Derivation of seawater pCO₂ from net community production identifies the South Atlantic Ocean as a CO₂ source

Daniel Ford[1,2], Gavin H. Tilstone[1], Jamie D. Shutler[2] and Vassilis Kitidis[1]

[1] Plymouth Marine Laboratory, Plymouth, UK

5 [2] College of Life and Environmental Sciences, University of Exeter, UK

Correspondence to: Daniel Ford (dfo@pml.ac.uk)

Abstract. A key step in assessing the global carbon budget is the determination of the partial pressure of CO_2 in seawater ($pCO_{2 (sw)}$). Spatially complete observational fields of $pCO_{2 (sw)}$ are routinely produced for regional and global ocean carbon budget assessments by extrapolating sparse *in situ* measurements of $pCO_{2 (sw)}$ using satellite observations. Within these

10 schemes, satellite chlorophyll *a* (Chl *a*) is often used as a proxy for the biological drawdown or release of CO_2. Chl *a* does not however quantify carbon fixed through photosynthesis and then respired, which is determined by net community production (NCP).

In this study, $pCO_{2 (sw)}$ over the South Atlantic Ocean is estimated using a feed forward neural network (FNN) scheme and either satellite derived NCP, net primary production (NPP) or Chl *a* to compare which biological proxy is the most accurate.

15 Estimates of $pCO_{2 (sw)}$ using NCP, NPP or Chl *a* were similar, but NCP was more accurate for the Amazon Plume and upwelling regions, which were not fully reproduced when using Chl *a* or NPP. Reducing the uncertainties in the satellite biological parameters to estimate $pCO_{2 (sw)}$, illustrated further improvement and greater differences for NCP compared to NPP or Chl *a*. Using NCP to estimate $pCO_{2 (sw)}$ showed that the South Atlantic Ocean is a CO_2 source, whereas if no biological parameters are used in the FNN (following existing annual carbon assessments), this region becomes a sink for

20 CO_2. These results highlight that using NCP improved the accuracy of estimating $pCO_{2 (sw)}$, and changes the South Atlantic Ocean from a CO_2 sink to a source. Reducing the uncertainties in NCP derived from satellite parameters will further improve our ability to quantify the global ocean CO_2 sink.

1. Introduction

25 Since the industrial revolution, anthropogenic CO_2 emissions have resulted in an increase in atmospheric CO_2 concentrations (Friedlingstein et al., 2020; IPCC, 2013). By acting as a sink for CO_2, the oceans have buffered the increase in anthropogenic atmospheric CO_2, without which the atmospheric concentration would be 42-44 % higher (DeVries, 2014). The long-term absorption of CO_2 by the oceans is altering the marine carbonate chemistry of the ocean, resulting in a lowering of pH; a process known as ocean acidification (Raven et al., 2005). Observational fields of the partial pressure of CO_2 in seawater

30 ($pCO_{2\,(sw)}$) are one of the key datasets needed to routinely assess the strength of the oceanic CO_2 sink (Friedlingstein et al., 2020; Landschützer et al., 2014, 2020; Rödenbeck et al., 2015; Watson et al., 2020b). These methods are reliant on the extrapolation of sparse *in situ* observations of $pCO_{2\,(sw)}$ using satellite observations of parameters which account for the variability of, and the controls on, $pCO_{2\,(sw)}$ (Shutler et al., 2020). These parameters include sea surface temperature (SST; e.g. Landschützer et al., 2013; Stephens et al., 1995), salinity and Chl *a* (Rödenbeck et al., 2015). SST and salinity control

35 $pCO_{2\,(sw)}$ by changing the solubility of CO_2 in seawater (Weiss, 1974), whilst biological processes such as photosynthesis and respiration contribute by modulating its concentration.

Chl *a* is routinely used as a proxy for this biological activity (Rödenbeck et al., 2015), but it does not distinguish between carbon fixation through photosynthesis and the carbon respired by the plankton community. Net primary production (the net carbon fixation rate; NPP) is determined by the standing stock of phytoplankton, for which the Chl *a* concentration is used as

40 a proxy, and modified by the photosynthetic rate and the available light in the water column (Behrenfeld et al., 2016). Photosynthetic rates are, in turn, modified by ambient nutrient and temperature conditions (Behrenfeld and Falkowski, 1997; Marañón et al., 2003). Elevated Chl *a* does not always equate to elevated NPP (Poulton et al., 2006), and for the same Chl *a* concentrations, NPP can vary depending on the health and metabolic state of the plankton community. All of these controls are captured by the net community production (NCP), which is the metabolic balance of the plankton community resulting

45 from the carbon fixed through photosynthesis and that lost through respiration. When NCP is positive the plankton community is autotrophic which implies that there is a drawdown of CO_2 from seawater (since the plankton reduce the CO_2 in the water column), and when it is negative the community is heterotrophic implying a release of CO_2 into the ocean (as the plankton produce or release CO_2) which can then be released into the atmosphere (Jiang et al., 2019; Schloss et al., 2007). Using NCP to estimate $pCO_{2\,(sw)}$ compared to Chl *a* should therefore theoretically lead to an improvement in the derivation

50 of $pCO_{2\,(sw)}$.

Many studies have used satellite Chl *a* to estimate $pCO_{2\,(sw)}$ at both regional (Benallal et al., 2017; Chierici et al., 2012; Moussa et al., 2016), and global scales (Landschützer et al., 2014; Liu and Xie, 2017). Chierici et al. (2012) attempted to use satellite NPP to estimate $pCO_{2\,(sw)}$ in the southern Pacific Ocean, but there was no significant improvement over using satellite Chl *a*. This is not surprising as NPP captures more of the biological signal, but still lacks any inclusion of respiration which results in the release of CO_2 into the water column. To our knowledge the use of satellite NCP to estimate $pCO_{2\,(sw)}$

55 has not been attempted before and could be a means of improving estimates of $pCO_{2\,(sw)}$ as long as satellite NCP observations are accurate (Ford et al., 2021; Tilstone et al., 2015a). These satellite measurements may improve the estimation of $pCO_{2\,(sw)}$ as NCP includes the full biological control on $pCO_{2\,(sw)}$, which could be important in regions of sparse *in situ* $pCO_{2\,(sw)}$ observations where interpolation and neural network techniques are likely to struggle (Watson et al., 2020b).

60 The objective of this paper is therefore to compare the estimation of $pCO_{2\,(sw)}$ using either NCP, NPP or Chl *a* to determine which biological descriptor is the most accurate. A 16 year time series of $pCO_{2\,(sw)}$ was generated for the South Atlantic Ocean using satellite NCP, NPP or Chl *a*, as the biological input, alongside a baseline approach with no biological parameters as input. Regional differences in the generated $pCO_{2\,(sw)}$ fields are assessed. The seasonal and interannual

variability in pCO$_2$ $_{(sw)}$ estimated from NCP, NPP, Chl a and the baseline approach were also compared. A perturbation

65 analysis was conducted to evaluate the potential reduction in the uncertainty in the pCO$_2$ $_{(sw)}$ fields when estimated from NCP, NPP or Chl a. This is discussed in the context of reducing uncertainties in these input variables for future improvements in producing spatially complete fields of pCO$_2$ $_{(sw)}$, and the effect on estimates of the oceanic carbon sink.
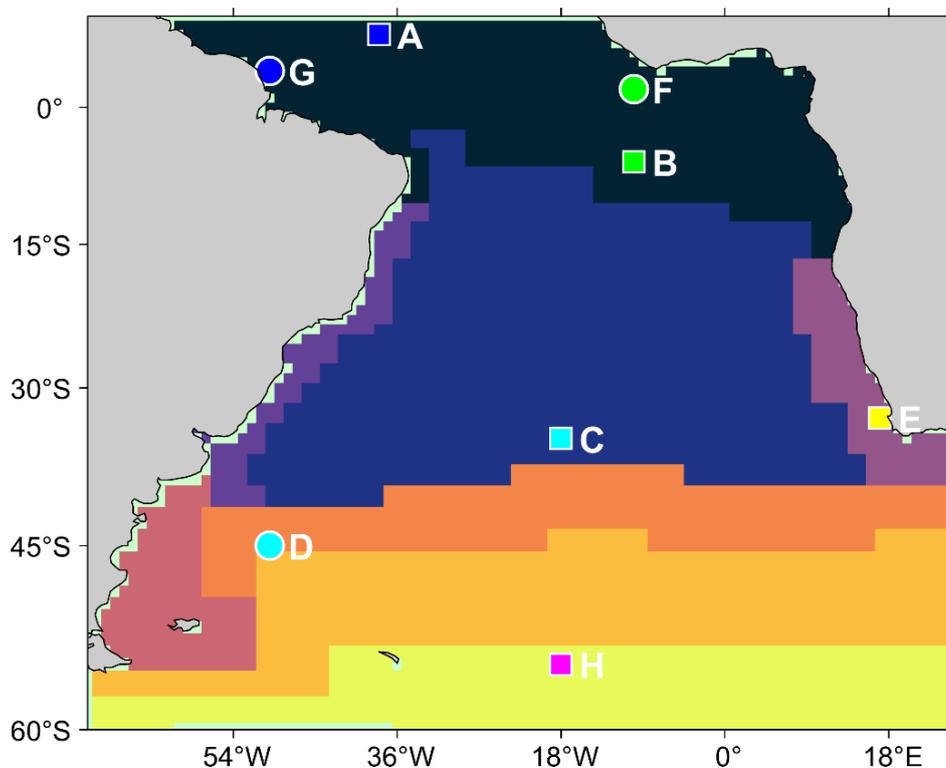
Fig. 1: Map of the 8 static biogeochemical provinces in the South Atlantic Ocean. Markers and letters indicate the locations of
70 **timeseries extracted in Fig. 3.**

2. Methods

2.1. Surface Ocean Carbon Atlas (SOCAT) $pCO_{2\,(sw)}$ and atmospheric CO_2

SOCATv2020 (Bakker et al., 2016; Pfeil et al., 2013) individual fugacity of CO_2 in seawater ($fCO_{2\,(sw)}$) observations were
75 downloaded from https://www.socat.info/index.php/data-access/. Data were extracted from 2002 to 2018 for the South
Atlantic Ocean (10º N-60º S, 25º E–80º W; Fig. 1). The individual cruise observations were collected from different depths,
and are not representative of the $fCO_{2\,(sw)}$ in the top ~100 μm of the ocean, where gas exchange occurs (Goddijn-Murphy et
al., 2015; Woolf et al., 2016). Therefore, the SOCAT observations were re-analysed to a standard temperature dataset and
depth (Reynolds et al., 2002) that is considered representative of the bottom of the mass boundary layer (Woolf et al., 2016).
80 This was achieved using the 'fe_reanalyse_socat' utility in the open source FluxEngine toolbox (Holding et al., 2019;
Shutler et al., 2016), which follows the methodology described in Goddijn-Murphy et al. (2015). The reanalysed $fCO_{2\,(sw)}$
observations were converted to $pCO_{2\,(sw)}$, and gridded onto 1º monthly grids following SOCAT protocols (Sabine et al.,
2013). The uncertainties in the *in situ* data were taken as the standard deviation of the observations in each grid cell, or
where a single observation exists were set as 5 μatm following Bakker et al. (2016).
85 Monthly 1º grids of atmospheric pCO_2 ($pCO_{2\,(atm)}$) were extracted from v5.5 of the global estimates of $pCO_{2\,(sw)}$ dataset
(Landschützer et al., 2016, 2017). $pCO_{2\,(atm)}$ was estimated using the dry mixing ratio of CO_2 from the NOAA-ESRL marine
boundary layer reference (https://www.esrl.noaa.gov/gmd/ccgg/mbl/), Optimum Interpolated SST (Reynolds et al., 2002)
and sea level pressure following Dickson et al. (2007).

2.2. Moderate Resolution Spectroradiometer on Aqua (MODIS-A) satellite observations

4 km resolution monthly mean Chl *a* were calculated from MODIS-A Level 1 granules, retrieved from National Aeronautics
and Space Administration (NASA) Ocean Colour website (https://oceancolor.gsfc.nasa.gov/) using SeaDAS v7.5, and
applying the standard OC3-CI Chl *a* algorithm (https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/). In addition, monthly mean
MODIS-A SST and photosynthetically active radiation (PAR) were also downloaded from the NASA Ocean Colour website.
95 Mean monthly NPP were generated from MODIS-A Chl *a*, SST and PAR using the Wavelength Resolving Model (Morel,
1991) with the look up table described in Smyth et al. (2005). Coincident mean monthly NCP using the algorithm NCP-D
described in Tilstone et al. (2015a) were generated using the MODIS-A NPP and SST data. Further details of the satellite
algorithms are given in O'Reilly et al. (1998; 2019), Hu et al. (2012) for Chl *a*, Smyth et al. (2005), Tilstone et al. (2005,
2009) for NPP and Tilstone et al. (2015a) for NCP. All monthly mean data were generated between July 2002 and December
100 2018 and were re-gridded onto the same 1º grid as the $pCO_{2\,(sw)}$ observations. The assessed uncertainties from the literature
for each of the input parameters used are given in Table 1.

Table 1: Uncertainties in the input parameters of the Feed Forward Neural Network used in Monte Carlo uncertainty propagation and perturbation analysis.

Parameter	Algorithm Uncertainty	Reference
Chlorophyll *a*	0.15 $\log_{10}$(mg m^{-3})	Ford, et al (2021)
Net Primary Production	0.20 $\log_{10}$(mg C m^{-2} d^{-1})	Ford, et al (2021)
Net Community Production	45 mmol O$_2$ m^{-2} d^{-1}	Ford, et al (2021)
SST	0.41 ℃	Ford, et al (2021)
pCO$_{2\ (atm)}$	1 µatm	

105

2.3. Feed forward neural network scheme

The South Atlantic Ocean was partitioned into 8 biogeochemical provinces (Fig. 1), following Longhurst et al. (1995). The pCO$_{2\ (sw)}$ observations in the eastern Equatorial Atlantic were sparse, and therefore the Equatorial region was merged into 1 province. In each province the available monthly pCO$_{2\ (sw)}$ observations were matched to temporally and spatially coincident

110 pCO$_{2\ (atm)}$, MODIS-A, NCP and SST, to provide training data for the feed forward neural network (FNN). Observations in coastal regions (< 200 m water depth) were removed from the analysis, due to the increased uncertainty in ocean colour observations in these areas (e.g. Lavender et al., 2004). Due to constraints on the coverage of ocean colour data, no data were available in austral winter below ~50º S.

The coincident observations in each province were randomly split into 3 datasets: 1.) A training dataset (50 % of the

115 observations) used to train the FNNs; 2.) A validation dataset (30 % of the observations) used to assess the performance of the FNN and to prevent the networks from overfitting; 3.) An independent test dataset (20 % of the observations) to assess the final performance of the FNN, with observations that are independent of the network training. The optimal split (r$_{opt}$) method of Amari et al. (1997) was used to partition the data into these three sets, as follows:

$$r_{opt} = 1 - \frac{1}{\sqrt{2m}} \tag{1}$$

120 where m is number of input parameters. For our three input parameters, a split of 60 % training to 40 % validation datasets would occur, where we removed 10 % from each dataset to provide a further independent test dataset. A pre-training step was used to determine the optimum number of hidden neurons in the FNN (Benallal et al., 2017; Landschützer et al., 2013; Moussa et al., 2016), to provide the best fit for the observations, whilst preventing over fitting (Demuth et al., 2008).

The FNN was trained using the optimum number of hidden neurons, in an iterative process until the Root Mean Square

125 Difference (RMSD) remained unchanged for 6 iterations. The best performing FNN, with the lowest RMSD was then used to estimate pCO$_{2\ (sw)}$. The uncertainties in the input parameters were propagated through the FNN, using a Monte Carlo uncertainty propagation, where 1000 calculations were made perturbing the input parameters, using random noise for their uncertainty (Table 1). The output from the 8 province FNNs were then combined and weighted statistics, which account for

both the satellite and *in situ* uncertainty, were used to assess the overall performance of the FNN (as used within Ford et al.,
130   2021). The combined 8 FNNs approach will hereafter be referred to as SA-FNN.

The approach to training the FNNs was repeated replacing NCP with Chl *a* or NPP sequentially, to determine if there was an improvement by using NCP. A baseline SA-FNN with no biological parameters as input was trained using $pCO_{2\ (atm)}$, MODIS-A SST, alongside sea surface salinity and mixed layer depth from the Copernicus Marine Environment Modelling Service (https://resources.marine.copernicus.eu/) global ocean physics reanalysis product (GLORYS12V1). This parameter
135   combination has recently been included within a neural network scheme to estimate global fields of $pCO_{2\ (sw)}$ (Watson et al., 2020b).

Following these methods, a monthly mean time-series of $pCO_{2\ (sw)}$ was generated in the South Atlantic Ocean, applying the SA-FNN approach using NCP (SA-FNN$_{NCP}$), NPP (SA-FNN$_{NPP}$), Chl *a* (SA-FNN$_{CHLA}$) or no biological parameters (SA-FNN$_{NO-BIO}$). The $pCO_{2\ (sw)}$ fields were spatially averaged using a $3{\times}3$ pixel filter, but were not averaged temporally as in
140   previous studies (Landschützer et al., 2014, 2016) because averaging temporally could mask features that occur within single months of the year. The uncertainties in the input parameters (Table 1) were propagated through the neural network on a per pixel basis, and combined in quadrature with the RMSD of the test dataset, to produce a combined uncertainty budget for each pixel, assuming all sources of uncertainty are independent and uncorrelated (BIPM, 2008; Taylor, 1997).

2.4. Atlantic Meridional Transect *in situ* data

145   To assess the accuracy of the SA-FNN, coincident *in situ* measurements of NCP, NPP, Chl *a*, SST, $pCO_{2\ (atm)}$ and $pCO_{2\ (sw)}$, with uncertainties, were provided by Atlantic Meridional Transects 20, 21, 22 and 23 in 2010, 2011, 2012 and 2013, respectively. All the Atlantic Meridional Transect data described in this section can be obtained from the British Oceanographic Data Centre (https://www.bodc.ac.uk/). Chl *a* was estimated following the methods of Brewin et al. (2016), using underway continuous spectrophotometric measurements, and uncertainties were estimated as ~0.06 $\log_{10}$(mg m^{-3})
150   (Ford et al., 2021). ^{14}C based NPP measurements were made based on dawn to dusk simulated *in situ* incubations, following the methods given in Tilstone et al. (2017), at 56 stations with a per station uncertainty. Uncertainties ranged between 8 and 213 mg C m^{-2} d^{-1} and were on average 53 mg C m^{-2} d^{-1}. NCP was estimated using *in vitro* changes in dissolved O$_2$, following the methods of Gist et al. (2009) and Tilstone et al. (2015a) at 51 stations with a per station uncertainty calculated. Uncertainties ranged between 5 and 25 mmol O$_2$ m^{-2} d^{-1} and were on average 14 mmol O$_2$ m^{-2} d^{-1}.
155   Underway measurements of $pCO_{2\ (sw)}$ and $pCO_{2\ (atm)}$ were performed continuously, following the methods of Kitidis et al. (2017). SST was continuously measured alongside all observations (SeaBird SBE45), with a factory calibrated uncertainty of ±0.01 °C. The mean of underway $pCO_{2\ (sw)}$, $pCO_{2\ (atm)}$, SST and Chl *a* were taken ±20 minutes around each station where NCP and NPP were measured.

Biogeosciences
Discussions

2.5. Perturbation analysis

160 Following the approach of Saba et al. (2011), a perturbation analysis was conducted, to evaluate the potential reduction in SA-FNN $pCO_{2\ (sw)}$ RMSD that could be attributed to the input parameters. The analysis indicates the maximum reduction in RMSD that could be achieved if uncertainties in the input parameters were reduced to ~0. Each of the input parameters; NCP, SST and $pCO_{2\ (atm)}$ can have three possible values for each *in situ* $pCO_{2\ (sw)}$ estimate (original value, original ± uncertainty; Table 1), enabling 27 perturbations of the input data. RMSD and bias were used to assess the performance of

165 SA-FNN under different scenarios. For each *in situ* $pCO_{2\ (sw)}$ observation, the 27 perturbations of SA-FNN $pCO_{2\ (sw)}$ were examined, and the perturbation that produced the lowest RMSD and bias combination in two scenarios was selected; (1) uncertainty in individual input parameters (NCP, SST and $pCO_{2\ (atm)}$) and (2) uncertainty in all input parameters together. The approach was conducted on all three training datasets, and on the Atlantic Meridional Transect *in situ* data. The analysis was repeated sequentially replacing NCP with Chl *a* and NPP, to determine if there was a greater maximum reduction in

170 RMSD using NCP. The analysis was also conducted allowing for a 10 % reduction in input parameter uncertainties, to indicate the short-term reduction in $pCO_{2\ (sw)}$ RMSD that could be achieved by reducing the input parameter uncertainties.

2.6. Comparison of the SA-FNN$_{NCP}$ with the SA-FNN$_{NO-BIO}$, SA-FNN$_{CHLA}$, SA-FNN$_{NPP}$ and 'state of the art' data for the South Atlantic

The most comprehensive $pCO_{2\ (sw)}$ fields to date are from Watson et al. (2020b, 2020a). The 'standard method' $pCO_{2\ (sw)}$

175 fields within the Watson et al. (2020b, 2020a) data were produced by extrapolating the *in situ* reanalysed SOCATv2019 $pCO_{2\ (sw)}$ observations using a self-organising map feed forward neural network approach (Landschützer et al., 2016), and will be referred to as 'W2020'. A time-series was extracted from the W2020 data, coincident to SA-FNN$_{NCP}$, SA-FNN$_{NPP}$, SA-FNN$_{CHLA}$ and SA-FNN$_{NO-BIO}$. For the five methods, a monthly climatology referenced to the year 2010 was computed, assuming an atmospheric CO_2 increase of 1.5 $\mu atm\ yr^{-1}$ (Takahashi et al., 2009; Zeng et al., 2014). The climatology should

180 be insensitive to the chosen atmospheric CO_2 rise due to the reference year being central to the time series. The standard deviation of this climatology was also computed on a per pixel basis.

For each station in Fig. 1, the monthly climatology of $pCO_{2\ (sw)}$, representing the average seasonal cycle of $pCO_{2\ (sw)}$, and the standard deviation of the climatology, as an indication of the interannual variability, were extracted from the five approaches. The $pCO_{2\ (sw)}$ value for each station was the statistical mean of the four nearest data points weighted by their

185 respective proximity to the station coordinate.

The station climatologies for the SA-FNN$_{NO-BIO}$, W2020, SA-FNN$_{CHLA}$, and SA-FNN$_{NPP}$ were compared to the SA-FNN$_{NCP}$, by testing for significant differences in the seasonal cycle and annual $pCO_{2\ (sw)}$ (offset). The seasonal cycles were compared using a non-parametric Spearman's correlation and deemed statistically different where the correlation was not significant ($\alpha < 0.05$). A non-parametric Kruskal-Wallis tested for significant ($\alpha < 0.05$) differences in the annual $pCO_{2\ (sw)}$, indicating an

190 offset between the two tested climatologies. The Southern Ocean station (station H) was excluded from the statistical analysis due to missing data in the SA-FNN.

2.7. Estimation of the bulk CO_2 flux

The flux of CO_2 (F) between the atmosphere and ocean (air-sea) can be expressed in a bulk parameterisation as:

$$F = k \left(\alpha_W \, pCO_{2\,(sw)} - \alpha_s \, pCO_{2\,(atm)} \right) \qquad (2)$$

195 Where k is the gas transfer velocity, and α_w and α_s are the solubility of CO_2 at the base and top of the mass boundary layer at the sea surface respectively (Woolf et al., 2016). k was estimated from ERA5 monthly reanalysis wind speed (downloaded from the Copernicus Climate Data Store; https://cds.climate.copernicus.eu/) following the parameterisation of Nightingale et al. (2000). α_w was estimated as a function of SST and sea surface salinity (Weiss, 1974) using the monthly Optimum Interpolated SST (Reynolds et al., 2002) and sea surface salinity from the Copernicus Marine Environment Modelling

200 Service global ocean physics reanalysis product (GLORYS12V1). α_s was estimated using the same temperature and salinity datasets but included a gradient from the base to the top of mass boundary layer of -0.17 K (Donlon et al., 1999) and +0.1 salinity units (Woolf et al., 2016). $pCO_{2\,(atm)}$ was estimated using the dry mixing ratio of CO_2 from the NOAA-ESRL marine boundary layer reference, Optimum Interpolated SST (Reynolds et al., 2002) applying a cool skin bias (0.17K; Donlon et al., 1999) and sea level pressure following Dickson et al. (2007). Spatially and temporally complete $pCO_{2\,(sw)}$ fields, which are

205 representative of $pCO_{2\,(sw)}$ at the base of the mass boundary layer, were extracted from the SA-FNN$_{NCP}$, SA-FNN$_{NPP}$, SA-FNN$_{CHLA}$, SA-FNN$_{NO-BIO}$ and W2020.

The monthly CO_2 flux was calculated using the open source FluxEngine toolbox (Holding et al., 2019; Shutler et al., 2016) between 2003 and 2018 for the five $pCO_{2\,(sw)}$ inputs, using the 'rapid transport' approximation (described in Woolf et al., 2016). The net annual flux was determined for the South Atlantic Ocean (10° N-44° S; 25° E-70° W) using the

210 'fe_calc_budgets.py' utility within FluxEngine with the supplied area and land percentage masks. The mean net annual flux was calculated as the mean of the 15 year net annual fluxes. Positive net fluxes indicate a net source to the atmosphere, and negative net fluxes a sink.

Table 2: The percentage reduction in $pCO_{2\,(sw)}$ RMSD by reducing NCP, net primary production and chlorophyll a uncertainties
215 **to ~0 as described in Section 2.5. The full results can be found in Appendix Table A1.**

Parameter	Training	Validation	Independent Test	AMT *in situ*
NCP	32 %	40 %	36 %	25 %
Net Primary Production	31 %	37 %	36 %	13 %
Chlorophyll *a*	17 %	21 %	20 %	7 %
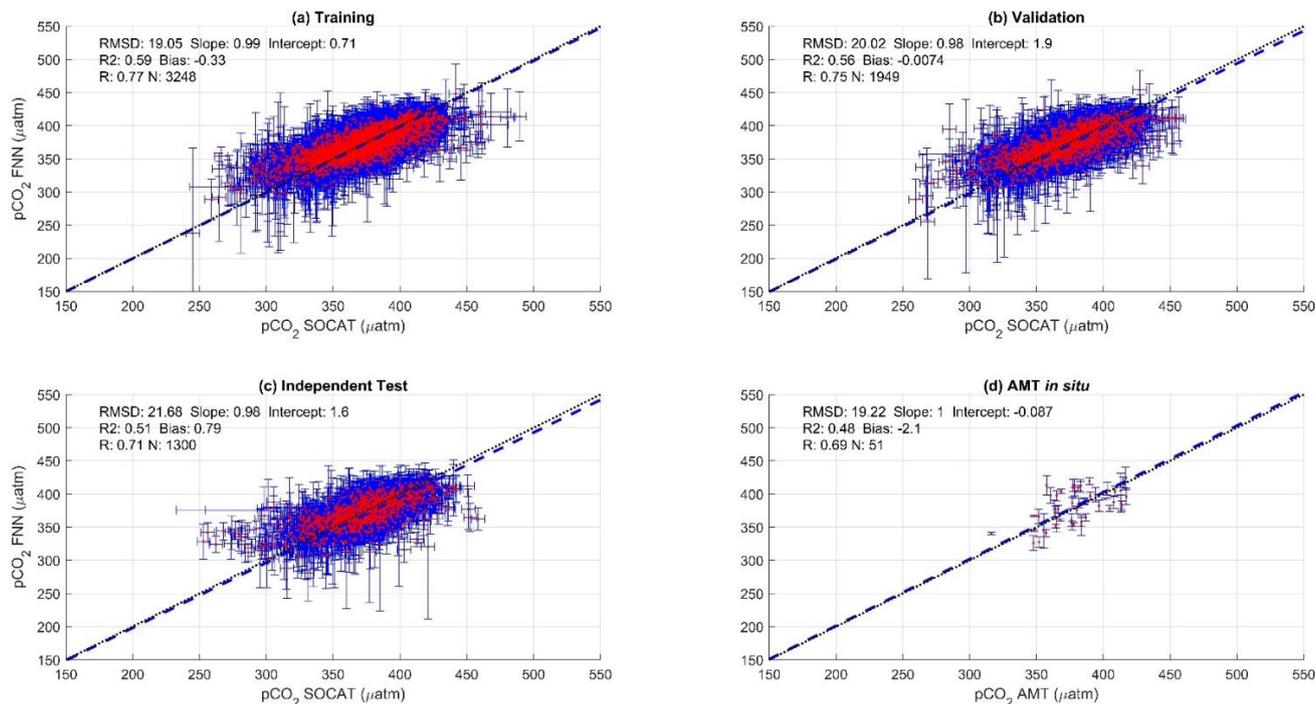
Fig. 2: Scatter plots showing the combined performance of the 8 feed forward neural networks trained using NCP for each biogeochemical province (Fig. 1) using 4 separate training and validation datasets; (a) Training, (b) Validation, (c) Independent Test and (d) Atlantic Meridional Transect (AMT) in situ. The blue dashed line is the Type II regression and the black dashed line is the 1:1 line. Horizontal errorbars indicate the uncertainty of the SOCATv2020 $pCO_{2\ (sw)}$. Vertical errorbars indicate the uncertainty attributed to the input parameter uncertainty propagated through the feed forward neural networks. The statistics within each plot are; Root Mean Square Difference (RMSD), Slope and Intercept of the Type II regression, Coefficient of Determination (R2), Pearson's Correlation Coefficient (R), Bias and number of samples (N).

Table 3: The percentage reduction in $pCO_{2\ (sw)}$ RMSD by reducing NCP, net primary production and chlorophyll a uncertainties by 10 % as described in Section 2.5.

| Parameter	Training	Validation	Independent Test	AMT *in situ*
NCP	7 %	8 %	8 %	3 %
Net Primary Production	5 %	6 %	5 %	1.5 %
Chlorophyll *a*	2 %	2 %	2 %	0.5 %

3.   Results

230   ### 3.1. SA-FNN performance and perturbation analysis

The performance of the SA-FNN trained using $pCO_{2 (atm)}$, SST and NCP for the three training datasets is given in Fig. 2. The SA-FNN$_{NCP}$ had an accuracy (RMSD) of 21.67 µatm and a precision (bias) of 0.87 µatm, which was determined with the independent test data (N = 1300). Training the SA-FNN using Chl *a* or NPP instead of NCP, resulted in a similar performance (Appendix A Fig. A1, Fig. A2). The RMSD for the independent test data was within ~1.5 µatm for Chl *a* (19.88

235   µatm), NPP (20.48 µatm) and NCP (21.68 µatm) and bias near zero.

The reduction in $pCO_{2 (sw)}$ RMSD that could be achieved if input parameter uncertainties were reduced to ~0 was assessed using the previously described perturbation analysis (Table 2, Appendix A Table A1). This showed that satellite NCP uncertainties lead to an on average 36 % reduction in $pCO_{2 (sw)}$ RMSD, with NPP a 34 % reduction and Chl *a* a 19 % reduction, indicating that improving NCP uncertainties will have the largest impact on improving the estimated $pCO_{2 (sw)}$

240   fields. The bias remained near zero for all parameters indicating good precision of the SA-FNN approach (not shown). Applying the Atlantic Meridional Transect *in situ* data as input to the SA-FNN and resulting perturbation analysis, a decrease in $pCO_{2 (sw)}$ RMSD of 25 % for NCP, 13 % for NPP and 7 % for Chl *a* was observed.

The reduction in $pCO_{2 (sw)}$ RMSD from reducing input parameter uncertainties by 10 % was also assessed through the perturbation analysis (Table 3). This indicated a decrease in $pCO_{2 (sw)}$ RMSD of 8 % for NCP, 5 % for NPP and 2 % for Chl

245   *a*, again indicating that improving NCP uncertainties will have the largest impact on improving the estimated $pCO_{2 (sw)}$ fields.
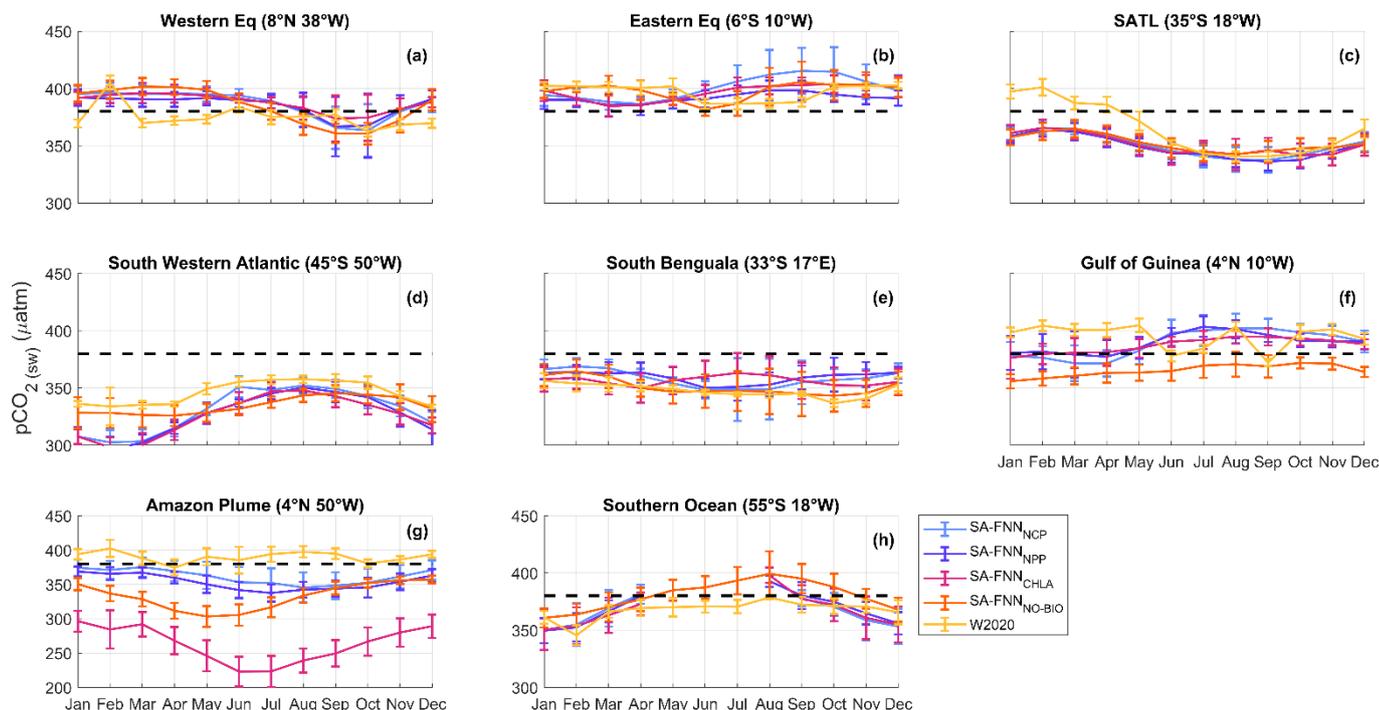
Fig. 3: Monthly climatologies of pCO$_2$ $_{(sw)}$ referenced to the year 2010 for the 8 stations marked in Fig. 1 from the SA-FNN$_{NCP}$, SA-FNN$_{NPP}$, SA-FNN$_{CHLA}$, SA-FNN$_{NO-BIO}$ and W2020 (Watson et al., 2020b). The atmospheric CO$_2$ increase was set as 1.5 µatm yr^{-1}. Black dashed line indicates the atmospheric pCO$_2$ (~380 µatm). Errorbars indicate the standard deviation of the climatology, where larger errorbars indicate a larger interannual variability. Note the different y-axis limits in Fig. 3g.

3.2. Comparison of SA-FNN$_{NCP}$ with other methods

The monthly climatology of pCO$_2$ $_{(sw)}$ generated using the SA-FNN$_{NCP}$ and referenced to the year 2010 showed differences with two published climatologies, especially in the Equatorial region (Appendix B). The monthly climatology for 8 stations (Fig. 1) were extracted from the SA-FNN$_{NCP}$, SA-FNN$_{NPP}$, SA-FNN$_{CHLA}$, SA-FNN$_{NO-BIO}$ and the W2020, to assess differences between the pCO$_2$ $_{(sw)}$ estimates (Fig. 3). The SA-FNN$_{NCP}$ and SA-FNN$_{NO-BIO}$ showed significant divergences in the Equatorial Atlantic (Figs. 3b, f, g; Fig. 4). At the eastern equatorial station, the interannual variability in pCO$_2$ $_{(sw)}$ from the SA-FNN$_{NCP}$ was high and a minimum occurred between January and April, which slowly increased to a maximum in September and October (Fig. 3b). The SA-FNN$_{NO-BIO}$ showed the opposite pattern, with a pCO$_2$ $_{(sw)}$ minima between May to July and a maxima for the remaining months with little interannual variability. The Gulf of Guinea station showed a similar variability in the SA-FNN$_{NCP}$ pCO$_2$ $_{(sw)}$ except the maxima was lower at this station (Fig. 3f). The SA-FNN$_{NO-BIO}$ indicated pCO$_2$ $_{(sw)}$ below the SA-FNN$_{NCP}$ throughout the year. The greatest divergence occurred near the Amazon plume (Fig. 3g)

11

where SA-FNN$_{NCP}$ pCO$_{2\ (sw)}$ was below or at pCO$_{2\ (atm)}$ for all months and there was a large interannual variability in pCO$_2$ $_{(sw)}$. The SA-FNN$_{NO-BIO}$ displayed lower pCO$_{2\ (sw)}$ between January and July and a lower interannual variability (Fig. 3g).

265   The SA-FNN$_{NCP}$ and SA-FNN$_{NO-BIO}$ showed no significant difference in the seasonal patterns of pCO$_{2\ (sw)}$ at stations south of 20 °S (Figs. 3c, d, e; Fig. 4). There was, however, a significant offset at some stations where the SA-FNN$_{NCP}$ generally exhibited lower pCO$_{2\ (sw)}$ in austral summer and a higher interannual variation. The SA-FNN$_{NCP}$ was significantly different to the W2020 at similar stations as the SA-FNN$_{NO-BIO}$ (Fig. 3, Fig. 4).

The SA-FNN$_{NCP}$ and SA-FNN$_{CHLA}$ showed significant differences in pCO$_{2\ (sw)}$ values in the South Benguela and Amazon

270   Plume. In the South Benguela (Fig. 3e; Fig. 4), SA-FNN$_{NCP}$ gave a pCO$_{2\ (sw)}$ maxima in austral summer, whereas the SA-FNN$_{CHL}$ maximum occurred in austral winter. In the Amazon Plume there was significant offset between the two methods and the SA-FNN$_{CHL}$ gave lower pCO$_{2\ (sw)}$ compared to the SA-FNN$_{NCP}$ (Fig. 3g; Fig. 4). The SA-FNN$_{NCP}$ and SA-FNN$_{NPP}$ had a significant offset at the Eastern Equatorial station (Fig. 3c; Fig. 4), where the SA-FNN$_{NPP}$ indicated lower pCO$_{2\ (sw)}$. For the other stations, no significant differences were observed.

275

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

4. Discussion

4.1. Assessment of biological parameters to estimate pCO$_2$ $_{(sw)}$

In this paper, the differences in estimating pCO$_2$ $_{(sw)}$ using satellite NCP, NPP or Chl a were assessed. The SA-FNN$_{NCP}$ had an overall accuracy that is consistent with other approaches that have been developed for the Atlantic (22.83 µatm; Landschützer et al., 2013), and slightly lower than the published global result of 25.95 µatm (Landschützer et al., 2014). Training the SA-FNN using Chl a, NPP or NCP, there was no overall improvement in the broad-scale accuracy of pCO$_2$ $_{(sw)}$ compared to literature values in the South Atlantic. When the uncertainties in the input parameters were investigated however, differences in the estimates of pCO$_2$ $_{(sw)}$ were apparent. The perturbation analysis indicated that up to a 36 % improvement in estimating pCO$_2$ $_{(sw)}$ could be achieved if NCP data uncertainties were reduced. A similar improvement could be obtained if the NPP uncertainties were reduced. Ford et al. (2021) showed that up to 40 % of the uncertainty in satellite NCP is attributed to the uncertainty in satellite NPP, which is an input to the NCP approach. This suggests that improvements in estimating NPP from satellite data will lead to a further improvement in estimating pCO$_2$ $_{(sw)}$ from NCP. These improvements could be achieved through advances in the water column light field (e.g. Sathyendranath et al., 2020) or assignment of photosynthetic parameters (e.g. Kulk et al., 2020), for example. For a discussion on improving satellite NPP estimates we refer the reader to Lee et al. (2015).

Satellite NCP is reliant on NPP as input (Tilstone et al., 2015a; Ford et al., 2021), and NPP requires Chl a as the primary input (e.g. Lobanova et al., 2018; Platt et al., 1991; Tilstone et al., 2015b). To uncouple the Chl a, NPP and NCP estimates and their uncertainties, the perturbation analysis was also conducted on Atlantic Meridional Transect *in situ* observations. This showed that reducing *in situ* NCP uncertainties provided the greatest reduction in pCO$_2$ $_{(sw)}$ RMSD, which was three times the reduction achievable using Chl a. This indicates that the optimal predictive power of Chl a to estimate pCO$_2$ $_{(sw)}$ has been reached and to achieve further improvements in estimates of pCO$_2$ $_{(sw)}$ and reduction in its associated uncertainty, requires the use of NCP.

A reduction of input uncertainties to ~0 is near impossible, but a reduction by 10 % could be feasible (e.g. NCP uncertainty reduced from 45 to 40.5 mmol O$_2$ m^{-2} d^{-1}; Table 1). A perturbation analysis conducted for this showed similar results, with NCP producing the greatest reduction in pCO$_2$ $_{(sw)}$ RMSD of 8 % compared to 2 % for Chl a. Thus reducing NCP uncertainties will provide a greater improvement in pCO$_2$ $_{(sw)}$ compared to reducing the uncertainties in Chl a.

These improvements in estimating NCP could be achieved through many components. Ford et al. (2021) showed 40 % of satellite NCP uncertainties were attributed to *in situ* NCP uncertainties. The *in situ* bottle incubation measurements could be improved using the principles of Fiducial Reference Measurements (FRM; Banks et al., 2020), which are traceable to metrology standards, referenced to inter-comparison exercises, with a full uncertainty budget. This however, becomes complicated when considering the number of different methods to measure NCP and the large divergence between them (Robinson et al., 2009). A review of these methods has already been conducted (Duarte et al., 2013; Ducklow and Doney, 2013; Williams et al., 2013). The methods broadly fall into the following categories: a.) *in vitro* incubations of samples

under light/dark treatments (Gist et al., 2009) and b.) *in situ* observations of oxygen to argon (O_2/Ar) ratios (Kaiser et al., 2005) or the observed isotopic signature of oxygen (Kroopnick, 1980; Luz and Barkan, 2000). All of these methods are

315 subject to, but do not account for, the photochemical sink which may lead to underestimation of *in vitro* NCP by up to 22 % (Kitidis et al., 2014). Independent ground measurements that use accepted protocols for the *in vitro* method are currently made on the Atlantic Meridional Transect, however a community consensus should consider a consistent methodology for NCP. Increasing the number of such observations for the purpose of algorithm development, would further constrain the NCP, but also provide observations across the lifetime of newly launched satellites. The uncertainties on each *in vitro*

320 measurement are assessed through replicate bottles which could be used to calculate a full uncertainty budget for each NCP measurement when combined with analytical uncertainties.

Serret et al. (2015) indicated that NCP is controlled by both the heterogeneity in NPP and respiration. The satellite NCP algorithm applied in this study accounts for some of the heterogeneity in respiration, through an empirical SST to NCP relationship (Tilstone et al., 2015a). Quantifying the variability in respiration could further improve NCP estimates when

325 coupled with NPP rates from satellite observations.

4.2. Accuracy of SA-FNN$_{NCP}$ pCO$_{2 (sw)}$ at seasonal and interannual scales

The seasonal and interannual variability of pCO$_{2 (sw)}$ estimated using the SA-FNN$_{NCP}$ was compared with the SA-FNN$_{NO\text{-}BIO}$, W2020 (Watson et al., 2020b), SA-FNN$_{CHL}$ and SA-FNN$_{NPP}$ at 8 stations. The stations (Fig. 1) represent locations of previous studies into *in situ* pCO$_{2 (sw)}$ variability in the South Atlantic Ocean and allow comparisons with literature values.

330 Significant differences between the SA-FNN$_{NCP}$ and SA-FNN$_{NO\text{-}BIO}$ were observed at four stations (Fig. 4), especially in the Equatorial Atlantic.

At 8° N 38° W (Fig. 3a), Lefèvre et al. (2020) reported pCO$_{2 (sw)}$ to be stable at ~400 µatm, between June and August 2013, and to decrease in September to ~360 µatm, which is attributed to the Amazon Plume propagating into the western Equatorial Atlantic (Coles et al., 2013). Bruto et al. (2017) indicated however, that elevated pCO$_{2 (sw)}$ at ~430 µatm exist

335 throughout the year from 2008 to 2011. For the station in the Amazon Plume at 4° N 50° W (Fig. 3g), where the effects of the plume extend northwest towards the Caribbean (Coles et al., 2013; Varona et al., 2019), Lefèvre et al. (2017) indicated that this region acts as a sink for CO_2 (pCO$_{2 (sw)}$ < pCO$_{2 (atm)}$), especially between May to July, coincident with maximum discharge from the Amazon River (Dai and Trenberth, 2002). Valerio et al. (2021) indicated pCO$_{2 (sw)}$ varied above and below pCO$_{2 (atm)}$ at 4° N 50° W consistent with the SA-FNN$_{NCP}$. The interannual variability of pCO$_{2 (sw)}$ has been shown to be

340 high in this region in all months (Lefèvre et al., 2017). The SA-FNN$_{NCP}$ provided a better representation of the seasonal and interannual variability induced by the Amazon River discharge and associated plume at these two stations compared to the SA-FNN$_{NO\text{-}BIO}$, although differences were small at 8° N 38° W.

The station in the Eastern Tropical Atlantic at 6° S 10° W (Fig. 3b), is under the influence of the equatorial upwelling (Lefèvre, Guillot, Beaumont, & Danguy, 2008), which is associated with upwelling of CO_2 rich waters between June and

345 September. Lefèvre et al. (2008) indicated that peak pCO$_{2 (sw)}$ of ~440 µatm was observed in September, and remained stable

14

until December, before decreasing to a minima of ~360 μatm in May (Parard et al., 2010). Lefèvre et al. (2016) showed however, that the influence of the equatorial upwelling does not reach the buoy in all years, and in some years lower pCO_2 (sw) is observed. Further north at the station at 4° N 10° W (Fig. 3f), Koffi et al. (2010) suggested that this region follows a similar seasonal cycle as the station at 6° S 10° W, but that $pCO_{2\,(sw)}$ is ~30 μatm lower (Koffi et al., 2016). The interannual

350   variability in SA-FNN$_{NCP}$ $pCO_{2\,(sw)}$ clearly shows the influence of the equatorial upwelling at these stations, with latitudinal gradients in $pCO_{2\,(sw)}$ during the upwelling period (Lefèvre et al., 2016). By contrast, the SA-FNN$_{NO-BIO}$ indicated little influence from the equatorial upwelling, little interannual variability, and a depressed $pCO_{2\,(sw)}$ during the upwelling season. The two methods converge on the seasonal cycle at the remaining stations although significant offsets in the mean annual $pCO_{2\,(sw)}$ remain. The station at 35° S 18° W (Fig. 3c) has consistently been implied as a sink for CO_2. Lencina-Avila et al.

355   (2016) showed the region to have $pCO_{2\,(sw)}$ at 340 μatm and to be a sink for CO_2 between October to December. Similarly, Kitidis et al. (2017) implied that the region is a sink for CO_2 during March to April. The region has depressed $pCO_{2\,(sw)}$ due to high biological activity that originates from the Patagonian shelf and the South Subtropical Convergence Zone. The station at 45° S 50° W (Fig. 3d), has also been implied as a strong, but highly variable sink, where $pCO_{2\,(sw)}$ can be between ~280 μatm and ~380 μatm during austral spring, and is constant at ~310 μatm during austral autumn (Kitidis et al., 2017).

360   The SA-FNN$_{NCP}$ and SA-FNN$_{NO-BIO}$ methods reproduced the seasonal variability in the $pCO_{2\,(sw)}$ at these two stations accurately, but only the SA-FNN$_{NCP}$ captures the magnitude of the depressed $pCO_{2\,(sw)}$ at 45° S.

Within the southern Benguela upwelling system, $pCO_{2\,(sw)}$ at the station 33° S 17° E (Fig. 3e) is influenced by gradients in the seasonal upwelling (Hutchings et al., 2009). Santana-Casiano et al. (2009) showed that $pCO_{2\,(sw)}$ varies from ~310 μatm in July to ~340 μatm in December and that the region is a CO_2 sink through the year. González-Dávila et al. (2009)

365   suggested however, that this CO_2 sink is highly variable during upwelling events, and that recently upwelled waters act as a source ($pCO_{2\,(sw)} > pCO_{2\,(atm)}$) of CO_2 to the atmosphere (Gregor and Monteiro, 2013). The SA-FNN$_{NCP}$ and SA-FNN$_{NO-BIO}$ were able to reproduce the seasonal cycle, although the SA-FNN$_{NCP}$ correctly represented the seasonal differences in pCO_2 (sw) as reported by Santana-Casiano et al. (2009).

Overall, compared to the SA-FNN$_{NO-BIO}$ at these stations, the SA-FNN$_{NCP}$ better represents the seasonality and the

370   interannual variability of $pCO_{2\,(sw)}$ in the South Atlantic Ocean, especially in the Equatorial Atlantic. The SA-FNN$_{NO-BIO}$ and W2020 both displayed significant differences to the SA-FNN$_{NCP}$ at similar stations (Fig. 4) although their $pCO_{2\,(sw)}$ estimates were not always consistent. The SA-FNN method uses only *in situ* $pCO_{2\,(sw)}$ observations from the South Atlantic Ocean to train the FNNs. The W2020 uses global *in situ* $pCO_{2\,(sw)}$ observations to train FNNs for 16 provinces with similar seasonal cycles (Landschützer et al., 2014; Watson et al., 2020b). The W2020 will therefore be weighted to $pCO_{2\,(sw)}$ variability in

375   regions of relatively abundant *in situ* observations (i.e. Northern Hemisphere) and may not be fully representative of the South Atlantic Ocean. This would explain the SA-FNN$_{NO-BIO}$ and W2020 differences, when driven using the same input variables.

Comparing the SA-FNN$_{NCP}$ and SA-FNN$_{CHLA}$ there were two significant differences (Fig. 4). A difference in the seasonal cycle in the southern Benguela (Fig. 3e) was observed. Santana-Casiano et al. (2009) showed that the minima $pCO_{2\,(sw)}$ in

380   July and maxima in December, consistent with the SA-FNN$_{NCP}$ and SA-FNN$_{NPP}$ whereas the SA-FNN$_{CHL}$ estimated the opposite scenario. Lamont et al. (2014) reported Chl a concentrations to remain consistent in May and October, but NPP rates were significantly higher in October, associated with increased surface PAR and enhanced upwelling. The disconnect between Chl a, NPP and NCP limits the ability of Chl a to estimate $pCO_{2\,(sw)}$, which is highlighted by the failure of the SA-FNN$_{CHLA}$ to identify the seasonal $pCO_{2\,(sw)}$ cycle.

385   A Chl a to NPP disconnect, due to light limitation caused by suspended sediments, has also been reported in the Amazon Plume (Smith and Demaster, 1996), where a significant offset between the SA-FNN$_{NCP}$ and SA-FNN$_{CHLA}$ was observed (Fig. 3g; Fig. 4). Lefèvre et al. (2017) reported $pCO_{2\,(sw)}$ values ranging from 400 ± ~10 µatm in January to ~240 ± ~70 µatm in May. Although, the SA-FNN$_{NCP}$ January estimates are consistent, the May estimates are higher than these *in situ* observations. These observations were made further north (6° N) where the turbidity within the plume has decreased

390   sufficiently for irradiance to elevate NPP rates (Smith and Demaster, 1996), which decrease $pCO_{2\,(sw)}$. Chl a remains relatively consistent across the plume (not shown), suggesting a disconnect between Chl a and NPP at 4° N 50° W which would lead to lower $pCO_{2\,(sw)}$ estimates by the SA-FNN$_{CHLA}$, where NPP rates are low due to light limitation (Chen et al., 2012; Smith and Demaster, 1996). Respiration would be elevated from the decomposition of riverine organic material reducing NCP further (Cooley et al., 2007; Jiang et al., 2019; Lefèvre et al., 2017). It is noted that the Amazon Plume is a

395   dynamic region with transient, localised biological and $pCO_{2\,(sw)}$ features (Cooley et al., 2007; Ibánhez et al., 2015; Lefèvre et al., 2017; Valerio et al., 2021) that maybe masked by the coarse resolution of estimates available using satellite data. The SA-FNN$_{NCP}$ however, agreed with *in situ* $pCO_{2\,(sw)}$ observations at 4° N 50° W where $pCO_{2\,(sw)}$ varied above and below $pCO_{2\,(atm)}$ (Valerio et al., 2021).

Though the differences between the SA-FNN$_{NCP}$ and SA-FNN$_{CHLA}$ may appear small, the Amazon Plume and Benguela

400   Upwelling have a higher intensity in the CO_2 flux per unit area compared to the open ocean, illustrating a disproportionate contribution to the overall global CO_2 sink than their small areal coverage implies (Laruelle et al., 2014). The differences in the $pCO_{2\,(sw)}$ estimates result in a 22 Tg C yr^{-1} alteration in the annual CO_2 flux for the South Atlantic Ocean (SA-FNN$_{NCP}$ = +14 Tg C yr^{-1}; SA-FNN$_{CHLA}$ = -9 Tg C yr^{-1}; Fig. 5). This unequivocally reinforces the use of NCP to improve basin scale estimates of $pCO_{2\,(sw)}$, especially in regions where Chl a, NPP and NCP become disconnected.

405   Recent assessments of the strength of the global oceanic CO_2 sink have been made using $pCO_{2\,(sw)}$ fields estimated using no biological parameters as input (Watson et al., 2020b). Our results indicate that the SA-FNN$_{NCP}$ more accurately represented the $pCO_{2\,(sw)}$ variability in the South Atlantic Ocean compared to the SA-FNN$_{NO-BIO}$. Estimating the South Atlantic Ocean net CO_2 flux with the SA-FNN$_{NCP}$ $pCO_{2\,(sw)}$ produced a 14 Tg C yr^{-1} source compared to a 10 Tg C yr^{-1} sink indicated by the SA-FNN$_{NO-BIO}$ (Fig. 5). The incremental inclusion of parameters to properly account for the biological signal starting with

410   Chl a (-9 Tg C yr^{-1}) to NPP (-7 Tg C yr^{-1}) to NCP (+14 Tg C yr^{-1}) switched the South Atlantic Ocean from a CO_2 sink to a source, driven by differences in the $pCO_{2\,(sw)}$ estimates in regions that are biologically controlled, such as the Equatorial Atlantic.
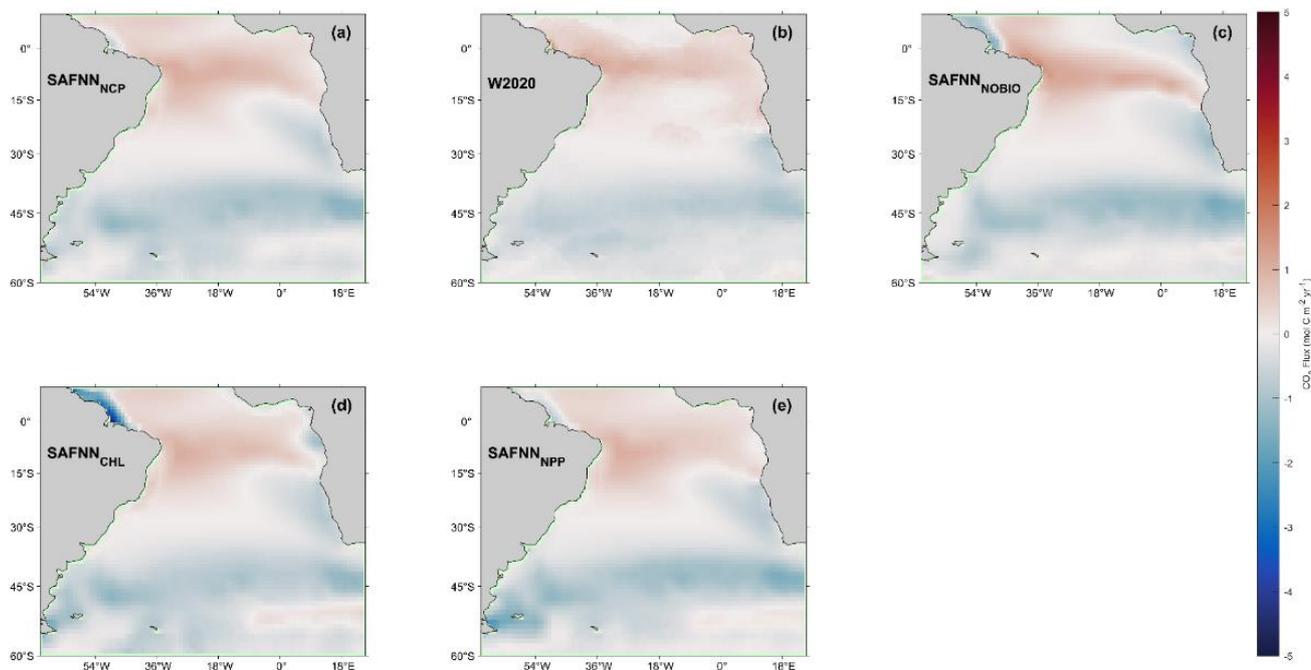
Fig. 5: Long term average annual mean CO₂ flux for the South Atlantic Ocean, using $pCO_{2\,(sw)}$ estimates from (a) SA-FNN$_{NCP}$, (b) W2020 (Watson, et al., 2020a), (c) SA-FNN$_{NOBIO}$, (d) SA-FNN$_{CHLA}$ and (e) SA-FNN$_{NPP}$

The W2020 identified the South Atlantic Ocean as a 15 Tg C yr^{-1} source for CO₂ consistent with the SA-FNN$_{NCP}$ (Fig. 5). The SA-FNN$_{NCP}$ however, indicated the Equatorial Atlantic (10° N to 20° S) as a 20 Tg C yr^{-1} stronger source and south of 20° S (20° S to 44° S) as a 20 Tg C yr^{-1} stronger sink. These differences indicate that biologically induced variability in $pCO_{2\,(sw)}$ would not be captured by the W2020 and may reduce the variability in the global ocean CO₂ sink reinforcing the improvement that NCP provides.

5. Conclusions

In this paper, we compare using three biological proxies, Chl *a*, NPP or NCP available from earth observation data to train a neural network scheme to estimate $pCO_{2\,(sw)}$. The results suggest that using NCP improved the estimation of $pCO_{2\,(sw)}$. The differences between satellite Chl *a*, NPP or NCP were initially small, but the use of a perturbation analysis to account for uncertainties in these parameters, showed that NCP has the greatest potential reduction in $pCO_{2\,(sw)}$ uncertainty of up to ~36 % of the RMSD, compared to a ~19 % reduction for Chl *a*. These results were verified using *in situ* observations from the Atlantic Meridional Transect, which resulted in a 25 % improvement in $pCO_{2\,(sw)}$ RMSD when the *in situ* NCP uncertainties

430 are reduced, compared to 7 % for Chl *a* and 13 % for NPP. Practical approaches to reduce the uncertainties in both the *in situ* and satellite NCP observations are discussed.

Monthly climatological estimates of $pCO_{2 (sw)}$ calculated using satellite NCP were compared with the NPP and the Chl *a* approaches and a baseline approach that does not use biological parameters, at 8 stations in the South Atlantic Ocean. The NCP approach significantly improved on the baseline approach at 4 stations in reconstructing the seasonal and interannual

435 variability, compared to *in situ* $pCO_{2 (sw)}$ observations. At the remaining 4 stations, differences were also observed although these were not statistically significant. A significant difference between the NCP and NPP approaches occurred in the eastern Equatorial Atlantic, in the equatorial upwelling region. Significant differences between the NCP and Chl *a* approaches were also observed in the Benguela upwelling and Amazon Plume, where $pCO_{2 (sw)}$ from Chl *a* suggested that photosynthetic rates were not solely controlled by Chl *a*. Using $pCO_{2 (sw)}$ estimated from NCP identified the South Atlantic Ocean as a net source

440 of CO_2, whereas methods that only include physical controls have indicated it to be a small sink for CO_2. Sequentially using firstly Chl *a* to estimate $pCO_{2 (sw)}$, then NPP incrementally reduced the South Atlantic CO_2 sink and finally using NCP switched the area to a source of CO_2. These results indicate that in regions where biological activity is important in controlling the variability in $pCO_{2 (sw)}$, the use of NCP available from satellite data is important for quantifying the ocean carbon pump, and for providing data in areas that are sparsely covered by observations such as the Southern Ocean.
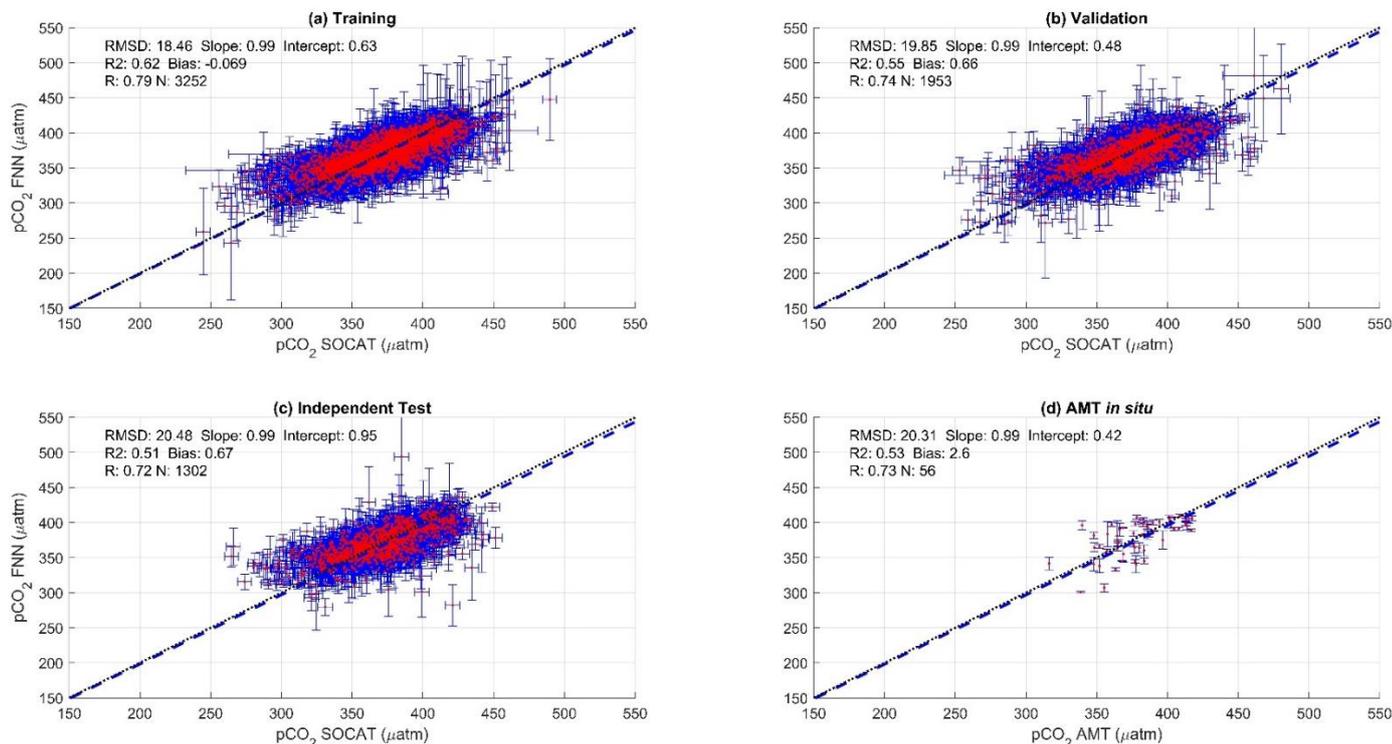
445

Appendices

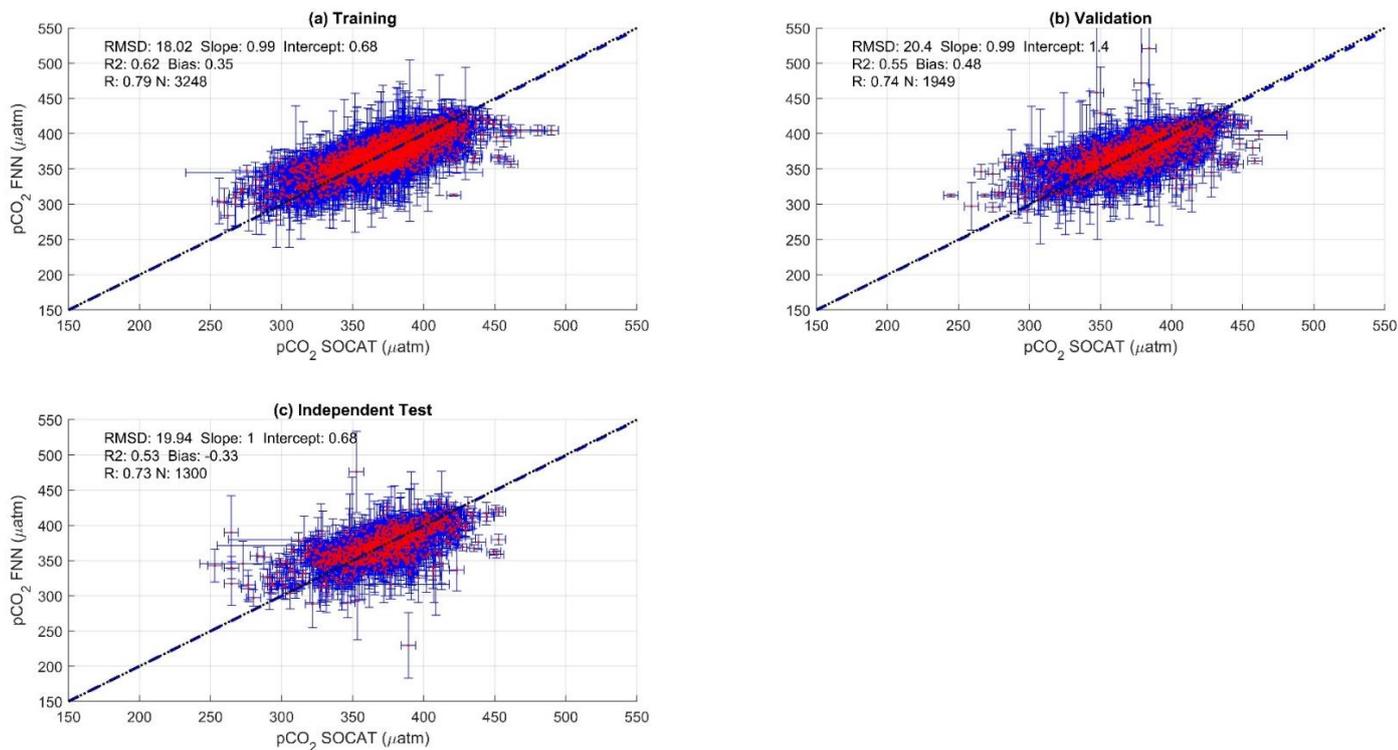### Appendix A - Feed forward neural network training and perturbation analysis
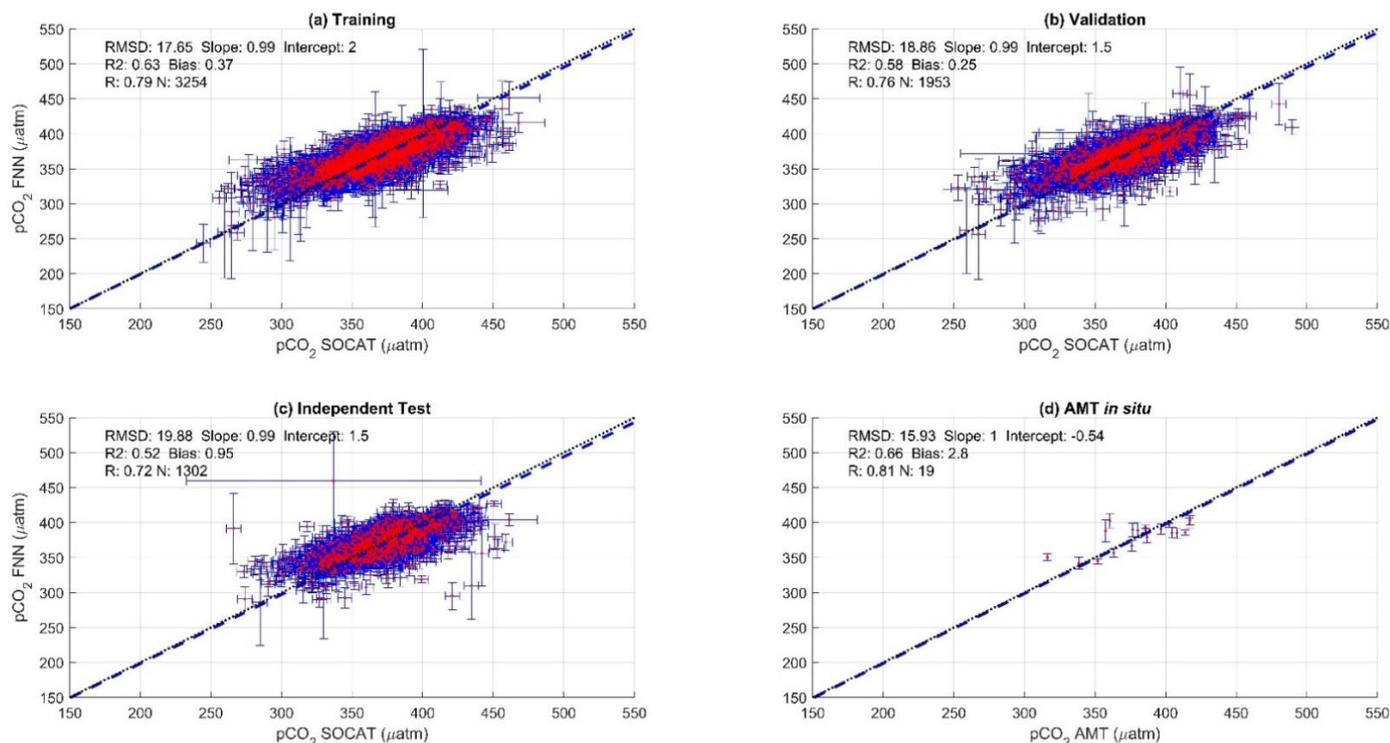
Fig. A1: Scatter plots showing the combined performance of the 8 feed forward neural networks trained using chlorophyll a for 4
450   separate training and validation datasets; (a) Training, (b) Validation, (c) Independent Test and (d) Atlantic Meridional Transect
(AMT) in situ. The blue dashed line is the Type II regression and the black dashed line is the 1:1 line. Horizontal errorbars
indicate the uncertainty of the SOCATv2020 pCO_2 (sw). Vertical errorbars indicate the uncertainty attributed to the input
parameter uncertainty propagated through the feed forward neural networks. The statistics within each plot are; Root Mean
Square Difference (RMSD), Slope and Intercept of the Type II regression, Coefficient of Determination (R2), Pearson's
455   Correlation Coefficient (R), Bias and number of samples (N).

Fig. A2: Scatter plots showing the combined performance of the 8 feed forward neural networks trained using net primary production for 4 separate training and validation datasets; (a) Training, (b) Validation, (c) Independent Test and (d) Atlantic Meridional Transect (AMT) in situ. The blue dashed line is the Type II regression and the black dashed line is the 1:1 line. Horizontal errorbars indicate the uncertainty of the SOCATv2020 $pCO_{2\,(sw)}$. Vertical errorbars indicate the resulting uncertainty attributed to the input parameter uncertainty propagated through the feed forward neural networks. The statistics within each plot are; Root Mean Square Difference (RMSD), Slope and Intercept of the Type II regression, Coefficient of Determination (R2), Pearson's Correlation Coefficient (R), Bias and number of samples (N).

465

Fig. A3: Scatter plots showing the combined performance of the 8 feed forward neural networks trained using no biological parameters for 3 separate training and validation datasets; (a) Training, (b) Validation and (c) Independent Test. The blue dashed line is the Type II regression and the black dashed line is the 1:1 line. Horizontal errorbars indicate the uncertainty of the SOCATv2020 $pCO_{2\ (sw)}$. Vertical errorbars indicate the resulting uncertainty attributed to the input parameter uncertainty
470 **propagated through the feed forward neural networks. The statistics within each plot are; Root Mean Square Difference (RMSD), Slope and Intercept of the Type II regression, Coefficient of Determination (R2), Pearson's Correlation Coefficient (R), Bias and number of samples (N).**

Table A1: The percentage reduction in Root Mean Square Difference (RMSD) attributable to the uncertainties in the input parameter for each training and validation datasets determined from a perturbation analysis as described in Sect. 2.5.

|	Parameter	Training	Validation	Independent Test	AMT *in situ*
NCP	ALL	33 %	42 %	38 %	28 %
	SST	10 %	12 %	10 %	0.5 %
	Net Community Production	32 %	40 %	36 %	25 %
	pCO$_{2\,(atm)}$	6 %	7 %	6 %	9 %
Net Primary Production	ALL	34 %	40 %	40 %	17 %
	SST	9 %	10 %	10 %	0.4 %
	Net Primary Production	31 %	37 %	36 %	13 %
	pCO$_{2\,(atm)}$	6 %	6 %	6 %	9 %
Chlorophyll *a*	ALL	22 %	26 %	25 %	29 %
	SST	9 %	10 %	9 %	0.4 %
	Chlorophyll *a*	17 %	21 %	20 %	7 %
	pCO$_{2\,(atm)}$	8 %	9 %	9 %	16 %

Appendix B - Climatology comparison

A monthly climatology was generated from the SA-FNN$_{NCP}$ monthly timeseries (Fig. B1), referenced to the year 2010, assuming an atmospheric CO$_2$ increase of 1.5 µatm yr^{-1} (Takahashi et al., 2009; Zeng et al., 2014). The standard deviation of the monthly climatology was computed, as an indication of the interannual variations in the climatology. The ability of the SA-FNN$_{NCP}$ to estimate the spatial distribution of pCO$_{2\,(sw)}$ was compared to two methods.

Firstly, the SA-FNN$_{NCP}$ climatology was compared to the climatology from Woolf et al. (2019), produced following the statistical 'ordinary block kriging' approach described in Goddijn-Murphy et al. (2015), using the SOCATv4 reanalysed data. The method provides an interpolation uncertainty where in regions of sparse data this becomes larger. Fig. B2 shows the methods produce similar climatological pCO$_{2\,(sw)}$ values for the South Atlantic Ocean, with some clear differences along the African coastline, and equatorial region.

Secondly, the SA-FNN$_{NCP}$ was compared to a climatology calculated from the 'standard method', a Self Organising Map Feed Forward Neural Network presented in Watson et al. (2020b; W2020). Fig. B3 shows the methods produce similar

climatological pCO$_2$ $_{(sw)}$ values for the South Atlantic Ocean, however, clear differences in the Equatorial region occur across all months. In the central South Atlantic Ocean, artefacts form the self organising map can be seen during January and February.
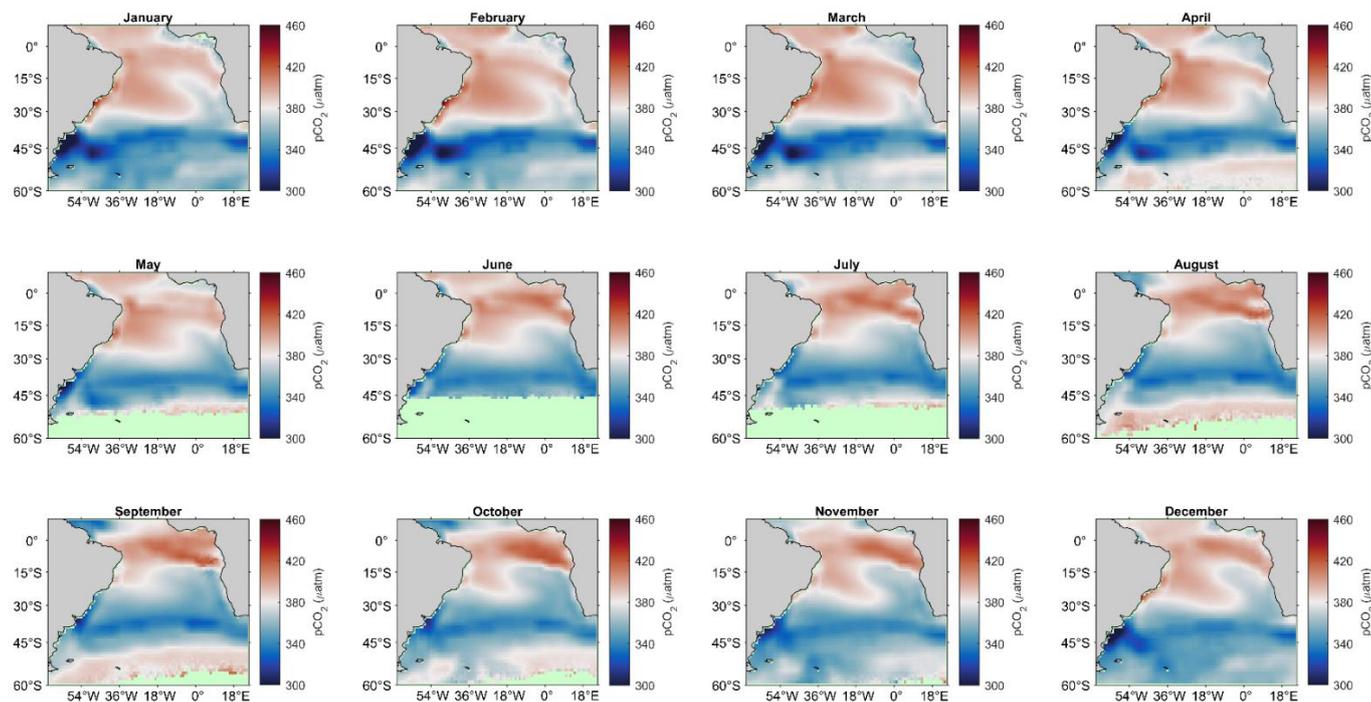
Fig. B1: Monthly climatologies of pCO$_2$ $_{(sw)}$ between July 2002 and December 2018 estimated by the SA-FNN$_{NCP}$ approach referenced to 2010. The atmospheric CO$_2$ increase was set as 1.5 µatm yr^{-1}. The colour scale is centred on the atmospheric concentration for 2010 (~380 µatm). Red shaded areas indicate oversaturated regions, and blue shaded areas indicate under saturated regions. Light green areas indicate where no input data to compute pCO$_2$ $_{(sw)}$ are available.
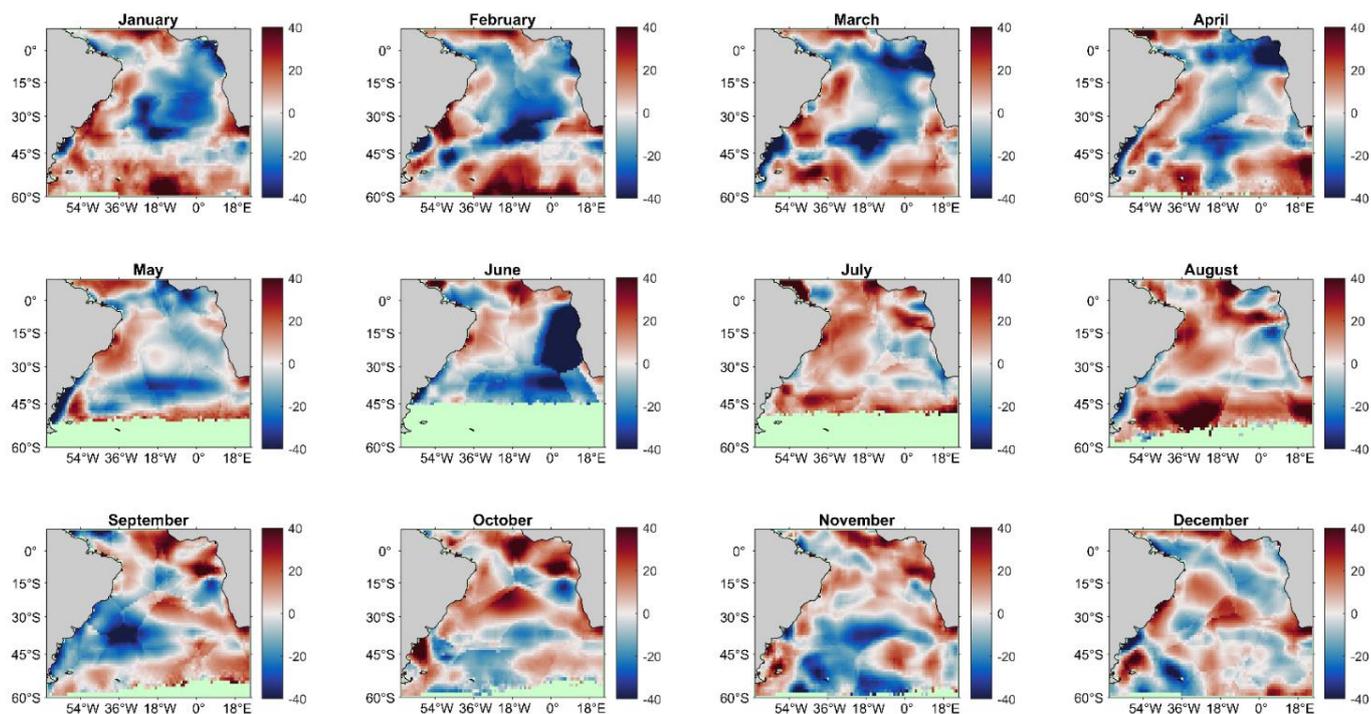
500 **Fig. B2: Monthly comparison between pCO$_2$ $_{(sw)}$ climatology estimated by the SA-FNN$_{NCP}$ and Woolf et al (2019) climatology referenced to 2010 (SA-FNN$_{NCP}$ pCO$_2$ − Woolf pCO$_2$). Red (Blue) shades indicate regions where SA-FNN is greater (less) than the Woolf climatology.**
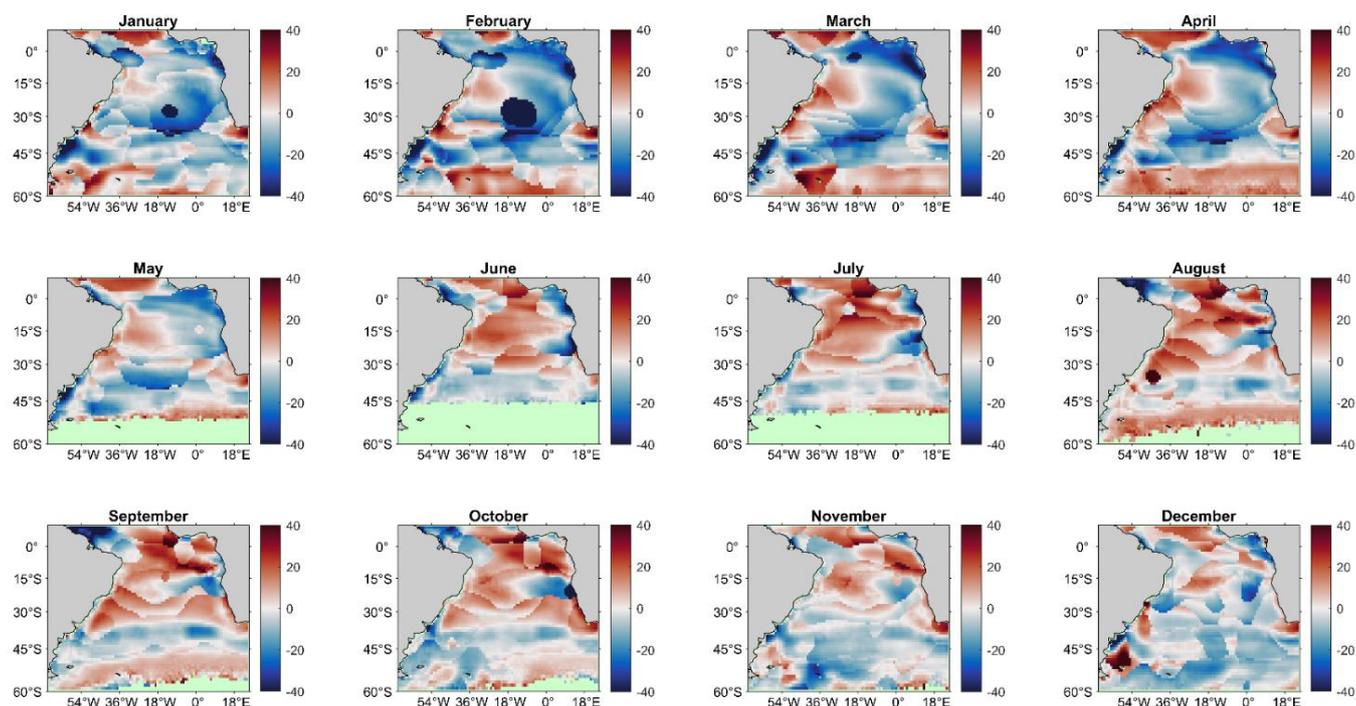
Fig. B3: Monthly comparison between pCO$_2$ $_{(sw)}$ climatologies estimated by the SA-FNN$_{NCP}$ and W2020 (Watson et al, 2020)
climatology referenced to 2010 (SA-FNN$_{NCP}$ pCO$_2$ – W2020 pCO$_2$). Red (Blue) shades indicate regions where SA-FNN$_{NCP}$ is
greater (less) than the W2020 climatology.

Data Availability

Moderate Resolution Imaging Spectroradiometer on Aqua (MODIS-A) estimates of chlorophyll-a, photosynthetically active
radiation and sea surface temperature are available from the National Aeronautics Space Administration (NASA) ocean
colour website (https://oceancolor.gsfc.nasa.gov/). Modelled sea surface salinity and mixed layer depth from the Copernicus
Marine Environment Modelling Service global ocean physics reanalysis product (GLORYS12V1) are available from
https://resources.marine.copernicus.eu/. ERA5 monthly reanalysis wind speeds are available from the Copernicus Climate
Data Store (https://cds.climate.copernicus.eu/) pCO$_2$ $_{(atm)}$ data are available from v5.5 of the global estimates of pCO$_2$ $_{(sw)}$
dataset (Landschützer et al., 2016, 2017). *In situ* observations of fCO$_2$ $_{(sw)}$ from v2020 of the Surface Ocean Carbon Atlas
(SOCAT) are available from https://www.socat.info/index.php/data-access/. *In situ* Atlantic Meridional Transect data can be
obtained from the British Oceanographic Data Centre (https://www.bodc.ac.uk/). pCO$_2$ $_{(sw)}$ estimates from the W2020 are

available from Watson et al. (2020b). $pCO_{2\ (sw)}$ estimates generated by the SA-FNN$_{NCP}$, SA-FNN$_{NPP}$, SA-FNN$_{CHLA}$ and SA-FNN$_{NO\text{-}BIO}$ are available from Pangaea (Ford et al., submitted).

520

Author Contribution

DF, GT, JS and VK conceived and directed the research. DF developed the code and prepared the manuscript. GT, JS and VK provided comments that shaped the final manuscript.

Competing Interests

525 The authors declare that they have no conflict of interest.

References

Amari, S. I., Murata, N., Müller, K. R., Finke, M. and Yang, H. H.: Asymptotic statistical theory of overtraining and cross-validation, IEEE Trans. Neural Networks, 8(5), 985–996, doi:10.1109/72.623200, 1997.
Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. 545 D., Nakaoka, S. I., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof,

R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W. J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibánhez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A.,

550   Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., Van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J. and Xu, S.: A multi-decade record of high-quality fCO_2 data in version 3 of the

555   Surface Ocean CO_2 Atlas (SOCAT), Earth Syst. Sci. Data, 8(2), 383–413, doi:10.5194/essd-8-383-2016, 2016.

Banks, A. C., Vendt, R., Alikas, K., Bialek, A., Kuusk, J., Lerebourg, C., Ruddick, K., Tilstone, G., Vabson, V., Donlon, C. and Casal, T.: Fiducial reference measurements for satellite ocean colour (FRM4SOC), Remote Sens., 12(8), doi:10.3390/RS12081322, 2020.

Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, Limnol.
560   Oceanogr., 42(1), 1–20, doi:10.4319/lo.1997.42.1.0001, 1997.

Behrenfeld, M. J., O'Malley, R. T., Boss, E. S., Westberry, T. K., Graff, J. R., Halsey, K. H., Milligan, A. J., Siegel, D. A. and Brown, M. B.: Revaluating ocean warming impacts on global phytoplankton, Nat. Clim. Chang., 6(3), 323–330, doi:10.1038/nclimate2838, 2016.

Benallal, M. A., Moussa, H., Lencina-Avila, J. M., Touratier, F., Goyet, C., El Jai, M. C., Poisson, N. and Poisson, A.:
565   Satellite-derived CO_2 flux in the surface seawater of the Austral Ocean south of Australia, Int. J. Remote Sens., 38(6), 1600–1625, doi:10.1080/01431161.2017.1286054, 2017.

BIPM: Evaluation of measurement data—Guide to the expression of uncertainty in measurement. [online] Available from: http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf, 2008.

Brewin, R. J. W., Dall'Olmo, G., Pardo, S., van Dongen-Vogels, V. and Boss, E. S.: Underway spectrophotometry along the
570   Atlantic Meridional Transect reveals high performance in satellite chlorophyll retrievals, Remote Sens. Environ., doi:10.1016/j.rse.2016.05.005, 2016.

Bruto, L., Araujo, M., Noriega, C., Veleda, D. and Lefèvre, N.: Variability of CO_2 fugacity at the western edge of the tropical Atlantic Ocean from the 8°N to 38°W PIRATA buoy, Dyn. Atmos. Ocean., 78, 1–13, doi:10.1016/j.dynatmoce.2017.01.003, 2017.

575   Chen, C. T. A., Huang, T. H., Fu, Y. H., Bai, Y. and He, X.: Strong sources of CO_2 in upper estuaries become sinks of CO_2 in large river plumes, Curr. Opin. Environ. Sustain., 4(2), 179–185, doi:10.1016/j.cosust.2012.02.003, 2012.

Chierici, M., Signorini, S. R., Mattsdotter-Björk, M., Fransson, A. and Olsen, A.: Surface water fCO_2 algorithms for the high-latitude Pacific sector of the Southern Ocean, Remote Sens. Environ., 119, 184–196, doi:10.1016/j.rse.2011.12.020, 2012.

580   Coles, V. J., Brooks, M. T., Hopkins, J., Stukel, M. R., Yager, P. L. and Hood, R. R.: The pathways and properties of the Amazon river plume in the tropical North Atlantic Ocean, J. Geophys. Res. Ocean., 118(12), 6894–6913, doi:10.1002/2013JC008981, 2013.

Cooley, S. R., Coles, V. J., Subramaniam, A. and Yager, P. L.: Seasonal variations in the Amazon plume-related atmospheric carbon sink, Global Biogeochem. Cycles, 21(3), 1–15, doi:10.1029/2006GB002831, 2007.

585   Dai, A. and Trenberth, K. E.: Estimates of freshwater discharge from continents: Latitudinal and seasonal variations, J. Hydrometeorol., 3(6), 660–687, doi:10.1175/1525-7541(2002)003<0660:EOFDFC>2.0.CO;2, 2002.

Demuth, H., Beale, M. and Hagan, M.: Neural Network Toolbox 6 Users Guide, 3 Apple Hill Drive, Natick, MA., 2008.

DeVries, T.: The oceanic anthropogenic CO2 sink: Storage, air-sea fluxes, and transports over the industrial era, Global Biogeochem. Cycles, 28(7), 631–647, doi:10.1002/2013GB004739, 2014.

590   Dickson, A. G., Sabine, C. L. and Christian, J. R.: Guide to Best Practices for Ocean CO_2 measurements., 2007.

Donlon, C. J., Nightingale, T. J., Sheasby, T., Turner, J., Robinson, I. S. and Emergy, W. J.: Implications of the oceanic thermal skin temperature deviation at high wind speed, Geophys. Res. Lett., 26(16), 2505–2508, doi:10.1029/1999GL900547, 1999.

Duarte, C. M., Regaudie-de-Gioux, A., Arrieta, J. M., Delgado-Huertas, A. and Agustí, S.: The Oligotrophic Ocean Is
595   Heterotrophic, Ann. Rev. Mar. Sci., 5(1), 551–569, doi:10.1146/annurev-marine-121211-172337, 2013.

Ducklow, H. W. and Doney, S. C.: What Is the Metabolic State of the Oligotrophic Ocean? A Debate, Ann. Rev. Mar. Sci., 5(1), 525–533, doi:10.1146/annurev-marine-121211-172331, 2013.

Ford, D., Tilstone, G. H., Shutler, J. D., Kitidis, V., Lobanova, P., Schwarz, J., Poulton, A. J., Serret, P., Lamont, T., Chuqui, M., Barlow, R., Lozano, J., Kampel, M. and Brandini, F.: Wind speed and mesoscale features drive net autotrophy in the
600 South Atlantic Ocean, Remote Sens. Environ., 260, 112435, doi:10.1016/j.rse.2021.112435, 2021.

Ford, D., Tilstone, G. H., Shutler, J. D. and Kitidis, V.: Interpolated surface ocean carbon dioxide partial pressure for the South Atlantic Ocean (2002-2018) using different biological parameters, PANGAEA, submitted

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates,
605 N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L.,
610 Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X. and Zaehle, S.: Global Carbon Budget 2020, Earth Syst. Sci. Data, 12(4), 3269–3340, doi:10.5194/essd-12-3269-2020, 2020.

Gist, N., Serret, P., Woodward, E. M. S., Chamberlain, K. and Robinson, C.: Seasonal and spatial variability in plankton
615 production and respiration in the Subtropical Gyres of the Atlantic Ocean, Deep. Res. Part II Top. Stud. Oceanogr., 56(15), 931–940, doi:10.1016/j.dsr2.2008.10.035, 2009.

Goddijn-Murphy, L. M., Woolf, D. K., Land, P. E., Shutler, J. D. and Donlon, C.: The OceanFlux Greenhouse Gases methodology for deriving a sea surface climatology of CO_2 fugacity in support of air-sea gas flux studies, Ocean Sci., 11(4), 519–541, doi:10.5194/os-11-519-2015, 2015.

620 González-Dávila, M., Santana-Casiano, J. M. and Ucha, I. R.: Seasonal variability of fCO_2 in the Angola-Benguela region, Prog. Oceanogr., 83(1–4), 124–133, doi:10.1016/j.pocean.2009.07.033, 2009.

Gregor, L. and Monteiro, P. M. S.: Is the southern benguela a significant regional sink of CO_2?, S. Afr. J. Sci., 109(5–6), 1–5, doi:10.1590/sajs.2013/20120094, 2013.

Holding, T., Ashton, I. G., Shutler, J. D., Land, P. E., Nightingale, P. D., Rees, A. P., Brown, I., Piolle, J.-F., Kock, A.,
625 Bange, H. W., Woolf, D. K., Goddijn-Murphy, L., Pereira, R., Paul, F., Girard-Ardhuin, F., Chapron, B., Rehder, G., Ardhuin, F. and Donlon, C. J.: The FluxEngine air–sea gas flux toolbox: simplified interface and extensions for in situ analyses and multiple sparingly soluble gases, Ocean Sci., 15(6), 1707–1728, doi:10.5194/os-15-1707-2019, 2019.

Hu, C., Lee, Z. and Franz, B.: Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference, J. Geophys. Res. Ocean., 117(1), 1–25, doi:10.1029/2011JC007395, 2012.

630 Hutchings, L., van der Lingen, C. D., Shannon, L. J., Crawford, R. J. M., Verheye, H. M. S., Bartholomae, C. H., van der Plas, A. K., Louw, D., Kreiner, A., Ostrowski, M., Fidel, Q., Barlow, R. G., Lamont, T., Coetzee, J., Shillington, F., Veitch, J., Currie, J. C. and Monteiro, P. M. S.: The Benguela Current: An ecosystem of four components, Prog. Oceanogr., 83(1–4), 15–32, doi:10.1016/j.pocean.2009.07.046, 2009.

Ibánhez, J. S. P., Diverrès, D., Araujo, M. and Lefèvre, N.: Seasonal and interannual variability of sea-air CO_2 fluxes in the
635 tropical Atlantic affected by the Amazon River plume, Global Biogeochem. Cycles, 29(10), 1640–1655, doi:10.1002/2015GB005110, 2015.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. B. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, Cambridge University Press, Cambridge, UK., 2013.

640 Jiang, Z. P., Cai, W. J., Lehrter, J., Chen, B., Ouyang, Z., Le, C., Roberts, B. J., Hussain, N., Scaboo, M. K., Zhang, J. and Xu, Y.: Spring net community production and its coupling with the CO_2 dynamics in the surface water of the northern Gulf of Mexico, Biogeosciences, 16(18), 3507–3525, doi:10.5194/bg-16-3507-2019, 2019.

Kaiser, J., Reuer, M. K., Barnett, B. and Bender, M. L.: Marine productivity estimates from continuous O_2/Ar ratio measurements by membrane inlet mass spectrometry, Geophys. Res. Lett., 32(19), n/a-n/a, doi:10.1029/2005GL023459,
645 2005.

Kitidis, V., Tilstone, G. H., Serret, P., Smyth, T. J., Torres, R. and Robinson, C.: Oxygen photolysis in the Mauritanian upwelling: Implications for net community production, Limnol. Oceanogr., 59(2), 299–310, doi:10.4319/lo.2014.59.2.0299, 2014.

650 Kitidis, V., Brown, I., Hardman-mountford, N. and Lefèvre, N.: Surface ocean carbon dioxide during the Atlantic Meridional Transect (1995 – 2013); evidence of ocean acidification, Prog. Oceanogr., 158, 65–75, doi:10.1016/j.pocean.2016.08.005, 2017.

Koffi, U., Lefèvre, N., Kouadio, G. and Boutin, J.: Surface CO_2 parameters and air-sea CO_2 flux distribution in the eastern equatorial Atlantic Ocean, J. Mar. Syst., 82(3), 135–144, doi:10.1016/j.jmarsys.2010.04.010, 2010.

Koffi, U., Kouadio, G. and Kouadio, Y. K.: Estimates and Variability of the Air-Sea CO_2 Fluxes in the Gulf of Guinea 655 during the 2005-2007 Period, Open J. Mar. Sci., 06(01), 11–22, doi:10.4236/ojms.2016.61002, 2016.

Kroopnick, P.: Isotopic fractionations during oxygen consumption and carbonate dissolution within the North Atlantic Deep Water, Earth Planet. Sci. Lett., 49(2), 485–498, doi:10.1016/0012-821X(80)90089-8, 1980.

Kulk, G., Platt, T., Dingle, J., Jackson, T., Jönsson, B. F., Bouman, H. A., Babin, M., Brewin, R. J. W., Doblin, M., Estrada, M., Figueiras, F. G., Furuya, K., González-Benítez, N., Gudfinnsson, H. G., Gudmundsson, K., Huang, B., Isada, T., Kovač, 660 Ž., Lutz, V. A., Marañón, E., Raman, M., Richardson, K., Rozema, P. D., van de Poll, W. H., Segura, V., Tilstone, G. H., Uitz, J., van Dongen-Vogels, V., Yoshikawa, T. and Sathyendranath, S.: Primary production, an index of climate change in the ocean: Satellite-based estimates over two decades, Remote Sens., 12(5), doi:10.3390/rs12050826, 2020.

Lamont, T., Barlow, R. G. and Kyewalyanga, M. S.: Physical drivers of phytoplankton production in the southern Benguela upwelling system, Deep. Res. Part I Oceanogr. Res. Pap., 90(1), 1–16, doi:10.1016/j.dsr.2014.03.003, 2014.

665 Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., Sasse, T. P. and Zeng, J.: A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink, Biogeosciences, 10(11), 7793–7815, doi:10.5194/bg-10-7793-2013, 2013.

Landschützer, P., Gruber, N., Bakker, D. C. E. and Schuster, U.: Recent variability of the global ocean carbon sink, Global Biogeochem. Cycles, 28(9), 927–949, doi:10.1002/2014GB004853, 2014.

670 Landschützer, P., Gruber, N. and Bakker, D. C. E.: Decadal variations and trends of the global ocean carbon sink, Global Biogeochem. Cycles, 30(10), 1396–1417, doi:10.1002/2015GB005359, 2016.

Landschützer, P., Gruber, N. and Bakker, D. C. E.: An observation-based global monthly gridded sea surface pCO_2 product from 1982 onward and its monthly climatology (NCEI Accession 0160558), NOAA Natl. Centers Environ. Information. Dataset, doi:10.7289/v5z899n6, 2017.

675 Landschützer, P., Laruelle, G. G., Roobaert, A. and Regnier, P.: A uniform pCO_2 climatology combining open and coastal oceans, Earth Syst. Sci. Data, 12(4), 2537–2553, doi:10.5194/essd-12-2537-2020, 2020.

Laruelle, G. G., Lauerwald, R., Pfeil, B. and Regnier, P.: Regionalized global budget of the CO_2 exchange at the air-water interface in continental shelf seas, Global Biogeochem. Cycles, 28(11), 1199–1214, doi:10.1002/2014GB004832, 2014.

Lavender, S. J., Pinkerton, M. H., Froidefond, J.-M., Morales, J., Aiken, J. and Moore, G. F.: SeaWiFS validation in 680 European coastal waters using optical and bio-geochemical measurements, Int. J. Remote Sens., 25(7–8), 1481–1488, doi:10.1080/01431160310001592481, 2004.

Lee, Z., Marra, J., Perry, M. J. and Kahru, M.: Estimating oceanic primary productivity from ocean color remote sensing: A strategic assessment, J. Mar. Syst., 149, 50–59, doi:10.1016/j.jmarsys.2014.11.015, 2015.

Lefèvre, N., Guillot, A., Beaumont, L. and Danguy, T.: Variability of fCO_2 in the Eastern Tropical Atlantic from a moored 685 buoy, J. Geophys. Res. Ocean., 113(1), doi:10.1029/2007JC004146, 2008.

Lefèvre, N., Veleda, D., Araujo, M. and Caniaux, G.: Variability and trends of carbon parameters at a time series in the eastern tropical Atlantic, Tellus, Ser. B Chem. Phys. Meteorol., 68(1), doi:10.3402/tellusb.v68.30305, 2016.

Lefèvre, N., Montes, M. F., Gaspar, F. L., Rocha, C., Jiang, S., De Araújo, M. C. and Severino Pino Ibánhez, J.: Net heterotrophy in the Amazon continental shelf changes rapidly to a sink of CO_2 in the outer Amazon plume, Front. Mar. Sci., 690 4(SEP), 1–16, doi:10.3389/fmars.2017.00278, 2017.

Lefèvre, N., Tyaquiçã, P., Veleda, D., Perruche, C. and van Gennip, S. J.: Amazon River propagation evidenced by a CO_2 decrease at 8°N, 38°W in September 2013, J. Mar. Syst., 211(July), 103419, doi:10.1016/j.jmarsys.2020.103419, 2020.

Lencina-Avila, J. M., Ito, R. G., Garcia, C. A. E. and Tavano, V. M.: Sea-air carbon dioxide fluxes along 35°S in the South Atlantic Ocean, Deep. Res. Part I Oceanogr. Res. Pap., 115, 175–187, doi:10.1016/j.dsr.2016.06.004, 2016.

695 Liu, W. T. and Xie, X.: Space Observation of Carbon Dioxide Partial Pressure at Ocean Surface, IEEE J. Sel. Top. Appl.

Earth Obs. Remote Sens., 10(12), 5472–5484, doi:10.1109/JSTARS.2017.2766138, 2017.

Lobanova, P., Tilstone, G. H., Bashmachnikov, I. and Brotas, V.: Accuracy assessment of primary production models with and without photoinhibition using Ocean-Colour climate change initiative data in the North East Atlantic Ocean, Remote Sens., 10(7), 1–24, doi:10.3390/rs10071116, 2018.

700 Longhurst, A., Sathyendranath, S., Platt, T. and Caverhill, C.: An estimate of global primary production in the ocean from satellite radiometer data, J. Plankton Res., 17(6), 1245–1271, doi:10.1093/plankt/17.6.1245, 1995.

Luz, B. and Barkan, E.: Assessment of Oceanic Productivity with the Triple-Isotope Composition of Dissolved Oxygen, Science (80-.)., 288(5473), 2028–2031, doi:10.1126/science.288.5473.2028, 2000.

Marañón, E., Behrenfeld, M. J., González, N., Mouriño, B. and Zubkov, M. V.: High variability of primary production in
705 oligotrophic waters of the Atlantic Ocean: Uncoupling from phytoplankton biomass and size structure, Mar. Ecol. Prog. Ser., 257, 1–11, doi:10.3354/meps257001, 2003.

Morel, A.: Light and marine photosynthesis: a spectral model with geochemical and climatological implications, Prog. Oceanogr., 26(3), 263–306, doi:10.1016/0079-6611(91)90004-6, 1991.

Moussa, H., Benallal, M. A., Goyet, C. and Lefèvre, N.: Satellite-derived CO_2 fugacity in surface seawater of the tropical
710 Atlantic Ocean using a feedforward neural network, Int. J. Remote Sens., 37(3), 580–598, doi:10.1080/01431161.2015.1131872, 2016.

Nightingale, P. D., Malin, G., Law, C. S., Watson, A. J., Liss, P. S., Liddicoat, M. I., Boutin, J. and Upstill-Goddard, R. C.: In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers, Global Biogeochem. Cycles, 14(1), 373–387, doi:10.1029/1999GB900091, 2000.

715 O'Reilly, J. E. and Werdell, P. J.: Chlorophyll algorithms for ocean color sensors - OC4, OC5 & OC6, Remote Sens. Environ., 229(May), 32–47, doi:10.1016/j.rse.2019.04.021, 2019.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M. and Mcclain, C.: Ocean color chlorophyll algorithms for SeaWiFS encompassing chlorophyll concentrations between, J. Geophys. Res., 103(C11), 24937–24953, 1998.

720 Parard, G., Lefévre, N. and Boutin, J.: Sea water fugacity of CO_2 at the PIRATA mooring at 6°S, 10°W, Tellus, Ser. B Chem. Phys. Meteorol., 62(5), 636–648, doi:10.1111/j.1600-0889.2010.00503.x, 2010.

Pfeil, B., Olsen, A., Bakker, D. C. E., Hankin, S., Koyuk, H., Kozyr, A., Malczyk, J., Manke, A., Metzl, N., Sabine, C. L., Akl, J., Alin, S. R., Bates, N., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W. J., Chavez, F. P., Chen, A., Cosca, C., Fassbender, A. J., Feely, R. A., González-Dávila, M., Goyet, C., Hales, B., Hardman-Mountford, N., Heinze, C.,
725 Hood, M., Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., Jones, S. D., Key, R. M., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lourantou, A., Merlivat, L., Midorikawa, T., Mintrop, L., Miyazaki, C., Murata, A., Nakadate, A., Nakano, Y., Nakaoka, S., Nojiri, Y., Omar, A. M., Padin, X. A., Park, G. H., Paterson, K., Perez, F. F., Pierrot, D., Poisson, A., Ríos, A. F., Santana-Casiano, J. M., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Schneider, B., Schuster, U., Sieger, R., Skjelvan, I., Steinhoff, T., Suzuki, T., Takahashi, T., Tedesco, K.,
730 Telszewski, M., Thomas, H., Tilbrook, B., Tjiputra, J., Vandemark, D., Veness, T., Wanninkhof, R., Watson, A. J., Weiss, R., Wong, C. S. and Yoshikawa-Inoue, H.: A uniform, quality controlled Surface Ocean CO_2 Atlas (SOCAT), Earth Syst. Sci. Data, 5(1), 125–143, doi:10.5194/essd-5-125-2013, 2013.

Platt, T., Caverhill, C. and Sathyendranath, S.: Basin-scale estimates of oceanic primary production by remote sensing: The North Atlantic, J. Geophys. Res. Ocean., 96(C8), 15147–15159, doi:10.1029/91JC01118, 1991.

735 Poulton, A. J., Holligan, P. M., Hickman, A., Kim, Y. N., Adey, T. R., Stinchcombe, M. C., Holeton, C., Root, S. and Woodward, E. M. S.: Phytoplankton carbon fixation, chlorophyll-biomass and diagnostic pigments in the Atlantic Ocean, Deep. Res. Part II Top. Stud. Oceanogr., 53(14–16), 1593–1610, doi:10.1016/j.dsr2.2006.05.007, 2006.

Raven, J., Caldeira, K., Elderfield, H., Hoegh-Guldberg, O., Liss, P., Riebesell, U., Shepherd, J., Turley, C. and Watson, A.: Ocean acidification due to increasing atmospheric carbon dioxide, The Royal Society, London., 2005.

740 Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C. and Wang, W.: An improved in situ and satellite SST analysis for climate, J. Clim., 15(13), 1609–1625, doi:10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2, 2002.

Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., Landschützer, P., Metzl, N., Nakaoka, S., Olsen, A., Park, G.-H., Peylin, P., Rodgers, K. B., Sasse, T. P., Schuster, U., Shutler, J. D., Valsala, V., Wanninkhof, R. and Zeng, J.: Data-based estimates of the ocean carbon sink variability – first results of the Surface Ocean pCO2 Mapping
745 intercomparison (SOCOM), Biogeosciences, 12(23), 7251–7278, doi:10.5194/bg-12-7251-2015, 2015.

Saba, V. S., Friedrichs, M. A. M., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M. J., Ciotti, A. M., Dowell, M., Hoepffner, N., Hyde, K. J. W., Ishizaka, J., Kameda, T., Marra, J., Mélin, F., Morel, A., O'reilly, J., Scardi, M., Smith, W. O., Smyth, T. J., Tang, S., Uitz, J., Waters, K. and Westberry, T. K.: An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe, Biogeosciences, 8, 489–503, doi:10.5194/bg-8-489-2011, 2011.

750

Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C. E., 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. J., Chavez, F. P., Chen, A., Cosca, C., Feely, R. A., González-Dávila, M., Goyet, C., Hardman-Mountford, N., Heinze, C., Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., Key, R. M., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lourantou, A., Merlivat, L., Midorikawa, T., Mintrop, L., Miyazaki, C., Murata, A., Nakadate, A., Nakano, Y., Nakaoka, S., Nojiri, Y., Omar, A. M., Padin, X. A., Park, G. H., Paterson, K., Perez, F. F., Pierrot, D., Poisson, A., Ríos, A. F., Salisbury, J., Santana-Casiano, J. M., S. Sarma, V. V. S., Schlitzer, R., Schneider, B., Schuster, U., Sieger, R., Skjelvan, I., Steinhoff, T., Suzuki, T., Takahashi, T., Tedesco, K., Telszewski, M., Thomas, H., Tilbrook, B., Vandemark, D., Veness, T., Watson, A. J., Weiss, R., Wong, C. S. and Yoshikawa-Inoue, H.: Surface Ocean CO_2 Atlas (SOCAT) gridded data products,

760

Earth Syst. Sci. Data, 5(1), 145–153, doi:10.5194/essd-5-145-2013, 2013.

Santana-Casiano, J. M., González-Dávila, M. and Ucha, I. R.: Carbon dioxide fluxes in the Benguela upwelling system during winter and spring: A comparison between 2005 and 2006, Deep Sea Res. Part II Top. Stud. Oceanogr., 56(8–10), 533–541, doi:10.1016/j.dsr2.2008.12.010, 2009.

Sathyendranath, S., Platt, T., Kovač, Ž., Dingle, J., Jackson, T., Brewin, R. J. W., Franks, P., Marañón, E., Kulk, G. and

765

Bouman, H. A.: Reconciling models of primary production and photoacclimation [Invited], Appl. Opt., 59(10), C100, doi:10.1364/AO.386252, 2020.

Schloss, I. R., Ferreyra, G. A., Ferrario, M. E., Almandoz, G. O., Codina, R., Bianchi, A. A., Balestrini, C. F., Ochoa, H. A., Pino, D. R. and Poisson, A.: Role of plankton communities in sea-air variations in pCO_2 in the SW Atlantic Ocean, Mar. Ecol. Prog. Ser., 332, 93–106, doi:10.3354/meps332093, 2007.

770

Serret, P., Robinson, C., Aranguren-Gassis, M., García-Martín, E. E., Gist, N., Kitidis, V., Lozano, J., Stephens, J., Harris, C. and Thomas, R.: Both respiration and photosynthesis determine the scaling of plankton metabolism in the oligotrophic ocean, Nat. Commun., 6, 1–10, doi:10.1038/ncomms7961, 2015.

Shutler, J. D., Land, P. E., Piolle, J. F., Woolf, D. K., Goddijn-Murphy, L., Paul, F., Girard-Ardhuin, F., Chapron, B. and Donlon, C. J.: FluxEngine: A flexible processing system for calculating atmosphere-ocean carbon dioxide gas fluxes and

775

climatologies, J. Atmos. Ocean. Technol., 33(4), 741–756, doi:10.1175/JTECH-D-14-00204.1, 2016.

Shutler, J. D., Wanninkhof, R., Nightingale, P. D., Woolf, D. K., Bakker, D. C., Watson, A., Ashton, I., Holding, T., Chapron, B., Quilfen, Y., Fairall, C., Schuster, U., Nakajima, M. and Donlon, C. J.: Satellites will address critical science priorities for quantifying ocean carbon, Front. Ecol. Environ., 18(1), 27–35, doi:10.1002/fee.2129, 2020.

Smith, W. O. and Demaster, D. J.: Phytoplankton biomass and productivity in the Amazon River plume: Correlation with

780

seasonal river discharge, Cont. Shelf Res., 16(3), 291–319, doi:10.1016/0278-4343(95)00007-N, 1996.

Smyth, T. J., Tilstone, G. H. and Groom, S. B.: Integration of radiative transfer into satellite models of ocean primary production, J. Geophys. Res. C Ocean., 110(10), 1–11, doi:10.1029/2004JC002784, 2005.

Stephens, M. P., Samuels, G., Olson, D. B., Fine, R. A. and Takahashi, T.: Sea-air flux of CO_2 in the North Pacific using shipboard and satellite data, J. Geophys. Res., 100(C7), 13571, doi:10.1029/95JC00901, 1995.

785

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C. S., Delille, B., Bates, N. R. and de Baar, H. J. W.: Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, Deep. Res. Part II Top. Stud. Oceanogr.,

790

56(8–10), 554–577, doi:10.1016/j.dsr2.2008.12.009, 2009.

Taylor, J. R.: An introduction to error analysis, University Science Books, Sausalito, Calif., 1997.

Tilstone, G. H., Smyth, T. J., Gowen, R. J., Martinez-Vicente, V. and Groom, S. B.: Inherent optical properties of the Irish Sea and their effect on satellite primary production algorithms, J. Plankton Res., 27(11), 1127–1148, doi:10.1093/plankt/fbi075, 2005.

795

Tilstone, G. H., Smyth, T., Poulton, A. and Hutson, R.: Measured and remotely sensed estimates of primary production in

the Atlantic Ocean from 1998 to 2005, Deep. Res. Part II Top. Stud. Oceanogr., 56(15), 918–930, doi:10.1016/j.dsr2.2008.10.034, 2009.

Tilstone, G. H., Taylor, B. H., Blondeau-Patissier, D., Powell, T., Groom, S. B., Rees, A. P. and Lucas, M. I.: Comparison of new and primary production models using SeaWiFS data in contrasting hydrographic zones of the northern North Atlantic, Remote Sens. Environ., 156, 473–489, doi:10.1016/j.rse.2014.10.013, 2015a.

Tilstone, G. H., Xie, Y. yuan, Robinson, C., Serret, P., Raitsos, D. E., Powell, T., Aranguren-Gassis, M., Garcia-Martin, E. E. and Kitidis, V.: Satellite estimates of net community production indicate predominance of net autotrophy in the Atlantic Ocean, Remote Sens. Environ., 164, 254–269, doi:10.1016/j.rse.2015.03.017, 2015b.

Tilstone, G. H., Lange, P. K., Misra, A., Brewin, R. J. W. and Cain, T.: Micro-phytoplankton photosynthesis, primary production and potential export production in the Atlantic Ocean, Prog. Oceanogr., 158, 109–129, doi:10.1016/j.pocean.2017.01.006, 2017.

Valerio, A. M., Kampel, M., Ward, N. D., Sawakuchi, H. O., Cunha, A. C. and Richey, J. E.: CO_2 partial pressure and fluxes in the Amazon River plume using in situ and remote sensing data, Cont. Shelf Res., 215(May 2020), 104348, doi:10.1016/j.csr.2021.104348, 2021.

Varona, H. L., Veleda, D., Silva, M., Cintra, M. and Araujo, M.: Amazon River plume influence on Western Tropical Atlantic dynamic variability, Dyn. Atmos. Ocean., 85(September 2018), 1–15, doi:10.1016/j.dynatmoce.2018.10.002, 2019.

Watson, A. J., Schuster, U., Shutler, J. D., Holding, T., Ashton, I. G. C., Landschützer, P., Woolf, D. K. and Goddijn-Murphy, L.: Interpolated Global surface ocean carbon dioxide partial pressure and ocean-atmosphere fluxes 1992-2018, corrected for surface temperature deviations, PANGAEA, doi:10.1594/PANGAEA.922985, 2020a.

Watson, A. J., Schuster, U., Shutler, J. D., Holding, T., Ashton, I. G. C., Landschützer, P., Woolf, D. K. and Goddijn-Murphy, L.: Revised estimates of ocean-atmosphere CO_2 flux are consistent with ocean carbon inventory, Nat. Commun., 11(1), 1–6, doi:10.1038/s41467-020-18203-3, 2020b.

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

Williams, P. J. le B., Quay, P. D., Westberry, T. K. and Behrenfeld, M. J.: The Oligotrophic Ocean Is Autotrophic, Ann. Rev. Mar. Sci., 5(1), 535–549, doi:10.1146/annurev-marine-121211-172335, 2013.

Woolf, D. K., Land, P. E., Shutler, J. D., Goddijn-Murphy, L. M. and Donlon, C. J.: On the calculation of air-sea fluxes of CO_2 in the presence of temperature and salinity gradients, J. Geophys. Res. Ocean., 121(2), 1229–1248, doi:10.1002/2015JC011427, 2016.

Woolf, D. K., Shutler, J. D., Goddijn-Murphy, L., Watson, A. J., Chapron, B., Nightingale, P. D., Donlon, C. J., Piskozub, J., Yelland, M. J., Ashton, I., Holding, T., Schuster, U., Girard-Ardhuin, F., Grouazel, A., Piolle, J. F., Warren, M., Wrobel-Niedzwiecka, I., Land, P. E., Torres, R., Prytherch, J., Moat, B., Hanafin, J., Ardhuin, F. and Paul, F.: Key Uncertainties in the Recent Air-Sea Flux of CO_2, Global Biogeochem. Cycles, 33(12), 1548–1563, doi:10.1029/2018GB006041, 2019.

Zeng, J., Nojiri, Y., Landschützer, P., Telszewski, M. and Nakaoka, S.: A global surface ocean fCO_2 climatology based on a feed-forward neural network, J. Atmos. Ocean. Technol., 31(8), 1838–1849, doi:10.1175/JTECH-D-13-00137.1, 2014.