Global analysis of surface ocean CO_2 fugacity and air-sea fluxes with low latency

Thi-Tuyet-Trang Chau[1], Frédéric Chevallier[1], and Marion Gehlen[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

Key Points:

- We demonstrate the capacity of statistical models to generate global maps of fCO_2 and air-sea flux with a latency reduced to one month.
- A decrease in the CO_2 source for January to August 2023 diagnosed in the tropical Pacific coheres with the retreat of the La Niña event.
- An unusual northeastern Atlantic sink reduction diagnosed for June 2023 is linked to record heat and exceptionally low winds.

Corresponding author: Thi-Tuyet-Trang Chau, thi.tuyet.trang.chau@gmail.com

Abstract

The Surface Ocean CO_2 Atlas (SOCAT) of CO_2 fugacity (fCO_2) observations is a key resource supporting annual assessments of CO_2 uptake by the ocean and its side effects on the marine ecosystem. SOCAT data are usually released with a lag of up to 1.5 years which hampers timely quantification of recent variations of carbon fluxes between the Earth System components, not only with the ocean. This study uses a statistical ensemble approach to analyse fCO_2 with a latency of one month only based on the previous SOCAT release and a series of predictors. A retrospective prediction for the years 2021-2022 is made to test the model skill, followed by the generation of fCO_2 and fluxes from January to August in 2023. Results indicate a modest degradation of the model skill in prediction mode and open the possibility to provide robust information about marine carbonate system variables with low latency.

Plain Language Summary

There is a growing need to monitor carbon emissions and removals over the globe in near real time in order to correctly interpret changes in CO_2 concentrations as they unfold. For the oceans, the best information comes from measurements of the surface ocean CO_2 fugacity (fCO_2) by the international marine carbon research community. So far, this data is mostly available 6 to 18 months behind real time after collection, qualification, harmonization, and processing. Here, we show that a set of biological, chemical, and physical predictors available in near-real time, allows the information contained in the "old" fCO_2 measurements to be transferred over time. Based on a statistical technique, we combine all these data sources to estimate global monthly maps of fCO_2 and of CO_2 fluxes at the air-sea interface within one month behind real time and with good accuracy.

1 Introduction

The ocean is a sink taking up about 26% of atmospheric carbon dioxide (CO_2) and 90% of the heat-induced largely by anthropogenic greenhouse gas emissions (Canadell et al., 2021; Friedlingstein et al., 2022). A side effect of the ocean's role as a global climate modulator is the increase in seawater acidity, which dramatically affects marine ecosystems (Hopkins et al., 2020; Doney et al., 2020; Cooley et al., 2022). The global ocean carbon sink is proportional to CO_2 human emissions only at the decadal scale. On shorter time scales, it varies with the climate (mostly temperature and winds), with a dependency that

also varies from basin to basin given their respective geographical, dynamic, and biological specificities (Rödenbeck et al., 2015; Landschützer et al., 2016; Gruber et al., 2023).

Measurements of surface ocean CO_2 fugacity (fCO_2) from ships, drifters, moorings, and autonomous surface platforms are the main reference to document the actual variation of air-sea fluxes ($fgCO_2$) in space and time (Friedlingstein et al., 2022) because the two are linearly related. Long-term efforts in maintaining and expanding international observing networks together with a coordinated data aggregation of the Surface Ocean CO_2 Atlas database - SOCAT (Bakker et al., 2016, 2023) have provided millions of individual fCO_2 observations since the 1950s and associated gridded products. However, fCO_2 data are poorly sampled leaving out most areas for some or all of the year. Statistical data-based reconstructions of fCO_2 (Rödenbeck et al., 2013; Landschützer et al., 2016; Gregor & Gruber, 2021; Chau et al., 2022b) have emerged to gap-fill the SOCAT database using auxiliary data, resulting in reconstructions of fCO_2 global monthly maps. They are still the topic of active research to improve the reconstruction quality, but these maps lag behind real time by 0.5 to 1.5 years: the update of the SOCAT archive follows an annual pace with a public release usually in June after measurement collection, quality control, and processing. This lag is problematic for the documentation of the carbon cycle as it evolves, while the main variables of the carbon cycle are progressively integrated within operational programmes with much faster data releases. A prominent example of operational programmes in need of a reduced time lag is the operational observation-based anthropogenic CO_2 emissions monitoring and verification support capacity (CO_2MVS) that the European Commission is building under its Copernicus Earth Observation programme (e.g., Janssens-Maenhout et al. (2020)). As its observational component relies heavily on satellite observations of CO_2 in the atmosphere, which is affected by the ocean as well as terrestrial emissions and removals, better estimates of fCO_2 would result in efficient estimates of air-sea fluxes and thence benefit air-land flux accuracy, in addition to being directly interesting to users. The CO_2MVS fits within the Global Greenhouse Gas Watch, an even larger greenhouse gas monitoring infrastructure that the World Meteorological Organization (WMO) is setting up (`https://public.wmo.int/en/media/press-release/world-meteorological-congress-approves-global-greenhouse-gas-watch`, last access: 20/9/2023).

Here, we demonstrate the capability to retrieve global monthly maps of fCO_2 from SOCAT data and then to generate the corresponding fields of air-sea fluxes with a lag reduced to one month. To do that, we extend the work of Chau et al. (2022b) who have been

gap-filling SOCAT gridded data within the framework of the Copernicus Marine Environment Monitoring Service (CMEMS) based on an ensemble of feed-forward neural network models (also referred to as CMEMS-LSCE-FFNN) and a set of biological, chemical, and physical predictors. While Chau et al. (2022b) made the dates of the predictors and the date of the gridded SOCAT data coincide, we turn to a prediction mode in which the relationship found between the predictors and the SOCAT data more than 6 months before is kept. Section 2 below describes the method. We test the approach in the years 2021-2022 by examining the retrospective prediction skill based on the available SOCAT data. Then we expand model prediction of fCO_2 and generate $fgCO_2$ up to present with a latency of 1 month: data access via the Institut Pierre-Simon Laplace (LSCE/IPSL) data center, `https://dods.lsce.ipsl.fr/invsat/FFNN_low-latency/`. The results include the finding of anomalous variations in regional CO_2 uptake and release by the ocean predicted in January to August 2023, as described in Section 3. Section 4 draws the main conclusions of the study.

2 Materials and Methods

CMEMS-LSCE-FFNN (Chau et al., 2022b) is built on machine-learning techniques. It consists of an ensemble of feed-forward neural network (FFNN) models. This ensemble approach was developed at LSCE in order to reconstruct surface ocean carbonate system variables and to support the operational distribution of such datasets by CMEMS since 2019 (Product identity: MULTIOBS_GLO_BIO_CARBON_SURFACE_REP_015_008, `https://doi.org/10.48670/moi-00047`, last access: 22/9/2023). The CMEMS-LSCE-FFNN fields cover the global ocean at a resolution of $1° \times 1°$ currently and for the period since the year 1985 at monthly resolution.

Under the hood, these FFNN models represent nonlinear mappings of fCO_2 against a set of predictors. Monthly gridded observation-based products of fCO_2 from SOCAT (Bakker et al., 2016) are used as the target data in model fitting. fCO_2 predictors are environmental variables: sea surface temperature (SST), sea surface salinity (SSS), sea surface height (SSH), chlorophyll-a (Chl-a), mix-layer-depth (MLD), CO_2 surface mole fractions (xCO_2), climatological fCO_2 ($fCO_2{}^{\mathrm{clim}}$), and geographical coordinates (latitude and longitude). Product resources of input datasets are detailed in Table S1. CMEMS-LSCE-FFNN comprises monthly adaptive FFNN models for which the fCO_2 and predictor datasets available within a time span of 3 months for all the years since 1985 (the reconstruction month

excepted) are used in the fitting phase. SOCAT fCO_2 in the reconstruction month is only used in model evaluation. The ensemble of multi-FFNN models was designed by randomly splitting two-thirds of the 3-month sliding datasets for training and the rest for model test (Chau et al., 2022b). From the ensemble reconstructions, the model best estimate (ensemble mean) and 1σ - model uncertainty (ensemble standard deviation) of fCO_2 are derived at the desired resolution.

Here we revisit the two versions of CMEMS-LSCE-FFNN referred to as FFNNv2021 and FFNNv2022. These two models respectively used SOCATv2021 and SOCATv2022 datasets (Bakker et al., 2021, 2022) as the target input data of fCO_2. Note that SOCAT has been annually published mid-June. Due to the delay mode for data collection, reprocessing, and qualify control, SOCAT provides gridded data up to the year before the publication date (see Bakker et al. (2016, 2023) for instance). For the period 1985-2021, SOCATv2022 offers an amount of roughly 311700 monthly 1-degree gridded data, 5000 more than SOCATv2021 (Table S3a). The data increase in SOCATv2022 is mostly distributed within the last three years due to the late availability of some data sources (Figure 1). However, SOCATv2021 has more data before 2018, up to at least 1000 more in some years (e.g., 2011 and 2012) due to an erroneous flagging of some data (Bakker et al., 2021). Despite this feature, the two corresponding FFNN reconstructions do not exhibit large systematic offsets in their fCO_2 estimates (Chau et al., 2022a).

For all experiments in this study, the ensemble size (i.e., number of FFNN model runs) is set to 50. FFNN with 50 ensemble members has less computational complexity than with the usual size of 100 but it shows similar reconstruction skill (Chau et al. (2022b); Figure S2). The same input data of predictors is fed to the two FFNN model runs (Table S1). The FFNNv2021 (respectively FFNNv2022) model relies on SOCATv2021 (respectively SOCATv2022) and predictor datasets in 1985-2020 (respectively 1985-2021). This allows deriving the ensemble global reconstructions of fCO_2 over the 36-year and 37-year periods, accordingly. The ensemble of FFNN models is then applied to predict fCO_2 given the set of predictors in the years 2021-2022 for version 2021 and in the year 2022 for the latter. The quality assessments are made for (1) the two global reconstructions in the period 1985-2020, (2) FFNNv2021 one-year prediction against FFNNv2022 one-year reconstruction in 2021, and (3) FFNNv2021 two-year prediction against FFNNv2022 one-year prediction in 2022. Model performances will be qualified with the latest SOCAT data, i.e., SOCATv2023 (Bakker et al., 2023). The number of evaluation data for prediction in the years 2021 and

2022 over the global ocean is 10908 and 8602, respectively (Table S3a), which is statistically sufficient for significant validation.

Model skills are examined from global to sub-basin scale. Here we consider the sub-basins defined by the REgional Carbon Cycle Assessment and Processes2 project (`https://github.com/RECCAP2-ocean/RECCAP2-shared-resources/tree/master/data/regions`, last access: 20/3/2023). Due to a lack of evaluation data in several RECCAP2 biomes, we aggregate some of them, yielding 14 provinces in total (see Table S2 and Figure S1). These ocean provinces, therefore, differ from the original biomes proposed by Fay and McKinley (2014). Apart from the Northern Indian Ocean (11.NIO), the number of data for prediction evaluation ranges from 133 (12.SIO, i.e., Southern Indian Ocean) to 2350 (2.NA-SS, i.e., North Atlantic seasonally stratified) in the year 2021 and from 73 to 2265 in the year 2022.

For the actual prediction in 2022 and 2023, the latest model (FFNNv2022) has been run given monthly data of predictors (Table S1) in the year 2022 to present. We choose to release the maps of fCO_2 and $fgCO_2$ for the previous month on the 15th of each month.

3 Evaluation and Discussions

3.1 Reconstruction and Prediction of CO_2 fugacity in 1985-2022

3.1.1 Global qualification

FFNNv2021 and FFNNv2022 share consistent global RMSD and determination coefficient r^2 (Figure 1 and Table S3). Between 1985 and 2020, the two reconstructions inherit the same RMSD of 19.1 μatm and r^2 of 0.78 (Table S3b). Improvement in the global reconstruction skill of FFNNv2022 in recent years (Figure 1b) is moderate despite 5000 additional fCO_2 data in the model training (Figure 1a). In detail, these 1.7% additional data in SOCATv2022 (311694 in total) in 1985-2021 correspond to 9615 data added in 2021 and 4278 data removed from SOCATv2021 in 1985-2020 (see the spatial distribution of removal data in Figure S2c).

The RMSD variability before 2018 (Figure 1b) is likely linked to changes in the data sampling in regions with high spatiotemporal variability of fCO_2 (see Gregor et al. (2019); Chau et al. (2022b) for further analysis). However, the difference between the RMSD of the two reconstructions is negligible then, as it fluctuates within $[-0.1, 0.1]$ μatm. During the last four years, a monotonous increase in RMSD (Figure 1b) coexists with a decrease in

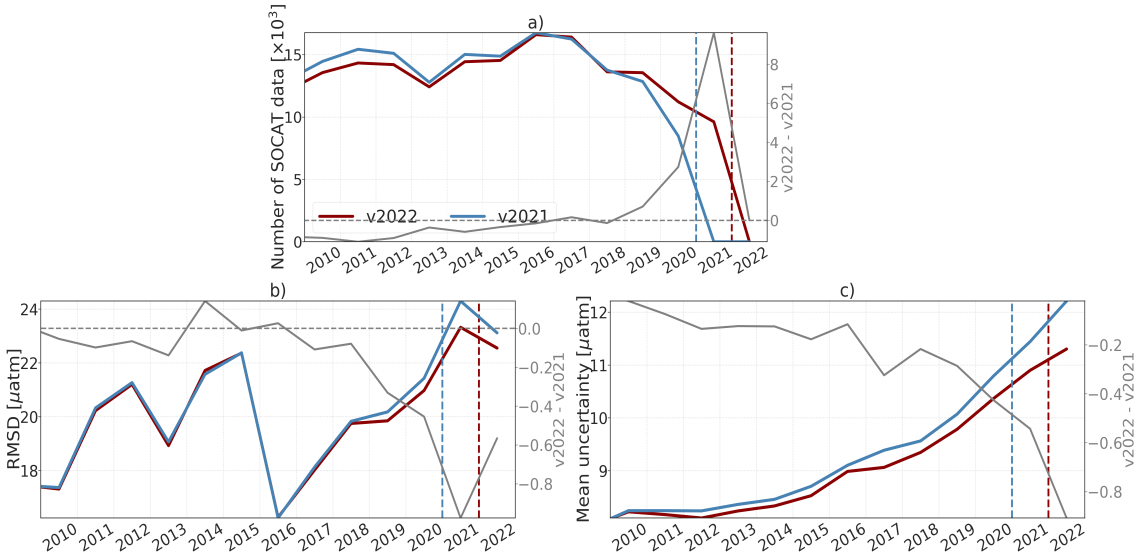

Figure 1. (a) Number of data per year in SOCATv2021 and SOCATv2022, (b) RMSD of FFNNv2021 and FFNNv2022 against SOCATv2023 fCO_2, (c) yearly global mean uncertainty (1σ). Differences between the two versions are shown with a grey solid curve with values on the right y-axis whereas the grey solid curve below 0 (grey dashed horizontal line). The blue and red vertical lines mark the start of the prediction mode for FFNNv2021 and FFNNv2022, respectively.

the number of SOCAT data (Figure 1a), and the FFNNv2021 reconstruction slightly, but increasingly, underperforms compared to FFNNv2022. In 2021 and 2022, the FFNNv2021 prediction RMSD is 24.3 μatm and 23.1 μatm, respectively, roughly $0.5-1$ μatm higher than that of the FFNNv2022 reconstruction and prediction (Table S3). Likewise, the variation of SOCAT fCO_2 is reproduced with high r^2 values (0.74 and 0.75), close to the one-year reconstruction and prediction of FFNNv2022 (0.76) for the years 2021-2022.

The yearly-mean uncertainty over the global ocean (Figure 1c) is computed by weighting the model estimated uncertainty (ensemble spread) per grid cell (σ) with the geographical area. The two reconstructions before the year 2015 are rather stable with an uncertainty about 8.5 μatm. The increase in FFNNv2021 [v2022] model uncertainty from 8.7 μatm [8.5 μatm] to 10.8 μatm [10.4 μatm] between 2015-2020 follows a decrease in observation-based data from 14877 [14533] to 8482 [11217] (Figure 1a). In the year 2021, the FFNNv2021 uncertainty of predicted fCO_2 (11.4 μatm) is slightly higher than that of the FFNNv2022 reconstruction but the offset between the two values is as small as 0.5 μatm (Figure 1c). The

prediction uncertainty in 2022 increases by $0.4-0.8$ μatm for the two models (FFNNv2021: 12.2 μatm, FFNNv2022: 11.3 μatm).

3.1.2 Regional assessment

Model reconstruction and prediction skills are assessed over 14 ocean provinces (Figure S1 and Table S2) in the years 1985-2020 and 2021-2022 (1985-2021 and 2022) for FFNNv2021 (FFNNv2022). Results of the regional evaluation are summarized in Figure 2 and Table S4. The two FFNN models perform with a similar skill in reconstruction mode (1985-2020) over all ocean provinces. Evidently, their reconstructions share consistent patterns in regional-mean fCO_2 (Figure 2b) and in the spatial and temporal variations (Figures S4abc and S7) with systematic biases below 1 μatm for most of the basins (Table S4). Differences in uncertainty estimates and RMSD do not exceed 0.5 μatm while those in r^2 are nearly the same (Figure 2cde and Table S4).

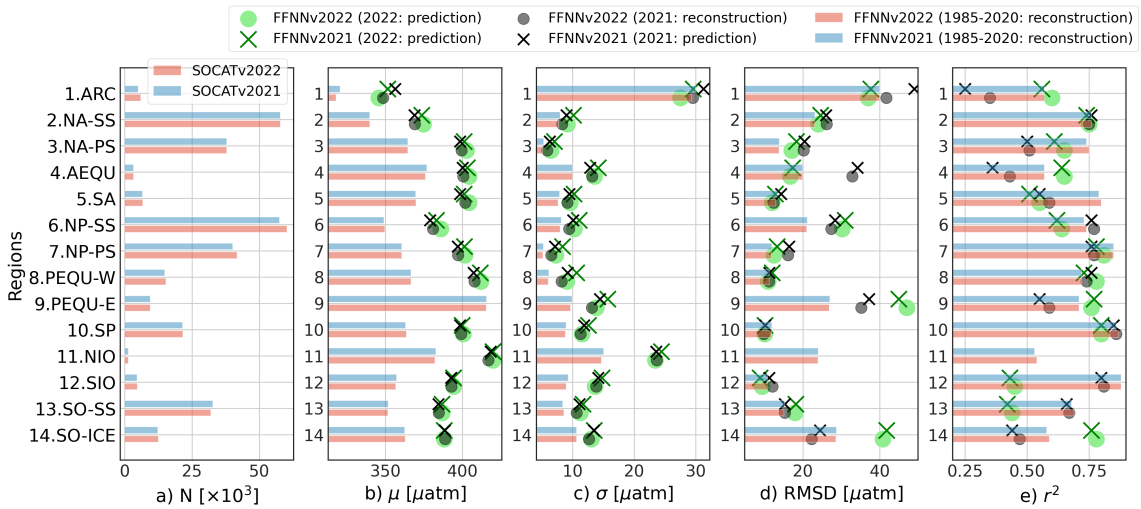

Figure 2. Regional comparisons of the two FFNN reconstructions in 1985-2020 (bars) and of the FFNNv2021 prediction versus the FFNNv2022 reconstruction [prediction] in 2021 [2022] (objects) in terms of (a) N- number of SOCAT monthly gridded data used in model fitting, (b) μ- mean fCO_2, (c) σ-mean uncertainty, (d) RMSD model-data deviation, and (e) r^2 model-data correlation.

In the years 2021-2022, RMSD (r^2) of the FFNN prediction does not change from the full-period reconstruction by more than about 5 μatm (0.1) over many sub-basins (e.g., 2.NA-SS, 7.NP-PS, 8.PEQU-W, 10.SP, 12.SIO, and 13.SO-SS). As expected, FFNNv2022 (one-year prediction) performs slightly better than FFNNv2021 (two-year prediction) in the

2022 prediction for many regions (Figures 2de and Table S4). However, the differences in regional skill scores of the two models are substantially small, i.e., below 3 μatm for RMSD and 0.05 for r^2. These results suggest a high confidence level in FFNN prediction for a few years ahead. The analysis of the spatial distribution and of the time series (Figures 2, S4, and S7) also reveals consistent features (horizontal gradients of fCO_2 and seasonality to long-term variations) from the reconstruction years to the prediction years. fCO_2 increases over time (see f.i., 7.NP-PS, 8.PEQU-W, 12.SIO) following the trend in atmospheric CO_2 concentration. Among the fCO_2 predictors, xCO_2 stands out with its large increasing trend that brings some xCO_2 data used in the prediction above the range of those used in the training. The growth of atmospheric CO_2 is the primary factor driving the increase in sea surface fCO_2 (Bates et al., 2014; Gruber et al., 2019; Landschützer et al., 2019; Friedlingstein et al., 2022). The prediction skill, however, does not degrade compared to the reconstruction as the annual increment of fCO_2 is typically smaller than its intra-annual variability (Figure S6). The latter is dominantly driven by temperature-dependent CO_2 solubility and biological processes (Takahashi et al., 2002; Gallego et al., 2018; Rustogi et al., 2023). The range of the pre-2021 [pre-2022] training datasets of physical and biological predictors (e.g., SST, Chl-a) remains similar to that including input data in the next year, seasonality to multi-month variations of fCO_2 in the years 2021-2022 can be, therefore, propagated with these covariates overall. The majority of SOCAT fCO_2 data for 2021 [2022] stays within the full range of training data which also supports FFNNs to achieve a skillful prediction (Figure S3). Further analysis of FFNN prediction skills over ocean basins is presented in the Supporting Information document.

3.2 Prediction of air-sea CO_2 fluxes in 2022-2023

The previous results emphasize the skill and reliability of FFNN models in both reconstruction and prediction of CO_2 fugacity (fCO_2). In this section, we will use the FFNNv2022 predicted fCO_2 field to generate corresponding air-sea fluxes ($fgCO_2$) and analyze preliminary results for 20 months, from January 2022 to August 2023. $fgCO_2$ is given in molC.m^{-2}.yr^{-1} for a flux density and in PgC.yr^{-1} for integration over ocean basins (see Supporting Information for details of flux calculation and analysis). FFNNv2022 predicts a reduction in the global ocean uptake of CO_2 for 2022 (2.25 ± 0.5 PgC.yr^{-1}) compared to the previous year (2.36 ± 0.43 PgC.yr^{-1}). When adjusting the estimated global net fluxes with the riverine outgassing of CO_2 of 0.65 PgC.yr^{-1} (Regnier et al., 2022) and the total

ocean surface area (FFNNv2022 data covers 95% of the global ocean), one obtains the estimates of anthropogenic ocean carbon uptake consistent with the 2022 projection proposed by Friedlingstein et al. (2022): the anthropogenic ocean sink in 2021 was 2.9 ± 0.4 PgC.yr^{-1} remains unchanged for the year 2022. This evidence supports their hypothesis that the persistence of cooling climate patterns (La Niña conditions) weakened CO_2 ocean uptake in 2021-2022 (high peaks appeared mid-2022, Figure S9). FFNNv2022 predicts a global net flux of 2.45 ± 0.56 PgC.yr^{-1} for January to August 2023, the enhancement of global ocean uptake compared to that in 2022 (2.17 ± 0.50 PgC.yr^{-1}) is synchronous with the retreat of La Niña.

The model prediction retains the seasonal to interannual variations of fCO_2 and $fgCO_2$ in the pre-2022 reconstruction over many ocean basins (Figures S6 and S8). One of the remarkable changes is observed at the equatorial Atlantic (4.AEQU), where the regional mean fCO_2 increases by 4.2 μatm from the year 2021 to 2022 (Figure S6). However, such a high increment in the AEQU fCO_2 is negligible in terms of its contribution to the global net ocean sink variations between the two years (Figure S8 and Table S5). In Rödenbeck et al. (2015) [Figures A2 and A4], it is also illustrated that $pCO_2{}^{sea}$ ranges from 350 μatm to 400 μatm over an 18-year period while the AEQU net flux has performed with nearly constant magnitude. Its low interannual variability is in contrast with the eastern equatorial Pacific (9.PEQU-E) showing the strong impact on temporal variations of the global net sink (Figure S8). The signature of fCO_2 dampening (-9.4 μatm) over PEQU-E in Jan to August of 2022-2023 is opposed to its increasing (1.8 μatm) with respect to 2021-2022 (Figure S6). As illustrated in Figures S8 and S9, FFNNv2022 prediction marks an anomalous decline of CO_2 source in the first eight months of 2023 (-0.30 ± 0.04 PgC.yr^{-1}) compared to that of 2022 (-0.37 ± 0.04 PgC.yr^{-1}). This reduced source of 0.07 PgC.yr^{-1} in PEQU-E contributes to 25% of the increase in the global ocean sink mentioned above. The reduction in the PEQU-E CO_2 source marks the transition from La Niña to El Niño announced by e.g., WMO (`https://public.wmo.int/en/media/press-release/world-meteorological-organization-declares-onset-of-el-ni%C3%B1o-conditions`, last access: 05/9/2023).

While the onset of El Niño over the tropical Pacific (Figure S9a) had been driving the reduction of ocean CO_2 emission La Niña anomalies (Figure S8), an exceptional warming event occurred and spread over the north Atlantic since May-June 2023 (Copernicus Climate Change Service: `https://climate.copernicus.eu/copernicus-record-north`

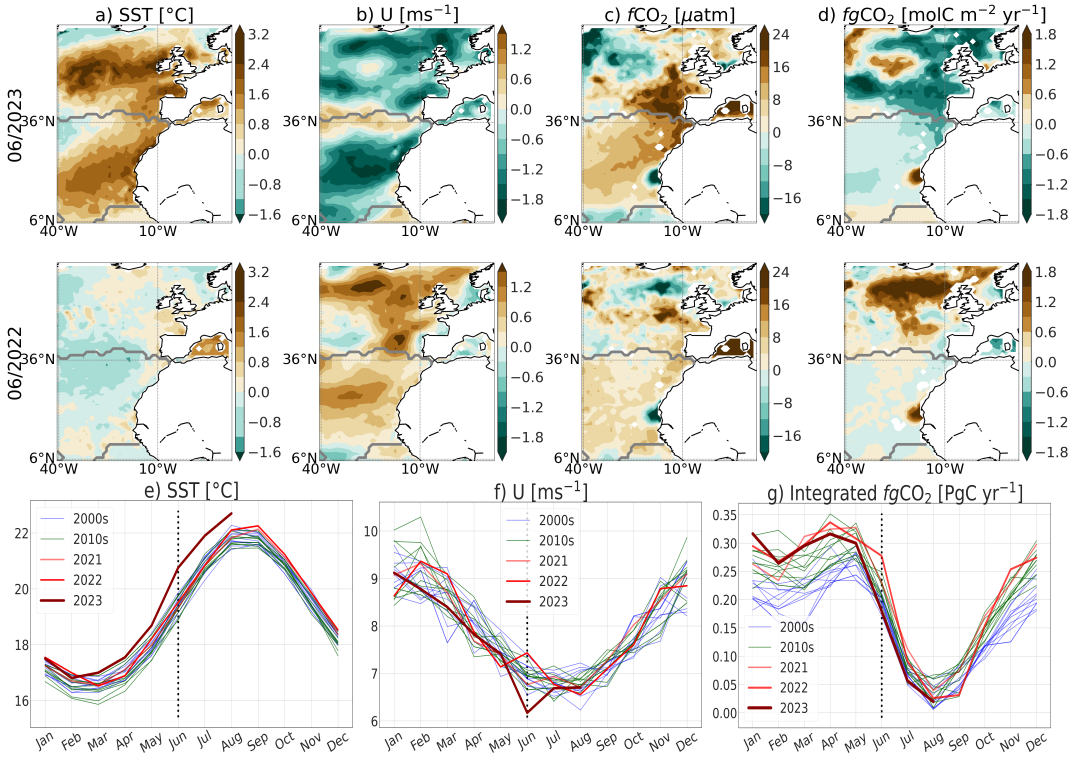

Figure 3. Top panels (a-d): anomalies observed in FFNNv2022 prediction of fCO$_2$ and fgCO$_2$ (c,d) follow an extreme marine heatwave event (a,b) over the northeastern Atlantic in June 2023 relative to June 2022 (top panels). Anomalies of surface temperature (SST), wind speed (U), fCO$_2$, and fgCO$_2$ are computed by subtracting long-term trends and seasonal climatologies relative to the years 1985-2022. Grey curve represents regional division (Figure S1). Bottom panels (e-g): regional seasonal cycles of SST, U, and integrated air-sea fluxes since 2000s.

-atlantic-warmth-hottest-june-record-globally, last access: 20/9/2023). It substantially lessened the ocean CO$_2$ uptake (Figure 3). Based on the CMEMS SST analyses (Table S1), June 2023 corresponds to the first marine extreme heatwave in the northeastern Atlantic (40°W-12°E, 5°N-65°N) with an average SST anomaly about 1.1°C (Figure 3ae). As a comparison, the June anomaly had been typically in a range of −0.5°C to 0.5°C for the past three decades. In 2023, SST anomalies even exceeded 1.5°C over the northeastern Atlantic seasonally stratified biome (NA-SS, 36°N northward). FFNNv2022 predicts an enhancement in fCO$_2$ (Figure 3c) following the anomalous warmth in the northeastern Atlantic which is not seen in June 2022 (Figure 3a). As other environmental factors (e.g., salinity and chlorophyll-a) have no remarkable anomalies over this ocean basin (Figure S10), warming primarily reduces CO$_2$ solubility and that leads to substantially high surface par-

tial pressure of CO_2 (Figure 3c). fCO_2 anomalies were mostly between 4 μatm and 12 μatm in the subtropics, i.e., north Altlantic permanently stratified region (NA-PS) and increased eastward. FFNNv2022 records the largest fCO_2 anomalies in the southeast of NA-SS towards the European coast with values above 16 μatm. Consequently, the predicted air-sea fluxes in June 2023 (Figure 3d) suggest lower-than-average CO_2 uptake capability. While $fgCO_2$ slightly decreased throughout the NA-PS, an anomalous drawdown is found in the NA-SS exceeding -0.6 molC.m^{-2}.yr^{-1} (equivalent to roughly a reduction in ocean CO_2 uptake of 0.11 PgC.yr^{-1}). It is noteworthy that a decline in ocean CO_2 uptake is strengthened if surface wind speeds (U) are lowering and fCO_2 increases. Accompanied by the largest positive SST anomaly in June 2023, there is an unusual reduction in wind intensity, i.e., U anomalies potentially below -1.2 m.s^{-1} as illustrated in Figure 3b. Overall, regional seasonal cycles plotted for each year show the 2023 SST mostly on top of those in the past (Figure 3e). The most striking warmth recorded in June 2023 was at 1.24°C above that in June 2022. July and August 2023 followed up with SST increasing but the SST values are less different from 2022 then (1.06°C and 0.59°C respectively). Also in June 2023, wind speed dropped out of the lower bound of all seasonal cycles and the difference from the previous year was about -1.26 m.s^{-1} (Figure 3f). The combined anomalies in June 2023 marine extreme heat waves set the northeastern Atlantic ocean sink from an enhanced sink in 2022 (0.29 PgC.yr^{-1}) back to its magnitude in the 2000s (0.18 PgC yr^{-1}) (Figure 3g).

4 Conclusions and Perspectives

This study first examined the skill of CMEMS-LSCE-FFNN, an ensemble approach of feed-forward neural networks (FFNN) developed by Chau et al. (2022b), in a retrospective prediction of CO_2 fugacity (fCO_2) over the global ocean. The assessment was done for two FFNN models. While the latest version (FFNNv2022) trained on SOCATv2022 data for the period 1985-2021 was used to predict fCO_2 in 2022, FFNNv2021 trained on SOCATv2021 in 1985-2020 was used to predict fCO_2 in 2021-2022 allowing the qualification of the two-year model prediction. SOCATv2023 with data available in the prediction years was used for the prediction assessment. Our evaluation confirms a robust performance of the FFNN prediction in comparison to independent observation-based data and to the FFNN reconstruction. The retrospective prediction for the years 2021-2022 retained intra-seasonal to interannual variations of fCO_2 as those in the reconstruction time series and no large systematic bias has been observed between the two across all ocean provinces. The closeness between the

predicted and reconstructed global net ocean budget implies that, when used as input to an atmospheric transport model, the prediction removes an appropriate mass of carbon from the simulated atmosphere: this is an important asset for greenhouse gas monitoring.

The latest model version, FFNNv2022, was ultimately used to predict fCO_2 from January 2022 to August 2023, i.e., up to 20 months beyond the coverage of its training dataset. This study also exemplified the assessment of air-sea CO_2 fluxes ($fgCO_2$) generated from the predicted fCO_2 in the years 2022-2023 over the eastern tropical Pacific, where regional CO_2 gas exchanges greatly vary with El Niño-Southern Oscillation (ENSO) conditions and thus affect substantially on interannual variability of the global net sink. The year 2022 has been predicted with persistently high fCO_2 (strong CO_2 outgassing to the atmosphere) in response to the maintenance of La Niña since summer 2020. A remarkable reduction in the tropical Pacific CO_2 source in August 2023 relative to the year before coincides with the weakening of the cooling phase. Recent discussions about the interaction between the ocean and climate have largely put attention on the El Niño revisits, their high possibility in triggering more extreme heat worldwide, and further impacts on the marine carbon cycle early at the end of 2023 onwards. However, already in June 2023 as exceptional surface ocean warming and extraordinarily low wind intensity fall out historical records over the northeastern Atlantic ocean, we have found an anomalous reduction in CO_2 uptake setting this regional sink back to its magnitude in the 2000s. These results emphasise critical needs and open the possibility to derive monthly predictions for global surface ocean maps of numerous variables driven by fCO_2, including air-sea fluxes, seawater pH, and dissolved inorganic carbon, as the reconstruction quality of fCO_2 drives that of the other variables (Chau et al., 2022a, 2022b). The new datasets for the year 2022 (January) to 2023 (August) are available via the LSCE/IPSL data center (see Section Data availability) and are updated each month. This demonstration of an operational service will be extended at an increased horizontal resolution, following the current development of the reference CMEMS-LSCE-FFNN reconstructions (Chau et al., 2023).

Data availability

Data provided in this research are available for use with open access granted by the French LSCE/IPSL Data Center (`https://dods.lsce.ipsl.fr/invsat/FFNN_low-latency/`).

Acknowledgments

We acknowledge funding support by the MOB TAC project of the European Copernicus Marine Environment Monitoring Service (CMEMS, `https://marine.copernicus.eu/about/producers/mob-tac`, last access: 20/9/2023). SOCAT is an international effort endorsed by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS) and the Integrated Marine Biogeochemistry and Ecosystem Research program (IMBER), to deliver a uniformly quality-controlled surface ocean CO_2 database. The data center LSCE/IPSL has offered data repository and access. We thank François-Marie Bréon for the tool development of xCO_2 extrapolation and comments on the first manuscript version and we thank Cédric Bacour for his help in running the extrapolation tool.

References

Bakker, D., Alin, S., Becker, M., Bittig, H., Castaño-Primo, R., Feely, R. A., ... others (2022). *Socat version 2022 for quantification of ocean co$_2$ uptake.* Retrieved from `https://socat.info/index.php/version-2022/` doi: https://doi.org/10.25921/1h9f-nb73

Bakker, D., Alin, S., Castaño-Primo, R., Margot, C., Gritzalis, T., Kozyr, A., ... others (2021). *Socat version 2021 for quantification of ocean co$_2$ uptake.* Retrieved from `https://socat.info/index.php/version-2021/`

Bakker, D., Alin, S. R., Bates, N., Becker, M., Feely, R. A., Gkritzalis, T., ... others (2023). *Surface ocean co$_2$ atlas database version 2023 (socatv2023).* doi: https://doi.org/10.25921/r7xa-bt92

Bakker, D., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., ... Xu, S. (2016). A multi-decade record of high-quality $f$co$_2$ data in version 3 of the surface ocean co$_2$ atlas (socat). *Earth System Science Data*, *8*(2), 383–413. Retrieved from `https://essd.copernicus.org/articles/8/383/2016/` doi: 10.5194/essd-8-383-2016

Bates, N. R., Astor, Y. M., Church, M. J., Currie, K., Dore, J. E., González-Dávila, M., ... Santana-Casiano, J. M. (2014). A time-series view of changing surface ocean chemistry due to ocean uptake of anthropogenic co$_2$ and ocean acidification. *Oceanography*, *27*(1), 126–141.

Canadell, J. G., Monteiro, P. M., Costa, M. H., Da Cunha, L. C., Cox, P. M., Alexey, V., ... others (2021). Global carbon and other biogeochemical cycles and feedbacks in

climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change. Cambridge University Press. doi: 10.1017/9781009157896.007

Chau, T. T. T., Gehlen, M., & Chevallier, F. (2022a, December). *Global ocean surface carbon product* (Research Report No. CMEMS-MOB-QUID-015-008). Le Laboratoire des Sciences du Climat et de l'Environnement. Retrieved from `https://hal.archives-ouvertes.fr/hal-02957656` (Quality Information Document) doi: 10.48670/moi-00047

Chau, T. T. T., Gehlen, M., & Chevallier, F. (2022b). A seamless ensemble-based reconstruction of surface ocean pco_2 and air–sea co_2 fluxes over the global coastal and open oceans. *Biogeosciences*, *19*(4), 1087–1109. Retrieved from `https://bg.copernicus.org/articles/19/1087/2022/` doi: 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 Science Data Discussions*, *2023*, 1–52. doi: 10.5194/essd-2023-146

Cooley, S., Schoeman, D., Bopp, L., Boyd, P., Donner, S., Ito, S.-i., ... others (2022). Oceans and coastal ecosystems and their services. In *Ipcc ar6 wgii*. Cambridge University Press.

Doney, S. C., Busch, D. S., Cooley, S. R., & Kroeker, K. J. (2020). The impacts of ocean acidification on marine ecosystems and reliant human communities. *Annual Review of Environment and Resources*, *45*(1).

Fay, A. R., & McKinley, G. A. (2014). Global open-ocean biomes: mean and temporal variability. *Earth System Science Data*, *6*(2), 273–284. Retrieved from `https://essd.copernicus.org/articles/6/273/2014/` doi: 10.5194/essd-6-273-2014

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*, *14*(11), 4811–4900. Retrieved from `https://essd.copernicus.org/articles/14/4811/2022/` doi: 10.5194/essd-14-4811-2022

Gallego, M., Timmermann, A., Friedrich, T., & Zeebe, R. E. (2018). Drivers of future seasonal cycle changes in oceanic pco 2. *Biogeosciences*, *15*(17), 5315–5327.

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. Retrieved from `https://essd.copernicus`

.org/articles/13/777/2021/ doi: 10.5194/essd-13-777-2021

Gregor, L., Lebehot, A. D., Kok, S., & Scheel Monteiro, P. M. (2019). A comparative assessment of the uncertainties of global surface ocean CO_2 estimates using a machine-learning ensemble (csir-ml6 version 2019a) – have we hit the wall? *Geoscientific Model Development*, *12*(12), 5113–5136. Retrieved from https://gmd.copernicus.org/articles/12/5113/2019/ doi: 10.5194/gmd-12-5113-2019

Gruber, N., Bakker, D. C., DeVries, T., Gregor, L., Hauck, J., Landschützer, P., ... Müller, J. D. (2023). Trends and variability in the ocean carbon sink. *Nature Reviews Earth & Environment*, *4*(2), 119–134.

Gruber, N., Landschützer, P., & Lovenduski, N. S. (2019). The variable southern ocean carbon sink. *Annual review of marine science*, *11*, 159–186.

Hopkins, F. E., Suntharalingam, P., Gehlen, M., Andrews, O., Archer, S. D., Bopp, L., ... others (2020). The impacts of ocean acidification on marine trace gases and the implications for atmospheric chemistry and climate. *Proceedings of the Royal Society A*, *476*(2237), 20190769.

Janssens-Maenhout, G., Pinty, B., Dowell, M., Zunker, H., Andersson, E., Balsamo, G., ... others (2020). Toward an operational anthropogenic co 2 emissions monitoring and verification support capacity. *Bulletin of the American Meteorological Society*, *101*(8), E1439–E1451.

Landschützer, P., Gruber, N., & Bakker, D. C. (2016). Decadal variations and trends of the global ocean carbon sink. *Global Biogeochemical Cycles*, *30*(10), 1396–1417.

Landschützer, P., Ilyina, T., & Lovenduski, N. S. (2019). Detecting regional modes of variability in observation-based surface ocean pco_2. *Geophysical Research Letters*, *46*(5), 2670–2679.

Regnier, P., Resplandy, L., Najjar, R. G., & Ciais, P. (2022). The land-to-ocean loops of the global carbon cycle. *Nature*, *603*(7901), 401–410.

Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., ... Zeng, J. (2015). Data-based estimates of the ocean carbon sink variability – first results of the surface ocean pco_2 mapping intercomparison (socom). *Biogeosciences*, *12*(23), 7251–7278. Retrieved from https://bg.copernicus.org/articles/12/7251/2015/ doi: 10.5194/bg-12-7251-2015

Rödenbeck, C., Keeling, R. F., Bakker, D. C., Metzl, N., Olsen, A., Sabine, C., & Heimann, M. (2013). Global surface-ocean pco_2 and sea–air co_2 flux variability from an

438 observation-driven ocean mixed-layer scheme. *Ocean Science*, *9*(2), 193–216.
439 Rustogi, P., Landschützer, P., Brune, S., & Baehr, J. (2023). The impact of seasonality
440 on the annual air-sea carbon flux and its interannual variability. *npj Climate and*
441 *Atmospheric Science*, *6*(1), 66.
442 Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., ...
443 others (2002). Global sea–air co2 flux based on climatological surface ocean pco2,
444 and seasonal biological and temperature effects. *Deep Sea Research Part II: Topical*
445 *Studies in Oceanography*, *49*(9-10), 1601–1622.