<u>Reviewer #2 Evaluations:</u>

Thank you for the opportunity to review "A novel sea surface partial pressure of carbon dioxide (pCO2) data product for the global coastal ocean resolving trends over the 1982-2020 period"

Summary and overall impression

The paper is well written. It highlights one of the potential sources of differences in existing global pCO2-products as they do not explicitly include the coastal ocean and the ones that do, cannot yet sufficiently capture the specific and changing conditions occurring along the coastal domain and only provide a climatology covering a relatively short period of time. The authors propose a resolution by addressing these shortcomings of the original global coastal pCO2-product by Laruelle et al. (2017) which was limited to the 1998-2015 period and expanding it to a much longer period (1982-2020) while updating the methodology to resolve long-term trends in global pCO2.

I strongly endorse the utilisation of two-step machine learning approaches for estimating sea surface pCO2 such as SOM-FFN where the authors first created biogeochemical clusters or provinces using SOM, and secondly, within each province identified in the step 1, established FFN-based nonlinear relationships between the observed sea surface pCO2 and independent environmental variables or drivers. The methodology setup is well explained.

However, I have made a few comments about some confusing terms, which I believe should be addressed quickly before publication. Overall, I enthusiastically recommend publication of the manuscript.

Author's response: We would like to express our gratitude for your thorough review of our manuscript. We are particularly grateful for your constructive feedback regarding potentially confusing terms. We will in the following address these concerns to improve the clarity and precision of our work. Changes to the original manuscript were made using Word's "track changes" option, and the line numbers indicated in our responses refer to the revised version of the manuscript.

Please also note that we updated the ULB-SOM-FFN-coastalv2 pCO_2 and FCO_2 products to address a small issue related to the sea-ice product, which was identified while working on these revisions. All figures, tables, and numbers in the text were thus updated to account for this correction but our results are only marginally modified, and our discussion and conclusions thus remain unchanged.

On behalf of the co-authors,

Alizée Roobaert

General Comments:

R2C1: I understand that 1998-2015 (Laruelle et al., 2017) is relatively short to evaluate the long-term trends of the coastal air-sea CO2 fluxes, but how can this period not be suitable to evaluate inter-annual variability as you mentioned? Given that it is 19 years of observations, can you elaborate more on this point? For example, how do you define "long-term trend"?

R2R1: We agree with this point raised both by R2R1 and R1R1 regarding the need for a more precise definition of the temporal scales used. Here, we first repeat our definition supplied in response to R1R1 as it strongly relates to the comment raised here: In this study, we have established a "long-term trend" or "multidecadal trend" as a linear trend that spans a period exceeding 10 years. In our case, this trend encompasses the years 1982 to 2020, resulting in a total of 39 years of observations. The term "decadal trend" is employed to describe a linear trend derived from a time series spanning 10 years, while "interannual variability" refers to the year-to-year fluctuations. Based on the given definitions, we concur with R2C1 that the period covered by the Laruelle et al. (2017) product (1998-2015) is appropriate for evaluating interannual variability and, to a lesser extent, for examining trends over a single decade. However, it is important to note that the product published and discussed by Laruelle et al. (2017) only provides a monthly climatology averaged over the period 1998-2015 and does not analyze interannual variability. This limitation is a key motivation for updating the Laruelle et al. (2017) product in our study. The purpose of this update is to facilitate investigations into changes in sea surface pCO_2 and FCO_2 in the global coastal ocean across a wide range of temporal scales, encompassing both short-term interannual variability and longer-term variations. While our present study primarily offers a concise overview of long-term variations, it does not delve into specific analyses focused on decadal and interannual timescales.

Based on this comment and R1C1, we modified the text with a more precise definition of the different temporal scales used in this study.

Lines 18-21: 'This product however has shortcomings because it only provides a climatology covering a relatively short period (1998-2015), thus hindering its application to the evaluation of the interannual variability, decadal changes and the long-term trends of the coastal air-sea CO_2 exchange, a temporal evolution that is still poorly understood and highly debated.'

Lines 28-30: 'Our results also show that the temporal trend changes in the air-sea pCO_2 gradient plays a significant role in the decadal long-term evolution of the coastal CO_2 sink, along with wind speed and sea-ice coverage changes that can also play an important role in some regions, particularly at high latitudes.'

Lines 35-39: 'Advancements in understanding the coastal air-sea CO_2 exchange (FCO_2) have been made, but change over different temporal scales (i.e., interannual and decadal changes as well as long-term temporal trends) remain unclear. Our research, based on observations and a machine learning approach, reconstructs the longest global time series of coastal FCO_2 (1982 to 2020). Results show the coastal ocean acts as a CO_2 sink, with increasing intensity over time. This new coastal FCO_2 product allows establishing regional carbon budgets and provides new constraints for closing the global carbon cycle.'

Lines 101-107: 'Using ULB-SOM-FFN-coastalv2 that relies on ~ 18 million coastal direct observations from the SOCATv2022 database, we recalculate the coastal air-sea CO_2 exchange (FCO_2) for the 1982-2020 period and briefly describe the long-term trend temporal evolution of the global coastal CO_2 sink over this timeframe. The long-term trend "or "multidecadal trend" in this study is defined as a linear trend that spans a period exceeding 10 years that is in our case a trend that encompasses the years 1982 to 2020, resulting in a total of 39 years of observations. This study does not discuss the decadal change (period of 10 years) and interannual variabilities (years-to-years fluctuations) of the global coastal sink.'

Line 447: '3.3.2 Decadal long-term trends in the coastal CO_2 sink'

Lines 449-453: 'Our results reveal significant spatial heterogeneities between the long-term temporal FCO_2 trends (linear trends that spans over 39 years) observed within different coastal regions, a finding consistent the range of varying slopes (including changes in sign of the slopes) already reported in local regional and discontinuous global studies (e.g., Becker et al., 2021; Laruelle et al., 2018; Wang et al., 2017). Our results also show that the decadal rates of changes in ΔpCO_2 and FCO_2 follow each other (compare Figs. 6a with 6b).'

Lines 463-466: 'Although our results suggest that the decadal long-term change in FCO_2 intensity mainly results from that of the ΔpCO_2 (compare Fig. 6a with 6b), the rate of change in FCO_2 can be amplified or dampened in some regions by changes in wind speed patterns and/or sea-ice coverage (through their effect on Eq. (1)), in agreement with recent findings by Resplandy et al. (2023).'

Lines 468-500: 'However, tTaking also the large heterogeneity in decadal the long-term FCO_2 trends, a quantitative analysis of the respective contributions of different coastal systems to the global strengthening of the coastal CO_2 sink should also be performed in the future, using a regionalized approach.'

Lines 446-448: 'However, this product was not designed or evaluated regarding its ability to resolve the interannual and decadal variabilitiesy and the long-term evolution of the coastal air-sea CO_2 exchange, which are still poorly understood (e.g., Bauer et al., 2013; Lacroix et al., 2021a; Laruelle et al., 2018; Regnier et al., 2013; Regnier et al., 2022; Resplandy et al., in rev.; Wang et al., 2017).'

Lines 552-555: 'We also provided a new coastal air-sea CO_2 exchange product for the same period and examined the long-term trends that is to say the temporal evolution of the global coastal CO_2 sink over the past four decades. This analysis reveals that the long-term temporal trend of the air-sea pCO_2 gradient drives most of the decadal long-term evolution of the coastal CO_2 sink, wind speed and sea-ice coverage playing a significant role regionally.'

Lines 562-564: 'In the future, our machine-learning approach could also be used to diagnose the main drivers of change in the global coastal ocean sink and more specifically, changes in the decadallong-term trends evolution of the coastal pCO_2 field.'

In Fig.4's caption: '…For each region, we report the bias (µatm), RMSE (µatm) and number of cells for the calculation between the reconstructed pCO_2-product and SOCAT_b as well as their respective long-term pCO_2 trend (in µatm decade^{-1}, which is calculate first as the slope of a linear trend using the monthly median values of all the deseasonalized data).'

In Fig. 6's caption: 'Figure 6. Long-term trend ofRate of change in (a) the coastal FCO_2 (in mol C m^{-2} yr^{-1} decade^{-1}), (b) the air-sea pCO_2 gradient (ΔpCO_2, in µatm decade^{-1}), (c) the wind speed at 10 meters above the sea surface (m s^{-1} decade^{-1}) and (d) the sea-ice coverage (decade^{-1}) from 1982 to 2020. For each panel, the rate changelong-term trend is calculated as the slope of a linear regression on the monthly median values of all the deseasonalized data from 1982 to 2020. We only present grid cells where a significant trend is detected based on a Mann-Kendall statistical test.'

Specific and Minor Comments:

R2C2: Line 131: "… each 0.25° cell is allocated to one of the 10 provinces (or neurons)." The content of the parenthesis, "or neurons" does not line up with the full sentence. It looks as if biogeochemical provinces/clusters were also neurons. A province is a self-organised map (SOM), a lattice of neurons or a single-layer neural network. I understand you referenced Landschutzer et al. (2013, 2014) which provide more details, but I suggest a revision of this segment to avoid confusion.

R2R2: We agree with R2C2 that the content of the parenthesis does not line up and to avoid confusion we modified the text to '… to one of the 10 provinces defined by the SOM'.

R2C3: Lines 297-298: "... since the algorithm minimizes the Root Mean Square Error (RMSE) between measurements and target observations." There seems to be some confusion here. Isn't the algorithm supposed to minimize the RMSE between "reconstructed values" and "target observations"?

R2R3: We agree with R2C3 and modified the term 'measurements' with 'reconstructed values. This has also been considered in other parts of the text.

Lines 312-315: 'A global mean of the residuals (bias) value of 0 µatm and a coefficient of determination (r^2) of 0.7 are calculated, as expected since the algorithm minimizes the Root Mean Square Error (RMSE) between ~~measurements~~ the reconstructed pCO_2 and target pCO_2 observations.'

R2C4: In Fig 3a-b, the similarity of the shape and spread of the four histograms of the residuals between decades raises questions on how you obtained the two sets of data SOCAT_a (80%) and SOCAT_b(20%). Since you randomly divided the original dataset to obtain them (Lines 179-182), how can you explain the "perfect" representation of data across the four decades?

R2R4: In this study, we use two distinct subsets of surface fCO_2 measurements from the SOCATv2022 database. The first set (SOCAT_a) is used for training the FFN whereas the second set (SOCAT_b) is used to validate our reconstructed pCO_2 derived from the FFN. Both these two sets have been created by randomly dividing the ~ 32 million CO_2 measurements into a group of 80 % of the original dataset (SOCAT_a) and a second group (SOCAT_b, 20 % of the original dataset). The large number of SOCAT data available to create the two subsets is enough to ensure that each is large enough to be fairly representative of the entire SOCAT dataset. These two sets of data are then gridded for each month at 0.25° using the average of all fCO_2 values in each cell. Values are then converted from fCO_2 to pCO_2 using the equation of Takahashi et al. (2019, page 67) and a coastal mask is applied on both gridded pCO_2 products. We agree with R2C4 that the shape and spread of the four histograms of the residual between Fig. 3a and Fig. 3b looks perfectly identical. One conclusion that could be drawn from it is that the neural network reconstruction is able to generalize the available measurements and thus is not overfitting the 80 % training data. One may also conclude that random division of the data does not guarantee independence and thus the 20 % independent measurements are reconstructed well as a result of the known autocorrelation length scales among measurements (Jones et al., 2012). However, we would also like to note that there are distinct features visible between these two figures that show differences that are visible for the first decade in the 80' 90'. Nevertheless, this example highlights the importance of independent data testing and - following the suggestion of referee 1, we have added a comparison to 3 buoy timeseries - see comment R1C1.

We now also modified the text in Sect. 2.1 to clarify how SOCAT_a and SOCAT_b have been calculated:

Lines 174-187: 'The surface pCO_2 data ~~used by the FFN~~ are extracted from the SOCATv2022 database (Bakker et al., 2022) that originally contains ~ 40 million pCO_2 measurements for the entire global ocean (open and coastal seas combined). We randomly divide this dataset into two independent datasets: a group of data used for the FFN algorithm (SOCAT_a, see below) and a group of data that we use to validate our reconstructed pCO2 (SOCAT_b). To do so, from the SOCATv2022 database, ~~W~~we follow the recommendation of the SOCAT community and use their accuracy criteria to only retain the data with the highest accuracy. To do so, we first select sea surface measurements expressed in fugacity of CO_2 (fCO_2) with a quality flag ranging from A to D (which corresponds to an estimated accuracy better than 5 µatm) and a World Ocean Circulation Experiment (WOCE) flag of 2 (good dataset following SOCAT) for the 1982-2020 period. Following Laruelle et al. (2017), we also remove fCO_2 values < 30 µatm and > 1000 µatm that are likely derived from estuarine or fresh water systems that are not included in our coastal domain. We then randomly divide this dataset rich of ~ 32 million fCO_2 measurements into a group of data used for the FFN algorithm ('a', 80 % of the original dataset) and a group of data that we use to validate our reconstructed pCO_2 ('b', 20 % of the original dataset). The two sets of data (SOCAT_a and SOCAT_b) are then gridded for each month at 0.25° using the average of all fCO_2 values in each cell. Values are then converted from fCO_2 to pCO_2 using the equation of Takahashi et al. (201~~2~~9, page ~~6~~7) and a coastal mask is applied on both gridded pCO_2 products'

R2C5: Given that "the spatial extension of the provinces varies from one month to the other because of the seasonal variations of the environmental drivers", I suggest an update of the caption of Table 2 to be specific with the "biogeochemical provinces" on which spatial evaluation is performed.

R2R5: We agree with the reviewer that the caption of Table 2 was unclear considering that the spatial extension of the biogeochemical provinces varies from one month to the other. For each province, statistical analyses are performed each month using all of the cells of the province over the entire 1982-2020 period thus accounting for temporal variations of the spatial extent of the provinces. We updated the caption of Table 2 to stress out this point.

In Table 2's caption: '**Table 2: Statistical analyses (bias, RMSE and r²) of the reconstructed coastal pCO_2-product against pCO_2 observations from SOCAT_a and SOCAT_b -for the different biogeochemical provinces. For each province, bias, RMSE and r² are calculated using all of the monthly cells of the province for the period 1982-2020.**'

R2C6: Lines 338-339: "This dataset consists of a pool of 404,206 gridded cells that are uniformly distributed between both hemispheres". From reading this, it now seems clear that you randomly divided the gridded cells of the original dataset (pCO2 observations). If this is the case, provide a better explanation in Sect. 2.1 because this would clarify my earlier comments on Fig. 3a-b.

R2R6: Indeed, we have now provided more clarity in the text (see R2C4 above).

R2C7: Line 333: "can likely also explain". The term "can explain" already implies a level of likelihood or possibility, so adding "likely" before it is unnecessary and redundant

R2R7: We agree with R2C7 and modified the text accordingly.

R2C8: Check the units of $\Delta pCO2$ and pCO2 throughout the manuscript. You put "atm" instead of "µatm". See Lines 226-227, for example.

R2R8: The units of the atmospheric and oceanic pCO_2 are always expressed in µatm in the Text, Figures and Tables. The units of ΔpCO_2 are also expressed in µatm when its long-term change is discussed (µatm decade^{-1}, Fig. 6). We just expressed the units of ΔpCO_2 in atm when used in Eq. (1) to calculate a FCO_2 expressed in mol m^{-2} yr^{-1}. We thus consider that no change must be done for the units of pCO_2 and ΔpCO_2 in the text.

R2C9: Line 140: "South Hemisphere" should be read "Southern Hemisphere".

R2R9: We agree with R2C9 and modified the text accordingly.

R2C10: Line 147: "a target variable" should be read "the target variable" given that it is known.

R2R10: We agree with R2C10 and modified the term accordingly.

R2C11: Line 157: instead of "calculate", I suggest you use "estimate" as it sounds more appropriate.

R2R11: Following R2C3, we decided to use 'reconstruct' instead of 'calculate' to be consistent throughout the manuscript. This change is also considered in other places in the text to be consistent.

Lines 148-161: 'In a second step, within each biogeochemical province identified in step 1 (SOM), a FFN algorithm establishes nonlinear relationships between the observed sea surface pCO_2 and independent variables, or drivers, that are known to control its spatial and temporal variability. For each province, the FFN algorithm calculates relationships between ~~a~~the observed target variable (here pCO_2 using pCO_2 observations from the SOCAT_a dataset - see below) and inputs (environmental drivers - see below and Table 1) by adjusting weighting factors of a sigmoid activation function (one sigmoid function per neuron in the hidden layer) following an iterative procedure, i.e., a Levenberg-Marquardt backpropagation algorithm. At the first iteration, the weights of neurons are randomly assigned and the ~~estimated~~ reconstructed pCO_2 is compared with the actual pCO_2 observations. Based on the resulting mismatch, the network weights are iteratively updated in a way that the error function - in our case the mean squared error between network output and actual observations - gets minimized. For each iteration, the FFN algorithm uses a fraction of the pCO_2 observations for the actual training of the network (i.e., the adjustment of the neuron weights), while another randomly selected fraction of the dataset is used to independently evaluate the performance of the algorithm. The final coefficients are obtained when the reconstructed pCO_2 simulated from the validation data does not significantly improve relative to the pCO_2 observations, to prevent overfitting. The final neuron weights and thus the resulting input-output relationships are used to reconstruct ~~calculate~~ pCO_2 in each cell and for each month during the 1982-2020 period.'

R2C12: Line 288: "see section 3.3.3" should be written "Sect. 3.3.3" for consistency

R2R12: we agree with R2C12 and modified the text accordingly.

R2C13: Line 396: "southern Hemisphere" should be written "Southern Hemisphere"..

R2R13: we agree with the reviewer and modified the text accordingly.

R2C14: Line 406 and Fig. 6's caption: "RMS" is used instead of "RMSE". I suggest you check these also throughout the manuscript.

R2R14: The Root Mean Square (RMS) is used on line 406 of the original manuscript and in Fig. 6's caption since we evaluate and discuss the seasonal amplitude signal of the air-sea CO_2 exchange (and thus not an error term). The Root Mean Square Error (RMSE) in contrast is used when comparing two datasets and their respective differences (e.g., in this study when comparing our reconstructed pCO_2 and FCO_2 against the known truth or observed values). Therefore, no changes have been made to the manuscript in response to this comment.

Lines 443-446: 'Finally, the seasonal coastal FCO_2 variability (expressed as the Root Mean Square (RMS) of the seasonal amplitude,) agrees with the few studies performed at global scale (see e.g., Dai et al., 2022; Roobaert et al., 2019) with high seasonal FCO_2 amplitudes (RMS~~rms~~ values > 1.5 mol C m^{-2} yr^{-1}) at temperate and high latitudes and a low amplitude over the subtropical band (Fig. 5b).'

R2C15: Line 407: "rms values" should be written "RMSE values".

R2R15: see R2C14

R2C16: Line 487: "and can be display large variations the regional scale …" This sentence needs revision.

R2R15: we agree with the reviewer and revised the text:

Lines 529-531: 'It should be noted that all these uncertainties are calculated globally and can be ~~display large variations the regional scale~~larger at the regional scale (see e.g., Roobaert et al., 2019) as exemplified by the uncertainty associated with the choice of wind speed product on the FCO_2 calculation (see Roobaert et al., 2018).'

R2C17: Line 495: "depend" instead of "depending".

R2R17: we agree with the reviewer and modified the text accordingly.

R2C18: Line 496: "use" instead of "used".

R2R18: We suppose that the reviewer refers to 'used' instead of 'use' and thus modified the text accordingly.

Lines 535-540: 'Using the narrow definition of the coastal domain (i.e., the shelf break as the outer limit), we calculate a global value of 0.01 Pg C yr^{-1} for σ_{FCO_2} (7 % uncertainty on the global FCO_2, which is consistent with the global FCO_2 uncertainty calculated by Roobaert et al. (2019, 10 %)), 7 % FCO_2 difference depend on the k-formulation used (σ_k value of 0.01), 2 % difference on the FCO_2 calculation depend~~ing~~ on the wind product choice (σ_{wind} = 0.002 Pg C yr^{-1}), 8 % for the sea-ice choice (σ_{ice} = 0.01 Pg C yr^{-1}), 36 µatm for θ_{map}, 8 µatm for θ_{grid} and 3 µatm for θ_{obs}. '

REFERENCES

Jones, S. D., C. Le Quere, and C. Rödenbeck (2012), Autocorrelation characteristics of surface ocean pCO2 and air-sea CO2 fluxes, Global Biogeochem. Cycles, 26, GB2042, doi:10.1029/2010GB004017

Laruelle, G. G., Landschützer, P., Gruber, N., Tison, J. L., Delille, B., & Regnier, P. (2017). Global high-resolution monthly pCO2 climatology for the coastal ocean derived from neural network interpolation. Biogeosciences, 14(19), 4545–4561. https://doi.org/10.5194/bg-14-4545-2017

Takahashi, T., Sutherland, S. C., & Kozyr, A. (2019). Global ocean surface water partial pressure of CO2 database: Measurements performed during 1957–2018 (version 2018). NOAA/NCEI/OCADS NDP-088 (V2018) Rep., 25 pp., https://www. ncei. noaa. gov/access/ocean-carbon-data-system/oceans/LDEO_Underway_Database/NDP-088_V2018. pdf.