A novel sea surface pCO_2-product for the global coastal ocean resolving trends over the 1982-2020 period

Alizée Roobaert[1,2], Pierre Regnier[1], Peter Landschützer[2] and Goulven G. Laruelle[1]

[1]Department of Geosciences, Environment & Society-BGEOSYS, Université Libre de Bruxelles, Brussels, CP160/02, Belgium

[2]Flanders Marine Institute (VLIZ), Ostend, Belgium

Correspondence to: Alizée Roobaert (alizee.roobaert@ulb.be)

Abstract

In recent years, advancements in machine learning based interpolation methods have enabled the production of high-resolution maps of sea surface partial pressure of CO_2 (pCO_2) derived from observations extracted from databases such as the Surface Ocean CO_2 Atlas (SOCAT). These pCO_2-products now allow quantifying the oceanic air-sea CO_2 exchange based on observations. However, most of them do not yet explicitly include the coastal ocean. Instead, they simply extend the open ocean values onto the nearshore shallow waters, or their spatial resolution is simply so coarse that they do not accurately capture the highly heterogeneous spatiotemporal pCO_2 dynamics of coastal zones. Until today, only one global pCO_2-product was specifically designed for the coastal ocean (Laruelle et al., 2017). 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. Here we aim at closing this knowledge gap and update the coastal product of Laruelle et al. (2017) to investigate the longest global monthly time series available for the coastal ocean from 1982 to 2020. The method remains based on a 2-step Self Organizing Maps and Feed Forward Network method adapted for coastal regions, but we include additional environmental predictors and use a larger pool of training and validation data with ~ 18 million direct observations extracted from the latest release of the SOCAT database. Our study reveals that the coastal ocean has been acting as an atmospheric CO_2 sink of -0.40 Pg C yr^{-1} (-0.182 Pg C yr^{-1} with a narrower coastal domain) on average since 1982, and the intensity of this sink has increased at a rate of 0.061 Pg C yr^{-1} decade^{-1} (0.023 Pg C yr^{-1} decade^{-1} with a narrower coastal domain) over time. 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

30   play an important role in some regions, particularly at high latitudes. This new reconstructed coastal pCO_2-product (Roobaert et al., 2023, https://www.ncei.noaa.gov/archive/accession/0279118) allows establishing regional carbon budgets requiring high-resolution coastal flux estimates and provides new constraints for closing the global carbon cycle.

Short summary

The quantification of the coastal air-sea CO_2 exchange (FCO_2) has improved in recent years, but its multiannual variability remains unclear. This study, based on interpolated observations, 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.

40   **1 Introduction**

The exchange of carbon dioxide (CO_2) between the atmosphere and the ocean mainly depends on the gradient between the partial pressure of CO_2 (pCO_2) at the surface of the ocean and that of the overlying air on the global average. Over the past decade, the number of high-quality measurements of sea surface pCO_2 collected by research field programs and ships of opportunities hasve considerably increased. Moreover, large-scale community efforts have led to the compilation of tens of

45   millions of sea surface pCO_2 measurements into uniform quality-controlled databases such as SOCAT (for Surface Ocean CO_2 Atlas, Bakker et al., 2014), allowing for the quantification of the global oceanic CO_2 sink. However, in spite of this tremendous increase in data coverage, once gridded monthly at a typical spatial resolution of $1°$ for the open ocean and $0.25°$ for the coastal ocean, pCO_2 measurements remain largely discontinuous in time and space. The remaining regions and periods of time devoid of data thus prevent one from fully quantifying the air-sea CO_2 exchange and its full spatiotemporal variability based on

50   measurements alone.

Therefore, in parallel to the on-going measurement synthesis efforts, another research branch aiming at developing robust interpolation techniques to circumvent the spatial and temporal gaps in the data products has emerged. These techniques allow creating maps of pCO_2 that are continuous in space and time, typically at the monthly resolution (e.g., Chau et al., 2022; Gloege

55   et al., 2022; Gregor & Gruber, 2021; Landschützer et al., 2014; Rödenbeck et al., 2014, 2015). The resulting observation-based continuous products (called hereafter 'pCO$_2$-products') however differ in their spatial resolutions (e.g., from $0.25°$x $0.25°$ in Chau et al. (2023) over $1°$ x $1°$ in Landschützer et al., 2014 to $4°$ x $5°$ in Majkut et al., 2014), their temporal coverage and their method of interpolation. Several studies have relied on direct interpolations of available pCO_2 measurements (e.g., Jones et al., 2015; Rödenbeck et al., 2014; Shutler et al., 2016) while others have first established linear (e.g., Iida et al., 2015; Park et

60   al., 2010; Schuster et al., 2013) or nonlinear (e.g., Landschützer et al., 2014; Nakaoka et al., 2013; Zeng et al., 2014) predictive regression equations between a set of environment parameters (available everywhere and at everytime within the domain of

interest) and observed pCO_2 to perform the spatiotemporal extrapolation. These complementary pCO_2-products provide a better quantification of the spatial and temporal variability of the global oceanic CO_2 sink and its associated uncertainty on different time scales, going from seasonal fluctuations to decadal trends through interannual variability, while providing much improved observation-based benchmarks against which outputs from global model results can be evaluated (e.g., Hauck et al., 2020).

While significant efforts have been invested by the community to develop pCO_2-products for the global ocean, leading to a growing number of assessments of the CO_2 sink, most of these pCO_2-products ignore the coastal ocean (e.g., Landschützer et al., 2014) or resolve it by simply combining the coast with the open ocean (Chau et al., 2023). Indeed, the spatiotemporal investigations are performed for the entire ocean using the full set of observed pCO_2 data (coast and open ocean) in such a way that the specific conditions characteristic of coastal settings are not accurately accounted for in these products (e.g., Chau et al., 2022; Rödenbeck et al., 2013). In response to this shortcoming, other continuous pCO_2-products have been developed at the regional scale for several well monitored coastal seas (e.g., Bai et al., 2015; Hales et al., 2012; Jamet et al., 2007; Ono et al., 2004; Sarma et al., 2006) such as the California Current system (Sharp et al., 2022), European shelves (Becker et al., 2021) or the West Florida shelf (Chen et al., 2016). At the global scale, a significant step forward was made by Laruelle et al. in 2017 when the first global coastal pCO_2-product at high spatial resolution (0.25°) was released for the entire coastal domain. This product, which is still, to date, the only one available specifically developed for the global coastal ocean, is based on gridded coastal pCO_2 observations and nonlinear predictive regression equations between a set of environmental variables (drivers) and observed pCO_2 to perform the spatiotemporal extrapolation (the Self Organizing Map and Feed Forward Network coastal pCO_2-product, ULB-SOM-FFN-coastalv1, Laruelle et al., 2017). This global coastal pCO_2-product provided a climatological mean (period 1998-2015) which allowed unprecedented investigation of the spatial distribution of the CO_2 sources and sinks in the global coastal ocean, especially for regions lacking data or regional assessments. It also allowed resolving the seasonal variability of the air-sea CO_2 exchange in the coastal domain (Roobaert et al., 2019). Moreover, it was recently merged with an open ocean product to obtain a global reconstruction of the ocean CO_2 sink (Landschützer et al., 2020) and has been subsequently used to reduce the spread in global reconstructions (Fay et al., 2021). However, the ULB-SOM-FFN-coastalv1 pCO_2-product remains limited in its applications because it only provides a climatology covering a relatively short period (1998-2015) and is thus not suitable to evaluate the interannual variability or the long-term trends of the coastal air-sea CO_2 exchange. Such questions currently are at the forefront of the coastal research community's preoccupations (Bauer et al., 2013; Lacroix et al., 2021a; Laruelle et al., 2018; Regnier et al., 2013; ~~Resplandy et al., in rev.~~Resplandy et al., 2023; Wang et al., 2017) but, because of the lack of adequate product, our confidence in the extent to which humans have perturbed the coastal air-sea CO_2 exchange since pre-industrial times remains low (Regnier et al., 2022). Moreover, the limitations of the ULB-SOM-FFN-coastalv1 do not yet allow producing robust trends in coastal pCO_2 fields against which global model outputs can be evaluated (e.g., ~~Resplandy et al., in rev.~~Resplandy et al., 2023).

To address these limitations, this study expands and improves upon the version of the global coastal pCO$_2$-product of Laruelle et al. (2017) by extending its temporal coverage to four decades (1982-2020) and updating the methodology to resolve longer-terms changes in pCO$_2$, as described in the following section. The evaluation of this new product (ULB-SOM-FFN-coastalv2, Roobaert et al., 2023, https://www.ncei.noaa.gov/archive/accession/0279118) is done both spatially and for each decade individually, which represents an improvement compared to Laruelle et al. (2017) where the ULB-SOM-FFN-coastalv1 evaluation was limited to spatial and climatological seasonal cycles only. 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. In the future, these updated pCO$_2$ and FCO$_2$ products can be used as benchmarks for global oceanic models resolving trends in the coastal CO$_2$ dynamics, fulfilling a key knowledge gap identified in the latest Regional Carbon Cycle Assessment and Processes coastal synthesis (RECCAP2, Resplandy et al., in rev.Resplandy et al., 2023).

2. Methods

This section first describes the 2-step interpolation method used to generate the new version of the coastal pCO$_2$-product (Sect. 2.1) and the different datasets involved in this two steps procedure (Sect. 2.2). We then describe how the coastal air-sea CO$_2$ exchange is calculated (Sect. 2.3) and finally explain the approach used to quantify the uncertainties associated with our new pCO$_2$ and FCO$_2$ products (Sect. 2.4).

2.1 Self Organizing Maps and Feed Forward Network

We build upon the method described in Laruelle et al. (2017) to construct an updated observation-based continuous monthly pCO$_2$-product for the coastal ocean (ULB-SOM-FFN-coastalv2) at a 0.25° spatial resolution over the 1982 to 2020 period. The method is based on the application of two artificial neural networks (the Self Organizing Map, SOM and the Feed Forward Network, FFN). The SOM first clusters the global coastal ocean into provinces characterized by similar environmental properties. In each province, the FFN then establishes nonlinear relationships between the observed pCO$_2$ and a set of environmental drivers of the coastal pCO$_2$ dynamics (which may be different from those used by the SOM). These relationships are then used to perform the spatiotemporal extrapolation of pCO$_2$ in each region defined by the SOM. This method was originally developed for the open ocean and is extensively described in Landschützer et al. (2013, 2014). It was later adapted for the global coastal ocean by Laruelle et al. (2017). We thus provide only a brief description of the methodology and focus here on the modifications introduced in this study.

In a the first step, the global coastal ocean is divided into 10 biogeochemical provinces using the SOM clustering algorithm. Each resulting province is characterized by similar spatiotemporal patterns of a set of environmental variables, or drivers. In this study, we use the same drivers as in Laruelle et al. (2017), which consist of the wind speed calculated at 10 meters above the sea surface (U_{10}), the Sea Surface Temperature (SST), the Sea Surface Salinity (SSS), the bathymetry, and the rate of change of the sea ice coverage (see Sect. 2.2 for a description of the datasets). The SOM uses a neural network to detect similarities within multivariate datasets and uses an iterative procedure to distribute them into a predefined number of clusters. For each environmental driver, continuous monthly maps at the spatial resolution of 0.25° are used as inputs for the neural network and each 0.25° cell is allocated to one of the 10 provinces defined by the SOM (or neurons). This procedure aims at minimizing the Euclidean distance between all points within each neuron of the network (see Landschützer et al., 2013 for more details). The spatial extension of these provinces varies from one month to the other because of the seasonal variations of the environmental drivers in such a way that a fixed grid cell in space may be assigned to several provinces over the course of a year. The choice of 10 provinces in the SOM stems from a sensitivity analysis that minimizes the average deviation between the observed pCO_2 and those simulated by the FFN algorithm (see second step below) while ensuring the presence of a minimum number of grid cells (>100) that can be used for the validation in each province (Laruelle et al., 2017). While their spatial extent varies seasonally, each province remains associated with specific regions over the course of the entire 1982-2020 period and the province occurring most often in each grid cell is shown in Fig. 1. Broadly, these provinces represent: Province 1 (P1) the Antarctic shelf, P2-P3 two subpolar/temperate coastal provinces of the Southern Hemisphere, P4 and P6 the large tropical coastal provinces, P5 a temperate province of the Northern Hemisphere which includes the Mediterranean Sea and the Norwegian Sea. P7, P8 and P10 represent high latitudes of the Northern Hemisphere provinces that are seasonally partly covered by sea-ice (with the Baltic Sea and the Hudson Bay in P8) while P9 represents a permanent and cold polar province.

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

160 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.

The predictors used for the FFN are U$_{10}$, SST, SSS, the atmospheric pCO$_2$, the rate of change in sea ice coverage (except in regions not covered by sea ice, i.e., in P2, P3, P4 and P6), the bathymetry and the chlorophyll-a concentration (Chl-a). The

165 Chl-a is expressed as log$_{10}$(Chl-a) to minimize the influence of its skewed distribution (Wrobel-Niedzwiecka et al., 2022)~~.~~. In P1, P8 and P9, we do not use Chl-a as a driver because of the poor data coverage resulting from recurring cloud and/or sea-ice coverage in those provinces (see Sect. 2.2). This incomplete data coverage for Chl-a is incidentally the reason why this predictor is not used at the SOM stage because it requires complete global datasets. Atmospheric pCO$_2$, which was not included in the ULB-SOM-FFN-coastalv1, is also used as a driver of multi-decadal changes induced by the increasing atmospheric

170 pCO$_2$ concentration. Finally, we smooth spatially the monthly-resolved coastal pCO$_2$ field generated by the FFN using a moving 3 by 3 pixel window to remove abrupt pCO$_2$ transitions sometimes occurring at the boundaries between provinces. This smoothing procedure is described by Landschützer et al. (2014) and was also used in the ULB-SOM-FFN-coastalv1.

The surface pCO$_2$ data ~~used by the FFN~~ are extracted from the SOCATv2022 database (Bakker et al., 2022) that originally

175 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 pCO$_2$ (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

180 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

185 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. In 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

190 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).

2.2 Environmental variables

The observational SST and SSS fields used as inputs for the SOM-FFN algorithm are calculated as the monthly means of the daily NOAA OI SST V2 (Reynolds et al., 2007) and of the daily Hadley center EN4 SSS (Good et al., 2013), respectively (Table 1). 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., 2011Hersbach et al., 2020), which has a native temporal resolution of 6 hours. The monthly mean of the daily 0.25° dataset of Reynolds et al. (2007) is used for the sea-ice coverage. The rate of change in the sea-ice coverage for a given month x is then calculated as the difference between the sea-ice coverages of months x+1 and x-1. 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 referenceESRL surface marine boundary layer CO_2 product (https://www.esrl.noaa.gov/gmd/ccgg/mbl/data.php, Duglockencky et al., 2019https://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). 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. The use of spatially resolved dry air mixing ratio of CO_2 datasets such as the one from the NASA's Orbiting Carbon Observatory 2 Goddard Earth Observing System (OCO-2 GEOS, Eldering et al., 2017) instead of the product used in this study might be more appropriate to include this effect. However, OCO-2 GEOS only covers the 2015-2022 period, which is too short for the purpose of our study. It is also expected that the choice of the atmospheric pCO_2 does not considerably influence our FCO_2 calculations as the air-sea pCO_2 difference is mainly controlled by the oceanic pCO_2 (see Sect. 2.4). We use the bathymetry from the 2 min global ETOPO2 database (US Department of Commerce, 2006) and the Chl-a field derived from the monthly 4 km merged globcolour product from the U.U Copernicus Marine Service information (https://doi.org/10.48670/moi-00281), which is the product with the longest Chl-a temporal coverage (1998-2020 period). However, because of recurrent cloud coverage everywhere and sea-ice coverage at high latitudes, the monthly averaged Chl-a field at a 0.25° resolution is discontinuous with grid cells devoid of data. We fill these cells (9 % excluding high latitude coastal regions) using a cascade of interpolation methods, which order depends on data availability in time and space in the surrounding cells : for an empty cell of the month x, 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) the monthly mean value x of the cell for the entire 1998-2020 period, (4) the annual mean value of the cell for the entire 1998-2020 period. At high latitudes, where none of these options are feasible because of

large bands without any data, we assign the modal value of Chl-a as default value in order to ensure the continuity of the maps that is required for the FFN algorithm. From this continuous gridded Chl-a product over the recorded period, we then calculate for each cell a monthly seasonal climatology and attribute the climatological values to the unrecorded period (1982-1997).

230 This means that, in our calculations, Chl-a does not contribute to long-term changes in pCO_2 before 1997. All observational fields are converted from their original spatiotemporal resolution to monthly 0.25° gridded resolution for the 1982-2020 period (except for the bathymetry which is constant over time) to match the observational pCO_2 product (SOCAT_a) resolution (Table 1).

2.3 Air-sea CO_2 exchange

235 The pCO_2 field generated by the SOM-FFN algorithm (ULB-SOM-FFN-coastalv2) is used to calculate the air-sea CO_2 exchange for each grid cell at the monthly time scale over the 1982-2020 period following Eq. (1):

$$FCO_2 = k \, K_0 \, \Delta pCO_2 \, (1 - ice) \tag{1}$$

240 where FCO_2 (mol C m^{-2} yr^{-1}) represents the coastal air-sea CO_2 exchange, ΔpCO_2 (atm) is the gradient between the oceanic pCO_2 and the atmospheric pCO_2 and K_0 (mol m^{-3} atm^{-1}) is the CO_2 solubility in seawater which is a function of SST and SSS following the equation of Weiss (1974). k (m yr^{-1}) represents the gas exchange transfer velocity which is a function of the second moment of the wind speed at 10 meters above sea level and is calculated using the equation of Ho et al. (2011) and using the Schmidt number based on the equation of Wanninkhof (2014). The sea-ice coverage for each grid cell is represented

245 by the term ice and ranges from 0 (no ice cover) to 1 (100 % ice cover). By convention a positive FCO_2 value corresponds to a source of CO_2 for the atmosphere. We use the same U_{10}, SST, SSS, sea-ice coverage and atmospheric pCO_2 datasets (see Sect. 2.2) to calculate FCO_2 and perform the FFN pCO_2-reconstruction. Our reconstructed coastal FCO_2 is also compared with coastal FCO_2 estimates derived from a synthesis of 214 regional FCO_2 estimations (Dai et al., 2022) and from the FCO_2-product derived from the original ULB-SOM-FFN-coastalv1 pCO_2-product (Roobaert et al., 2019).

250 2.4 Uncertainties of the reconstructed coastal data-products

The uncertainties associated with our reconstructed pCO_2 and FCO_2 coastal products are estimated using the method proposed by Landschützer et al. (2014, 2018) and used by subsequent authors (e.g., Roobaert et al., 2019; Sharp et al., 2022). The FCO_2 uncertainty results from four sources of uncertainties, which are considered independent and thus summed quadratically:

255 $$\sigma_{FCO_2} = \sqrt{\sigma_{\Delta pCO_2}^2 + \sigma_k^2 + \sigma_{wind}^2 + \sigma_{ice}^2} \tag{2}$$

Where σ_{FCO_2} represents the total FCO_2 uncertainty (Petagrams of Carbon per year, Pg C yr^{-1}), $\sigma_{\Delta pCO_2}$ is the uncertainty of the air-sea pCO_2 gradient, σ_k is the uncertainty associated with the choice of the k-formulation in Eq. (1) and σ_{wind} is the uncertainty associated with the choice of the wind speed product (see Roobaert et al., 2018). 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.~~Resplandy et al., 2023). 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 uncertaint~~yies 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.

σ_k is estimated as the standard deviation of global FCO_2 fields calculated with four different global-scale k-parametrizations with the same wind speed (ERA-interim). We use the formulations of Ho et al. (2011), Sweeney et al. (2007), Takahashi et al. (2009), and Wanninkhof (2014); all suited for global scale applications (see e.g., Roobaert et al., 2018). σ_{ice} is calculated as the standard deviation of global FCO_2 fields calculated with 2 different sea-ice products: The NOAA dataset of Reynolds et al. (2007) and the sea-ice dataset of Rayner et al. (2003). $\sigma_{\Delta pCO_2}$ mainly results from the oceanic pCO_2 uncertainty since the atmospheric pCO_2 uncertainty is significantly lower (Landschützer et al., 2018). For instance, Roobaert et al. (2019) quantified that uncertainties in atmospheric pCO_2 only contribute to 6 % in the overall FCO_2 uncertainty.

Following Landschützer et al. (2014), the uncertainty over the oceanic pCO_2 can be obtained from the quadratic sum of 3 sources of uncertainties:

$$\theta_{pCO_2} = \sqrt{\left(\frac{\theta_{obs}}{\sqrt{N}}\right)^2 + \left(\frac{\theta_{grid}}{\sqrt{N}}\right)^2 + \left(\frac{\theta_{map}}{\sqrt{N_{eff}}}\right)^2} \qquad (3)$$

Where θ_{pCO_2} represents the total uncertainty of the oceanic pCO_2 (µatm), θ_{obs} is the experimental uncertainty associated with the sampling on the field of the observations from the SOCAT database (µatm), θ_{grid} is the uncertainty associated with the gridding of the observations from SOCAT into 0.25° monthly meshed maps (µatm) and θ_{map} is the uncertainty derived from the comparison between the reconstructed pCO_2 and the observed gridded pCO_2 from the SOCAT database (µatm). Following Sharp et al. (2022), we use accuracies that are attributed to each fCO_2 measurement by the SOCAT community (flags A to D) to calculate θ_{obs}. Flags 'A' and 'B' represent an estimated accuracy of 2 µatm while the accuracy of flags 'C' and 'D' is 5 µatm. We first calculate the mean of all fCO_2 flags in each grid cell for each month. We then calculate the global average

gridded flags uncertainty of all cells for the year x (or the entire period). For θ_{grid}, we first calculate in each grid cell of the month x the standard deviations of all fCO$_2$ values from the SOCAT database used for the gridding. We then calculate the average of these standard deviations of all grid cells for the year x (or the entire period). 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). We also divide each source of uncertainty (i.e., θ_{obs}, θ_{grid} and θ_{map}) by the square root of the number of pCO$_2$ samples (N, see Landschützer et al., 2018; Roobaert et al., 2019 for details). For θ_{map}, the value of N is corrected to account for the fact that all individual errors are not spatially independent. To this end, we calculate the effective sample size (N$_{eff}$, see Landschutzer et al., 2018) by randomly selecting 1000 samples (40 % of the samples if the total number of samples is < 1000) that cover our study period and calculating a lag 1 autocorrelation coefficient following equations (18) and (19) of Landschützer et al. (2018). As we only use a subset of 1000 samples, we perform Monte Carlo simulations where this procedure is repeated 10 times and our final N$_{eff}$ is calculated as the median of the 10 iterations. Finally, the total uncertainty on FCO$_2$ associated with the reconstructed pCO$_2$ ($\sigma_{\Delta pCO_2}$, Pg C yr^{-1}) is obtained by applying θ_{pCO_2} (µatm) in Eq. (1). All these procedures are performed globally for each year and for the entire period of our study.

3 Results and discussion

3.1 pCO₂-product evaluation

Globally, our reconstructed coastal pCO$_2$-product compares well with the observed pCO$_2$ used to train the FFN algorithm (SOCAT_a) and reproduces all the well-known global spatial pCO$_2$ patterns with generally low pCO$_2$ (< 360 µatm) at temperate as well as high latitudes and high pCO$_2$ (> 360 µatm) at low latitudes (contrast Fig. 2a with Fig. 2b). The spatial distribution of the temporal mean residuals (i.e., difference between the coastal pCO$_2$-product and SOCAT_a in each grid cell for every month where observations are available) reveals that, in some regions, underestimations (negative mean residual, blue colors in Fig. 2c) or overestimations (red colors) of the pCO$_2$ can be generated by the coastal pCO$_2$-product. However, most of the calculated residuals fall within the -20 to 20 µatm range, accounting for 6~~8~~9 % of the grid cells, while 4~~4~~5 % of the grid cells have absolute residuals < 10 µatm (Fig. 2c and Fig. 3a).

A global mean of the residuals (bias) value of 0 µatm and a coefficient of determination (r²) 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. The global RMSE is however, substantially larger (29 µatm) yet still comparable to those calculated, at the regional scale, in previous coastal pCO$_2$ studies based on statistical interpolations (see e.g., Chen et al., 2016) and slightly lower than the global RMSE calculated by Laruelle et al. (2017, 32 µatm). Large differences can be observed between our product and SOCAT_a locally in regions that are known to present large spatiotemporal variabilities in pCO$_2$ and/or in regions lacking data to train our FFN algorithm. For instance, residuals and/or standard deviations > 20 µatm are encountered in the

320   Baltic Sea (which is hence later treated as an independent biogeochemical province for some calculations), in upwelling regions (e.g., along the Peruvian upwelling), in coastal seas under the influence of seasonal changes in sea ice coverage (e.g., along the Antarctica shelf) as well as along the very nearshore coastal domain (e.g., along the California current coast, Figs. 2c, d).

The overall consistency between our coastal pCO_2-product and SOCAT_a is diagnosed over the entire timespan of our study,
325   as illustrated by the histograms of residuals calculated for each of the four decades of our calculation period (Fig. 3a) or when the calculations are performed for each individual year (Table S1). This is an important test as a study by Gloege et al. (2022) suggests that decadal trends in pCO_2 products may be obscured by changing residual distributions over time. In spite of the highly heterogeneous distribution of the number of pCO_2 observations available through time (< 500 gridded cells in the 80's vs > 10 000 in the 2010's), the shape and spread of the four histograms of the residuals are closely similar between decades
330   with a distribution centered on a global mean bias close to 0 µatm and most of the residuals falling in the -20 and 20 µatm range. This demonstrates the accuracy of the method over time despite the skewed distribution of the calibration data. The analyses performed for each individual year reveal that global biases do not exceed 5 µatm. Exceptions are observed in the 80's where biases (e.g., absolute bias of 6 µatm in 1987) and RMSE (e.g., 42 µatm in 1989) can be larger and partly attributed to an exceptionally low pCO_2 observational coverage during these periods (see e.g., Bakker et al 2014). In the first version of
335   the coastal product (ULB-SOM-FFN-coastalv1), the evaluation of the pCO_2-product of Laruelle et al. (2017) was restricted to spatial and climatological seasonal cycles. In this study, we successfully extended this analysis to the entire time period and evaluated each year and decade individually.

At the regional scale, 7 out of the 10 biogeochemical provinces yield RMSEs against SOCAT_a close to 20 µatm or lower
340   with the best fit in P2 (RMSE = 13 µatm, Table 2). This is a significant improvement over the ULB-SOM-FFN-coastalv1 product which only had 3 provinces with RMSE < 20 µatm (Laruelle et al., 2017) and none < 15 µatm. We attribute this improvement to our advanced setup of the method such as the inclusion of the atmospheric pCO_2 as a driver as well as an increased number of available observations to train our FFN algorithms. In three provinces (P1, P5 and P7) however, the RMSE exceeds 35 µatm. Such values can partly be explained by the complex dynamics of the sea ice in the Antarctic Shelf
345   (P1) and by the limited number of observational data combined with the inclusion of coastal regions that present large spatiotemporal variabilities and cover two disconnected temperate basins (P5 and P7) of the Northern Hemisphere. This discrepancy was also highlighted by Landschützer et al. (2020). High discrepancies are further observed in the Baltic Sea (Figs. 2c-d) which is analyzed as a separate province (RMSE value of 4~~3~~1 µatm, Table 2). Excluding the Baltic Sea from P8 considerably reduces RMSE from 4~~2~~1 to 23 µatm. The inclusion of the Baltic Sea in P8 can ~~likely~~ also explain the large RMSE
350   calculated by Laruelle et al. (2017) for the corresponding province of the ULB-SOM-FFN-coastalv1 coastal pCO_2-product (~ 47 µatm).

3.2 Validation against independent data

Our reconstructed coastal pCO$_2$-product is also validated against an independent dataset that is derived from pCO$_2$ observations from the SOCATv2022 that were not used for the training of the FFN algorithm (see Sect. 2.1). This dataset consists of a pool of 404,206 gridded cells that are uniformly distributed between both hemispheres (SOCAT_b, Fig. S1)~~-~~, ~~which is~~ an essential criterion for training the network ~~training~~ ~~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)~~. Globally, a good match is observed between the coastal pCO$_2$-product and SOCAT_b with a global bias and a RMSE of 0 and 29~~30~~ µatm, respectively. These values are similar to those derived from the statistical analysis performed against SOCAT_a. At the biogeochemical provinces scale, RMSEs generally do not exceed 23 µatm (maximum value in P6, Table 2), except where important RMSEs (~~-~~ 34~~5~~ µatm for P1, P5 and P7) had already been calculated during the comparison with SOCAT_a (i.e., in regions under the sea-ice coverage dynamics and poor data coverage and provinces which encompass regions with high spatiotemporal pCO$_2$ dynamics).

As with SOCAT_a, the analysis against SOCAT_b demonstrates a good performance of our reconstructed coastal pCO$_2$-product over time with the histograms of the residuals calculated for each of the four decades presenting the same shape and spread in spite of the marked decrease of grid cells numbers over time (1054 grid cells in 80' vs 248,626 grid cell in 2010', Fig. 3b). Each of these histograms shows a distribution centered on a value of 0 µatm with ~ 50 % of the grid cell residuals falling between -10 and 10 µatm. This is also true at the scale of the biogeochemical provinces with the four histograms of the residuals revealing global mean biases of 0 µatm and ~ 50 % of the residual falling in the -10 and 10 µatm range. Only P8 stands as an exception (where 50 % of the residuals are between -40 and 40 µatm, Fig. S3~~2~~), mainly due to the presence of the Baltic Sea in this province.

~~Finally,~~ We~~we~~ also present pCO$_2$ time series derived from our reconstructed coastal pCO$_2$-product and compare them to data extracted from SOCAT_b for 8 coastal sites (Fig. 4). The choice of these coastal regions is motivated by their data coverage extending over 30 years (Fig. 4a) and the fact that it is possible in these grid cells to reconstruct a spatially complete seasonal climatological cycle (i.e., data are available for all 12 months). For each region, we only extracted cells for which observations extended > 30 years and reconstructed their times series from the coastal pCO$_2$-product and SOCAT_b, respectively (Figs. 4b-i). For most of the regions, the pCO$_2$-product properly captures the temporal dynamics of pCO$_2$ derived from the observations, bearing slight underestimations or overestimations in specific areas such as along the Cascadian shelf and the East coast of Australia (Figs. 4d and 4g). For the 8 coastal sites, absolute biases are all < 10 µatm with a minimum absolute bias of 1 µatm in the Irminger Sea and a maximum absolute bias of 9 µatm along the New Zealand coast. Except along the Cascadian Shelf, all coastal sites present RMSEs lower than ~30 µatm (5 over the 8 regions show RMSEs < ~20 µatm) which falls in the range of our global and regional RMSE values. Largest RMSE is calculated along the Cascadian shelf (62 µatm) and can partly be

explained by the large spatial pCO₂ variability in SOCAT_b (as shown by the vertical bars in Fig. 4d) because of the riverine influence in the region.

Finally, the reconstructed pCO₂ 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 pCO₂ 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 pCO₂ 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 pCO₂ 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 pCO₂ is observed compared to the buoy observations. The reconstructed pCO₂ 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 pCO₂-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 pCO₂ 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 pCO₂ against both SOCAT_a, SOCAT_b as well as against buoys data gives confidence to identify the linear trends of both the pCO₂ and FCO₂ 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 pCO₂ evolution with the newly reconstructed pCO₂ product. For example, our results show that for all the 8 studied regions represented in Fig. 4, an increase in pCO₂ over time comprised between 12 and 20 µatm decade⁻¹ is calculated for the long-term trend with our pCO₂ 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 pCO₂-product, they both agree that this region displays the fastest trend in terms of pCO₂ rise (18 µatm decade⁻¹ and 20 µatm decade⁻¹ for SOCAT_b and the pCO₂-product, respectively).

3.3 Air-sea CO₂ exchange

This section describes the coastal air-sea CO₂ exchanges patterns that are calculated using our new reconstructed pCO₂-product over the 1982-2020 period. The spatial and seasonal FCO₂ patterns are only briefly discussed (Sect. 3.3.1) since those have been extensively discussed in previous studies (see e.g., Dai et al., 2022; Resplandy et al., in rev.Resplandy et al., 2023;

Roobaert et al., 2019). We thus focus on the long-term FCO_2 trends (Sect. 3.3.2), which are still poorly understood and highly debated (e.g., Lacroix et al., 2021a; Laruelle et al., 2018; ~~Resplandy et al., in rev.~~Resplandy et al., 2023).

3.3.1 Spatial and seasonal ~~dynamics~~ variations

420   The spatial distribution of the climatological mean coastal FCO_2 shows that coastal regions in temperate areas (between 40°-60° in both hemisphere_s) and at high latitude (beyond 60° in both hemisphere) mainly act as ~~a~~ CO_2 sinks while CO_2 sources are mainly encountered in the sub-tropical band (Fig. 5a) which is consistent with the global latitudinal pattern established by previous studies (e.g., Borges, 2005; Borges et al., 2005; Cai, 2011; Cao et al., 2020; Chen et al., 2013; Dai et al., 2022; Laruelle et al., 2010, 2014; Roobaert et al., 2019). Globally, with the coastal delineation used in this study ('wide coastal

425   ocean', 76 million km²), the coastal ocean absorbs on average 0.4_0 Pg C per year (with an uncertainty of ± 0.03 Pg C yr^{-1}, see ~~Sect.~~section 3.3.3) over the 1982-2020 period. Using the shelf break as the outer limit of the coastal domain ('narrow coastal ocean', 28 million km²) which is a more common delineation of the coastal ocean, the globally integrated coastal sink amounts to -0.1_82 ± 0.01 Pg C yr^{-1} which is consistent with the latest estimates (e.g., -0.2 Pg C yr^{-1} in Roobaert et al. (2019) and -0.25 Pg C yr^{-1} in Dai et al. (2022)). It should be noted that these comparisons are not straightforward because they do not cover the

430   same time periods (i.e., 1998-2015 period in Roobaert et al. (2019), 1998-present in Dai et al. (2022) and 1982-2020 in this study) and older assessments often do not report an explicit calculation period (Regnier et al., 2022).

Most of the intense CO_2 sinks (absolute FCO_2 value > 0.5 mol C m^{-2} yr^{-1}) are encountered at high latitudes of the Northern Hemisphere and in the temperate regions of the S_southern Hemisphere while CO_2 sources in the tropical bands are moderate

435   except along upwelling areas such as in the Arabian Sea (Figs. 5a and 5c). A large fraction (4_42 % and 5_30 % for the wide and narrow domain, respectively) of the global CO_2 uptake is taking place north of 60° N, which was already suggested in Laruelle et al. (2010) and further confirmed in subsequent studies (e.g., Cai, 2011; Dai et al., 2022, Laruelle et al., 2014; Roobaert et al., 2019). The spatial distribution of coastal CO_2 sources and sinks also closely follows the latitudinal FCO_2 profile calculated by Roobaert et al. (2019) which is based on ULB-SOM-FFN-coastalv1 (red and green lines in Fig. 5c). These global pCO$_2$-

440   products however predict less variability in flux density than a compilation of regional estimations as shown in Fig. 5c when comparing our climatological FCO_2 latitudinal profile with the synthesis of 214 regional FCO_2 estimates which was already pointed by Dai et al. (2022) when comparing their data synthesis with the latitudinal FCO_2 profile of Roobaert et al. (2019) suggesting strong FCO_2 heterogeneities for a same latitudinal band. 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.,

445   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).

3.3.2 ~~Decadal~~ long-term trends in the coastal CO_2 sink

The rate of change in coastal FCO_2 and the various parameters involved in the FCO_2 calculation (i.e., ΔpCO_2, wind speed and sea-ice coverage) from 1982 to 2020 are presented in Fig. 6. 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). Coastal regions with negative (positive) ΔpCO_2 slopes present negative (positive) FCO_2 slopes, which translate into a stronger sink/weaker source (weaker sink/stronger source). Most coastal regions (~ 6~~7~~0 % of the grid cells that present a significant trend using a Mann-Kendall statistical test with a significance threshold of 95 %) exhibit negative ΔpCO_2 and FCO_2 slopes (i.e., stronger sinks or weaker sources, blue colors in Figs. 6a, b) in agreement with past studies (e.g., Laruelle et al., 2018; ~~Resplandy et al., in rev.~~Resplandy et al., 2023; Wang et al., 2017). Positives ΔpCO_2 and FCO_2 slopes (weaker sink or stronger sources, red colors) can also be observed such as along the Mediterranean Sea or Southeast Asia. Stronger FCO_2 rates of change (absolute value > 0.6 mol m^{-2} yr^{-1} decade^{-1}) are mainly observed in mid-to high latitudes coastal regions and along upwelling regions (e.g., Moroccan upwelling current) while low latitudes coastal regions show weaker slopes.

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. (in rev.)~~Resplandy et al. (2023). For most of the coastal ocean, an increase in wind speed has been observed over the study period (positive slope, red colors in Fig. 6c) with a median value for the rate of change value of 0.04 m s^{-1} decade^{-1}. This increase in wind speed promotes the FCO_2 exchange through its effect on the gas exchange transfer velocity (stronger sinks/sources). Rate changes in sea-ice coverage reveal a general retreat of sea-ice in the Northern Hemisphere (negative slope, blue colors in Fig. 6d) and a gain along the Antarctic shelf (positive slope, red colors) in agreement with e.g., Serreze & Meier (2019). A decrease of sea-ice coverage favors air-sea CO_2 exchange over a larger coastal surface area and during longer periods of the year, both of which strengthen for instance the CO_2 sink in coastal regions at high latitudes of the Northern Hemisphere.

Globally integrated, our results indicate that today's coastal ocean has been acting as a CO_2 sink since the beginning of our study period (1982) both in the wide coastal ocean (Fig. 7a) and in the narrow domain (Fig. 7b). For both domains this CO_2 sink however increases over time. In the wide coastal ocean, the global CO_2 uptake amounted to 0.28~~5~~ Pg C yr^{-1} in the 1980's (median value over the 1982-1992 period) and reached 0.54~~2~~ Pg C yr^{-1} in the 2010's (mean value over the 2010-2020 period) with small interannual fluctuations (~ 0.01 Pg C yr^{-1}) of the CO_2 sink intensity diagnosed by our algorithm. The overall intensification of the coastal sink that we observe in this study (0.06~~1~~ Pg Cyr^{-1} decade^{-1} ($\pm$ 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. Our results are also in agreement with the conceptual approach of Bauer et al. (2013) as well as modeling studies, either using box models (e.g., Mackenzie et al., 2004, 2012; Rabouille et al., 2001; Ver et al., 1999) or, more recently, global ocean biogeochemical models (Bourgeois et al., 2016; Lacroix et al., 2021a, b) that all predict an increase in efficiency of the global coastal CO_2 sink at the century-scale.

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.

3.4 Uncertainties associated with the data-products

The global coastal CO_2 sink of -0.4~~0~~ Pg C yr^{-1} (wide coastal domain) that we calculate in this study using the ULB-SOM-FFN-coastalv2 pCO_2-product is associated with a relative uncertainty that amounts to ~ 10 % (value of 0.03 Pg C yr^{-1} for σ_{FCO_2}, see Eq. (2) and Table 3). This global uncertainty mainly results from the uncertainty associated with the oceanic pCO_2 ($\sigma_{\Delta pCO_2}$, uncertainty of 0.02 Pg C yr^{-1}). The choice of the gas exchange transfer velocity formulation yields a 7 % difference (σ_k, uncertainty of 0.01 Pg C yr^{-1}) on the global FCO_2 calculation while we calculate a ~ 4 % difference on FCO_2 depending

on the wind speed product choice (σ_{wind}, uncertainty of 0.01 Pg C yr^{-1}) or on the sea-ice product choice (σ_{ice}, uncertainty of 0.01 Pg C yr^{-1}) on the global FCO$_2$ calculation. Note though, that the uncertainty in the mean is substantially lower than that calculated for individual months (see Fig. 7) or regions (see e.g., discussion above regarding the Baltic Sea region) due to compensating errors as was also identified in a study by Gloege et al. (2022).

The total uncertainty of the oceanic pCO$_2$ (θ_{pCO_2}, value of 0.6~~2~~3 µatm, see Eq. (3)) mainly results from the SOM-FFN mapping method to reconstruct the coastal pCO$_2$-product (θ_{map} = 28.82~~29~~ µatm, Table 3). This uncertainty falls within the range of values reported in the literature from different statistical interpolation methods to generate coastal pCO$_2$ data-products (RMSE values generally between 10 and 35 µatm, see Chen et al. (2016)) which are calculated from regional studies and would be expected to be smaller than those calculated for global scale analysis (or even the performance of our algorithm at the scale of its provinces, which generally cover a much larger surface area than most regional studies). The θ_{map} uncertainty calculated in this study is however higher than reported for the open ocean (typical RMSE values < 20 µatm, e.g., Landschutzer et al., 2014), mainly because of the complex biogeochemical dynamics and larger variability observed in the coastal seas compared with the open ocean. We calculate a global value of 3.~~0~~08 µatm for θ_{obs}, the uncertainty on the sampling in the field of the observations from the SOCAT database, which is slightly higher than the value reported by Pfeil et al. (2013, value of 2 µatm). For θ_{grid}, the uncertainty associated with the meshing of the observations from SOCAT to gridded 0.25° monthly maps, we calculate a global mean value of 5.70~~6~~ µatm, which is close to the value reported by Sabine et al. (2013, 5 µatm) for the open ocean. 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). Moreover, due to the temporal heterogeneity of the data coverage in the SOCAT database, our FCO$_2$ uncertainties can also vary temporally. As shown in Fig. 7, the global FCO$_2$ uncertainties that we report for each year (dashed black lines) are largest in the 80's (e.g., global σ_{FCO_2} value of 0.1~~0~~1 Pg C yr^{-1} in 1987) because of the scarcity of pCO$_2$ measurements before 1990 in the SOCAT database and decrease over time. Our global uncertainties are also slightly larger along the nearshore domain of the coastal ocean. 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 use~~d~~ (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}), 3~~6~~5 µatm for θ_{map}, 8 µatm for θ_{grid} and 3 µatm for θ_{obs}.

4 Conclusions and directions for future research

The release of the global coastal pCO_2-product in 2017 by Laruelle et al. was a significant step forward for the investigation of the spatial distribution of CO_2 sources and sinks as well as their seasonal variabilities in the shallow portion of the ocean. It was also instrumental to the completion and harmonization of global ocean air-sea CO_2 fluxes (Fay et al., 2021), hence supporting global carbon budget analyses (Friedlingstein et al, 2022). However, this product was not designed or evaluated regarding its ability to resolve the interannual and decadal variabilit~~y~~ies 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.~~Resplandy et al., 2023; Wang et al., 2017). 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). 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. Trends analysis of the coastal FCO_2 has also been attempted using global ocean pCO_2 products that cover the coastal domain (~~Resplandy et al., in rev.~~Resplandy et al., 2023). However, these investigations have been inconclusive, likely because global ocean pCO_2-products cannot yet sufficiently well capture the specific and changing conditions occurring along the coastal domain (e.g., Chau et al, 2022; Rödenbeck et al., 2013 see ~~Resplandy et al., in rev.~~Resplandy et al., 2023). Our updated coastal pCO_2-product circumvents these limitations and provides a first robust assessment against which outputs from global oceanic modes results can be evaluated (e.g., ~~Resplandy et al., in rev.~~Resplandy et al., 2023). It will thus help better constrain the anthropogenic perturbation of the global ocean carbon cycle. 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. This approach, in conjunction with process-based simulations, is critically needed to evaluate and mitigate the impact of multiple anthropogenic perturbations (e.g., atmospheric pCO_2 increase, physical climate, eutrophication and hypoxia) on the global coastal carbon cycle and associated biodiversity loss and other marine stressors.

Data availability

The ULB-SOM-FFN-coastalv2 pCO_2 and FCO_2 products can be found here: https://www.ncei.noaa.gov/archive/accession/0279118 (Roobaert et al., 2023). The bathymetry derived from the 2 min global ETOPO2 database (US Department of Commerce, 2006, Table 1), the Chl-a from the monthly 4 km merged globcolour product for the 1998-2020 period from the U.U Copernicus Marine Service information (https://doi.org/10.48670/moi-00281), the SST and SSS from the daily NOAA OI SST V2 (Reynolds et al., 2007) and of the daily Hadley center EN4 SSS (Good et al., 2013), respectively. We use the atmospheric pCO_2 product 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 ESRL surface marine boundary layer CO_2 product (https://www.esrl.noaa.gov/gmd/ccgg/mbl/data.php, Duglockencky et al., 2019) with a vapor correction according to Dickson et al. (2007) and using the ERA5 sea level pressure (Herbach et al., 2020).~~dry air mixing ratio of CO_2 (xCO_2) provided by the NOAA Marine Boundary Layer reference product (https://www.esrl.noaa.gov/gmd/ccgg/mbl/) 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) to calculated the atmospheric pCO_2.~~ The sea-ice coverage derived from the monthly mean of the daily 0.25° dataset of Reynolds et al. (2007), the wind speed from the 6 hours first moment of the 0.25° resolution product of the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 (Hersbach et al., 2020)~~global atmospheric reanalysis ERA-interim wind product (Dee et al., 2011)~~. The pCO_2 observations derived from the SOCAT database v2022 (Bakker et al., 2022). The wind speed is calculated from the monthly mean of the 0.25° resolution product of the European Center for Medium-Range Weather Forecasts (ECMWF) ERA5 wind product (Hersbach et al., 2020), which has a native temporal resolution of 6 hours.

Author contribution

AR and GGL designed the study. AR prepared the manuscript with contributions from all co-authors.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

PL acknowledges support from the European Community's Horizon 2020 Project under grant agreement no. 821003 (4C). GGL is research associate of the F.R.S-FNRS at the Université Libre de Bruxelles. PR acknowledges funding from the European Union's Horizon 2020 research and innovation program under Grant agreement no. 101003536 (ESM2025— Earth System Models for the Future) and BELSPO (project ReCAP funded under the FEd-tWIN programme).

References

Atlas, R., Hoffman, R. N., Ardizzone, J., Leidner, S. M., Jusem, J. C., Smith, D. K., & Gombos, D. (2011). A cross-calibrated, multiplatform ocean surface wind velocity product for meteorological and oceanographic applications. *Bulletin of the American Meteorological Society*, *92*(2), 157–174. https://doi.org/10.1175/2010BAMS2946.1

Bai, Y., Cai, W. J., He, X., Zhai, W., Pan, D., Dai, M., & Yu, P. (2015). A mechanistic semi-analytical method for remotely sensing sea surface pCO2 in river-dominated coastal oceans: A case study from the East China Sea. *Journal of*

Geophysical Research: Oceans, *120*(3), 2331–2349. https://doi.org/10.1002/2014JC010632

605   Bakker, D. C. E., Pfeil, B., Smith, K., Hankin, S., Olsen, A., Alin, S. R., Cosca, C., Harasawa, S., Kozyr, A., Nojiri, Y.,
        O'Brien, K. M., Schuster, U., Telszewski, M., Tilbrook, B., Wada, C., Akl, J., Barbero, L., Bates, N. R., Boutin, J., …
        Watson, A. J. (2014). An update to the surface ocean CO2 atlas (SOCAT version 2). *Earth System Science Data*, *6*(1),
        69–90. https://doi.org/10.5194/essd-6-69-2014

        Bakker, D. C., Alin, S. R., Becker, M., Bittig, H. C., Castaño-Primo, R., Feely, R. A., ... & Wilson, D. (2022). Surface Ocean
610   CO2 Atlas Database Version 2022 (SOCATv2022) (NCEI Accession 0253659).

        Bauer, J. E., Cai, W., Raymond, P. A., Bianchi, T. S., Hopkinson, C. S., & Regnier, P. (2013). The changing carbon cycle of
        the coastal ocean. *Nature*, *504*(7478), 61. https://doi.org/10.1038/nature12857

        Becker, M., Olsen, A., Landschützer, P., Omar, A., Rehder, G., Rödenbeck, C., & Skjelvan, I. (2021). The northern European
        shelf as an increasing net sink for CO2. *Biogeosciences*, *18*(3), 1127–1147. https://doi.org/10.5194/bg-18-1127-2021

615   Borges, A. V. (2005). Do we have enough pieces of the jigsaw to integrate CO_2 fluxes in the coastal ocean? *Estuaries*, *28*(1),
        3–27. https://doi.org/10.1007/BF02732750

        Borges, A. V., Delille, B., & Frankignoulle, M. (2005). Budgeting sinks and sources of CO_2 in the coastal ocean: Diversity of
        ecosystem counts. *Geophysical Research Letters*, *32*(14), 1–4. https://doi.org/10.1029/2005GL023053

        Bourgeois, T., Orr, J. C., Resplandy, L., Terhaar, J., Ethé, C., Gehlen, M., & Bopp, L. (2016). Coastal-ocean uptake of
620   anthropogenic carbon. *Biogeosciences*, *13*(14), 4167–4185. https://doi.org/https://doi.org/10.5194/bg-13-4167-2016

        Cai, W. (2011). Estuarine and Coastal Ocean Carbon Paradox: CO2 Sinks or Sites of Terrestrial Carbon Incineration? *Annual
        Review of Marine Science*, *3*(1), 123–145. https://doi.org/10.1146/annurev-marine-120709-142723

        Cao, Z., Yang, W., Zhao, Y., Guo, X., Yin, Z., Du, C., Zhao, H., & Dai, M. (2020). Diagnosis of CO2 dynamics and fluxes in
        global coastal oceans. *National Science Review*, *7*(4), 786–797. https://doi.org/10.1093/NSR/NWZ105

625   Chau, T. T. T., Gehlen, M., & Chevallier, F. (2022). A seamless ensemble-based reconstruction of surface ocean pCO2 and
        air-sea CO2 fluxes over the global coastal and open oceans. *Biogeosciences*, *19*(4), 1087–1109.
        https://doi.org/10.5194/bg-19-1087-2022

        Chau, T. T. T., Gehlen M, Metzl N, & Chevallier F. (2023). CMEMS-LSCE: A global 0.25-degree, monthly reconstruction of
        the surface ocean carbonate system, Earth System Data Dissc., DOI: 10.5194/essd-2023-146, in review.

630   Chen, C., Huang, T. H., Chen, Y. C., Bai, Y., He, X., & Kang, Y. (2013). Air-sea exchanges of CO2 in the world's coastal
        seas. *Biogeosciences*, *10*(10), 6509–6544. https://doi.org/10.5194/bg-10-6509-2013

        Chen, S., Hu, C., Byrne, R. H., Robbins, L. L., & Yang, B. (2016). Remote estimation of surface pCO2 on the West Florida
        Shelf. *Continental Shelf Research*, *128*(July), 10–25. https://doi.org/10.1016/j.csr.2016.09.004

        Dai, M., Su, J., Zhao, Y., Hofmann, E. E., Cao, Z., Cai, W., Gan, J., Lacroix, F., Laruelle, G. G., Meng, F., Müller, J. D.,
635   Regnier, P., Wang, G., & Wang, Z. (2022). Carbon Fluxes in the Coastal Ocean: Synthesis, Boundary Processes and
        Future Trends. *Annual Review of Earth and Planetary Sciences*, *50*(1). https://doi.org/10.1146/annurev-earth-032320-
        090746

~~Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., … Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553-597. https://doi.org/10.1002/qj.828~~

Dickson, A. G., Sabine, C. L., & Christian, J. R. (2007). Guide to best practices for ocean CO_2 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).

Eldering, A., Wennberg, P. O., Crisp, D., Schimel, D. S., Gunson, M. R., Chatterjee, A., Liu, J., Schwandner, F. M., Sun, Y., O'Dell, C. W., Frankenberg, C., Taylor, T., Fisher, B., Osterman, G. B., Wunch, D., Hakkarainen, J., Tamminen, J., & Weir, B. (2017). The Orbiting Carbon Observatory-2 early science investigations of regional carbon dioxide fluxes. *Science*, *358*(6360). https://doi.org/10.1126/science.aam5745

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

Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., ... & Zheng, B. (2022). Global carbon budget 2022. Earth System Science Data Discussions, 2022, 1-159.

Gloege, L., Yan, M., Zheng, T., & McKinley, G. A. (2022). Improved Quantification of Ocean Carbon Uptake by Using Machine Learning to Merge Global Models and pCO2 Data. *Journal of Advances in Modeling Earth Systems*, *14*(2). https://doi.org/10.1029/2021MS002620

Good, S. A., Martin, M. J., & Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, *118*(12), 6704–6716. https://doi.org/10.1002/2013JC009067

Gregor, L., & Gruber, N. (2021). OceanSODA-ETHZ: A global gridded data set of the surface ocean carbonate system for seasonal to decadal studies of ocean acidification. *Earth System Science Data*, *13*(2), 777–808. https://doi.org/10.5194/essd-13-777-2021

Hales, B., Strutton, P. G., Saraceno, M., Letelier, R., Takahashi, T., Feely, R., Sabine, C., & Chavez, F. (2012). Satellite-based prediction of pCO2 in coastal waters of the eastern North Pacific. *Progress in Oceanography*, *103*, 1–15. https://doi.org/10.1016/j.pocean.2012.03.001

670 Hauck, J., Zeising, M., Le Quéré, C., Gruber, N., Bakker, D. C. E., Bopp, L., Chau, T. T. T., Gürses, Ö., Ilyina, T., Landschützer, P., Lenton, A., Resplandy, L., Rödenbeck, C., Schwinger, J., & Séférian, R. (2020). Consistency and Challenges in the Ocean Carbon Sink Estimate for the Global Carbon Budget. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.571720

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers,
675 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:
680 https://cds.climate.copernicus.eu/cdsapp#!/dataset/ reanalysis-era5-single-levels-monthly-means?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

685 Iida, Y., Kojima, A., Takatani, Y., Nakano, T., Sugimoto, H., Midorikawa, T., & Ishii, M. (2015). Trends in pCO2 and sea–air CO2 flux over the global open oceans for the last two decades. *Journal of Oceanography*, *71*(6), 637–661. https://doi.org/10.1007/s10872-015-0306-4

Jamet, C., Moulin, C., & Lefévre, N. (2007). Estimation of the oceanic pCO2 in the North Atlantic from VOS lines in-situ measurements: Parameters needed to generate seasonally mean maps. *Annales Geophysicae*, *25*(11), 2247–2257.
690 https://doi.org/10.5194/angeo-25-2247-2007

Jones, S. D., Le Quéré, C., Rödenbeck, C., Manning, A. C., & Olsen, A. (2015). A statistical gap-filling method to interpolate global monthly surface ocean carbon dioxide data. *Journal of Advances in Modeling Earth Systems*, *7*(4), 1554–1575. https://doi.org/10.1002/2014MS000416

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu,
695 Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., … Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Lacroix, F., Ilyina, T., Laruelle, G. G., & Regnier, P. (2021a). Reconstructing the Preindustrial Coastal Carbon Cycle Through a Global Ocean Circulation Model: Was the Global Continental Shelf Already Both Autotrophic and a CO2 Sink? *Global*
700 *Biogeochemical Cycles*, *35*(2), e2020GB006603. https://doi.org/10.1029/2020GB006603

Lacroix, F., Ilyina, T., Mathis, M., Laruelle, G. G., & Regnier, P. (2021b). Historical increases in land-derived nutrient inputs may alleviate effects of a changing physical climate on the oceanic carbon cycle. *Global Change Biology*, *27*(21), 5491–5513. https://doi.org/10.1111/gcb.15822

Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., Sasse, T., and Zeng, J.: A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink (2013), Biogeosciences, 10, 7793–7815, https://doi.org/10.5194/bg-10-7793-2013, 2013.

Landschützer, P., Gruber, N., Bakker, D. C. E., & Schuster, U. (2014). Recent variability of the global ocean carbon sink. *Global Biogeochemical Cycles*, *28*(9), 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

Landschützer, P., Gruber, N., Bakker, D. C. E., Stemmler, I., & Six, K. D. (2018). Strengthening seasonal marine CO2 variations due to increasing atmospheric CO2. *Nature Climate Change*, *8*, 146–150. https://doi.org/10.1038/s41558-017-0057-x

Landschützer, P., Laruelle, G. G., Roobaert, A., & Regnier, P. (2020). A uniform pCO2 climatology combining open and coastal oceans. *Earth System Science Data*, *12*(4), 2537–2553. https://doi.org/10.5194/essd-12-2537-2020

Laruelle, G. G., Cai, W., Hu, X., Gruber, N., Mackenzie, F. T., & Regnier, P. (2018). Continental shelves as a variable but increasing global sink for atmospheric carbon dioxide. *Nature Communications*, *9*(1), 454. https://doi.org/10.1038/s41467-017-02738-z

Laruelle, G. G., Dürr, H. H., Slomp, C. P., & Borges, A. V. (2010). Evaluation of sinks and sources of CO2 in the global coastal ocean using a spatially-explicit typology of estuaries and continental shelves. *Geophysical Research Letters*, *37*(15), 1–6. https://doi.org/10.1029/2010GL043691

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

Laruelle, G. G., Lauerwald, R., Pfeil, B., & Regnier, P. (2014). Regionalized global budget of the CO2 exchange at the air-water interface in continental shelf seas. *Global Biogeochemical Cycles*, *28*, 1199–1214. https://doi.org/10.1002/2014GB004832

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.

Mackenzie, F. T., De Carlo, E. H., & Lerman, A. (2012). Coupled C, N, P, and O Biogeochemical Cycling at the Land-Ocean Interface. In *Treatise on Estuarine and Coastal Science* (Vol. 5, Issue March 2019). https://doi.org/10.1016/B978-0-12-374711-2.00512-X

Mackenzie, F. T., Lerman, A., & Andersson, A. J. (2004). Past and present of sediment and carbon biogeochemical cycling models. *Biogeosciences*, *1*(1), 11–32. https://doi.org/10.5194/bg-1-11-2004

Majkut, J. D., Carter, B. R., Frölicher, T. L., Dufour, C. O., Rodgers, K. B., & Sarmiento, J. L. (2014). An observing system simulation for Southern Ocean carbon dioxide uptake. *Philosophical Transactions of the Royal Society A: Mathematical,*

Physical and Engineering Sciences, *372*(2019), 20130046. https://doi.org/10.1098/rsta.2013.0046

Nakaoka, S., Telszewski, M., Nojiri, Y., Yasunaka, S., Miyazaki, C., Mukai, H., & Usui, N. (2013). Estimating temporal and spatial variation of ocean surface pCO2 in the North Pacific using a self-organizing map neural network technique. *Biogeosciences*, *10*(9), 6093–6106. https://doi.org/10.5194/bg-10-6093-2013

Ono, T., Saino, T., Kurita, N., & Sasaki, K. (2004). Basin-scale extrapolation of shipboard pCO2 data by using satellite SST and Chla. *International Journal of Remote Sensing*, *25*(19), 3803–3815. https://doi.org/10.1080/01431160310001657515

Park, G. H., Wanninkhof, R., Doney, S. C., Takahashi, T., Lee, K., Feely, R. A., Sabine, C. L., Trińlanes, J., & Lima, I. D. (2010). Variability of global net sea-air CO2 fluxes over the last three decades using empirical relationships. *Tellus, Series B: Chemical and Physical Meteorology*, *62*(5), 352–368. https://doi.org/10.1111/j.1600-0889.2010.00498.x

Pfeil, B., Olsen, A., Baker, D. F., Hankin, S., Koyuk, H., Kozyr, A., Malczyk, J., Manke, A., Metzl, N., Sabine, C., Akl, J., Alin, S. R., Bates, N., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W., Chavez, F. P., … Yoshikawa-Inoue, H. (2013). A uniform, quality controlled Surface Ocean CO$_2$ Atlas (SOCAT). *Earth System Science Data*, *5*(1), 125–143. https://doi.org/10.5194/essd-5-125-2013

Rabouille, C., Mackenzie, F. T., & Ver, L. M. (2001). Influence of the human perturbation on carbon, nitrogen, and oxygen biogeochemical cycles in the global coastal ocean. *Geochimica et Cosmochimica Acta*, *65*(21), 3615–3641. https://doi.org/10.1016/S0016-7037(01)00760-8

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., & Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, *108*(D14), 4407. https://doi.org/10.1029/2002JD002670

Regnier, P., Friedlingstein, P., Ciais, P., Mackenzie, F. T., Gruber, N., Janssens, I. A., Laruelle, G. G., Lauerwald, R., Luyssaert, S., Andersson, A. J., Arndt, S., Arnosti, C., Borges, A. V, Dale, A. W., Gallego-Sala, A., Goddéris, Y., Goossens, N., Hartmann, J., Heinze, C., … Thullner, M. (2013). Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nature Geoscience*, *6*(8), 597–607. https://doi.org/10.1038/ngeo1830

Regnier, P., Resplandy, L., Najjar, R. G., & Ciais, P. (2022). The land-to-ocean loops of the global carbon cycle. In *Nature* (Vol. 603, Issue 7901, pp. 401–410). https://doi.org/10.1038/s41586-021-04339-9

Resplandy, L., Hogikyan, A, Bange, H et al. (2023in rev.). A Synthesis of Global Coastal Ocean Greenhouse Gas Fluxes. ESS Open Archive. April 18, 2023. DOI: 10.22541/essoar.168182303.39621839/v1

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, *20*(22), 5473–5496. https://doi.org/10.1175/2007JCLI1824.1

Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R. R., Jones, S., Landschützer, P., Metzl, N., Nakaoka, S., Olsen, A., Park, G. H. H., Peylin, P., Rodgers, K. B., Sasse, T. P. P., Schuster, U., Shutler, J. D. D., Valsala, V., Wanninkhof, R., Zeng, J., & Baker, D. F. (2015). Data-based estimates of the ocean carbon sink variability - First results of the Surface Ocean pCO2 Mapping intercomparison (SOCOM). *Biogeosciences*, *12*(23), 7251–7278.

https://doi.org/10.5194/bg-12-7251-2015

Rödenbeck, C., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., Cassar, N., Reum, F., Keeling, R. F., & Heimann, M. (2014). Interannual sea–air CO2 flux variability from an observation-driven ocean mixed-layer scheme. *Biogeosciences*, *11*(17), 4599–4613. https://doi.org/10.5194/bg-11-4599-2014

Rödenbeck, C., Keeling, R. F., Baker, D. F., Metzl, N., Olsen, A., Sabine, C., & Heimann, M. (2013). Global surface-ocean pCO2 and sea–air CO2 flux variability from an observation-driven ocean mixed-layer scheme. *Ocean Science*, *9*, 193–216. https://doi.org/10.5194/os-9-193-2013

Roobaert, A., Laruelle, G. G., Landschützer, P., Gruber, N., Chou, L., & Regnier, P. (2019). The Spatiotemporal Dynamics of the Sources and Sinks of CO2 in the Global Coastal Ocean. *Global Biogeochemical Cycles*, *33*, 1693–1714. https://doi.org/10.1029/2019GB006239

Roobaert, A., Laruelle, G. G., Landschützer, P., & Regnier, P. (2018). Uncertainty in the global oceanic CO_2 uptake induced by wind forcing: Quantification and spatial analysis. *Biogeosciences*, *15*(6), 1701–1720. https://doi.org/10.5194/bg-15-1701-2018

Roobaert, A.; Regnier, P.; Landschützer, P., & Laruelle, G. G. (2023). A novel sea surface partial pressure of carbon dioxide (pCO_2) data product for the global coastal ocean resolving trends over the 1982-2020 period (NCEI Accession 0279118). NOAA National Centers for Environmental Information. Unpublished Dataset. https://www.ncei.noaa.gov/archive/accession/0279118. Accessed [19/06/2023].

Sabine, C. L., Hankin, S., Koyuk, H., Baker, D. F., Pfeil, B., Olsen, A., Metzl, N., Kozyr, A., Fassbender, A., Manke, A., Malczyk, J., Akl, J., Alin, S. R., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W., Chavez, F. P., … Yoshikawa-Inoue, H. (2013). Surface Ocean CO_2 Atlas (SOCAT) gridded data products. *Earth System Science Data*, *5*(1), 145–153. https://doi.org/10.5194/essd-5-145-2013

Sarma, V. V. S. S., Saino, T., Sasaoka, K., Nojiri, Y., Ono, T., Ishii, M., Inoue, H. Y., & Matsumoto, K. (2006). Basin-scale pCO2 distribution using satellite sea surface temperature, Chl a, and climatological salinity in the North Pacific in spring and summer. *Global Biogeochemical Cycles*, *20*(3). https://doi.org/10.1029/2005GB002594

Schuster, U., McKinley, G. A., Bates, N., Chevallier, F., Doney, S. C., Fay, A. R., González-Dávila, M., Gruber, N., Jones, S., Krijnen, J., Landschützer, P., Lefèvre, N., Manizza, M., Mathis, J., Metzl, N., Olsen, A., Rios, A. F., Rödenbeck, C., Santana-Casiano, J. M., … Watson, A. J. (2013). An assessment of the Atlantic and Arctic sea–air CO2 fluxes, 1990–2009. *Biogeosciences*, *10*(1), 607–627. https://doi.org/10.5194/bg-10-607-2013

Serreze MC, Meier WN. The Arctic's sea ice cover: trends, variability, predictability, and comparisons to the Antarctic. Ann N Y Acad Sci. 2019 Jan;1436(1):36-53. doi: 10.1111/nyas.13856. Epub 2018 May 28. PMID: 29806697.

Sharp, J. D., Fassbender, A. J., Carter, B. R., Lavin, P. D., & Sutton, A. J. (2022). A monthly surface pCO_2 product for the California Current Large Marine Ecosystem. *Earth System Science Data*, *14*(4), 2081–2108. https://doi.org/10.5194/essd-14-2081-2022

Shutler, J. D., Land, P. E., Piolle, J. F., Woolf, D. K., Goddijn-Murphy, L., Paul, F., Girard-Ardhuin, F., Chapron, B., &

Donlon, C. J. (2016). FluxEngine: A flexible processing system for calculating atmosphere-ocean carbon dioxide gas fluxes and climatologies. *Journal of Atmospheric and Oceanic Technology*, *33*(4), 741–756. https://doi.org/10.1175/JTECH-D-14-00204.1

Sutton, A. J., Feely, R. A., Maenner-Jones, S., Musielwicz, S., Osborne, J., Dietrich, C., Monacci, N., Cross, J., Bott, R., Kozyr, A., Andersson, A. J., Bates, N. R., Cai, W.-J., Cronin, M. F., De Carlo, E. H., Hales, B., Howden, S. D., Lee, C. M., Manzello, D. P., McPhaden, M. J., Meléndez, M., Mickett, J. B., Newton, J. A., Noakes, S. E., Noh, J. H., Olafsdottir, S. R., Salisbury, J. E., Send, U., Trull, T. W., Vandemark, D. C., and Weller, R. A.(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

Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G. A., Sarmiento, J. L., & Wanninkhof, R. (2007). Constraining global air-sea gas exchange for CO_2 with recent bomb ^{14}C measurements. *Global Biogeochemical Cycles*, *21*(2), 1–10. https://doi.org/10.1029/2006GB002784

Takahashi, T., Sutherland, S. C., & Kozyr, A. (2012). Global ocean surface water partial pressure of CO2 database: Measurements performed during 1957–2011 (Version 2011). *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee,.*

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.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A. R. A., Chipman, D. W. D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Baker, D. F., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., … de Baar, H. J. W. (2009). Climatological mean and decadal change in surface ocean pCO₂, and net sea-air CO_2 flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(8), 554–577. https://doi.org/10.1016/j.dsr2.2008.12.009.Figure

US Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center, 2-minute Gridded Global Relief Data (ETOPO2v2), http://www.ngdc.noaa.gov/mgg/fliers/06mgg01.html (last access: 26 December 2022), 2006.

Ver, L. M. B., Mackenzie, F. T., & Lerman, A. (1999). Carbon cycle in the coastal zone: Effects of global perturbations and change in the past three centuries. *Chemical Geology*, *159*(1–4), 283–304. https://doi.org/10.1016/S0009-2541(99)00042-X

Wang, H., Hu, X., Cai, W. J., & Sterba-Boatwright, B. (2017). Decadal fCO2 trends in global ocean margins and adjacent boundary current-influenced areas. *Geophysical Research Letters*, *44*(17), 8962–8970. https://doi.org/10.1002/2017GL074724

Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited. *Limnology and*

840   *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

Wrobel-Niedzwiecka, I., Kitowska, M., Makuch, P., & Markuszewski, P. (2022). The Distribution of pCO2W and Air-Sea CO2 Fluxes Using FFNN at the Continental Shelf Areas of the Arctic Ocean. Remote Sensing, 14(2).

845   https://doi.org/10.3390/rs14020312

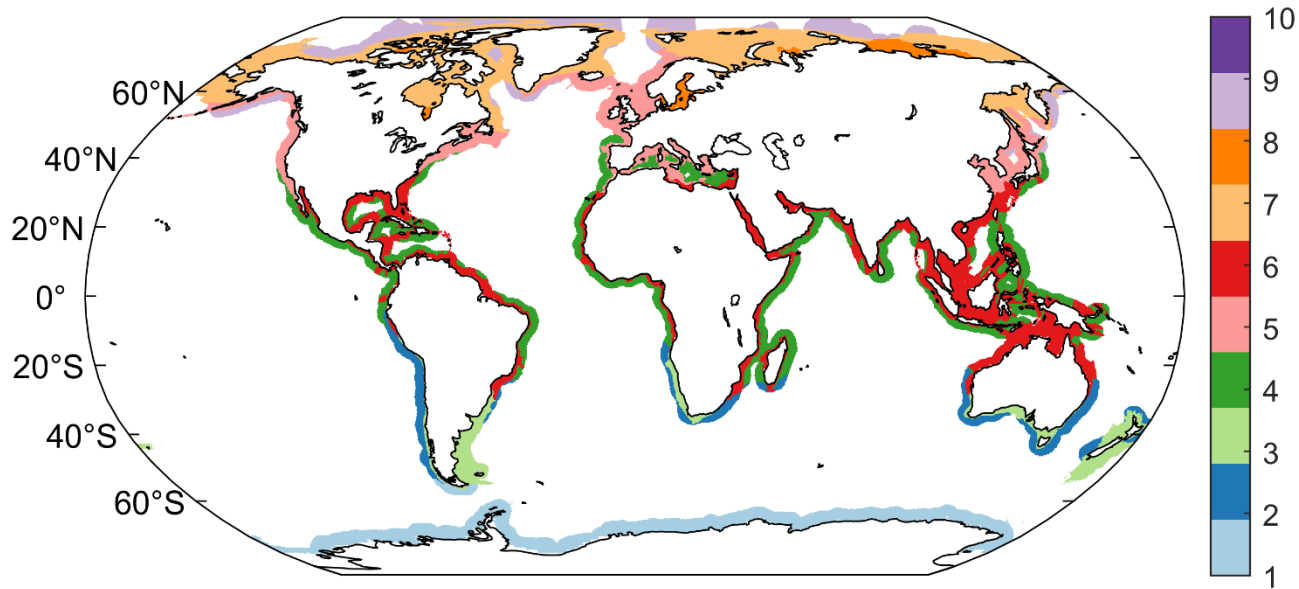
Zeng, J., Nojiri, Y., Landschützer, P., Telszewski, M., & Nakaoka, S. (2014). A global surface ocean fCO2 climatology based on a feed-forward neural network. *Journal of Atmospheric and Oceanic Technology*, *31*(8), 1838–1849. https://doi.org/10.1175/JTECH-D-13-00137.1

850

Figure 1: Coastal biogeochemical provinces generated by the Self Organizing Map clustering algorithm (SOM). The spatial extension of these provinces can vary from one month to another due to seasonal variations of the environmental drivers used during SOM. Here we present their modal spatial distribution (see Sect. 2.1 for further details).

855

Table 1: Environmental drivers used for the Self Organizing Map (SOM) and the Feed Forward Network (FFN) artificial neural networks to reconstruct the coastal pCO₂. These datasets are also used to calculate the coastal air-sea CO₂ exchange.

860

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 ~~Interim~~ (~~Dee et al., 2011~~ Hersbach et al., 2020)
Sea-ice coverage (ice, ~~no unit~~ 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 pCO₂ (in µatm)		X	1982-2020	~~0.25~~ 1°, monthly	aggregated to 0.25°	SeaFlux product (Fay et al., 2021) with xCO₂ derived from the ESRL surface MBL CO₂ 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) ~~NOAA Marine Boundary Layer reference product (https://www.esrl.noaa.gov/gmd/ccgg/mbl/) and Kalnay et al. (1996)~~
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)

865

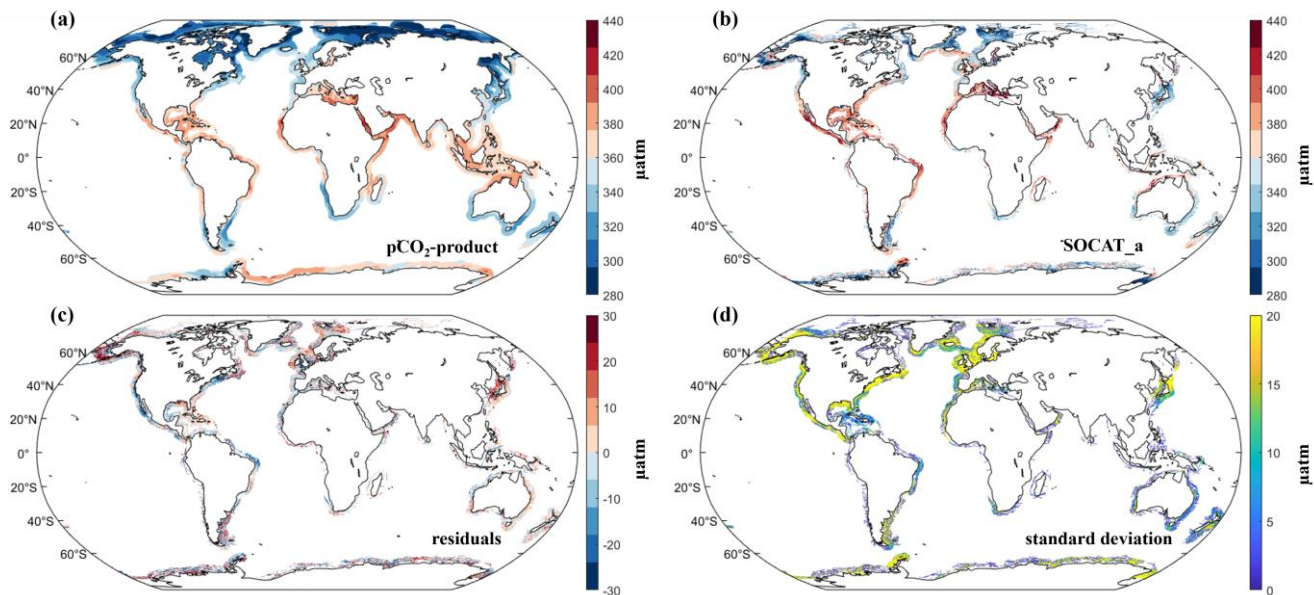



Figure 2: Global maps of the climatological (1982-2020) averaged (a) reconstructed coastal pCO$_2$-product, (b) gridded pCO$_2$ from the SOCATv2022 database and used for the FFN algorithm (SOCAT_a). (c) Temporal mean of the residuals between the coastal pCO$_2$-product and SOCAT_a and (d) standard deviation. All panels are expressed in μatm.

870

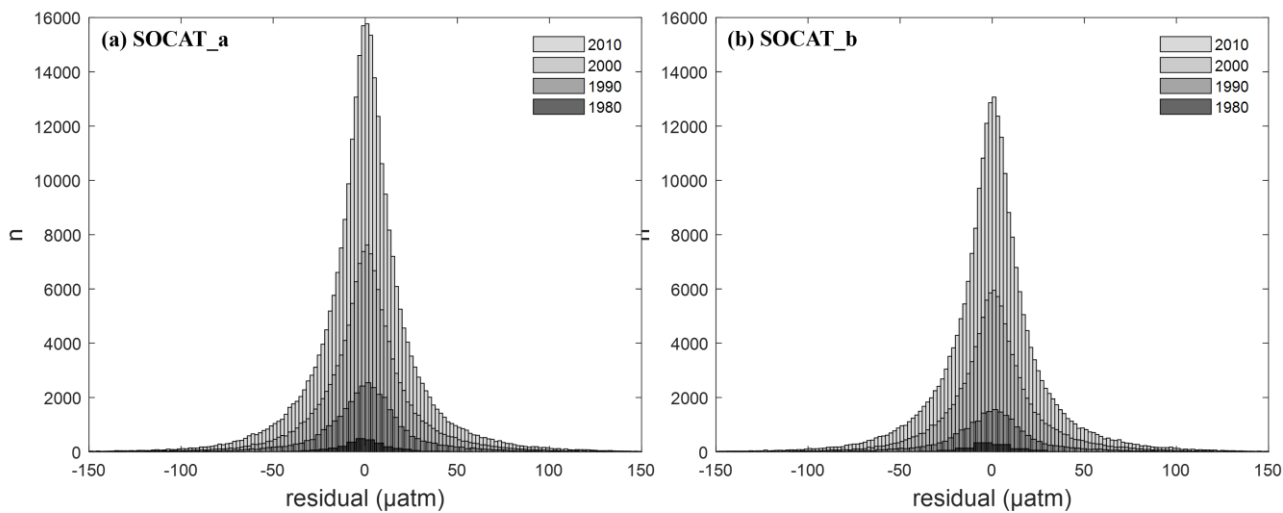



Figure 3: Histograms of the pCO$_2$ residuals (difference between the reconstructed coastal pCO$_2$-product with (a) SOCAT_a and (b) SOCAT_b) for four decades expressed in μatm.

875

Table S1: Bias (mean of the residuals between the reconstructed coastal pCO$_2$-product and SOCAT_a, in µatm), Root Mean Square Error (RMSE, µatm) and coefficient of determination (r^2) calculated for each year. The number of grid cells used for the comparison is also provided.

	bias (µatm)	RMSE (µatm)	r^2	cells
1982-12020	0	29	0.7	472134
1982	1	20	0.7	568
1983	5	15	0.7	209
1984	-1	13	0.7	341
1985	5	18	0.4	192
1986	-2	11	0.8	310
1987	-6	31	0.4	456
1988	-1	27	0.6	323
1989	-2	42	0.5	453
1990	2	26	0.2	181
1991	0	17	0.8	1094
1992	2	16	0.7	1176
1993	4	30	0.6	2067
1994	0	25	0.6	2939
1995	0	24	0.8	6220
1996	0	28	0.7	5130
1997	3	32	0.8	4928
1998	1	28	0.7	5259
1999	0	31	0.8	5183
2000	-2	32	0.7	6289
2001	-1	33	0.7	7055
2002	3	26	0.7	8833
2003	-3	24	0.8	8772
2004	-1	26	0.8	12276
2005	-1	25	0.7	14980
2006	0	28	0.7	22512
2007	2	29	0.7	21724
2008	2	26	0.8	19784
2009	1	30	0.7	18313
2010	0	28	0.7	21149
2011	0	34	0.7	27382
2012	-1	31	0.7	24474
2013	-1	30	0.7	21674
2014	0	32	0.7	27817
2015	0	30	0.7	28386
2016	1	25	0.8	29319
2017	2	26	0.8	34210
2018	1	28	0.7	30488
2019	-1	30	0.7	29283
2020	-2	30	0.8	20367

880

885

Table 2: Statistical analyses (bias, RMSE and r²) of the reconstructed coastal pCO₂-product against pCO₂ 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.

provinces	bias (µatm)		RMSE (µatm)		r²		cells	
	SOCAT_a	SOCAT_b	SOCAT_a	SOCAT_b	SOCAT_a	SOCAT_b	SOCAT_a	SOCAT_b
1	0.2	-0.1	34	34	0.7	0.7	53256	46943
2	0.0	0.1	13	14	0.8	0.8	11868	8898
3	-0.1	0.3	21	22	0.6	0.5	35981	26032
4	-0.2	-0.3	17	17	0.8	0.8	72811	54846
5	0.1	-0.1	34	34	0.5	0.5	137754	114017
6	0.7	0.2	23	23	0.7	0.6	69032	56322
7	0.5	0.8	33	34	0.7	0.7	48704	41308
8(8*)	0.4(0.7)	0.6(0.1)	41(22)	42(23)	0.9(0.9)	0.9(0.9)	18922(1086)	18039(952)
9	-0.2	-0.2	20	20	0.8	0.9	18779	13675
10	0.6	0.3	22	23	0.8	0.8	5009	4172
*Baltic Sea	0.3	0.6	41	42	0.9	0.9	1086	17911

890 *Numbers in parentheses for P8 represent statistics when the Baltic Sea is removed from P8 and defined as an isolated province.

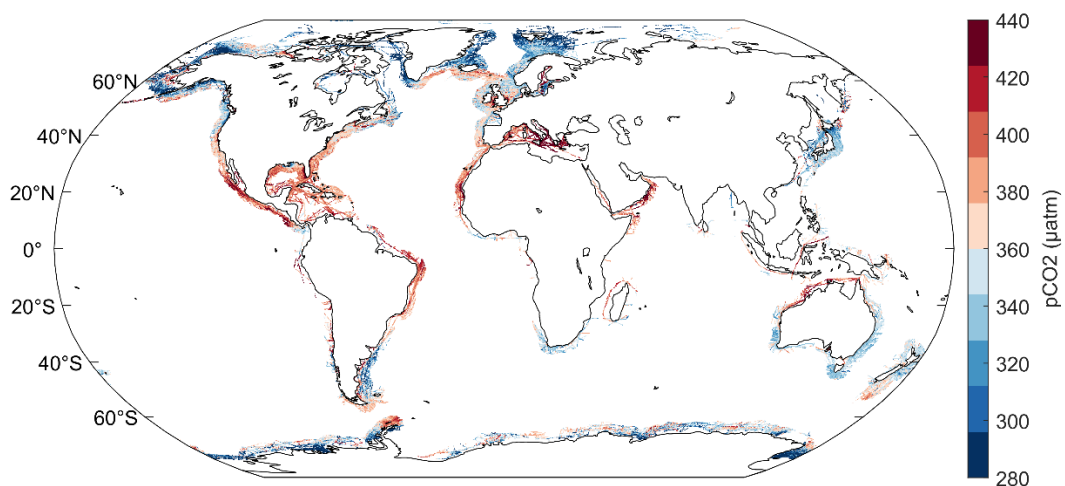
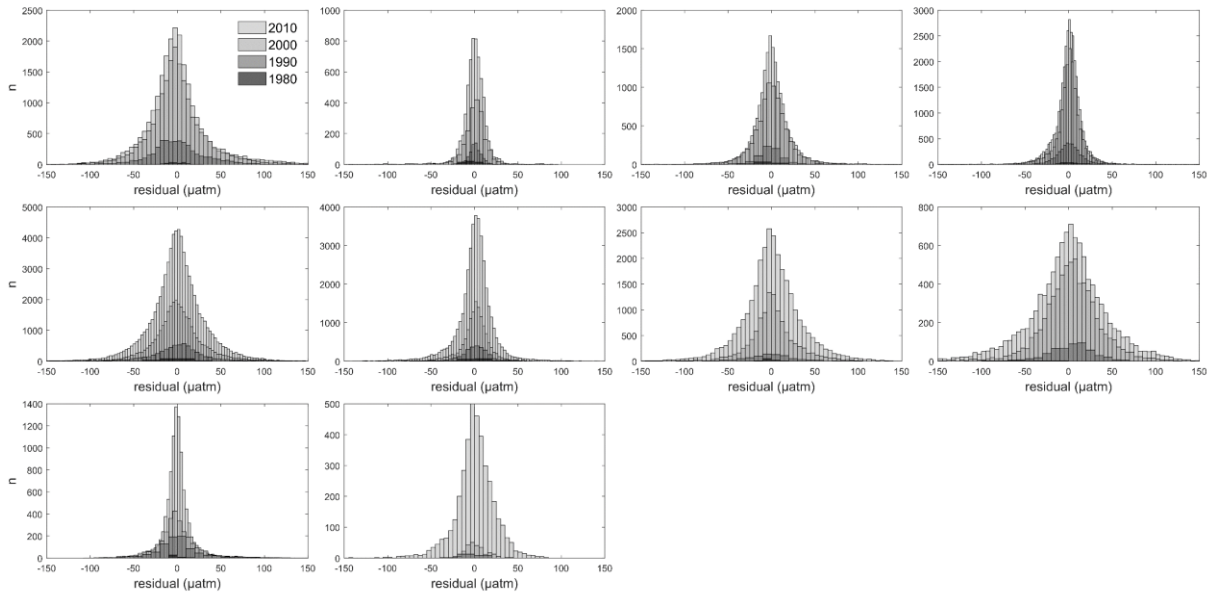



Figure S1. Global map of the climatological (1982-2020 period) averaged gridded sea surface coastal pCO₂ (µatm) from the SOCATv2022 database used as an independent dataset against for which the reconstructed coastal pCO₂-product is evaluated 895 (SOCAT_b).

900   Figure S3S2: Histograms of the residuals (in µatm) between the reconstructed coastal pCO$_2$-product and SOCAT_b for each of the four decades in each biogeochemical province.

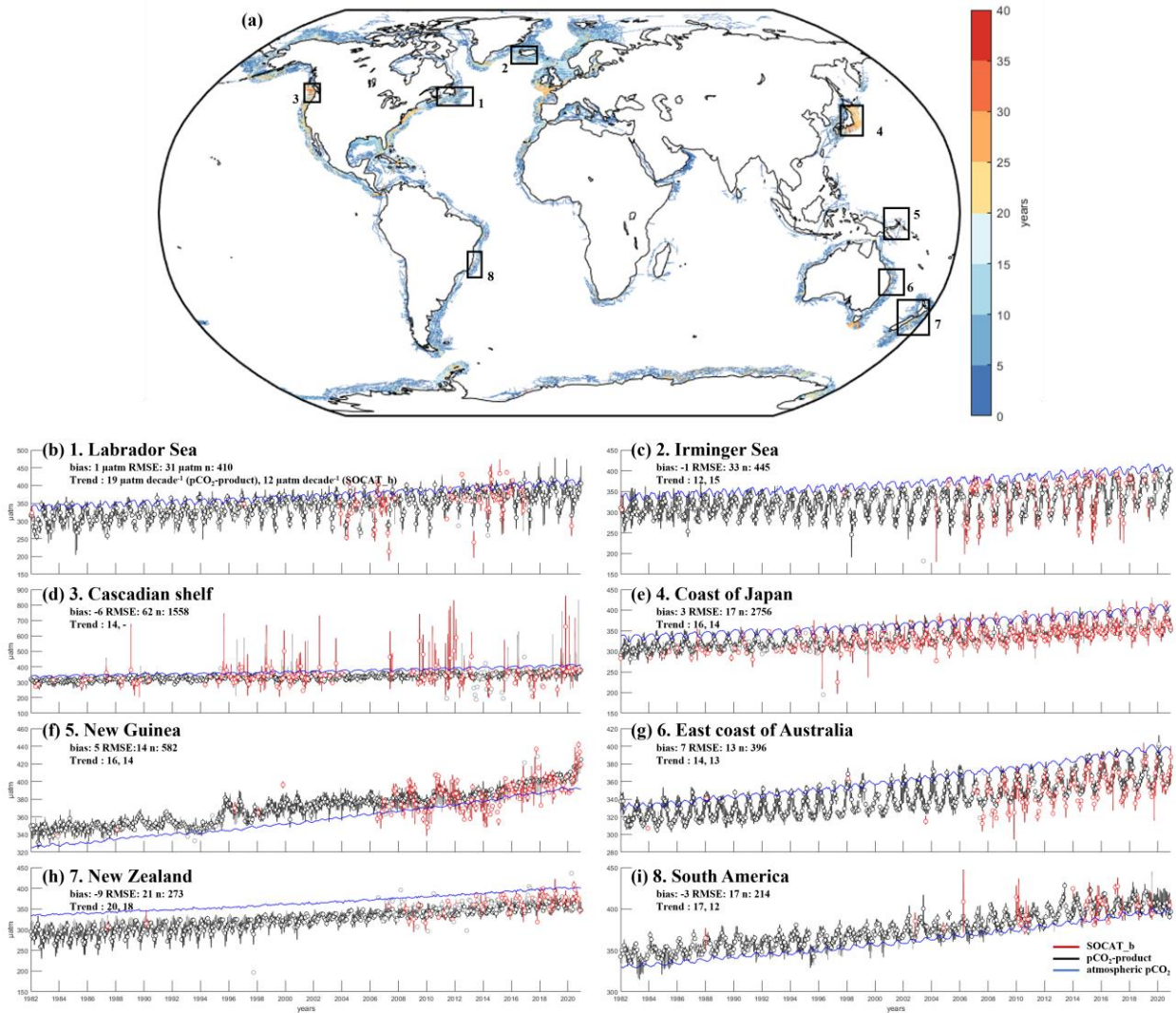



Figure 4. (a) Temporal coverage (in years) where pCO₂ measurements extend over x years in SOCAT_b. The location of the 8 coastal sites for which we present pCO₂ times series (black boxes) is also shown. (b-i) pCO₂ times series (in µatm) from the reconstructed pCO₂-product (in black), from SOCAT_b (in red), and from the atmospheric pCO₂ (in blue). For each region, we only select grid cells that extend over 30 years of observations in SOCAT_b. Medians are represented in circles and the vertical bars represent the monthly pCO₂ intra spatial variability in the region. For each region, we report the bias (µatm), RMSE (µatm) and number of cells for the calculation between the reconstructed pCO₂-product and SOCAT_b as well as their respective long-term pCO₂ 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). The Cascadian shelf has no value for the SOCAT_b trend since no significant trend is detected based on a Mann-Kendall statistical test.

905

910

915

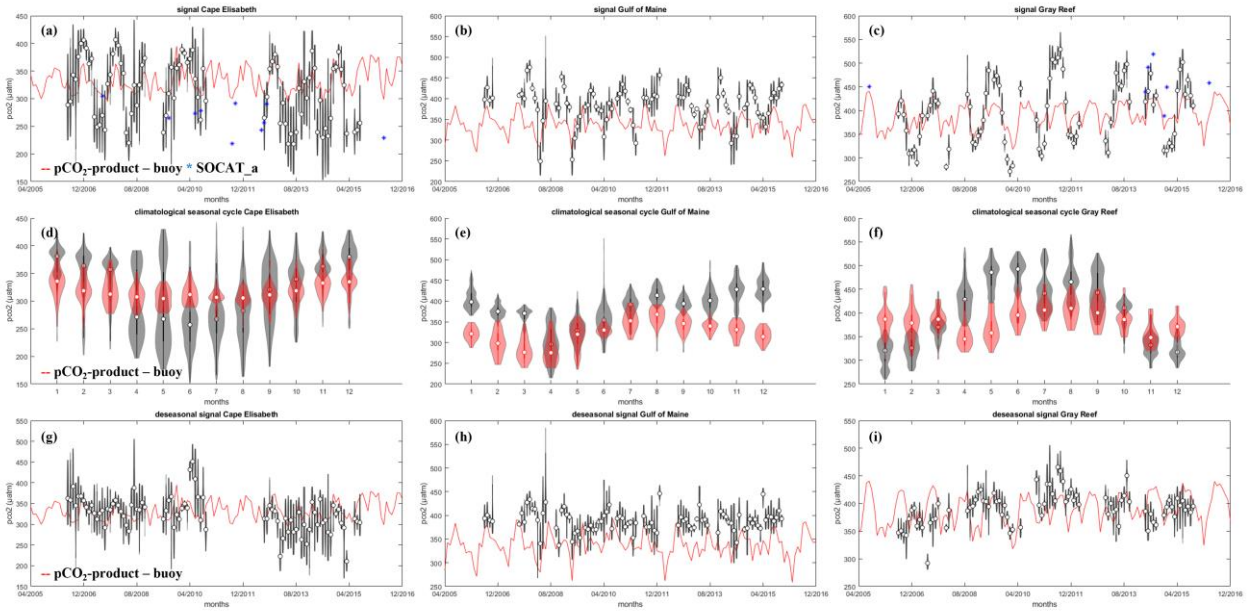



Figure. S3. Comparison between the reconstructed coastal pCO$_2$-product (red) against times series from 3 coastal buoys (black) namely in Cape Elizabeth (NDBC Buoy 46041), in Gray's Reef (NDBC Buoy 41008) and in the Gulf of Maine (Coastal Western Gulf of Maine mooring, Sutton et al., 2019). (a-c) Monthly 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)

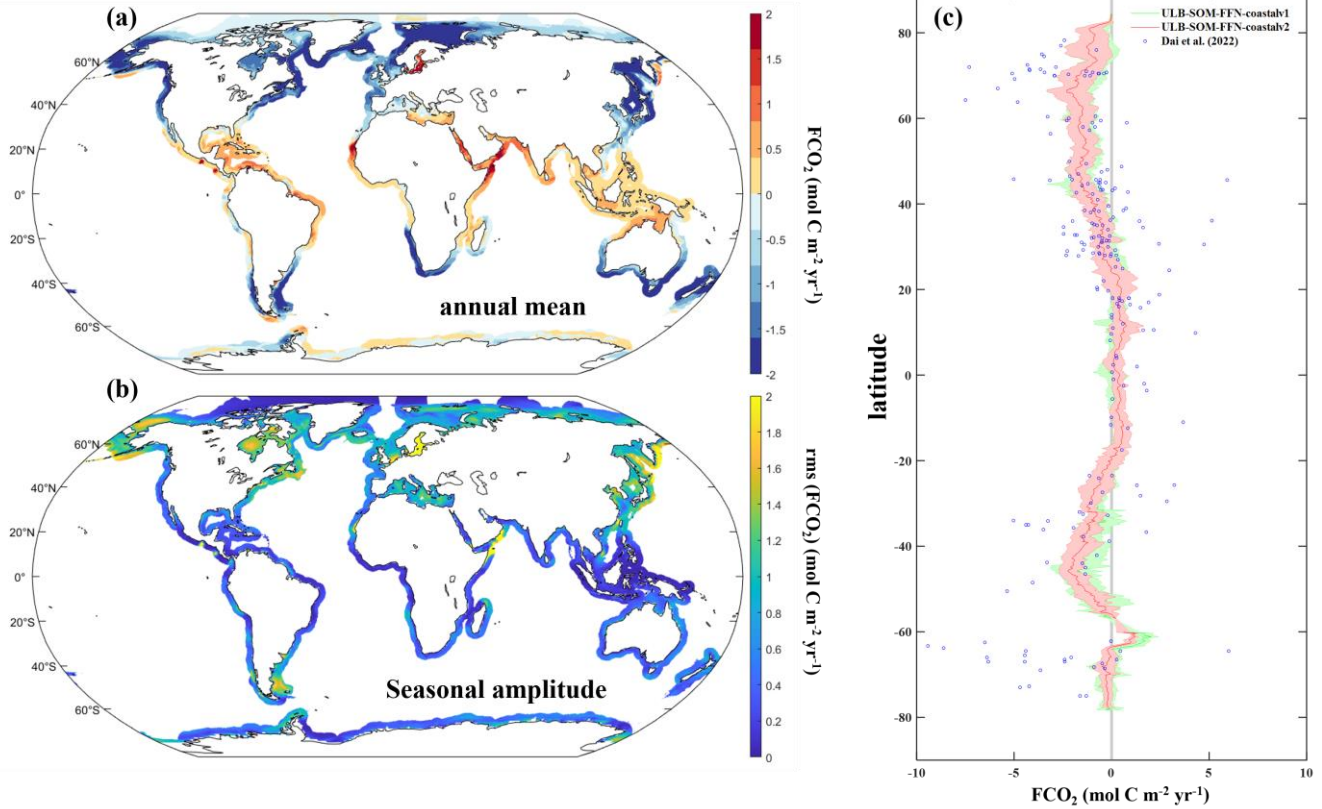
920 to (g) for the reconstructed pCO$_2$ represent the pCO$_2$ variability of the month x around its monthly climatological pCO$_2$ mean.

925

Figure 5. (a) Spatial distribution of the annual average air-sea CO_2 exchange (FCO$_2$, in mol C m^{-2} yr^{-1}) and **(b)** seasonal FCO$_2$ variability (expressed as the Root Mean Square (RMS, mol C m^{-2} yr^{-1})) calculated with the reconstructed coastal pCO$_2$-product (ULB-SOM-FFN-coastalv2, 1982-2020 climatology). The latitudinal mean FCO$_2$ distribution (red line) and its associated longitudinal variability (red shade) is presented in panel (c). This latter is compared to the FCO$_2$ calculated with the ULB-SOM-930   FFN-coastalv1 pCO$_2$-product (in green, Roobaert et al., 2019) and against a synthesis of 214 regional FCO$_2$ estimations (Dai et al., 2022; blue dots). For consistency in the comparison in panel (c) we applied the same coastal delimitation than in Dai et al. (2022) and Roobaert et al. (2019) to the FCO$_2$ ULB-SOM-FFN-coastalv2 product i.e., we used the shelf break as the outer limit of the coastal domain (narrow coastal ocean). Panel (c) is also reconstructed based on an overlap period between the three products (1998-2020; except FCO$_2$ ULB-SOM-FFN-coastalv1 which is limited to the 1998-2015 period).

935

940

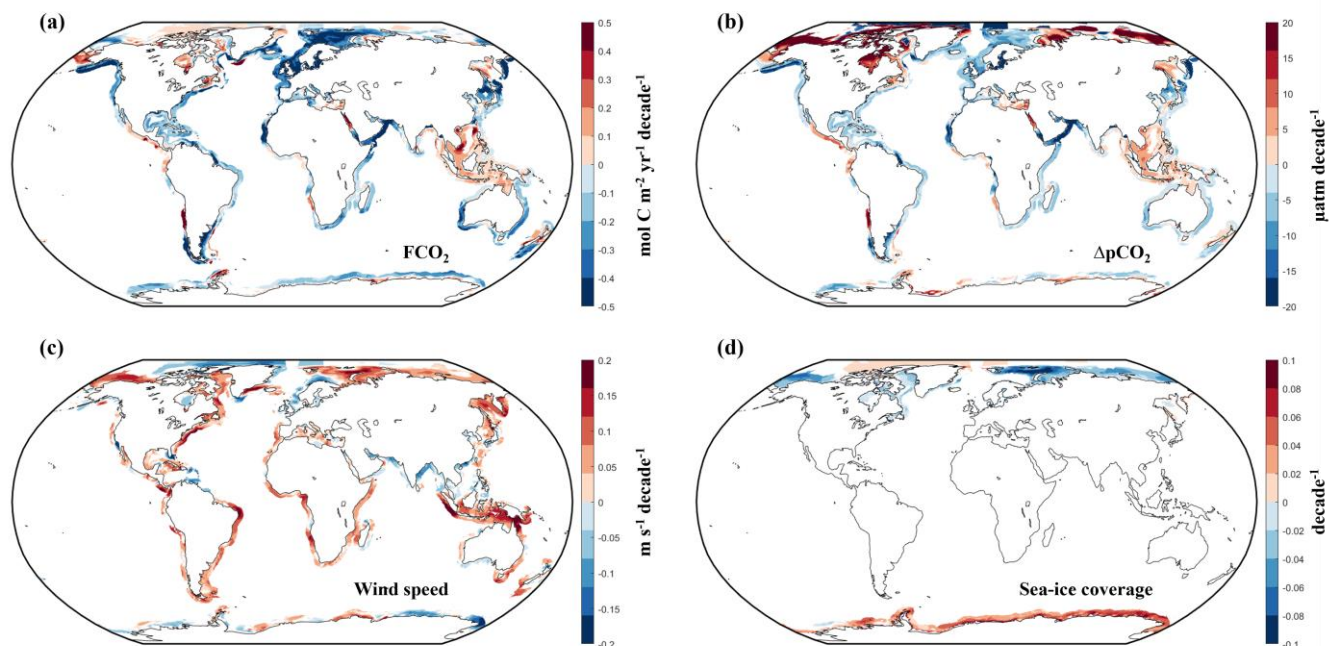



Figure 6. Long-term trend of~~Rate of change in~~ (a) the coastal FCO_2 (in mol C m^{-2} yr^{-1} decade^{-1}), (b) the air-sea pCO_2 gradient (ΔpCO_2, in µatm decade^{-1}), (c) the wind speed at 10 meters above the sea surface (m s^{-1} decade^{-1}) and (d) the sea-ice coverage (decade^{-1}) from 1982 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.

950

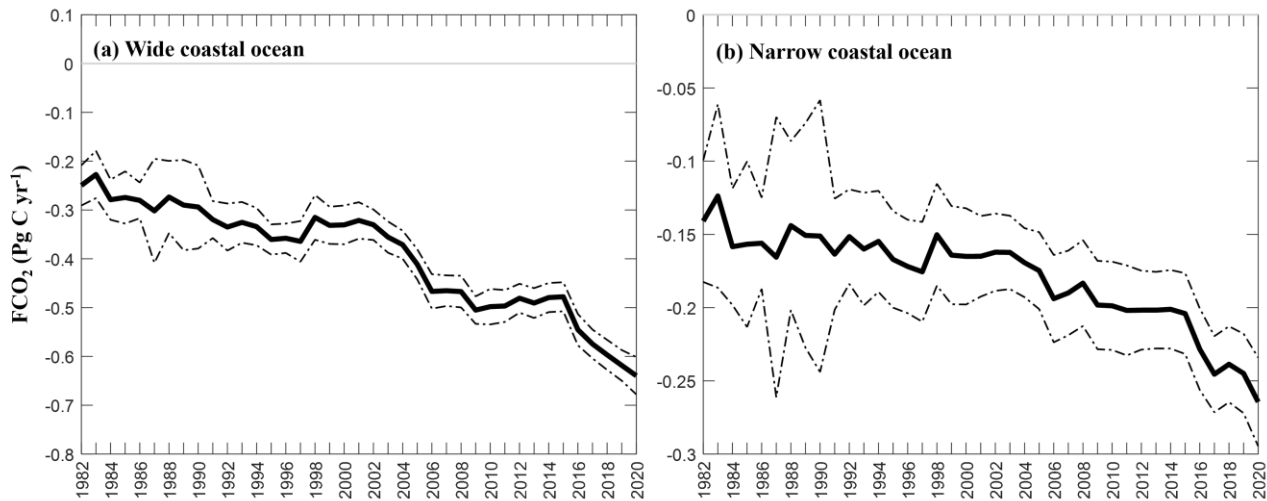



Figure 7. Evolution of the global coastal CO_2 sink (FCO_2, in Pg C yr^{-1}) over time using the reconstructed coastal pCO_2-product (black line) with its associated uncertainties (dashed black line, see Sects. 2.4 and 3.4 for further details). We use a 300 km distance from the coast as the outer limit of the coastal domain (wide coastal ocean) in panel (a) and (b) the shelf break as the outer limit of the coastal domain (narrow coastal ocean).

Table 3. Global FCO_2 and pCO_2 uncertainties calculated for the reconstructed data-products using the wide (narrow) definition of the coastal domain.

FCO₂ uncertainty (Pg C yr⁻¹)		$\sigma_{FCO_2} = \sqrt{\sigma^2_{\Delta pCO_2} + \sigma^2_k + \sigma^2_{wind} + \sigma^2_{ice}}$
σ_{FCO_2}	0.03 (0.01)*	Total FCO₂ uncertainty
$\sigma_{\Delta pCO_2}$	0.02 (0.01)	Uncertainty of the air-sea pCO₂ gradient
σ_k	0.01 (0.01)	Uncertainty associated with the choice of the k-formulation
σ_{wind}	0.01 (0.002)	Uncertainty associated with the choice of the wind speed product
σ_{ice}	0.01 (0.01)	Uncertainty associated with the choice of the sea-ice product
pCO₂ uncertainty (µatm)		$\theta_{pCO_2} = \sqrt{\left(\frac{\theta_{obs}}{\sqrt{N}}\right)^2 + \left(\frac{\theta_{grid}}{\sqrt{N}}\right)^2 + \left(\frac{\theta_{map}}{\sqrt{N_{eff}}}\right)^2}$
θ_{pCO_2}	0.63 (1.021)	Total uncertainty of the oceanic pCO₂
θ_{obs}	3.08 (3.18)	Uncertainty associated with the sampling on the field of the observations from the SOCAT
θ_{grid}	5.706 (8.01)	Uncertainty associated with the gridding of the observations from SOCAT into 0.25° monthly meshed maps
θ_{map}	28.8229 (35.276)	Uncertainty derived from the comparison between the reconstructed pCO₂ and the observed gridded pCO₂ from the SOCAT

* The numbers correspond to uncertainties calculated using a wide coastal delimitation, while those enclosed in brackets represent uncertainties calculated using a narrow coastal delimitation.

965