Reviewer #1 Evaluations:

General comments

This study updates the ULB-SOM-FFN-coastalv1 pCO2 data product, enhancing its capability to detect the long-term trends spanning from 1982 to 2020 (> 30 years) on a global scale. The authors achieved this objective by using additional environmental predictor (pCO2air) and new pCO2 observations from SOCAT v2022 for training and validation. The average bias between the pCO2 product and SOCAT observations are all close to 0 µatm in different decades (Figure 3), which means that this new version is suitable for investigating decadal trends in coastal pCO2 and CO2 fluxes on a global scale. To the best of my knowledge, this is the first coastal pCO2 product that tries to resolve long-term pCO2 trends. I tried to be picky to find out some potential mistakes, yet it's worth acknowledging that the author's calculations appear to be robust. At least my ability might not be sufficient to identify any big mistakes that might exist.

Author's response: We would like to express our gratitude to Zelun Wu for his examination of our study and the recognition of the interest of our work. We are also grateful for his constructive suggestions to improve the quality of our manuscript. We have considered all Zelun Wu's remarks and provide in the following sections detailed responses to each comment (in blue). 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

Nevertheless, as a data user, I would be very grateful if the author could clarify some of my concerns in the article so that I can use this data correctly in the future.

R1C1: My primary concern revolves around the lack of clarity in defining the term "long-term trend". This study focuses on assessing the "long-term trend" or specifically, the "decadal trend". The term "long-term" is a relative definition, and the term "decadal trend" means the linear or nonlinear tendency of a time series longer than 10 years. Given that Figure 4 and Section 3.2 employ a "30-year" constraint for comparison, It is interesting to ask whether the updated version can be used to identify the linear trends of the 10~30 years time scale.

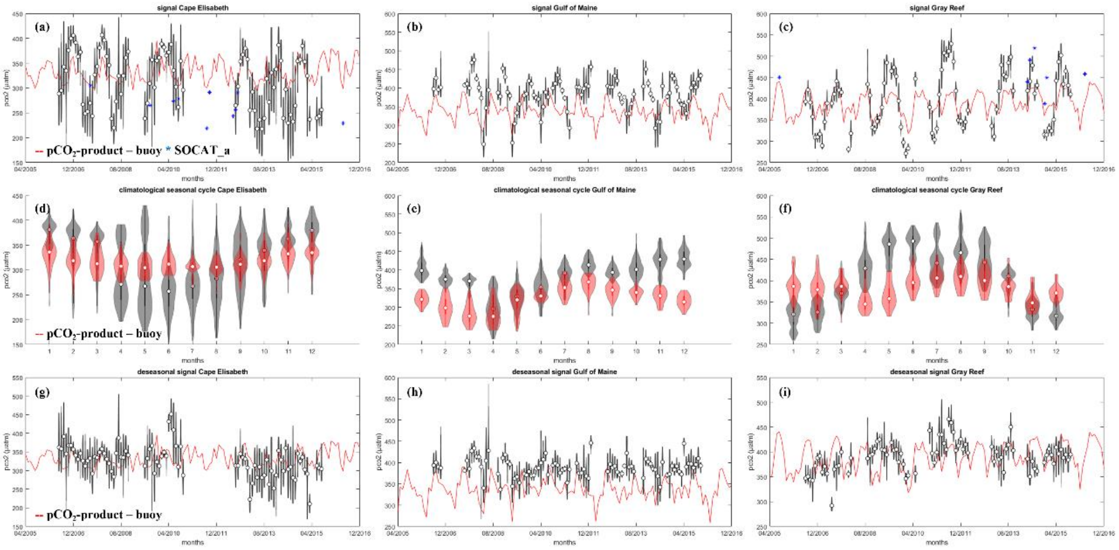
If the answer is yes, it is suggested to compare the pCO2 product with the buoy data, as mooring data are more continuous than SOCAT observations. Sutton et al. (2019) have comprehensively summarized these time series, with several buoys deployed in coastal areas and with continuous measurements longer than 10 years, including NDBC Buoy 46041 in Cape Elizabeth, NDBC Buoy 41008 in Gray's Reef, and Coastal Western Gulf of Maine Mooring in the Gulf of Maine. All three of these buoys have records since 2006, rendering them suitable for supplementary comparison, as proposed in Figure 4.

If the answer is no, i.e., the current version doesn't validate for trends < 30 years, I recommend clearly clarifying the definition of "trend" or "long-term" in the title, abstract, and conclusions, to ensure that the data can be used correctly in future studies.

Ref: *Sutton, et al., 2019. Autonomous seawater pCO2 and pH time series from 40 surface buoys and the emergence of anthropogenic trends. Earth Syst. Sci. Data 11, 421–439. https://doi.org/10.5194/essd-11-421-2019*

R1R1: We concur with this point which was raised by both reviewers in R1C1 and R2C1 and we agree that temporal scales investigated in our studies need to be better defined in the text. In our study, we consider the "long-term trend" or the "multidecadal trend" as a linear trend that spans at least over a period exceeding 10 years. In our study, long-term trends are calculated over the entire 1982 to 2020 period, resulting in a 39-year time series. The term "decadal trend" describes a linear trend derived from a time series spanning 10 years exactly, while the "interannual variability" refers to year-to-year fluctuations and is not derived from a linear trend analysis. We thus modified the text in several places with more precise definitions of the different temporal explored (see also R2C1).

The good evaluation of the reconstructed pCO$_2$ product against SOCAT (SOCAT_a and SOCAT_b) performed in this study lends confidence in our ability to properly capture the temporal evolution of coastal pCO$_2$ with the newly reconstructed pCO$_2$-product at the different timescales described above. We however agree with R1C1 that to complement the spatiotemporal evaluation against SOCAT, a comparison of the newly reconstructed pCO$_2$-product against pCO$_2$ times series of measurements sampled using buoys located in coastal areas provides a better picture and independent test of the performance of our data product. We thus followed R1 suggestion and compared the reconstructed pCO$_2$ times series against the 3 buoys with measurements longer than 10 years located in Cape Elizabeth (NDBC Buoy 46041), in Gray's Reef (NDBC Buoy 41008) and in the Gulf of Maine (Coastal Western Gulf of Maine mooring, Sutton et al., 2019). The three pCO$_2$ time series for each buoy location mentioned above are presented in FIG_R1C1.



FIG_R1C1: Comparison between the reconstructed coastal pCO$_2$-product (red) against times series from 3 coastal buoys (black) namely in Cape Elizabeth (NDBC Buoy 46041), in Gray's Reef (NDBC Buoy 41008) and in the Gulf of Maine (Coastal Western Gulf of Maine mooring, Sutton et al., 2019). (a-c) Monthly time series, (d-f) climatological mean seasonal cycles, (g-i) linear trends from deseasonalized data. The violins for the buoys represent the intra variability on the pCO$_2$ measurements. The violins in panels (d) to (g) for the reconstructed pCO$_2$ represent the pCO$_2$ variability of the month x around its monthly climatological pCO$_2$ mean.

Although smaller amplitude variabilities are generally observed, results show that the reconstructed pCO$_2$ times series follow those of the observational data with values that are mainly between the buoys errors (FIG_R1C1a-c). We speculate that the smaller amplitude stems from the coarser 0.25° x 0.25° degree grid resolution of our method compared to the point nature of the buoy data. Landschützer et al. (2016) drew a similar conclusion when they compared their open ocean pCO$_2$ data with open ocean at HOT and BATS and is further corroborated by the much smaller variabilities obtained when raw SOCAT data are averaged at the grid cell level (Fig. 4 of the Ms). The exception is the Gulf of Maine where a general underestimation of pCO$_2$ is observed compared to the buoy observations. The reconstructed pCO$_2$ product also reproduces the observed climatological seasonal cycles (including a relatively good timing of the seasonal maxima and minima) for the 3 buoys as shown in FIG_R1C1d-f. Absolute average biases values of 14, 4, 45 µatm and RMSE values of 50, 53 and 61µatm are calculated between the pCO$_2$-product and the observations for Cape Elisabeth, Gray's Reef and in the Gulf of Maine, respectively. These statistical error values are larger than those calculated when the comparison is performed against SOCAT, whether on a global or regional scale. This is not surprising since the reconstructed pCO$_2$ product is for a global application and is quite challenging to compare to specific coastal buoys that present high temporal and spatial variability such as shown by the violin in FIG_R1C1.

The good evaluation of the reconstructed pCO_2 against both SOCAT_a, SOCAT_b as well as against buoys data gives confidence to identify the linear trends of both the pCO_2 and FCO_2 over different temporal scales. This study, however, only presents an overview of the long-term variations of the coastal FCO_2 (Sect. 3.2) but does not discuss coastal pCO_2 and FCO_2 variabilities on decadal and interannual timescales.

We modified the text in the Ms to include the comparison against the buoys data:

Lines 387-411: 'Finally, the reconstructed pCO_2 times series are compared against 3 buoys with measurements longer than 10 years located in Cape Elizabeth (NDBC Buoy 46041), in Gray's Reef (NDBC Buoy 41008) and in the Gulf of Maine (Coastal Western Gulf of Maine mooring, Sutton et al., 2019). The three pCO_2 time series for each buoy location mentioned above are presented in Fig. S3. Although smaller amplitude variabilities are generally observed, results show that the reconstructed pCO_2 times series follow those of the observational data with values that are mainly between the buoys errors (Fig. S3a-c). We speculate that the smaller amplitude stems from the coarser 0.25° grid resolution of our method compared to the point nature of the buoy data. Landschützer et al. (2016) drew a similar conclusion when they compared their open ocean pCO_2 data with open ocean at HOT and BATS and is further corroborated by the much smaller variabilities obtained when raw SOCAT data are averaged at the grid cell level (Fig. 4). The exception is the Gulf of Maine where a general underestimation of pCO_2 is observed compared to the buoy observations. The reconstructed pCO_2 product also reproduces the observed climatological seasonal cycles (including a relatively good timing of the seasonal maxima and minima) for the 3 buoys as shown in Fig. S3d-f. Absolute average biases values of 14, 4, 45 µatm and RMSE values of 50, 53 and 61 µatm are calculated between the pCO_2-product and the observations for Cape Elisabeth, Gray's Reef and in the Gulf of Maine, respectively. These statistical error values are larger than those calculated when the comparison is performed against SOCAT, whether on a global or regional scale. This is not surprising since the reconstructed pCO_2 product is for a global application and is quite challenging to compare to specific coastal buoys that present high temporal and spatial variability such as shown by the violin in Fig. S3.

The good evaluation of the reconstructed pCO_2 against both SOCAT_a, SOCAT_b as well as against buoys data gives confidence to identify the linear trends of both the pCO_2 and FCO_2 over different temporal scales.~~The overall good agreement in terms of biases and RMSEs lends confidence to our ability to unravel the long-term temporal coastal pCO_2 evolution with the newly reconstructed pCO_2 product .~~ For example, our results show that for all the 8 studied regions represented in Fig. 4, an increase in pCO_2 over time comprised between 12~~0~~ and 20 µatm decade^{-1} is calculated for the long-term trend with our pCO_2 product, a range in good agreement with the 12-18 µatm decade^{-1} obtained with SOCAT_b. Although New Zealand shows the largest bias between SOCAT_b and the pCO_2-product, they both agree that this region displays the fastest trend in terms of pCO_2 rise (18 µatm decade^{-1} and 20 µatm decade^{-1} for SOCAT_b and the pCO_2-product, respectively).'

Some other comments:

R1C2: Line 41, "The exchange of carbon dioxide (CO2) … depends on … (ΔpCO2)", Quantification of fluxes also relies on gas transfer velocity and the solubility of CO2. Suggestion: "… mainly depends on … on the global average".

R1R2: We agree with R1C2 and we modified the text accordingly.

R1C3: Line 180, "fCO2to pCO2 using the equation of Takahashi et al. (2012, page 6)".

I guess you are using this equation?

pCO2 = fCO2 x [1.00436 – 4.669 x 10-5 x SST (°C)]

It is acceptable to utilize this empirical equation since it will only result in small numerical differences compared to the results computed using CO2SYS. But it is hard to get the 2012 edition now, I recommend referencing a more recent version, such as the 2017 edition, as it will be easier for readers to access the specific equation.

R1R3: We indeed use the equation pCO_2= fCO2 x [1.00436 – 4.669 x 10-5 x SST (°C)] and we agree that a more recent reference is more suitable when referring to this equation. We therefore cite the Takahashi (2018) edition in the updated manuscript (Takahashi et al., 2019).

R1C4: Line 183: "In this study, the coastal domain (total surface) … ".

The 200m isobath has been commonly accepted as the shelf break criterion. An alternative approach is to consider a broader domain characterized by a distance to the coast of less than 200 nautical miles (approximately 370 km or 400 km). A 300 km definition for the wide shelf could also be employed. In the results section, the authors use the term "narrow coastal domain" without providing a distinct definition. I suggest explicitly clarifying this definition in the method section.

R1R4: We agree with R1C4 that a clarification regarding the definition of the two coastal domains used in this study is needed in the method section. We thus changed the text accordingly:

Lines 187-195: 'this study, the coastal domain ('wide coastal ocean' with a total surface area of 76 million km², Laruelle et al., 2017) excludes the Black Sea, estuaries as well as inland water bodies and its outer limit is defined as whichever point is furthest from the shoreline between the 1000 m isobath and a fixed 300 km distance (roughly the outer edge of territorial waters), following the coarse SOCAT definition of the coastal oceanic domain. At the end of this entire procedure, a total of ~ 14 million and ~ 4 million discrete coastal data have been allocated to SOCAT_a and SOCAT_b, respectively. A more common delineation of the coastal ocean is also used in this study when discussing the air-sea CO_2 exchange (Sect. 3.1) using the shelf break as the outer limit of the coastal domain ('narrow coastal ocean', 28 million km²). The depth of the shelf break is calculated using a high-resolution global bathymetric database and estimated by calculating the slope of the sea floor. The isobath for which the increase in slope is the maximum over the 0–1000m interval, yet still inferior to 2 % define the outer limit of the shelf break (Laruelle et al., 2013).'

R1C5: Line 187: "a total of ~ 14 million and ~ 4 million coastal data", discrete samples or grids?

R1R5: We agree with R1C5 that a clarification was needed and modified the text accordingly:

Lines 191-192: 'At the end of this entire procedure, a total of ~ 14 million and ~ 4 million discrete coastal data have been allocated to SOCAT_a and SOCAT_b, respectively.'

R1C6: Line 193: "global atmospheric reanalysis ERA-interim wind product (Dee et al., 2011)", ERA-interim is the 3-gen reanalysis data offered by ECMWF, encompassing the period from 1979 to 2019. Given that this work covers the years 1982 to 2020, I'm curious about how you obtained the wind speed data for the final year of your analysis. Furthermore, why not use the latest version ERA5?

R1R6: We thank the reviewer for this comment since it provides us the opportunity to correct an oversight in our manuscript because we indeed used ERA5 and not ERA-interim for the wind product. Unfortunately, the old references (here and in the next comment) were copied from the original Laruelle et al. (2017) approach, whereas for this study we actually updated the wind product for our calculation but forgot to update references in the text. We now updated the text and Table 1 by changing ERA-interim with ERA5 and refer to Hersbach et al. (2020) instead of Dee et al. (2011).

Lines 200-202: 'For U_{10}, we use the monthly mean of the 0.25° resolution product of the European Center for Medium-Range Weather Forecasts (ECMWF) ERA5 ~~global atmospheric reanalysis ERA-interim~~ wind product (~~Dee et al., 2011~~Hersbach et al., 2020), which has a native temporal resolution of 6 hours.'

R1C7: Line 198, "using the NCEP reanalysis total pressure at sea level (Kalnay et al., 1996).", While I believe that the choice of pressure data products has a minimal impact on pCO2air calculations, I'm curious why not use the newer NCEP2 version. Additionally, since this work employed the wind speed from ERA-Interim, why not use pressure data from ERA5 or ERA-Interim, which could offer enhanced spatial resolution to align with the wind speed data. Could you elaborate on the reasons behind these choices?

R1R7: We thank the reviewer for this comment since we again referred to the wrong product for atmospheric pCO_2. In this study, we use the atmospheric pCO_2 from the SeaFlux product (Fay et al., 2021) which is calculated from the dry air mixing ratio of atmospheric CO_2 (xCO_2) derived from the ESRL surface marine boundary layer CO_2 product (Duglokencky et al., 2019) multiplied by the ERA5 sea level pressure (Herbach et al., 2020) with a vapor correction according to Dickson et al. (2007).

We thus modified the text and now refer to ERA5.

Lines 204-210: 'The atmospheric pCO_2 is from the SeaFlux product (Fay et al., 2021) which is calculated from the dry air mixing ratio of CO_2 (xCO_2) provided by the ~~NOAA Marine Boundary Layer reference~~ESRL surface marine boundary layer CO_2 product (https://www.esrl.noaa.gov/gmd/ccgg/mbl/data.php, Duglockencky et al., 2019~~https://www.esrl.noaa.gov/gmd/ccgg/mbl/~~) with a vapor correction according to ~~assuming 100 % humidity at the air-sea interface (following calculation specified in~~ Dickson et al. ~~,(~~2007), ~~and~~ using the ~~NCEP reanalysis total pressure at sea level (Kalnay et al., 1996)~~ERA5 sea level pressure (Herbach et al., 2020) and applying the method of Landschützer et al. (2014) and Fay et al. (2021).'

R1C8: Line 199: "due to the proximity to the continent of the coastal ocean, the latter might be more exposed to anthropogenic sources of CO2", I can't understand what "the latter" means, "coastal ocean"?

R1R8: We agree with R1C8 that the text was not clear, and we modified it as follows.

Lines 210-213: 'It should be noted that, due to the proximity to the continent~~ of the coastal ocean~~, ~~-~~the coastal ocean ~~the latter~~ might be more exposed to anthropogenic sources of CO_2 and thus might be exposed to higher atmospheric pCO_2 compared with the global oceanic average.'

R1C9: Line 212-214: "by computing (1) … (2) …", The final interpolated values use those options in the order of rank (i.e., using an "if... elseif" while programming), or take the average of all options?

R1R9: We agree with R1C9 that the text was not clear and modified it as follows.

Lines 223-225: 'the interpolation is performed by computing in the order of rank (1) the mean of the next neighboring cells of the month x, (2) the mean value of the month x+1 and month x-1 of the same cell, (3) …'

R1C10: Line 231: "Ho et al., (2011)", I'm interested to know if there is a specific reason for not utilizing the Wanninkhof 2014 method.

R1R10: We decided to use the empirical equation of Ho et al. (2011) because this particular equation is calibrated using an extensive collection of data sampled at different locations and is thus applicable at global scale for the coastal domain. Furthermore, we would like to note that both Ho et al. (2011) and Wanninkhof et al. (2014) are very similar and thus lead to the same results (due to the use of the same gas exchange coefficient). We nevertheless acknowledge that over the last decades many appropriate gas transfer formulations have been developed and published and that the choice remains subjective.

R1C11: Line 241, Equation 2. I have this question as well while reading Robaert et al., 2019. Notably, the uncertainty of CO2 solubility appears to be absent from the equation. While it may be plausible to argue that this uncertainty is minimal (0.2%, as suggested by Weiss, 1974), it remains essential to keep this term in the equation.

R1R11: We agree with R1C11 that all potential sources of uncertainties on the FCO_2 calculation should be discussed in the MS. Since we consider that the uncertainty of the solubility term (K_0) is very negligible such as suggested by Weiss (1974) compared to those from the wind speed product, sea-ice or k choice we didn't evaluate the influence of using different combination of SST and SSS products on the K_0 calculation and thus on the FCO_2 calculation. We however do not want to create the impression that we omit uncertainty terms, thus we added a sentence to discuss the uncertainty of the K_0 term:

Lines 259-264: 'We also include the effect of the choice of sea-ice product on the FCO_2 uncertainty (σ_{ice}) which was not included in the original calculations of Landschützer et al. (2014, 2018) but has been identified as a potential source of uncertainty in global coastal reconstructions (e.g., Resplandy et al., in rev.). All of these four sources of uncertainty are expressed in Pg C yr^{-1}. We do not include the uncertainty associated with the solubility term (K_0) in our uncertainty assessment since this contribution is minimal as suggested by Weiss et al. (1974, 0.2%).'

R1C12: Line 249-252, "σ_{wind} is calculated …", The wind speed data products cover different time periods, and please specify the time period for which the standard deviation is calculated.

R1R12: We agree with R1C12 and modified the text accordingly:

Lines 262-268: 'All of these four sources of uncertainty are expressed in Pg C yr^{-1}. We do not include the uncertainty associated with the solubility term (K_0) in our uncertainty~~ies~~ ~~calculation~~assessment since this ~~latter~~contribution is minimal as suggested by Weiss et al. (1974, 0.2%). σ_{wind} is calculated following the strategy described in Roobaert et al. (2018) which consists of using the standard deviation of global FCO_2 fields calculated with 3 different wind products: the ERA~~-interim~~5 (~~Dee et al., 2011~~Hersbach et al., 2020), the Cross-Calibrated Multi-Platform Ocean Wind Vector 3.0 (Atlas et al., 2011) and the NCEP/NCAR reanalysis 1 (Kalnay et al., 1996). Since these wind products cover different time periods, σ_{wind} is calculated for the overlap period (1991-2011) between products.'

R1C13: Line 275: "θmap is calculated as the RMSE …" RMSE of the training set (SOCAT_a vs. predicted) or independent validation set (SOCAT_b vs. predicted)?

R1R13: The RMSE is calculated against the training dataset. We modified the text according:

Lines 291-292: 'θ_{map} is calculated as the root mean squared deviation between the reconstructed pCO_2 and the gridded pCO_2 observation from the training dataset (SOCAT_a).'

R1C14: Line 297: "and a r^2 of 0.7 are calculated", We all know that R2 is the coefficient of determination, but please use the full name the first time it is mentioned in the manuscript.

R1R14: We agree with R1C14, and we modified the text accordingly.

R1C15: Line 338: "This dataset consists of a pool of 404,206 gridded cells that are uniformly distributed between both hemispheres (SOCAT_b, Fig. S1) and presents a good correspondence with SOCAT_a (93 % of the residuals between SOCAT_b and SOCAT_a are < 5 µatm and with a global RMSE value of 6 µatm, Fig. S2)." I tried to comprehend this sentence, but I failed. In my understanding, SOCAT_a and SOCAT_b are independent, they are non-repetitive random samples of the original data set; thus, I can't understand why they would "correspond." And what are the residuals between SOCAT_b and SOCAT_a mean?

R1R15: We apologize for the confusion. The reviewer is correct that both datasets are independent. What we tried to say here was that observations are homogeneously distributed, which is an essential criterion for the network training. Furthermore, we also noticed that - for both independent datasets - reconstruction errors and RMSE are very similar, thus our reconstruction performs equally well and is not biased towards one dataset, which would be a sign of overfitting.

For clarity, we now rephrased this to 'This dataset consists of a pool of 404,206 gridded cells that are uniformly distributed between both hemispheres (SOCAT_b, Fig. S1)~~.~~, an essential criterion for training the network ~~and presents a good correspondence with SOCAT_a (93 % of the residuals between SOCAT_b and SOCAT_a are < 5 µatm and with a global RMSE value of 6 µatm, Fig. S2)~~.'

R1C16: Line 381: The title of subsection "Spatial and seasonal dynamics", I think it is more appropriate to use "variations" than "dynamics" in the subtitle, as the discussion in this section pertains specifically to variations rather than dynamics.

R1R16: We modified the text accordingly.

R1C17: Line 388-389: "Using the shelf break as the outer limit of the coastal domain ('narrow coastal ocean', 28 million km²)", This marks the initial instance of employing the term "narrow coastal ocean" in the manuscript. Please clarify how you define the "shelf break", 200m?

R1R17: The 'narrow coastal domain' used in this study is based on the coastal definition of Laruelle et al. (2013). The outer limit is defined as the shelf break which is estimated by calculating the slope of the sea floor. The isobath for which the increase in slope is the maximum over the 0–1000 m interval, yet still inferior to 2 % is defined as the outer limit of the shelf by Laruelle et al. We clarified the definition of the 'narrow coastal domain' as suggested by R1C17 but also R1C4 in the method section of this study.

Lines 187-196: 'this study, the coastal domain ('wide coastal ocean' with a total surface area of 76 million km², Laruelle et al., 2017) excludes the Black Sea, estuaries as well as inland water bodies and its outer limit is defined as whichever point is furthest from the shoreline between the 1000 m isobath and a fixed 300 km distance (roughly the outer edge of territorial waters), following the coarse SOCAT definition of the coastal oceanic domain. At the end of this entire procedure, a total of ~ 14 million and ~ 4 million discrete coastal data have been allocated to SOCAT_a and SOCAT_b, respectively. A more common delineation of the coastal ocean is also used in this study when discussing the air-sea CO_2 exchange (Sect. 3.1) using the shelf break as the outer limit of the coastal domain ('narrow coastal ocean', 28 million km²). The depth of the shelf break is calculated using a high-resolution global bathymetric database and estimated by calculating the slope of the sea floor. The isobath for which the increase in slope is the maximum over the 0–1000m interval, yet still inferior to 2 % define the outer limit of the shelf break (Laruelle et al., 2013).'

R1C18: Line 387-393: The digits numbers are not consistent in the flux section. For example, "0.4 Pg C per year (with an uncertainty of ± 0.03 Pg C yr-1)", or "-0.2 ± 0.01 Pg C yr -1". I suggested using two digits number to keep consistent.

R1R18: We now use two digits numbers to keep the consistency of our formulation throughout the manuscript (the updated values are 0.40 Pg C yr^{-1} and 0.18 Pg C yr^{-1}, instead of 0.3 Pg C yr^{-1} and 0.2 Pg C yr^{-1} respectively).

R1C19: Line 440: "(0.1 Pg Cyr -1decade -1and 0.03 Pg C yr -1decade -1", What are the uncertainties and p-values? (i.e., Δa in y = (a ± Δa)x + b)

R1R19: Following R1C19, uncertainties have been added to the estimations. Our updated coastal trend of -0.06 Pg C yr^{-1} decade^{-1} has a p-value < 0.05 on a Mann-Kendall statistical test and an uncertainty (Δa) of 0.0009 Pg C yr^{-1} decade^{-1}. A p-value < 0.05 and an uncertainty (Δa) 0.0005 Pg C yr^{-1} decade^{-1} are calculated for our updated trend shelf values of -0.02 Pg C yr^{-1} decade^{-1}, respectively. These numbers have been included in the Ms.

Lines 479-484: 'The overall intensification of the coastal sink that we observe in this study (0.06~~1~~ Pg Cyr^{-1} decade^{-1} (± 0.0009 Pg C yr^{-1} decade^{-1} with a p-value < 0.05) and 0.02~~3~~ Pg C yr^{-1} decade^{-1} (± 0.0005 Pg C yr^{-1} decade^{-1} with a p-value < 0.05) for the wide and narrow coastal domain, respectively) supports the only two available observational coastal studies performed at the global scale (i.e., Laruelle et al., 2018; Wang et al, 2017) which were however significantly limited by the small fraction of the coastal ocean domain investigated (e.g., 6 % in Laruelle et al., 2018) and both predict an increase in efficiency of the global coastal CO_2 sink over the last three decades.'

R1C20: Line 455: "increase of the global coastal CO2 sink is to be found in the high latitudes of the northern hemisphere,", According to the product, yes. But there are still uncertainties for this conclusion, as the largest errors are observed in the high latitudes of the northern hemisphere (Figure 2c), notably within the Arctic Ocean.

R1R20: We agree with R1C20 and added a sentence to discuss this particular point.

Lines 489-504: 'The significant strengthening of this global coastal sink that we observe in this study, which has approximately doubled between 1982 and 2020 (wide coastal domain) results from a general tendency towards an increase of the coastal CO_2 sinks intensities (e.g., in the high latitude of the Northern Hemisphere, Fig. 6a) combined with decreases in intensity of several CO_2 sources such as along upwelling currents (e.g., in the Arabian sea). However, since a large fraction of the global CO_2 uptake results from coastal regions > 40° of the Northern Hemisphere and since these CO_2 sinks regions present strong negative rates of change in FCO_2 (Fig. 6a), our result suggests that the primary driver of this twofold increase of the global coastal CO_2 sink is to be found in the high latitudes of the northern hemisphere, which contribute disproportionately to the global-scale coastal FCO_2 trend. Further studies should however be carried out to support this conclusion, given the paucity of observational pCO_2 data in those high latitude regions that translate into high uncertainties in our pCO_2 product, as for example in the Arctic Ocean. ~~However, t~~Taking 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. Moreover, changes in wind speed and sea-ice coverage have likely not been constant over time and further analysis of their influence on the rate change of FCO_2 should be analyzed for each decade individually to better understand the interplay between these different drivers. Overall, our results highlight the complex nature of the coastal FCO_2 dynamics and emphasize~~s~~ the need for further investigation and understanding of the specific factors influencing the FCO_2 trends in different coastal regions.'

R1C21: Line 507: "latest release of the SOCAT", Please include the version information, such as "the latest release of SOCAT (v2022)," to provide further context.

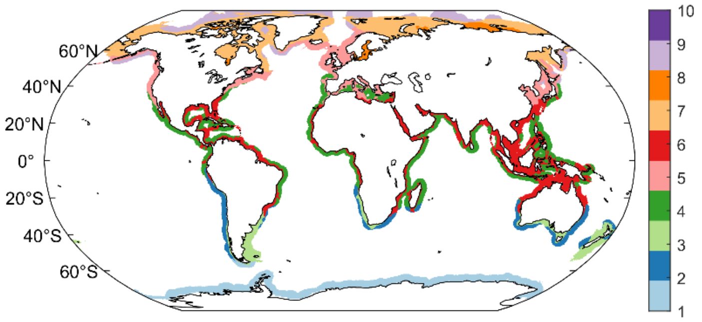
R1R21: We modified the text accordingly:

Lines 549-552: 'In this study, we presented a new coastal pCO_2-product for the 1982-2020 period using ~ 18 million direct coastal observations from the SOCATv2022 database (Bakker et al., 2022) ~~latest release of the SOCAT database~~ combined with an updated version of the coastal 2-step SOM and FFN method used by Laruelle et al. (2017).'

Figures and Tables:

R1C22: Figure 1. The color scale is hard for me to recognize, especially the P9 and P10. The font size of the latitude and width of the color bar can be smaller (like in Figure 2).

R1R22: Figure 1 has been updated (see Fig_R1C22 below) with suggestions from R1C22.



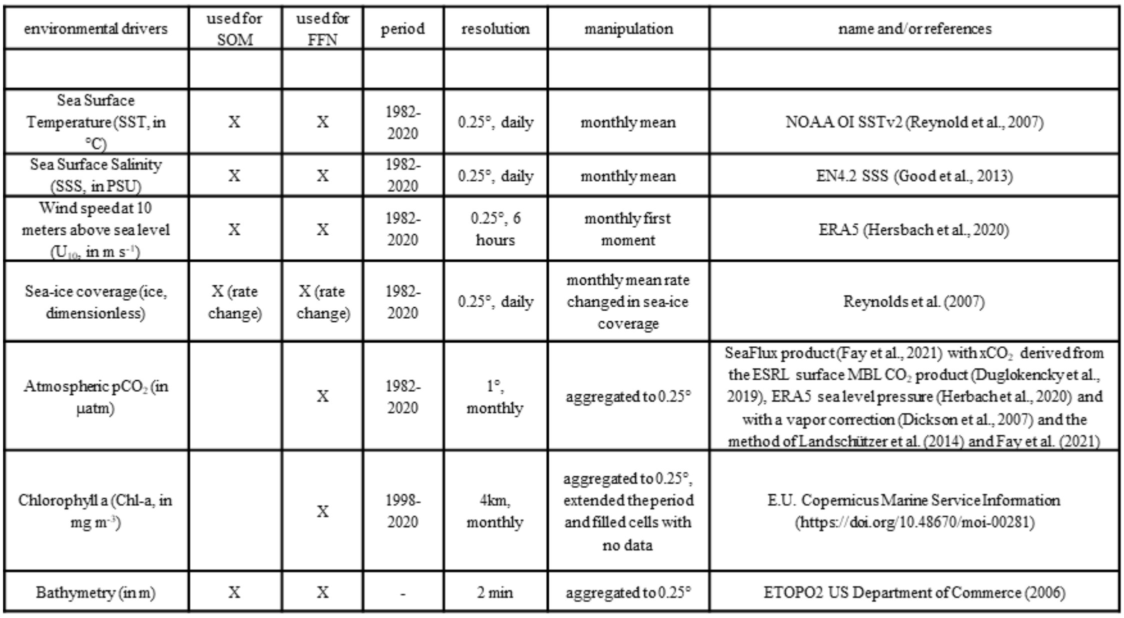
Fig_R1C22. New Fig. 1.

R1C23: Table 2. Row 4, column 1: Sea-ice coverage (ice, no unit) ◊ Sea-ice coverage (dimensionless) Row 5, column 5: MBL is using the zonal average or meridional average xCO2air? MBL doesn't provide 3-d gridded data; the resolution is 0.05 sin latitude. Also, because the SOM and FNN steps use different input variables, I suggest adding one more column or other ways to mention which variables are used in which step.

R1R23: Table 1 has been updated with suggestions from R1C23 (see Table_R1C23). The referee is correct that the MBL product does not provide 3D data. To derive the 3D atmospheric CO_2 within SeaFlux (Fay et al., 2021), the authors interpolate the data first from the original 0.05 sin latitude to 1° latitude and then use the same latitude value for each degree longitude to get the final 3D gridded atmospheric xCO_2 dry molar fraction. Multiplying this 3D field with the 3D ERA pressure fields and applying the Dickson et al. (2007) water vapor correction, the final pCO_2_wet can be calculated. We nevertheless believe the original reference to the MBL product should remain, however, we further added: '… applying the method of Landschützer et al. (2014) and Fay et al. (2021)'.

Lines 204-210: 'The atmospheric pCO_2 is from the SeaFlux product (Fay et al., 2021) which is calculated from the dry air mixing ratio of CO_2 (xCO_2) provided by the ~~NOAA Marine Boundary Layer reference~~ESRL surface marine boundary layer CO_2 product (https://www.esrl.noaa.gov/gmd/ccgg/mbl/data.php, Duglockencky et al., 2019~~https://www.esrl.noaa.gov/gmd/ccgg/mbl/~~) with a vapor correction according to ~~assuming 100 % humidity at the air-sea interface (following calculation specified in~~ Dickson et al. ,~~(~~2007), ~~and~~ using the ~~NCEP reanalysis total pressure at sea level (Kalnay et al., 1996)~~ERA5 sea level pressure (Herbach et al., 2020) and applying the method of Landschützer et al. (2014) and Fay et al. (2021).'

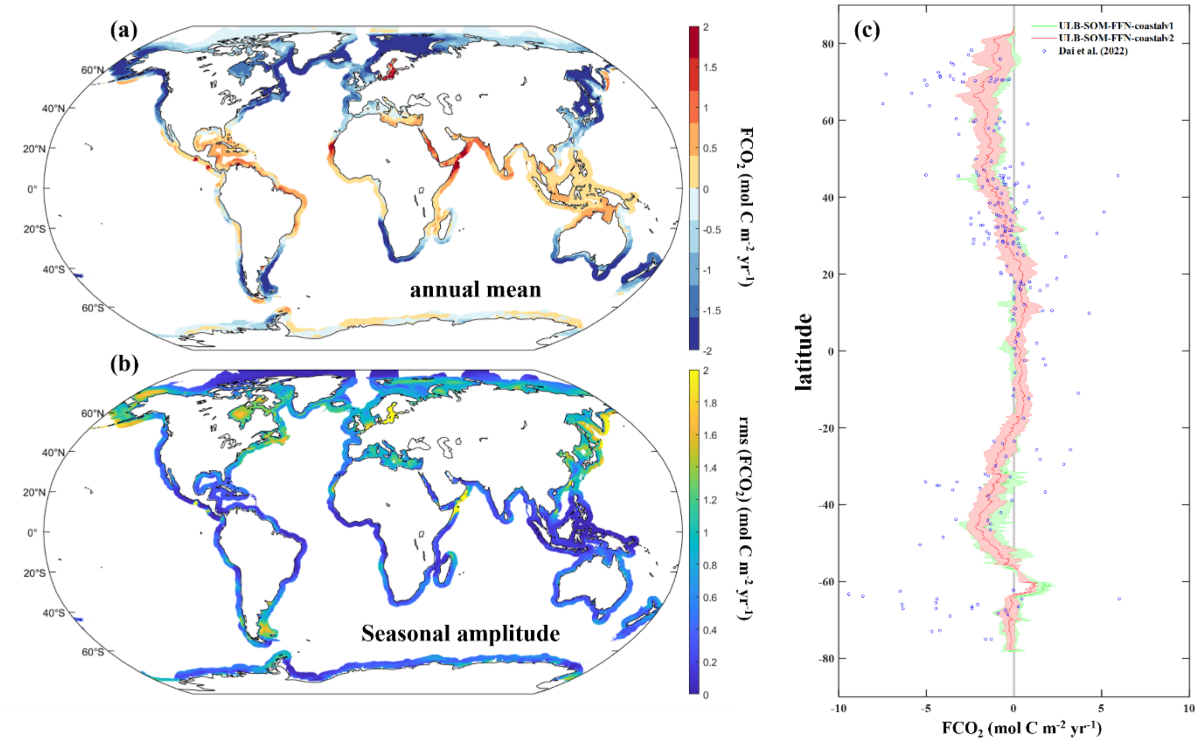
Table_R1C23. New Table 1.



environmental drivers	used for SOM	used for FFN	period	resolution	manipulation	name and/or references
Sea Surface Temperature (SST, in °C)	X	X	1982-2020	0.25°, daily	monthly mean	NOAA OI SSTv2 (Reynold et al., 2007)
Sea Surface Salinity (SSS, in PSU)	X	X	1982-2020	0.25°, daily	monthly mean	EN4.2 SSS (Good et al., 2013)
Wind speed at 10 meters above sea level (U_{10}, in m s^{-1})	X	X	1982-2020	0.25°, 6 hours	monthly first moment	ERA5 (Hersbach et al., 2020)
Sea-ice coverage (ice, dimensionless)	X (rate change)	X (rate change)	1982-2020	0.25°, daily	monthly mean rate changed in sea-ice coverage	Reynolds et al. (2007)
Atmospheric pCO_2 (in µatm)		X	1982-2020	1°, monthly	aggregated to 0.25°	SeaFlux product (Fay et al., 2021) with xCO_2 derived from the ESRL surface MBL CO_2 product (Duglokencky et al., 2019), ERA5 sea level pressure (Herbach et al., 2020) and with a vapor correction (Dickson et al., 2007) and the method of Landschützer et al. (2014) and Fay et al. (2021)
Chlorophyll a (Chl-a, in mg m^{-3})		X	1998-2020	4km, monthly	aggregated to 0.25°, extended the period and filled cells with no data	E.U. Copernicus Marine Service Information (https://doi.org/10.48670/moi-00281)
Bathymetry (in m)	X	X	-	2 min	aggregated to 0.25°	ETOPO2 US Department of Commerce (2006)

R1C24: Figure 5. Please adjust the colorbars and latitude font size in (a) and (b) like Figure 2 or Figure 4(a) to provide more information.

R1R24: Figure 5 has been updated following the reviewer's suggestion (see Fig_R1C24).



Fig_R1C24. Updated Fig. 5

Some grammatical mistakes:

R1C25: Line 43: “the number of … have considerably”, should be “has” I think

R1R25: We agree with R1C25 and modified the text accordingly.

R1C26: Line 104: benchmark should be plural

R1R26: We agree with R1C26 and modified the text accordingly.

R1C27: Line 124: “in a first step”, should be “in the first step”

R1R27: We agree with R1C27 and modified the text accordingly.

R1C28: Line 143: “P9 represent”, should be “represents”

R1R28: We agree with R1C28 and modified the text accordingly.

R1C29: Line 367: “Largest RMSE” should be “The largest RMSE”

R1R29: We agree with R1C29 and modified the text accordingly.

R1C30: Line 376: “CO2 exchanges” should be “CO2 exchange”

R1R30: We agree with R1C30 and modified the text accordingly.

R1C31: Line 383: “both hemisphere” should be “hemispheres”, “a CO2 sinks” should be “CO2 sinks”

R1R31: We agree with R1C31 and modified the text accordingly.

R1C32: Line 426: typo, “through”

R1R32: We agree with R1C32 and modified the text accordingly.

R1C33: Line 441: “at global scale” should be “at a/the global scale”

R1R33: We agree with R1C33 and modified the text accordingly.

R1C34: Line 461: “our results … and emphasizes” should be “emphasize”

R1R34: We agree with R1C34 and modified the text accordingly.

R1C35: Line 510: “However, these investigation ..” should be “investigations”

R1R35: We agree with R1C35 and modified the text accordingly.

R1C36: Line 532: “to calculated” should be “to calculate”

R1R36: We agree with R1C36 and modified the text accordingly.

REFERENCES

Dickson, A. G., Sabine, C. L., & Christian, J. R. (2007). Guide to best practices for ocean CO2 measurements. PICES Special Publication 3, 3(8), 191. https://doi.org/10.1159/000331784

Dlugokencky, E. J., Thoning, K. W., Lang, P. M., and Tans, P. P.: NOAA Greenhouse Gas Reference from Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle CooperativeGlobal Air Sampling Network, 2019 (data available at: https://www.esrl.noaa.gov/gmd/ccgg/mbl/data.php).

Fay, A. R., Gregor, L., Landschützer, P., McKinley, G. A., Gruber, N., Gehlen, M., Iida, Y., Laruelle, G. G., Rödenbeck, C., Roobaert, A., & Zeng, J. (2021). SeaFlux: Harmonization of air-sea CO2 fluxes from surface pCO2 data products using a standardized approach. In Earth System Science Data (Vol. 13, Issue 10, pp. 4693–4710). https://doi.org/10.5194/essd-13-4693-2021

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc. 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020 (data available at:https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthlymeans?tab=overview)

Ho, D. T., Wanninkhof, R., Schlosser, P., Ullman, D. S., Hebert, D., & Sullivan, K. F. (2011). Toward a universal relationship between wind speed and gas exchange: Gas transfer velocities measured with 3He/SF6 during the Southern Ocean Gas Exchange Experiment. Journal of Geophysical Research: Oceans, 116(7), C00F04. https://doi.org/10.1029/2010JC006854

Landschützer, P., Gruber, N., Bakker, D. C. E., & Schuster, U. (2014). Recent variability of the global ocean carbon sink. Global Biogeochemical Cycles, 28, 927–949. https://doi.org/10.1002/2014GB004853

Landschützer, P., Gruber, N., & Bakker, D. C. E. (2016). Decadal variations and trends of the global ocean carbon sink. Global Biogeochemical Cycles, 30, 1396–1417. https://doi.org/10.1002/2015GB005359

Laruelle, G. G., Dürr, H. H., Lauerwald, R., Hartmann, J., Slomp, C. P., Goossens, N., and Regnier, P. A. G. (2013): Global multi-scale segmentation of continental and coastal waters from the watersheds to the continental margins, Hydrol. Earth Syst. Sci., 17, 2029–2051, https://doi.org/10.5194/hess-17-2029-2013.

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

Sutton, et al., 2019. Autonomous seawater pCO2 and pH time series from 40 surface buoys and the emergence of anthropogenic trends. Earth Syst. Sci. Data 11, 421–439. https://doi.org/10.5194/essd-11-421-2019

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.

Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited. Limnology and Oceanography: Methods, 12(C5), 351–362. https://doi.org/10.1029/92JC00188

Weiss, R. F. (1974). Carbon dioxide in water and seawater: the solubility of a non-ideal gas. Marine Chemistry, 2(3), 203–215. https://doi.org/10.1016/0304-4203(74)90015-2

Reviewer #2 Evaluations:

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, t~~Taking 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 variabilities~~y~~ 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 ~~decadal~~long-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 of~~Rate 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 change~~long-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.