We have removed this comma.

Lines 45–48: I’d split this sentence into two; there’s a lot of information here and it’s a bit difficult to follow as written.

Response: We have now split this sentence to provide the information on autotrophic communities in the first sentence, and heterotrophic communities in the second sentence. These sentences now read as follows: “When NCP is positive, the plankton community is autotrophic which implies that there is a drawdown of CO₂ from seawater (since the plankton reduce the CO₂ in the water column). Where NCP is negative the community is heterotrophic implying a release of CO₂ into the ocean (as the plankton produce or release CO₂) which can then be released into the atmosphere (Jiang et al., 2019; Schloss et al., 2007).”

Figure 1: Do the colors and/or symbols represent different things in this figure? If so, I would mention it in the caption. If not, they could all be the same color/symbol since they’re labeled with letters anyway. Additionally, can you add the AMT stations or the transect lines to this figure?

Response: The different colours and symbols have no additional meaning to the stations, and therefore we have now made these the same symbol and colour. Additionally, and as requested, we have added the cruise tracks for AMT’s 20, 21, 22 and 23. The updated Fig. 1 is given below:

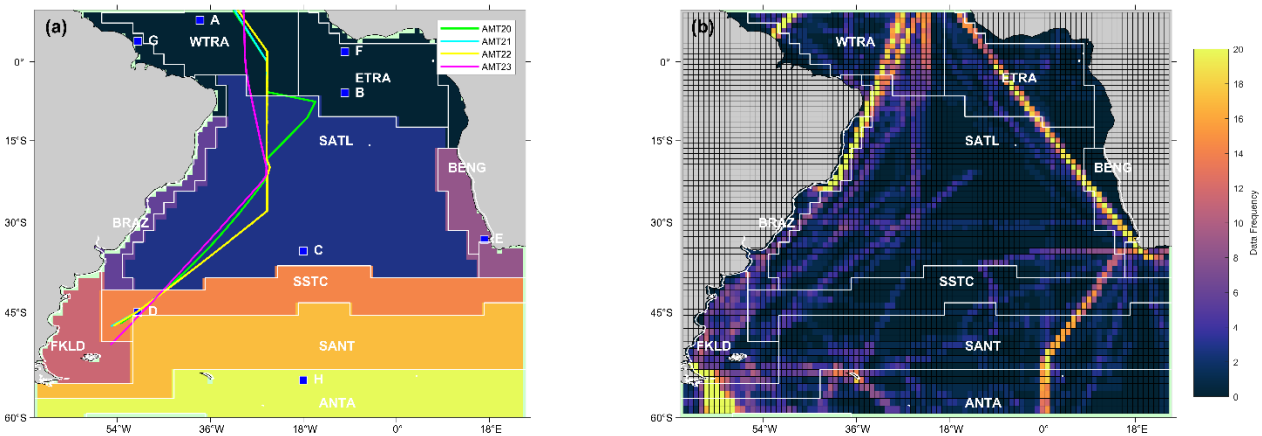


Table 1: What is the reference associated with the estimated uncertainty of 1 uatm in atmospheric pCO₂? Also, is it correct that the other uncertainty estimates all come from Ford et al. (2021)?

Response: We have now included the reference for the 1 uatm uncertainty in atmospheric pCO₂, as Takahashi et al. (2009). The use of Ford et al. (2021) for the remaining parameters is correct. Ford et al. (2021) performed an uncertainty analysis on the Moderate Resolution Imaging Spectroradiometer on Aqua (MODIS-A) estimates of chlorophyll-a, net primary production, net community production and sea surface temperatures for the South Atlantic Ocean which has provided all of the other uncertainty components.

Line 189: I’d rephrase this as “A non-parametric Kruskal-Wallis was used to test for...”



Registered Office:
Prospect Place
The Hoe, Plymouth
PL1 3DH, UK

T +44 (0)1752 633100
E forinfo@pml.ac.uk
W www.pml.ac.uk
@PlymouthMarine

Patron: James Cameron
Registered charity number 1091222.
PML is a company limited by guarantee,
registered in England & Wales,
company number 4178503

Response: Thank you for the suggested rephrasing. We have updated the sentence which now reads: “A non-parametric Kruskal-Wallis was used to test for significant ($\alpha < 0.05$) differences in the annual pCO_2 (sw), indicating an offset between the two tested climatologies.”.

Line 232: Should the accuracy here for SA-FNN_{NCP} be 21.68 matm, like in line 235 and Figure 2?

Response: Yes, that is correct. We have now corrected this typo.

Line 238: Based on Table 2, should these numbers be 36%, 36%, and 20%?

Response: The average reduction of the training, validation and independent datasets is described in the text. This average difference was 36% for NCP, 34% for NPP and 19% for Chl a. We have now clarified this in the manuscript, which now reads: “This showed that satellite NCP uncertainties lead to an on average 36 % reduction in pCO_2 (sw) RMSD, with NPP a 34 % reduction and Chl a a 19 % reduction across the three training and validation datasets.”.

Lines 259–260 (and elsewhere): Should be “minimum” instead of “minima” and “maximum” instead of “maxima”, or remove the article “a”. Minimum/maximum are singular whereas minima/maxima are plural.

Response: Thank you for highlighting this. We have now corrected this throughout the manuscript.

Line 283: Add the accuracy of the SA-FNN_{NCP} after you mention it here, so it can be easily compared with Landschützer et al. (2013) and Landschützer et al. (2014).

Response: We have now added the SA-FNN_{NCP} accuracy to the sentence as suggested. This now reads: “The SA-FNN_{NCP} had an overall accuracy (21.68 μ atm) which is consistent with other approaches that have been developed for the Atlantic (22.83 μ atm; Landschützer et al., 2013), and slightly lower than the published global result of 25.95 μ atm (Landschützer et al., 2014).”.

Lines 285–286: You’ve already mentioned in the previous sentence that the SA-FNN_{NCP} approach has a similar accuracy to other approaches in the literature, so I’d say here that training the SA-FNN with Chl a or NPP gave comparable broad-scale accuracy to training it with NCP.

Response: We have updated this sentence which now reads: “Training the SA-FNN using Chl a or NPP showed comparable broad-scale accuracy to NCP”

Line 295: “Satellite NCP is reliant on NPP as input” This has already been implied in line 288, so I’d remove the statement here or move it to the previous paragraph. The point is well made, it’s just that the writing is a bit repetitive here.

Response: We have removed this sentence and the paragraph as you suggest. It now reads: “To uncouple the Chl a, NPP and NCP estimates and their uncertainties, the perturbation analysis was also conducted on Atlantic Meridional Transect in situ observations.”



Line 298–299: “This showed that reducing in situ NCP uncertainties provided the greatest reduction in $p\text{CO}_{2(\text{sw})}$ RMSD, which was three times the reduction achievable using Chl a.” I’d make sure to refer the reader to Tables 2 and 3 so they’re not searching for this result.

Response: *We have now added the reference to Tables 2 and 3.*

Lines 328–329: “The stations (Fig. 1) represent locations of previous studies into in situ $p\text{CO}_{2(\text{sw})}$ variability in the South Atlantic Ocean and allow comparisons with literature values.” This point should be made earlier, in the Methods section, perhaps around line 182.

Response: *As suggested, we have now moved this sentence into the methods. This now reads: “The stations (Fig. 1) are representative of locations in the South Atlantic Ocean in the literature that have reported the variability of in situ $p\text{CO}_{2(\text{sw})}$. For each station, the monthly climatology of $p\text{CO}_{2(\text{sw})}$, representing the average seasonal cycle of $p\text{CO}_{2(\text{sw})}$, and the standard deviation of the climatology, as an indication of the interannual variability, were extracted from the five approaches.”*

Lines 338–339: “Valerio et al. (2021) indicated $p\text{CO}_{2(\text{sw})}$ varied above and below $p\text{CO}_{2(\text{atm})}$ at 4° N 50° W consistent with the SA-FNN_{NCP}.” It looks to me in Figure 3 like $p\text{CO}_{2(\text{sw})}$ from SA-FNN_{NCP} remains almost exclusively below $p\text{CO}_{2(\text{atm})}$ at this site?

Response: *We have corrected the sentence after rechecking the data presented in Valerio et al. (2021), which showed $p\text{CO}_{2(\text{sw})}$ to remain at or below $p\text{CO}_{2(\text{atm})}$. This now reads: “The SA-FNN_{NCP} however, agreed with in situ $p\text{CO}_{2(\text{sw})}$ observations at 4° N 50° W where $p\text{CO}_{2(\text{sw})}$ varied at or below $p\text{CO}_{2(\text{atm})}$ (Valerio et al., 2021).”*

Line 396: Should be “may be” instead of “maybe”

Response: *This has now been corrected.*



Response to Anonymous Reviewer (Reviewer #2)

This manuscript explored the importance of biological factors in modeling the seawater pCO₂ via the inclusion of different biological proxies (chl-a, NPP and NCP) in the model construction. The results reveal a similar model performance with the inclusion of different biological components but in some cases, projected pco₂ with the inclusion of NCP as a predictor works better. Overall, the paper emphasizes the significance of biological activity in controlling the magnitude and temporospatial pattern of sweater pCO₂ and underscores the urgency to further improve the accuracy in NCP prediction. However, I have some major concerns regarding the model test and data display.

Response: *Thank you for the appraisal of the manuscript. We have now addressed all of your concerns which have undoubtedly improved the work.*

Major comments:

1) Lack of comprehensive evaluation of multiple NPP and NCP algorithms. The key goal of your study is to elucidate the role of biological proxies in modeling seawater. Also, the author concluded the accuracy of NCP simulation is critical. If so, it is quite important to test the with the inclusion of commonly-used NPP and NCP algorithms (i.e, CbPM and VGPM-based NPP, NCP models by Li & Cassar, 2016; carbon export model by Laws et al., 2011 or (Henson et al., 2011)), rather than a single model output. Given that the focus is on the surface pCO₂, I would say NCP model developed by Li and Cassar seems a better fit than model by Tilstone (2015) as the former provides mixed-layer integration of NCP rates and later is corresponding to euphotic-zone integration. At the very least, the author needs to clarify the reason why they chose the present NPP and NCP models (i.e., they are tuned by the Atlantic dataset and therefore they are supposed to do the best job in your study region). But I am still looking forward to seeing if other NPP and NCP products can further improve pco₂ simulations.

Response: *The algorithms for Chl a, NPP and NCP are based on the work of Ford et al. (2021) which included an evaluation of the uncertainties (accuracy and bias) of the algorithms for the South Atlantic Ocean. In the Ford et al (2021) paper, five chl-a, three NPP and four NCP algorithms were evaluated, and the analysis is the first to account for-both satellite and in situ uncertainties. The NPP models assessed in Ford et al. (2021) were selected from models that had accurate performance in the Atlantic Ocean in previous studies (e.g. Campbell et al. 2002; Carr et al., 2006; Tilstone et al. 2009; Dogliotti et al., 2014; Lobanova et al. 2018), which included the VGPM. The NCP algorithms of Tilstone et al. (2015) were developed using bottle incubation data, and the algorithm used had the highest accuracy within the South Atlantic Ocean.*

Thanks for the suggestion about the Li and Cassar (2016) model. This NCP model was tuned using O₂/Ar ratio based measurements, where heterotrophic metabolic states (NCP<0) are not included. For the bottle incubations, which we used to validate the models, the NCP rates range from -50 to 50 mmol O₂ m⁻² d⁻¹. Since the O₂/Ar NCP measurements are not comparable over this range, using algorithms developed using bottle incubations and those using O₂/Ar NCP measurements are not directly comparable (please also see the discussions on this point within Duarte et al. 2013 and Williams et al. 2013). Therefore, evaluating the performance of the pCO₂(_{sw}) estimates using AMT data would not be applicable using the Li and Cassar (2016) approach. The export models of Laws et al. (2011) and Henson et al. (2011) do not derive NCP, but instead derive carbon export and are therefore not applicable.



Registered Office:
Prospect Place
The Hoe, Plymouth
PL1 3DH, UK

T +44 (0)1752 633100
E forinfo@pml.ac.uk
W www.pml.ac.uk
@PlymouthMarine

Patron: James Cameron
Registered charity number 1091222.
PML is a company limited by guarantee,
registered in England & Wales,
company number 4178503

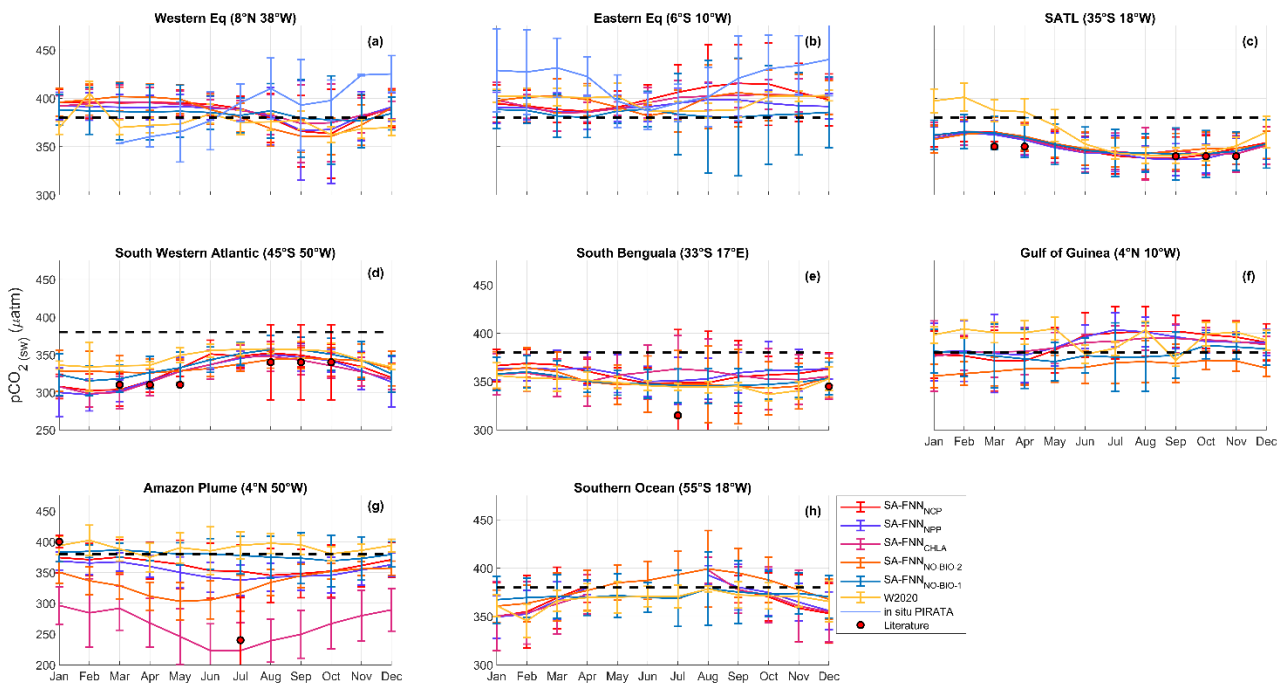
We have now clarified the algorithm choices in the text as follows: “These satellite algorithms were the most accurate for the South Atlantic Ocean in an algorithm inter-comparison which accounting for the uncertainties in both in situ, model and input data (Ford et al., 2021b).”

2) Unfair comparison between FNNbio and FNNNO_bio: The author used the model output without the inclusion of biological predictors (FNNNO_bio) as a reference to evaluate the improvement of FNNbio. However, FNNNO_bio was trained with inclusion of more physical parameters such as MLD, SST and salinity whereas FNN_bio just included SST. To assure a fair comparison, it should keep other predictors consistency except for the inclusion or exclusion of biological parameters

Response: We thank the Reviewer for their observation. To update the comparisons that we have conducted, we have now included a further SA-FNN_{NO-BIO-1} which is trained with just $pCO_2(atm)$ and SST.

For completeness, the original SA-FNN_{NO-BIO}, which is trained with $pCO_2(atm)$, SST, SSS and MLD, is also included in the manuscript (now labelled as SA-FNN_{NO-BIO-2}). This method therefore allows us to make a direct comparison between the SA-FNN_{NCP} and a SA-FNN trained with non-biological parameters consistent with the W2020 approach, but only using in situ $pCO_2(sw)$ observations from the South Atlantic Ocean.

We have updated Figs. 3, 4 and 5 to include the SA-FNN_{NO-BIO-1}. The updated Fig 3 is given below:



This new Fig 3 shows that the SA-FNN_{NO-BIO-1} and SA-FNN_{NO-BIO-2} have similar $pCO_2(sw)$ estimates at many of stations, but differences do occur within the Equatorial region. The pattern of the results is clearly shown in the updated Fig. 4, which is given below:



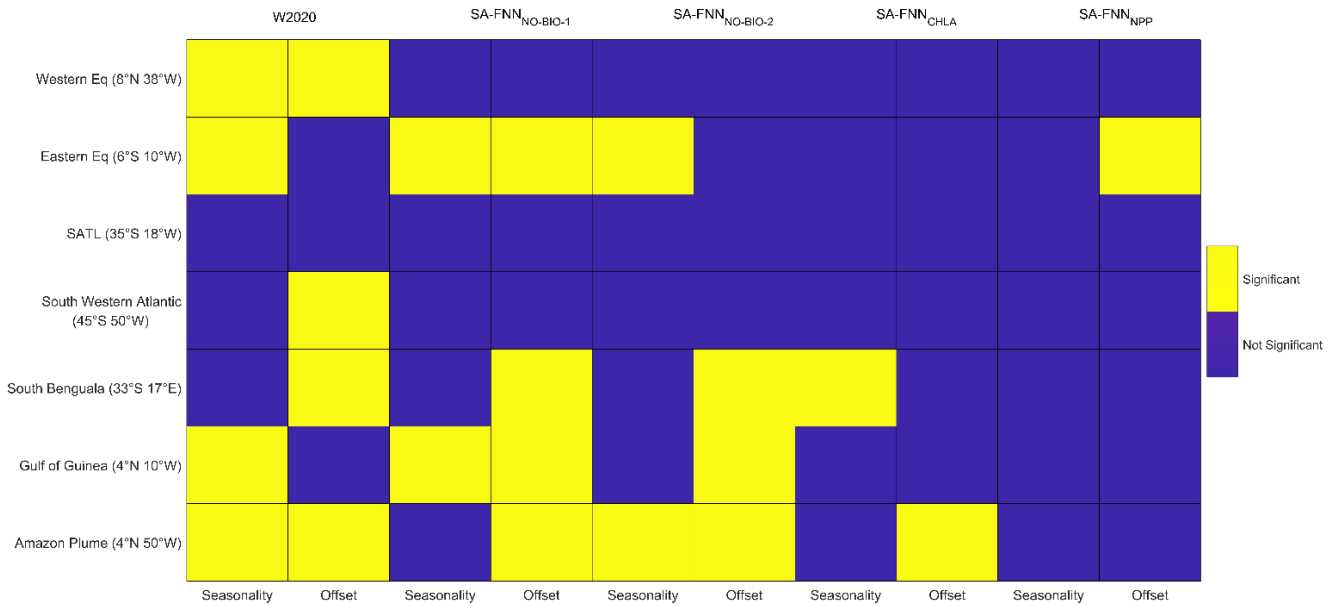


Fig. 4: Statistical comparison of the SA-FNN_{NCP} with the W2020, SA-FNN_{NO-BIO-1}, SA-FNN_{NO-BIO-2}, SA-FNN_{CHLA} and SA-FNN_{NPP} climatologies, where yellow blocks indicate a significant difference ($\alpha = 0.05$). Seasonality indicates a difference in the seasonal cycle and offset indicates a difference between the mean pCO_2 (sw) of the climatologies.

The SA-FNN_{NO-BIO-1} is also included in the mean annual CO_2 fluxes within the updated Fig. 5, which is given below:

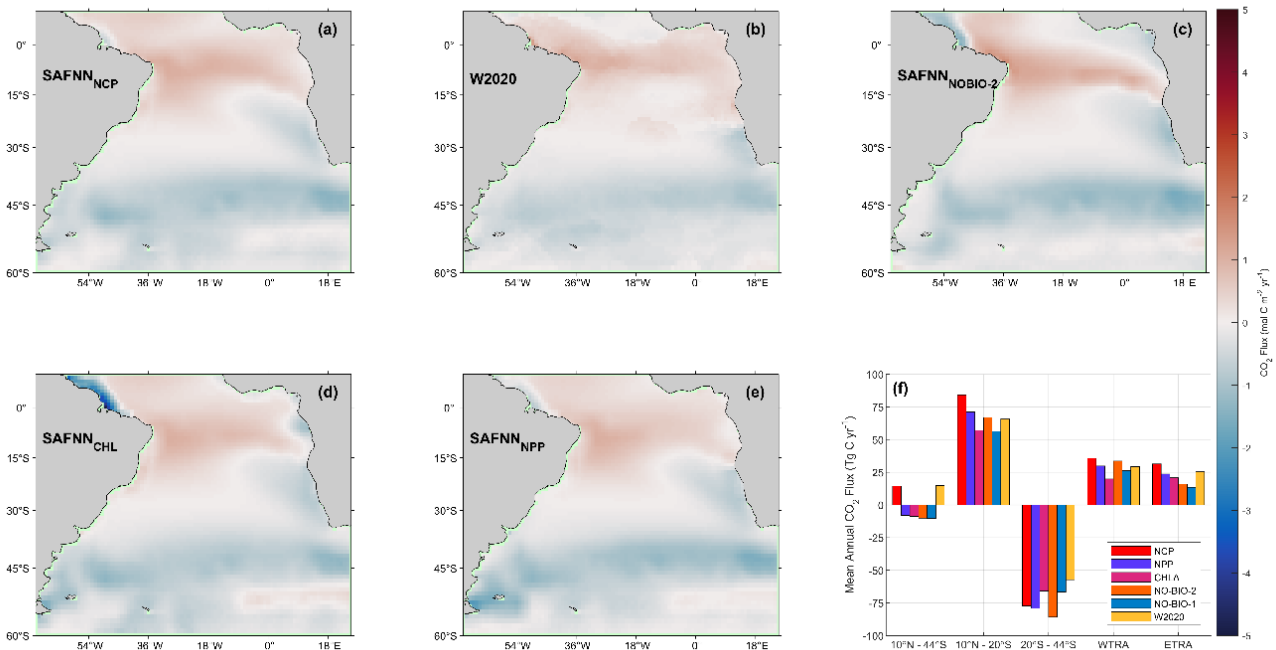


Fig. 5f indicates that the SA-FNN_{NO-BIO-1} and SA-FNN_{NO-BIO-2} have very similar mean annual CO_2 flux values for the whole South Atlantic (10° N – 44° S). The two approaches do exhibit regional

differences however, at $10^{\circ} N - 20^{\circ} S$ and $20^{\circ} S - 44^{\circ} S$. The result section 3.2 and discussion section 4.2 in the manuscript have now been modified to reflect the inclusion of the SA-FNN_{NO-BIO-1}.

For completeness and to address your comments, we also produced an SA-FNN trained using pCO_2 (_{atm}), SST, salinity, MLD and NCP. This showed similar mean annual fluxes to the SA-FNN_{NCP}, and this variant and result is only included in the discussion section to avoid confusion with the other SA-FNN variants. Section 4.2 includes this paragraph which now states: ‘A further SA-FNN trained with pCO_2 (_{atm}), SST, salinity, mixed layer depth and NCP indicated a similar CO_2 source of $12 Tg C yr^{-1}$ (data not shown) as the SA-FNN_{NCP} for the South Atlantic Ocean, highlighting that additional physical parameters cannot fully account for the biological contribution to the variability in pCO_2 (_{sw}). This further confirms the importance of using NCP within estimates of the global ocean CO_2 sink.’

3) The overstatement of the performance of FNNbio_NCP: In many places, the author argued that FNNbio_NC did the best among the models (i.e., Line 235, 360 and 365). However, it is hard to distinguish the visible improvement from the Figure. For example, I don’t think the difference between 34% and 36% is significant enough to say NCP is much better than NPP given that uncertainty.

Response: Thank you for pointing this out. We have now addressed this by either removing the statement (Line 235) or including further information in Fig. 3 to support the discussion at Lines 360 and 365. Please also see the responses to the minor comments below where this issue is also addressed.

Minor comments:

1.p in pCO_2 should be italic.

Response: We have made this change throughout the manuscript.

Line 15: it is hard to understand the method description regarding the “reducing uncertainty”. Please clarify or say more specifically,

Response: We have now clarified this sentence to make it clear that a perturbation analysis was used to explore the reduction in pCO_2 (_{sw}) uncertainties. The sentence now reads: “A perturbation analysis explored the maximum reduction in pCO_2 (_{sw}) uncertainties that could be achieved by reducing the uncertainties in the satellite biological parameters. This illustrated further improvement and greater differences for NCP compared to NPP or Chl a.”

Line35 : clarify the full name of chl-a when it appears in the main text for the first time

Response: We have now corrected this. Chlorophyll-a (Chl a) is now defined in the introduction.

Line 60: do you have a specific reason why you focus on the Southern Atlantic instead of the entire Atlantic basin or larger scale. You need some statement herein and also add some brief introduction about the setting of your study region.

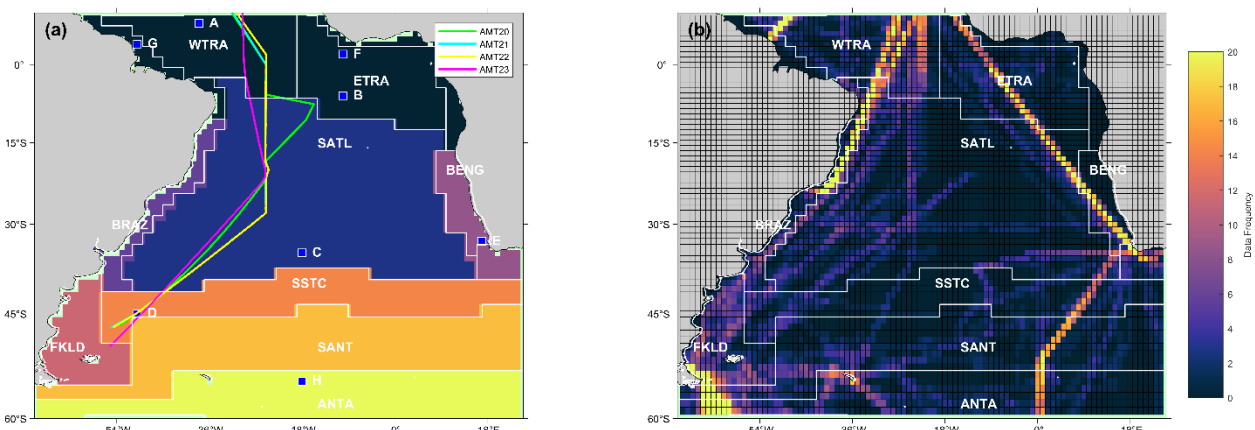


Response: The South Atlantic Ocean was chosen as the study region because current $pCO_{2(SW)}$ estimates have shown a large divergence in this region (e.g. Rodenbeck et al., 2015). This is because there are limited data available in most of the Southern Hemisphere. As we describe in the introduction, the use of NCP may aid neural network extrapolation techniques by constraining the biological control on $pCO_{2(SW)}$. We also selected the South Atlantic Ocean due to the comprehensive uncertainty analysis that Ford et al. (2021) performed on the satellite parameters, which included the consideration of uncertainties in both the satellite and in situ data. These estimates provide a robust foundation for the investigation into the uncertainties that are introduced into $pCO_{2(SW)}$ by the input parameters.

We have now made this clearer in the introduction, alongside a description of the major features in the South Atlantic Ocean. The paragraph in the introduction now reads: “The South Atlantic Ocean is under sampled with limited $pCO_{2(SW)}$ observations (e.g. Fay and McKinley, 2013; Watson et al., 2020b). The region is varied and dynamic as it includes the seasonal Equatorial upwelling, high biological activity on the south-western (Dogliotti et al., 2014) and south-eastern shelves (Lamont et al., 2014), as well as the propagation of the Amazon Plume into the western Equatorial Atlantic (Ibáñez et al., 2015). This in conjunction with a comprehensive database of satellite observation-based data with associated uncertainties (Ford et al., 2021b) provides the potential to identify the improvement to $pCO_{2(SW)}$ estimates that could be made from using NCP.”

Figure 1. Please add the data distribution (i.e., SOCAT and AMT cruise) you used for construction and validation of the model. It should be useful to add the name of the sub-regions in the map so the reader can easily navigate what subregions you are talking about in the following part. Probably you need two figures to get two issues done. Add the citation for the region division on the legend.

Response: Thank you for the suggestion. Fig. 1 has now been updated using two maps to address your comment. The first map shows the biogeochemical provinces, sub-region names, and the four Atlantic Meridional Transect cruise tracks as requested. The second map shows the frequency of the SOCAT dataset used in the training and validation of the SA-FNN. The updated Fig.1 is given below:



The Fig.1 caption has been updated to incorporate the changes, and now includes the citation of Longhurst et al. (1995) and Longhurst (1998). The legend now reads: “(a) Map of the 8 static biogeochemical provinces in the South Atlantic Ocean, following Longhurst et al. (1995) and Longhurst (1998). Markers and letters indicate the locations of timeseries extracted from Fig. 3. The four Atlantic Meridional Transect (AMT) cruise tracks are also overlaid (b) Map showing the spatial

distribution of the SOCATv2020 dataset used, where the data frequency is the number of available months of data within each 1° pixel. The province areas acronyms are: WTRA is Western Tropical Atlantic; ETRA is Eastern Equatorial Atlantic; SATL is South Atlantic Gyre; BRAZ is Brazilian current coastal; BENG is Benguela Current coastal upwelling; FKLD is Southwest Atlantic shelves; SSTC is South Subtropical Convergence; SANT is Sub Antarctic and ANTA is Antarctic.”.

Line 80: I would suggest changing pco2(atm) to pco2_air, which is more straightforward.

Response: *The use of pCO_2 (atm) is consistent with the nomenclature used for pCO_2 (sw) in this manuscript, and in previous manuscripts published in Biogeosciences (e.g. Landschützer et al., 2013; 2014). We have therefore left the notation as pCO_2 (atm). We hope that you find this acceptable.*

Table 1: does log10 means log10 (chl-a). It is very confusing. The error of 45 mmol O2 m-2 d-1 is very scaring because the typical range of NCP is from -50 and 50 mmol O2 m-2 d-1

Response: *The uncertainties in Chl a and NPP were assessed in log₁₀ space. We have now clarified that Chl a and NPP were log₁₀ transformed before input into the FNN, as follows: “Chl a and NPP estimates were log₁₀ transformed before input into the FNN, due to their respective uncertainties being determined in log₁₀ space.”.*

Line 115: quite confused about what ropt means. Does it mean the optimal data numbers for the training dataset?

Response: *The r_{opt} term indicates the optimal split in the input data to train and then validate the FNN, and this has been clarified in the methodology section 2.3.*

Line 125: The common practice for the data processing before training a neural network is to normalize the data to reduce the dynamics. Did you apply for it?

Response: *We applied no data processing to the input datasets (except for the log₁₀ transformation of Chl a and NPP)- Our approach is similar to Landschützer et al. (2013; 2014) who also applied no data processing, except for log₁₀ transformations of Chl a and Mixed Layer Depth. We have now included this detail within the methodology. This now reads (in section 2.3): “Chl a and NPP estimates were log₁₀ transformed before input into the FNN, due to their respective uncertainties which were determined in log₁₀ space.”.*

Line 126: please provide more details about the structure of FNN. How many layers and nodes do you set?

Response: *The FNN consist of 1 input layer, 1 hidden layer, comprising of between 3-32 nodes depending on the results of the pre-training step described in the methodology and 1 output layer. We have now added this information into the methodology, which now reads (section 2.3): “The FNN consisted of 1 hidden layer with between 2 and 30 nodes depending on the pre-training step and 1 output layer.”*

Line 145: Has pCO2 measurement in AMT cruise been included in the SOCAT dataset already?



Response: Yes, the AMT $pCO_{2(sw)}$ observations are included within the SOCAT dataset. The data points ± 20 minutes surrounding the in situ Chl-a, NPP or NCP observations ($N \approx 200$) were removed from the SOCATv2020 database. The removal of these data had minimal effect on the final gridded 1 degree monthly SOCAT observations used in training the SA-FNN. We have made this clear in the methodology describing the AMT datasets (Section 2.4). This reads: “These $pCO_{2(sw)}$ observations ($N \approx 200$) were removed from the SOCATv2020 dataset so that the Atlantic Meridional Transect data remained independent from the training and validation datasets.”

Table 2: you probably can bin Table 1 and Table 2 to make it easier for reading.

Response: Table 1 and Table 2 are important for interpreting the results. Table 1 indicates the uncertainties on the parameters used in the SA-FNN, which are key for understanding the perturbation analysis and the propagation of these uncertainties through the SA-FNN. Table 2 provides results on the perturbation analysis assuming uncertainties can be reduced to 0 and is directly referred to within the results and discussion sections. To facilitate understanding the results and logic of the analyses, we have kept Table 1 and 2 in the paper. We hope that you understand the reasoning on this.

Figure 2: the most useful information to evaluate the model performance is to look at the comparison with validation or independent dataset. I would suggest removing the figures of the training dataset and moving the validation results for models with inclusion of chla and NPP into the main texts.

Response: The SA-FNN_{NCP}, SA-FNN_{NPP} and SA-FNN_{CHLA} have similar broad-scale accuracies, and the addition of the SA-FNN_{NPP} and SA-FNN_{CHLA} plots within the main manuscript would add no additional information. The inclusion of the RMSD and bias within the text provides the necessary information in a more concise form, to show these similarities in their accuracy.

Figure 2: what does the color stand for? I don't think the blue line is useful. 1:1 line is enough. .

Response: The colours were chosen to highlight the error bars (blue) and the data points (red). If only one colour was chosen it became unclear where the data points were when the error bars were added. We have now added this description to the figure caption, which reads: “The data points are highlighted in red to distinguish them from the error bars in blue.”
The blue dashed line is the Type II weighted linear regression used in the validation of the SA-FNN outputs, as described in Ford et al. (2021). In cases where this line deviates from the 1:1, it would highlight the SA-FNN outputs had a systematic under/overestimation of $pCO_{2(sw)}$. This provides graphical evidence to the accuracy of the SA-FNN.

Line 2165: I have a very hard time understanding how you derived so-called maximum reduction. To achieve this goal, as I understand, you need to set a wide range of noise in the predictors (i.e., from -500% to 500%) and then find out the largest errors in pCO_2 compared to the white control (zero noise in predictors)?

Response: What the reviewer describes is similar to the maximum reduction that is determined by the perturbation analysis, for the individual parameter reductions. The “noise” in the input parameter is



set at the parameter's uncertainty (Table 1). The $p\text{CO}_2_{(sw)}$ estimates from the neural network are calculated for 3 perturbations: input value + uncertainty, input value and input value – uncertainty. These three $p\text{CO}_2_{(sw)}$ estimates are compared to the associated in situ $p\text{CO}_2_{(sw)}$ observation, and the estimate with the smallest difference is selected. This is repeated for all the available coincident observations in the dataset, for example $N = 1300$ for the independent test dataset. The RMSD is then calculated between the selected perturbations and the in situ $p\text{CO}_2_{(sw)}$. The percentage difference between this RMSD and the original RMSD (e.g. $21.68 \mu\text{atm}$ for independent test) is calculated, indicating the maximum reduction. We have now revised the methods section 2.5. to make this clearer.

Figure 4: how did you calculate the seasonality? Do you mean the amplitude of $p\text{CO}_2$? You can consider using Taylor diagrams to demonstrate the inter-model comparison, which provide more information.

Response: The seasonal $p\text{CO}_2_{(sw)}$ cycle at each station was estimated as the monthly climatology referenced to the year 2010. These were compared by calculating the Spearman Correlation coefficient ($n=12$) between the monthly climatology estimates from the SA-FNN_{NCP} and sequentially the five other methods. Where the Spearman Correlation was insignificant, the seasonality was deemed significantly different. We have now included the phrase 'seasonality' within the methodology, which was previously missing.

The use of a Taylor Diagram would be useful in the context of having temporally complete in situ fields of $p\text{CO}_2_{(sw)}$ to compare against. However, in the context of comparing multiple methods to the SA-FNN_{NCP}, without complete in situ data fields, there is no added benefit to using a Taylor Diagram, so it has not been added to the manuscript.

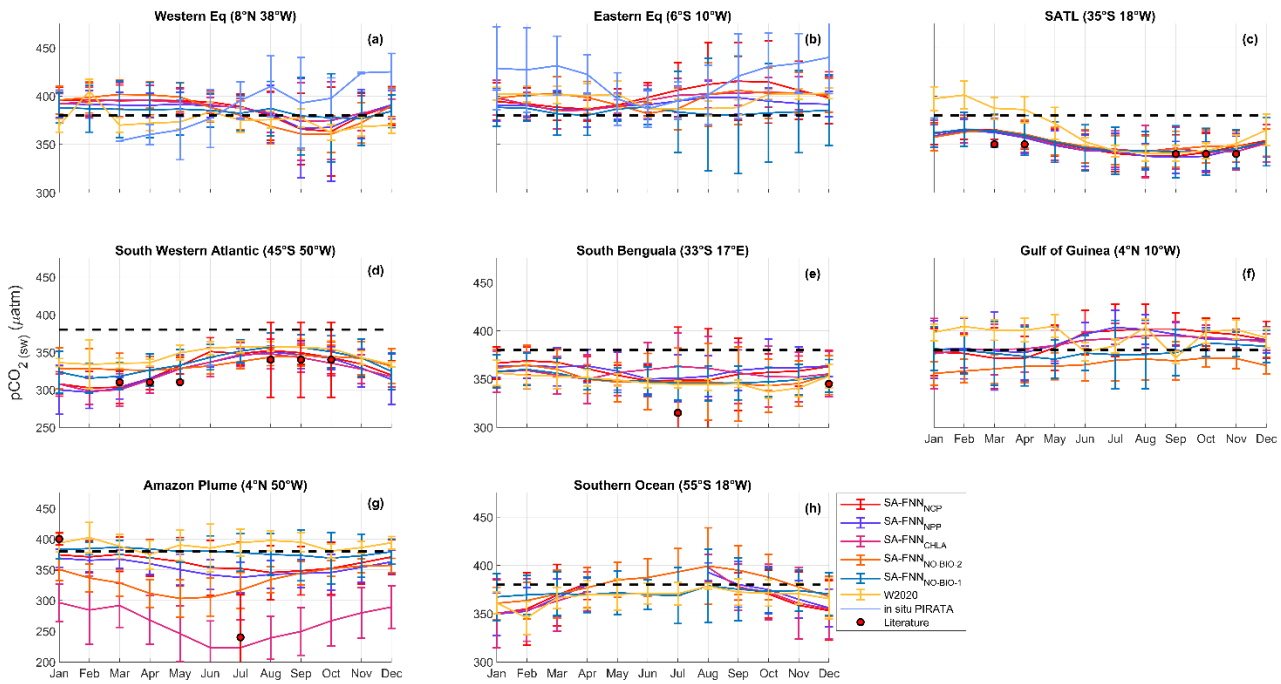
Line 235: I don't think the difference between 34% and 36% is significant enough to say that the NCP is much better than NPP.

Response: We have now removed the statement that reducing NCP uncertainties will have the largest impact on improving the $p\text{CO}_2_{(sw)}$ at this line. The similar percentage reductions between satellite NCP and NPP is due to the coupling of their uncertainties, which is discussed in section 4.1. Performing the perturbation analysis on the in situ AMT data removed this coupling of uncertainties and showed NCP would have the greatest improvement, which is also discussed in section 4.1.

Line 360 and Line 365: I had a hard time detecting significant improvement in SA-FNN_{NCP} compared to the other simulations.

Response: We have now updated Fig. 3 to include the literature values described in the text, to make it clearer about the variability discussed. The updated Fig. 3 is given below:





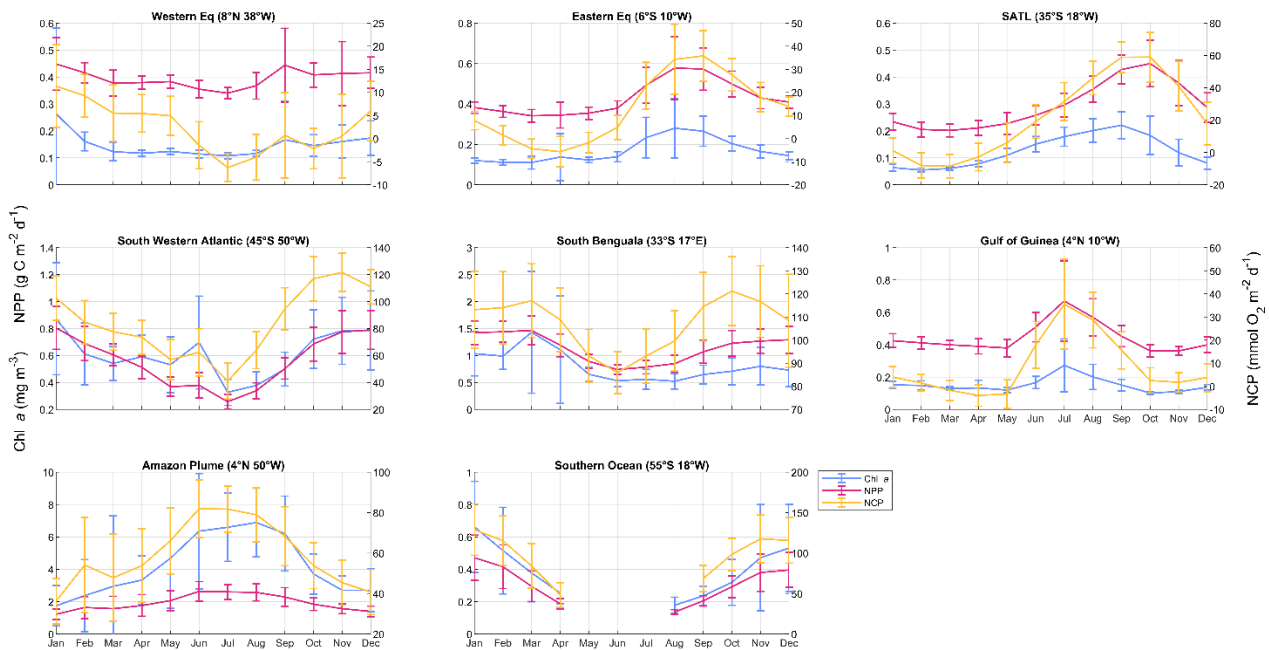
In reference to the sections highlighted. Fig. 3d now shows the improvement in SA-FNN_{NCP} pCO₂ (sw) compared to the literature values in austral autumn, where the SA-FNN_{NO-BIO-1} indicates higher pCO₂ (sw) similar to the W2020. We now highlight in the discussion (section 4.2) that there is no significant difference between the SA-FNN_{NCP} and the SA-FNN_{CHLA} and SA-FNN_{NPP} in Fig. 4.

Fig. 3e shows the literature values for the Benguela upwelling, which shows similarities between the SA-FNN_{NCP} and SA-FNN_{NO-BIO-1}. We suggest that the SA-FNN_{NCP} performs better and captures more of the variability due to upwelling through the larger interannual variability. We have further strengthened this section of the discussion by referring to the results of Arnone et al. (2017).

Line 380: Can you find more straightforward evidence by examining the seasonality of chl-a, NPP and NCP products to support your argument regarding the disconnection?

Response: *We have produced another figure for the appendix which displays the climatologies of Chl a, NPP and NCP at the 8 stations (Fig. C1) which is given below:*





The disconnect between Chl a and NPP reported by Lamont et al. (2014) can be seen in the satellite data, where NPP between April and October display similar values but Chl a is lower in October. The new Fig. C1 is referenced within section 4.2.

Line 430: cite the corresponding Table and Figure when you talked about the result.

Response: We have now corrected the omission of Table and Figure references throughout the manuscript, and at line 430.

Line 440: it should be useful to plot a bar chart to display the annual co2 flux among the regions

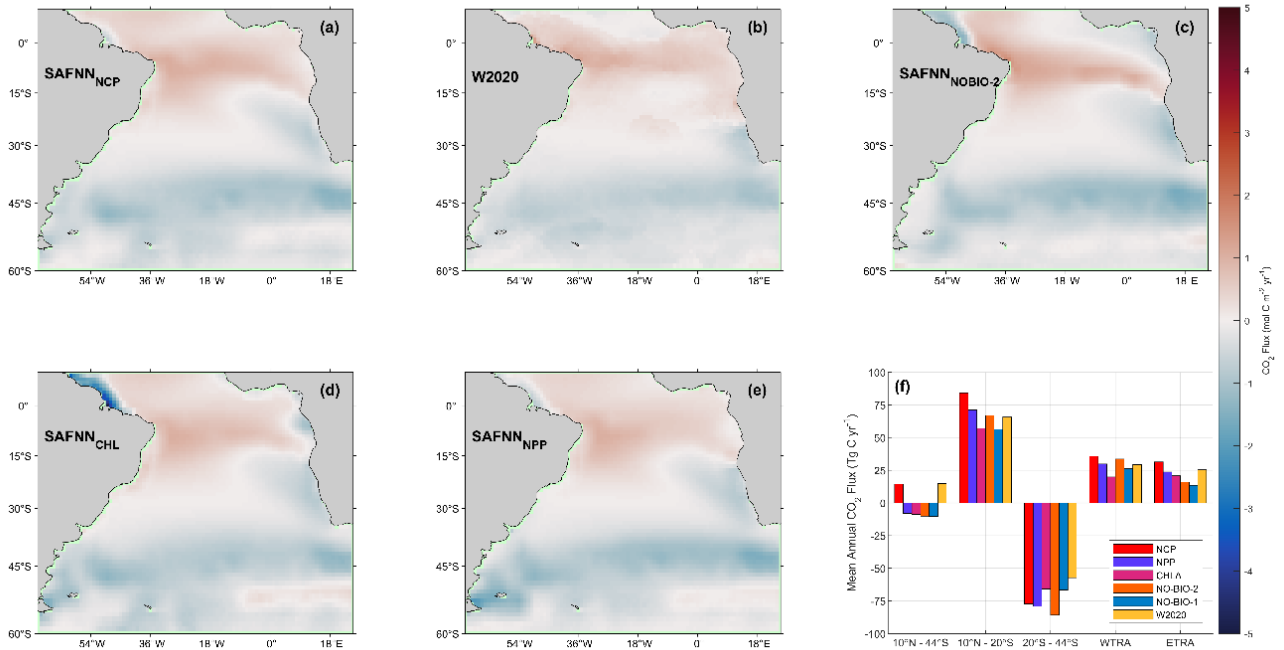
Response: We have now updated Fig. 5 to include a bar chart with the mean annual fluxes for the regions discussed in the text, and additionally for the WTRA and ETRA in line with another reviewer’s suggestion. The updated figure is given below:



Registered Office:
Prospect Place
The Hoe, Plymouth
PL1 3DH, UK

T +44 (0)1752 633100
E forinfo@pml.ac.uk
W www.pml.ac.uk
@PlymouthMarine

Patron: James Cameron
Registered charity number 1091222.
PML is a company limited by guarantee,
registered in England & Wales,
company number 4178503



The bar chart shows the overall CO_2 source estimated by the SA-FNN_{NCP}, mainly driven by increased outgassing in the Equatorial region. It also displays the similar overall CO_2 source estimate by the W2020, but the weaker sink between 20 °S – 44 °S and weaker source for 10 °N – 20 °S.



Registered Office:
Prospect Place
The Hoe, Plymouth
PL1 3DH, UK

T +44 (0)1752 633100
E forinfo@pml.ac.uk
W www.pml.ac.uk
@PlymouthMarine

Patron: James Cameron
Registered charity number 1091222.
PML is a company limited by guarantee,
registered in England & Wales,
company number 4178503