Biogeosciences, 10, 7793–7815, 2013
www.biogeosciences.net/10/7793/2013/
doi:10.5194/bg-10-7793-2013

Biogeosciences Open Access

A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink

P. Landschützer[1], N. Gruber[2], D. C. E. Bakker[1], U. Schuster[1,*], S. Nakaoka[3], M. R. Payne[2,4], T. P. Sasse[5], and J. Zeng[3]

[1]School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, UK
[2]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zürich, Switzerland
[3]Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan
[4]Section for Oceanography and Climate, National Institute for Aquatic Resources, Technical University of Denmark, Charlottenlund, Denmark
[5]Climate Change Research Centre, University of New South Wales, Sydney, Australia
[*]now at: College of Life and Environmental Sciences, Hatherly Laboratories, University of Exeter, Exeter, UK

Correspondence to: P. Landschützer (p.landschutzer@uea.ac.uk)

Received: 29 April 2013 – Published in Biogeosciences Discuss.: 29 May 2013
Revised: 18 September 2013 – Accepted: 30 October 2013 – Published: 29 November 2013

Abstract. The Atlantic Ocean is one of the most important sinks for atmospheric carbon dioxide (CO_2), but this sink has been shown to vary substantially in time. Here we use surface ocean CO_2 observations to estimate this sink and the temporal variability from 1998 through 2007 in the Atlantic Ocean. We benefit from (i) a continuous improvement of the observations, i.e. the Surface Ocean CO_2 Atlas (SOCAT) v1.5 database and (ii) a newly developed technique to interpolate the observations in space and time. In particular, we use a two-step neural network approach to reconstruct basin-wide monthly maps of the sea surface partial pressure of CO_2 (pCO_2) at a resolution of $1° \times 1°$. From those, we compute the air–sea CO_2 flux maps using a standard gas exchange parameterization and high-resolution wind speeds. The neural networks fit the observed pCO_2 data with a root mean square error (RMSE) of about 10 µatm and with almost no bias. A check against independent time-series data and new data from SOCAT v2 reveals a larger RMSE of 22.8 µatm for the entire Atlantic Ocean, which decreases to 16.3 µatm for data south of 40° N. We estimate a decadal mean uptake flux of -0.45 ± 0.15 Pg C yr^{-1} for the Atlantic between 44° S and 79° N, representing the sum of a strong uptake north of 18° N (-0.39 ± 0.10 Pg C yr^{-1}), outgassing in the tropics (18° S–18° N, 0.11 ± 0.07 Pg C yr^{-1}), and uptake in the subtropical/temperate South Atlantic south of 18° S (-0.16 ± 0.06 Pg C yr^{-1}), consistent with recent studies. The strongest seasonal variability of the CO_2 flux occurs in the temperature-driven subtropical North Atlantic, with uptake in winter and outgassing in summer. The seasonal cycle is antiphased in the subpolar latitudes relative to the subtropics largely as a result of the biologically driven winter-to-summer drawdown of CO_2. Over the 10 yr analysis period (1998 through 2007), sea surface pCO_2 increased faster than that of the atmosphere in large areas poleward of 40° N, while in other regions of the North Atlantic the sea surface pCO_2 increased at a slower rate, resulting in a barely changing Atlantic carbon sink north of the Equator (-0.01 ± 0.02 Pg C yr^{-1} decade^{-1}). Surface ocean pCO_2 increased at a slower rate relative to atmospheric CO_2 over most of the Atlantic south of the Equator, leading to a substantial trend toward a stronger CO_2 sink for the entire South Atlantic (-0.14 ± 0.02 Pg C yr^{-1} decade^{-1}). In contrast to the 10 yr trends, the Atlantic Ocean carbon sink varies relatively little on inter-annual timescales (± 0.04 Pg C yr^{-1}; 1σ).

1 Introduction

Over the last two decades, the Atlantic Ocean (44° S to 79° N and west of 19° E) has taken up about 0.49 ± 0.11 Pg C yr^{-1} on average (Schuster et al., 2013). This makes the Atlantic Ocean one of the most important sinks for atmospheric CO_2, and especially for anthropogenic CO_2 (Sabine et al., 2004; Mikaloff Fletcher et al., 2006, 2007; Gruber et al., 2009).

The Atlantic sink estimate was taken from a recent Regional Carbon Cycle Assessment and Processes (RECCAP) synthesis by Schuster et al. (2013), where the authors reviewed different methodologies to estimate the air–sea CO_2 fluxes and provided a "best" estimate. The methods included estimates derived (i) using ocean surface partial pressure of CO_2 (pCO_2) measurements (Takahashi et al., 2009), (ii) from ocean general circulation models that include a full representation of the oceanic carbon cycle (e.g. Le Quéré et al., 2007; Graven et al., 2012; Doney et al., 2009), (iii) inversions of ocean interior carbon measurements (e.g. Gruber et al., 2009) and (iv) from inversions of atmospheric CO_2 (e.g. Gurney et al., 2008). For the "best" estimate, only the observational-based estimates were used, however, i.e. those relying on pCO_2 observations and those derived from the ocean interior carbon observations through an inversion approach. These methods gave rather similar long-term mean values for the whole basin, although with substantial sub-basin differences.

The Atlantic Ocean sink varies substantially by season, which is in part driven by seasonal variations in surface ocean pCO_2 (Takahashi et al., 1993, 2002; Schuster et al., 2013), but is also affected by seasonal variations in surface ocean winds and atmospheric CO_2. Surface ocean pCO_2 varies over a wide range above and below the atmospheric pCO_2, with much of the seasonal amplitude dominated by temperature in the subtropical regions in both hemispheres (Gruber et al., 2002; Takahashi et al., 2002; Sarmiento and Gruber, 2006), explaining the summer maximum in surface ocean pCO_2. In contrast, biological processes acting in synergy with ocean mixing and circulation dominates the seasonal pCO_2 cycle in equatorial and high latitude regions (poleward of 40°) (Takahashi et al., 1993; Bennington et al., 2009), explaining the summer minimum in oceanic pCO_2. Due to their opposite phasing, these two drivers cancel each other out along the regime boundaries at around 40° (Takahashi et al., 2002), leading to a minimum in the seasonal amplitude there. Schuster et al. (2013) identified a broad agreement among independent seasonal flux estimates in the temperature driven subtropics, but not elsewhere.

Long-term trends and inter-annual variability of the Atlantic carbon sink represent a source of substantial disagreement between the different methodologies and studies. Using surface ocean pCO_2 observations, Schuster and Watson (2007) argued for a decrease in the North Atlantic carbon sink and a reduction in the seasonal amplitude in both the subtropical and temperate North Atlantic from the mid-1990s to the mid-2000s. They linked this reduction to the large changes that occurred in the climate mode of the North Atlantic over this period, i.e. a shift of the North Atlantic Oscillation (NAO) from very positive phases in the early 1990s to negative and near-zero phases in the mid-2000s. Corbière et al. (2007) supported this conclusion on the basis of their observations from the subpolar gyre over the 1993 to 2003 period, pointing out that the larger than expected increase in the observed pCO_2 is mainly a result of rapid warming. Further support for a decreasing North Atlantic sink comes from Lefèvre et al. (2004), Lüger et al. (2006), Olsen et al. (2006), and Schuster et al. (2009), although each study analysed different regions and periods and also used different methods to determine trends. Schuster et al. (2009), for example, analysed data from 1990 to 2006 in the eastern subpolar gyre and throughout most of the central North Atlantic, while Olsen et al. (2006) only focused on the Nordic Seas, but looked at a more extended period, i.e. from 1981 to 2002/2003.

Based on the results of a global ocean biogeochemistry model, Thomas et al. (2008) argued that this trend toward a smaller North Atlantic sink is transitory and is expected to rebound in the near-term future, i.e. that the decreasing sink strength is the result of "natural" variability, and should not be interpreted as a signal of anthropogenic climate change. They interpreted the decline in sink strength to be the result of a NAO-driven reduction in the transport of water by the North Atlantic Current into the eastern subpolar gyre. In a contrasting modelling study, Ullman et al. (2009) argued that the North Atlantic carbon sink actually increased from the mid-1990s to the mid-2000s. They proposed that the declining trend in the NAO from the early- to mid-1990s until the mid-2000s led to reduced convective mixing in the subpolar gyre, counteracting the impact of warming.

The initial year and period of data analysed for trends are crucial aspects to consider when resolving different perspectives (Gruber, 2009). McKinley et al. (2011) pointed out that when the surface ocean trends in pCO_2 are analysed over more than 25 yr, all regions in the North Atlantic exhibit trends that are not statistically different from the trend in atmospheric CO_2, implying no change in the sink/source strength. However, when the periods of analyses were shortened to 10 yr and the beginning and ends shifted, substantial trends emerged, largely reflecting inter-annual to decadal timescale variability.

A major challenge in detecting trends in the Atlantic Ocean carbon sink is due to the highly heterogeneous distribution of the surface ocean pCO_2 observations in time and space. Different approaches have been employed to overcome this limitation and to create basin or sub-basin wide estimates. These include the binning of data to $4° \times 5°$ bins in latitude and longitude followed by an advection-based interpolation method (Takahashi et al., 1999, 2003, 2009), binning of data to large-scale biogeochemical provinces (McKinley et al., 2011), multi-linear regression models (e.g. Chierici et al., 2009; Peng and Wanninkhof, 2010), and neural network-based methods (e.g. Lefèvre et al., 2005; Friedrich and Oschlies, 2009; Telszewski et al., 2009; Sasse et al., 2013). Each of these approaches has its strengths and weaknesses. For example, the binning and interpolation scheme employed by Takahashi et al. (1999) is well suited for constraining monthly climatologies, however, its coarse resolution tends to smooth out small-scale features. However, the method is very robust and it is not sensitive to outliers. The

binning to large-scale biogeochemical provinces works well to determine long-term trends (McKinley et al., 2011), but its resolution is even more coarse. The multi-linear regression models allow very finely resolved estimates, however, the explained variance in these statistical models is often relatively low, causing substantial uncertainties in the estimated fields.

In order to investigate the variability of the sea surface pCO_2 and the resulting carbon flux, we develop a novel neural network based approach that overcomes most of the aforementioned issues which have limited previous studies. Our method is capable of capturing a large amount of variability due to the non-linear predictor-observation relationship on a fine $1^\circ \times 1^\circ$ spatial grid. The method determines the non-linear relationships between the surface ocean pCO_2 observations and a set of independent observations to produce basin-wide sea surface maps of pCO_2 on a monthly basis. Our network gathers information from similar ocean biogeochemical provinces and provides us with regional pCO_2 estimates, which we then use to investigate the changing distribution of the sea surface pCO_2 in the Atlantic Ocean. Our method relies on the assumption that the Atlantic Ocean carbon sink and its variability can be estimated as a function of proxy variables, which are subjectively chosen. We further rely on ocean carbon measurements in order to establish a correct relationship. We therefore benefit from the recent publication of the Surface Ocean CO_2 Atlas (SOCAT), which provides the to-date largest global data set of surface ocean fugacity of CO_2 observations (Pfeil et al., 2013; Sabine et al., 2013), which we converted to pCO_2. To evaluate our estimates, we compare our results to independent data from time-series stations and additional independent pCO_2 measurements. Our basin-wide pCO_2 maps provide a basis to calculate the air–sea fluxes for the entire Atlantic and to quantify the seasonal to inter-annual variability in its sink strength.

2 Data and methods

We combine two neural network methods to reconstruct the sea surface partial pressure of CO_2 (pCO_2) for the period January 1998 to December 2007 on a monthly $1^\circ \times 1^\circ$ resolution. In particular, we use a two-stage approach to establish numerical relationships between pCO_2 and a suite of independent predictors that are known to drive its variability. In the first stage, we use a neural network clustering algorithm to define a discrete set of biogeochemical provinces that share a common relationship between the independent variables and climatological pCO_2. In the second stage, we derive for each biogeochemical province a non-linear and continuous relationship between pCO_2 and the input parameters on the basis of a feed-forward network (FFN) method. This input–output relationship is then used to estimate surface ocean pCO_2 for each month and each pixel. This two-stage approach has the advantage over previously employed one-stage approaches (e.g. Friedrich and Oschlies, 2009; Telszewski et al., 2009) that the pre-clustering of the data substantially reduces the variance of the pCO_2 that needs to be explained by neural network methods, resulting in substantially better fits.

For both stages we use global data, as this permits us to take advantage of the fact that biogeochemical provinces with limited coverage in a particular ocean basin can learn from observations in the same biogeochemical province in another ocean basin. Here, we evaluate and discuss only the results for the Atlantic Ocean. The resulting surface ocean pCO_2 distribution is then combined with corresponding atmospheric pCO_2 data and wind-speed based estimates of the gas transfer velocity to construct the mean and variability of the Atlantic Ocean carbon sink from 1998 through 2007.

2.1 Data

The gridded observations of the surface ocean fugacity of CO_2 (fCO_2) from SOCAT v1.5 (Sabine et al., 2013) form the basis for our computations. This data set includes global observations over the period 1970 to 2007 and was homogenized by an extensive series of automatic and manual secondary quality controls (Pfeil et al., 2013). The fCO_2 data distribution in the Atlantic Ocean is highly skewed in time and space. The number of $1^\circ \times 1^\circ$ pixels with fCO_2 measurements per year varies from as low as ~ 180 per year in 1999 and 2000, to over 4000 in 2006 and 2007, with the last two years accounting for 40 % of all measurements. The global data in the other ocean basins are more homogeneously distributed in time. In contrast, the Atlantic has a good spatial coverage, while this is not the case for many of the other ocean basins.

The reported fCO_2 estimates were converted to pCO_2 using the formulation (see e.g. Körtzinger, 1999)

$$pCO_2 = fCO_2 \cdot \exp\left(P_{\mathrm{atm}}^{\mathrm{surf}} \frac{B + 2 \cdot \delta}{R \cdot T}\right)^{-1}, \quad (1)$$

where $P_{\mathrm{atm}}^{\mathrm{surf}}$ is the total atmospheric surface pressure, B and δ are viral coefficients (Weiss, 1974), R is the gas constant and T is the absolute temperature. National Centers for Environmental Prediction (NCEP) monthly mean sea level pressure was used for $P_{\mathrm{atm}}^{\mathrm{surf}}$ (Kalnay et al., 1996).

A crucial choice concerns the selection of the independent input variables used for the training of the networks. We chose sea surface temperature (SST), chlorophyll a concentration (CHL), ocean mixed layer depth (MLD), sea surface salinity (SSS) and the atmospheric CO_2 mole fraction ($xCO_{2,\mathrm{atm}}$). These parameters represent physical, chemical and biological proxies determining the distribution of pCO_2 in the sea surface layer. For SST, we use the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation (OI) sea surface temperature v.2 (Reynolds et al., 2002), for CHL

the SeaWiFS mapped chlorophyll (SeaWiFSProject, http://oceancolor.gsfc.nasa.gov/cgi/l3), for MLD the mixed layer depth data from the Estimating the Circulation and Climate of the Ocean, Phase II (ECCO2) project (Menemenlis et al., 2008), for SSS the Simple Ocean Data Assimilation (SODA) sea surface salinity data (Carton and Giese, 2008) and for $xCO_{2,atm}$ the monthly atmospheric CO_2 from GLOBALVIEW-CO2 (2011). Furthermore, the monthly pCO_2 climatology of Takahashi et al. (2009) is used as an additional input parameter for defining the biogeochemical provinces. Due to their strongly skewed distribution, MLD and CHL were log-transformed before use as predictor values.

We restricted our analyses to the time period from January 1998 to December 2007 due to the temporal limitations of the data we chose for our study. No satellite chlorophyll data are available before the 1997 launch of the SeaWiFS mission, and the CO_2 observations in SOCAT v1.5 extend to the year 2007.

Data with an original resolution finer than the required $1^\circ \times 1^\circ$ were averaged onto the desired grid, whereas input data with a coarser resolution were interpolated using a bilinear interpolation. We further took monthly averages of all inputs with a finer temporal resolution.

To highlight anomalies and year-to-year trends within our data sets, we further produced deseasonalized sets of our input variables by removing their long-term mean seasonal cycle from 1990 to 2010 (1998 to 2010 in the case of chlorophyll *a* and 1992 to 2010 in case of MLD and SSS) from the original data set.

In the next step the monthly $1^\circ \times 1^\circ$ input data are rearranged into 3 major data sets. Each of these data sets consists of input vectors ($\boldsymbol{p}_n$) where the input data are organized as row vector elements, for example SST, log(MLD), SSS, and $pCO_{2,Takahashi}$ for the self-organizing map input (SINP) data set, sampled at the same space-time point (Table 1). Two of these sets, SINP and the feed-forward network input 2 set (FINP2) are global sets and do not have a corresponding target data set (Table 1). Input vectors with empty vector elements, e.g. where no salinity data are available, were removed from these data sets. The third major set, the feed-forward network input set (FINP), consists only of input vectors where corresponding SOCAT v1.5 observations, or targets (t), are available, i.e. they are subsampled in time at the locations where observations are available. In order to train the feed-forward network, two sub sets of the FINP set are created, namely the actual training (FITR) set and a validation (FIVAL) set (Table 1, Appendix A2).

Where no chlorophyll *a* satellite data are available, due e.g. to cloud cover or lack of sufficient light at high latitudes in winter time, we estimate the sea surface pCO_2 with the remaining input parameters. This applies to about 22 % of all pixels and mainly concerns the high latitude oceans in winter. Since chlorophyll concentrations tend to be low and unvarying during these months, we expect that this choice has a relatively small influence on our results.

We use data from two sources to evaluate and validate our results. First, we extracted time-series data from the combined record from BATS (Bermuda Atlantic Time Series Station) and Hydrostation "S" (Bates, 2007; Gruber et al., 2002) and the European Station for Time Series in the Ocean (ESTOC) (e.g. González-Dávila et al., 2007). Second, we use 3065 additional data points within our study region and period from the updated SOCAT v2 database (Bakker et al., 2013), which were not included in version v1.5 and therefore constitute independent data.

To evaluate the sensitivity of the results with regard to the chosen data product, we further performed 4 sensitivity runs, namely (i) SR1 (sensitivity run 1) where we replaced the SODA sea surface salinity with the World Ocean Atlas 2009 (Antonov et al., 2010) sea surface salinity climatology, (ii) SR2, where we replaced the ECCO2 MLD product with the de Boyer Montegut (de Boyer Montegut et al., 2004) MLD climatology, (iii) SR3, where we used the SODA (Carton and Giese, 2008) sea surface temperature and (iv) SR4, where we excluded chlorophyll *a* as an input parameter.

2.2 Methods

We use a self-organizing map (SOM) method (Kohonen, 1987, 2001) to partition the global ocean into 16 biogeochemical provinces. Such a biogeochemical province is characterized by each of its location having a similar relationship among all considered input variables. Based on trial and error, we finally chose the climatological pCO_2 from Takahashi et al. (2009), SST, log(MLD) and SSS as our input variables for the SOM (SINP data set). We did not include chlorophyll, i.e. log(CHL), due to its many missing values.

A SOM is a neural network based cluster algorithm that can detect regularities within the provided input data and then learns to group them together. Similar input data, arranged as input vectors, are identified via their Euclidean distance towards the nodes (or neurons) of the network. The choice of 16 provinces represents a subjectively determined optimum between too many provinces with too little data but a high degree of correlation between the provinces, and too few provinces with a lot of data, but too high variance in the data. Details on the SOM method can be found in Appendix A1.

The provinces change in shape from one month to the next and further change slightly between years, i.e. they are not static, unlike conventional provinces or biomes. As we did not provide any additional time or space information to the SOM, these temporal variations emerge solely from the temporal variability of the input data. Despite their strong seasonal dynamics in space (Fig. 1a) and time (Fig. 1b), the estimated biogeochemical provinces exhibit a coherent large-scale behaviour, reflecting the well known oceanic structures such as the gyres, the equatorial regions, and the high-latitude North Atlantic. Figure 1a shows the mode of the

Table 1. Input and target vector elements for each subset used within our method. The subscript ds describes deseasonalized data, which are computed by subtracting the long-term mean seasonal cycle from the original data set as explained in the text. Additionally, log(CHL) was excluded from sets FINP, FITR, FIVAL and FINP2 to estimate pCO_2 when no satellite chlorophyll a was available due to cloud cover.

Set name	Elements n of the jth input vectors $\boldsymbol{p}_n^j$	Targets (t^j)
SINP	SST, log(MLD), SSS, $pCO_{2,\text{Takahashi}}$	–
FINP, FITR, FIVAL	SST, log(CHL), log(MLD), SSS, $xCO_{2,\text{atm}}$, SST_{ds}, CHL_{ds}, MLD_{ds}, SSS_{ds}, $xCO_{2,\text{atm,ds}}$	$pCO_{2,\text{SOCAT}}$
FINP2	SST, log(CHL), log(MLD), SSS, $xCO_{2,\text{atm}}$, SST_{ds}, CHL_{ds}, MLD_{ds}, SSS_{ds}, $xCO_{2,\text{atm,ds}}$	–

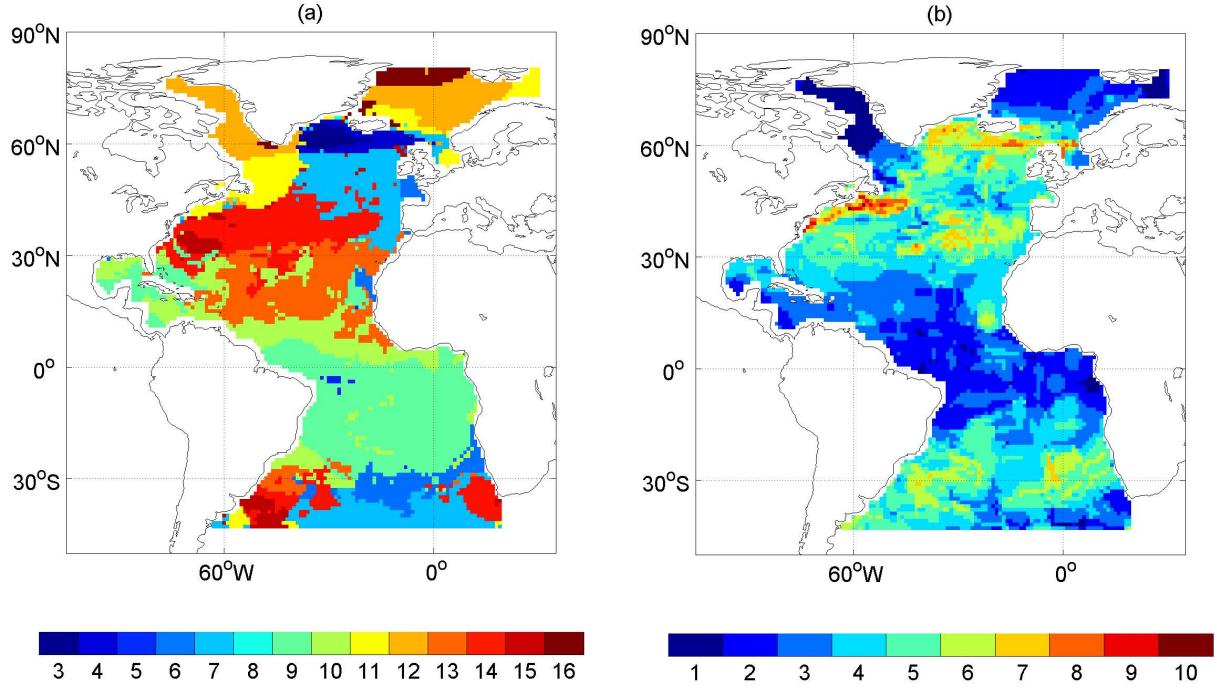

Fig. 1. Map of the biogeochemical provinces in the Atlantic Ocean identified by the self-organizing map (SOM) method: **(a)** province number of the mode (i.e. most frequent occurrence). Provinces 1 and 2 do not occur in the Atlantic Ocean. **(b)** The number of different provinces every pixel belongs to for each month from January 1998 to December 2007.

provinces, i.e. the province each pixel mainly belongs to from 1998 to 2007 and Fig. 1b shows the number of shifting provinces per pixel. These provinces vary in time and space mainly in accordance with the variability of the climatological pCO_2. In the tropics, and the high latitude North Atlantic, the climatological pCO_2 vary little seasonally and therefore the provinces remain fairly steady, with only minimal province shifts. In contrast, the gyre regions of both hemispheres exhibit much larger seasonal variability, hence pixels there undergo many more province changes. We find the largest shifts along the Gulf Stream, where certain regions change their province association up to 10 times.

As a second step we use a feed-forward network (FFN) method to reconstruct the non-linear relationship between our input variables and the target, i.e. pCO_2, separately for each of the 16 biogeochemical provinces. The FFN method is a type of backpropagation network method that is capable of approximating any function with a finite number of discontinuities (Demuth et al., 2008). Similar to multi-linear regressions, a feed-forward network adjusts coefficients to establish a relationship between inputs and targets. The adjustment of the coefficients follows an iterative process. The first iteration includes an initial guess, where the coefficients are randomly initialized, the estimates are computed and compared to the target observations. From there on the network goes backwards (hence the name backpropagation) and automatically re-adjusts the coefficients with the aim to reduce the mean squared error between estimates and targets. For each iteration, only a random subset of the data is used to train the network, while the remaining data are used for validation. The updating process of the coefficients is repeated until the network estimates derived from the validation set no longer improve significantly relative to the targets. The established relationship is further used to predict the pCO_2 for each point in time and space where no observations are

available. This process is explained in more detail in Appendix A2.

The feed-forward network was trained with the FINP data set that included all input variables including their deseasonalized representation (see Table 1). To this end the data set was split into the 16 ocean provinces (FINP^k, FINP2^k, with $k = 1, \ldots, 16$) and each of them was processed separately. Due to the temporal and spatial variability of the provinces and the heterogeneous spatiotemporal distribution of the pCO_2 data, large differences exist in the number of observations within the different provinces. However, our neural network fit does not show degeneration as a function of the data density, as shown in Sect. 3.1 for the temporal distribution, and the spatial heterogeneity of the data does not lead to any major hidden bias. Details on the settings used for the FFN can be found in Appendix A2.

The air–sea flux was then calculated from the estimated pCO_2 for each grid cell and month, using the quadratic wind speed dependence of the gas transfer velocity of Wanninkhof (1992) with the gas transfer coefficient from Sweeney et al. (2007) and monthly mean wind speeds from the cross-calibrated multi-platform (CCMP) product by Atlas et al. (2011). More details are provided in Appendix A3.

3 Results

3.1 Residuals and validation

The combined SOM-FFN method obtains very good fits with an overall mean r^2 between the fitted pCO_2 and the gridded Atlantic Ocean SOCAT v1.5 data of 0.87 and a RMSE of about 10 µatm (Table 2). The overall bias is small (-0.10 µatm). These results apply also to each year individually, indicating that the temporally inhomogeneous data distribution does not have a measurable effect on the estimates for each year.

The residuals are not entirely randomly distributed in space. As shown in Fig. 2a, the temporal mean residuals in each pixel show generally low values in the open ocean region, but tend to increase toward the fronts. The strongest model-observation discrepancies occur in the equatorial Atlantic, along the Gulf Stream and North Atlantic Current as well as in the Norwegian, Greenland and the North seas, i.e. mostly in regions with relatively strong horizontal gradients in surface ocean pCO_2.

The standard deviation of the residuals (Fig. 2b) shows that the strongest temporal errors occur again in the high latitudes of the North Atlantic, in particular the Norwegian and North seas, as well as along the North American coastline and in the eastern South Atlantic between 0 and 30° S. This indicates that the model input parameters are not able to predict all the temporal variability occurring in these regions with known biogeochemical complexity.

Table 2. Statistical measures of the comparison of the neural network-based estimates of pCO_2 with the SOCAT v1.5 gridded observations in the Atlantic Ocean from 44° S to 79° N and west of 30° E

Period	r^2	RMSE [µatm]	bias [µatm]	# data
1998–2007	0.87	9.89	−0.10	20 003
1998	0.93	7.15	−0.18	583
1999	0.89	9.35	1.62	186
2000	0.78	11.14	−0.19	178
2001	0.83	11.48	−0.87	566
2002	0.87	8.98	0.22	1941
2003	0.87	7.47	−0.11	1963
2004	0.87	8.13	−0.14	2729
2005	0.88	9.52	−0.43	3575
2006	0.85	11.51	−0.10	4280
2007	0.87	10.89	0.12	4002

To test the impact of the inhomogeneous distribution of the neural network input data and pCO_2 observations, we show the residuals calculated as the difference between the neural network pCO_2 estimates and the gridded SOCAT v1.5 pCO_2 observations (Fig. 3). The residuals are independent of the magnitude of the estimated pCO_2, and also do not show any dependence on the magnitude of the independent variables (Fig. 3). Each bin median of the residuals is close to zero, with the strongest spread occurring in the low pCO_2 bins around 275 µatm which coincides with low SST at around 5 °C and high log(CHL) concentrations at around -0.25 to 1.25 mg m^{-3}. Figure 3 further shows that large residuals, most of which stem from regions characterized by strong horizontal pCO_2 gradients, are independent of the data density.

In conclusion, the residuals indicate that the combined SOM-FFN method fulfils most tests for a good fit and does not contain any major hidden biases. In particular, there is no indication of a substantial degeneration of the fits as a function of data density, neither in time nor in space. Regions with high spatial or temporal variability are the least well fitted, while the fits for most of the less variable open ocean are very good.

3.2 Validation with independent observations

A second and more robust test of the model's ability to predict basin wide pCO_2 fields was conducted by comparing output values to independent data. To this end, we compare the network-based pCO_2 estimates with observations from two time-series stations in the subtropical North Atlantic, i.e. the combined record from BATS and Hydrostation "S" (Bates, 2007; Gruber et al., 2002), which is located in the northwestern Sargasso Sea near Bermuda (32°10′ N, 64°30′ W), and ESTOC (e.g. González-Dávila et al., 2007) located in the eastern subtropical gyre near the Canary Islands (29.04° N, 15.5° W). These stations provide

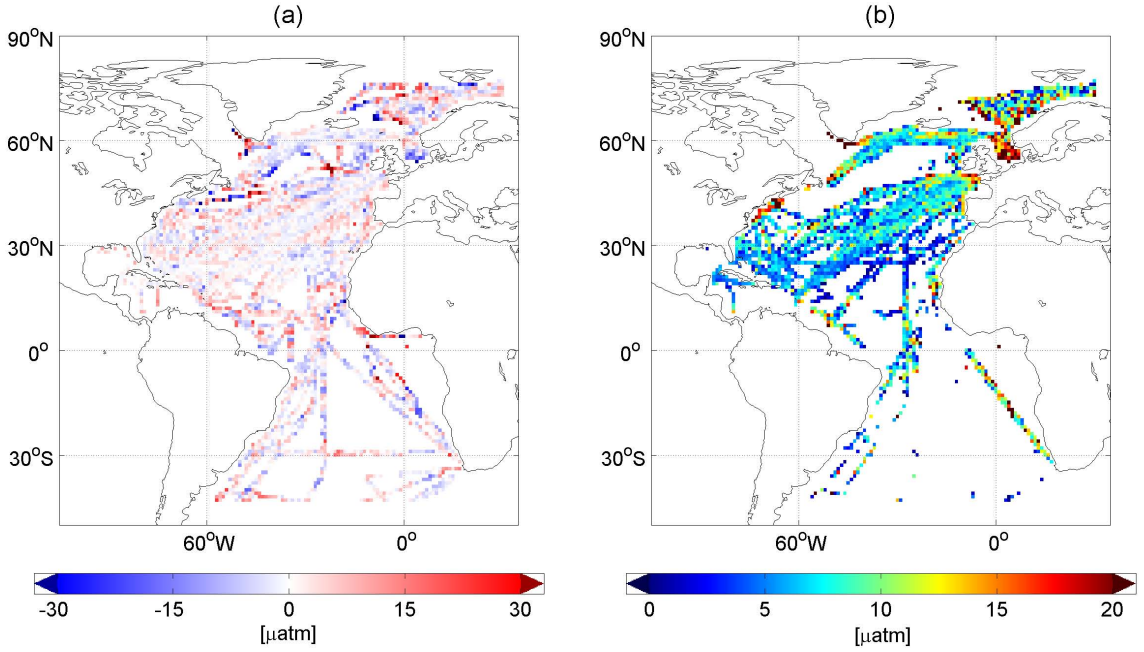

Fig. 2. (a) Temporal mean residuals and **(b)** standard deviation of the residuals in µatm between neural network estimates and SOCAT v1.5 gridded observations (Sabine et al., 2013) for the period from 1998 to 2007. Pixels that have a value in **(a)** but not in **(b)** indicate where only 1 observation in time is available.

near monthly coverage over the time period estimated by the model. As we do not have estimates centred at the exact geographical position of both time-series stations, we interpolate the 4 closest surrounding grid boxes, weighted by their distance, to compare our results.

Figure 4 shows the comparison between the neural network estimates with both time series for the period between 1998 through 2007 and additionally the mean seasonal cycle within this period. While the phase of the seasonal cycle is captured fairly well, Fig. 4 shows that the neural network estimates in general underestimate the winter minima at Bermuda from January to April and the autumn maxima at ESTOC from August to November. This underestimation of the seasonal amplitude is likely linked to the early stopping approach (see Appendix A2) which was implemented to prevent our neural network from overtraining. The neural network estimates further show a decrease in the summer sea surface pCO_2 from 2005 onwards which is not seen in the ESTOC data. The decadal mean difference between BATS data and neural network estimates is 7.56 µatm with a RMSE of 17.53 µatm. Similar to BATS, the decadal mean difference between the estimates in this study and the ESTOC data is −8.06 µatm with a RMSE of 16.85 µatm.

As a last test, we use data from the recently updated SOCAT v2 database (Bakker et al., 2013), which provides new independent data points within our study period to validate the results. A total of 3065 gridded observations, spread over the entire Atlantic Ocean have been added for our study region from 1998 to 2007 representing 15 % of the total amount of data used to train our network. Figure 5 shows the temporal mean and standard deviation of the residuals, similar to Fig. 2. The largest misfit between our estimates and the additional observations can again be identified along the Gulf Stream and North Atlantic Current, confirming that our method has difficulties to fully capture all variability within this region. Overall, the neural network estimates have a RMSE of 22.83 µatm and a bias of 4.85 µatm. When we exclude data north of 40° N, where we obtain the largest misfits, the results improve with a RMSE of 16.29 µatm and a mean difference of −1.12 µatm similar to the numbers obtained from the independent time-series stations. This suggests that over most of the ocean, our method succeeds in predicting the observed pCO_2 at any given time and place to within about 20 µatm, and a bias of a few µatm.

3.3 Uncertainty of the air–sea flux

The uncertainty of the flux product stems from the error in the estimated ΔpCO_2 and the uncertainty of the gas transfer coefficient (Takahashi et al., 2009). We estimate this uncertainty for the integrated flux over the 4 considered RECCAP/ocean inversion regions (see Table 3 for region borders) rather than for each $1° \times 1°$ grid cell.

The ΔpCO_2 estimate is subject to two main sources of errors, i.e. the error derived from discretizing the original observations into $1° \times 1°$ bins and the error of the neural network method to interpolate the data in time and space. For the discretizing error we use a value of 5 µatm as reported

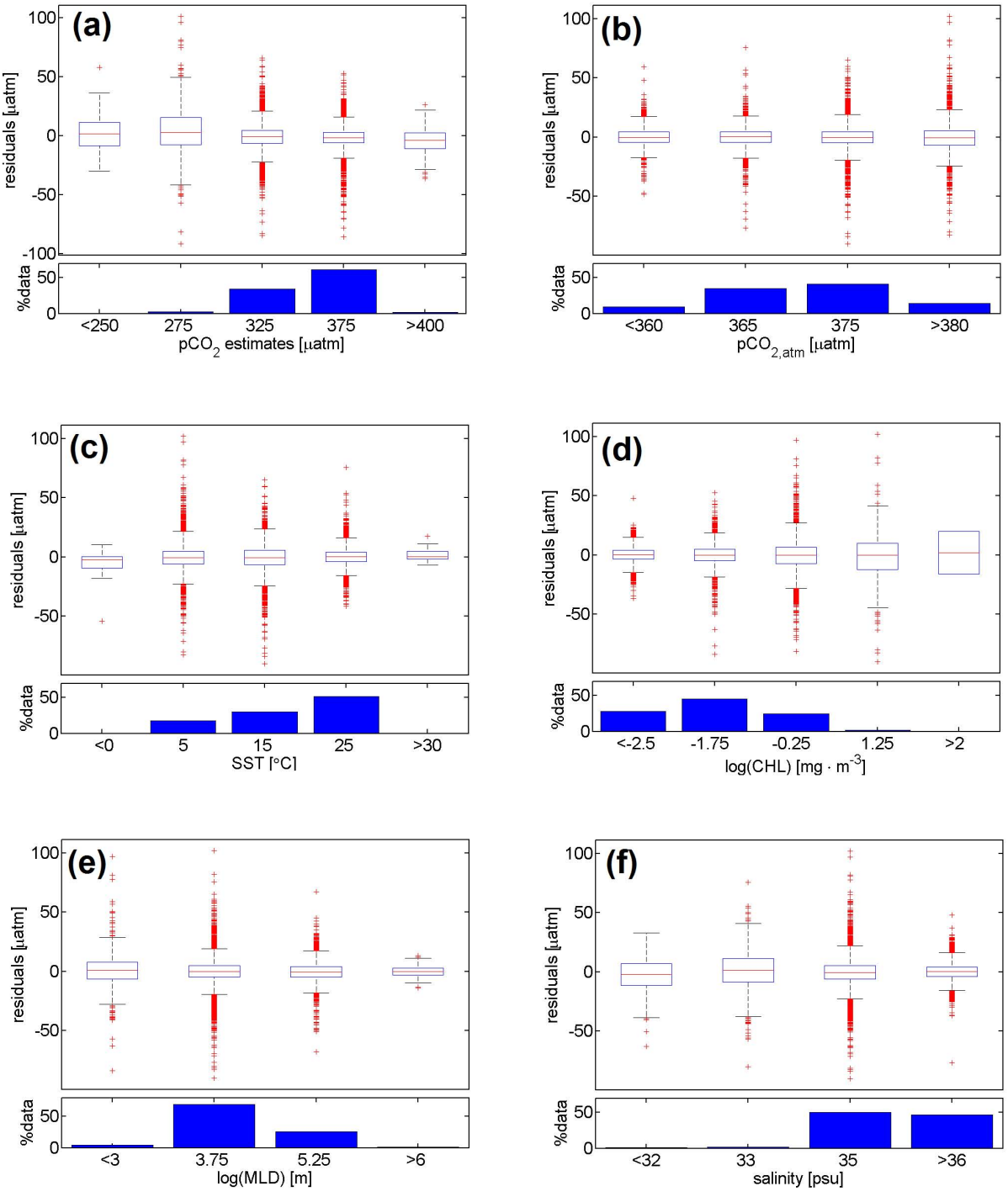

Fig. 3. Residuals as a function of **(a)** fitted pCO_2, **(b)** atmospheric pCO_2, **(c)** sea surface temperature (SST,) **(d)** natural log of surface chlorophyll (log(CHL)), **(e)** natural log of mixed layer depth (log(MLD)), and **(f)** sea surface salinity (SSS). The upper plot in each panel depicts the residuals, shown as a box-and-whiskers plot. The red line in the box show the median, the blue box indicates the 25 and 75 percentiles and red crosses mark residuals outside this interval. The lower plot shows the relative number of observations within each bin.

by Sabine et al. (2013), while we adopt our RMSE value of about 10 µatm (see Table 2) for the interpolation error. When computing next the error of the mean over a larger-scale region, it would be inappropriate to assume that each of the estimates is independent, as these errors are spatially correlated.

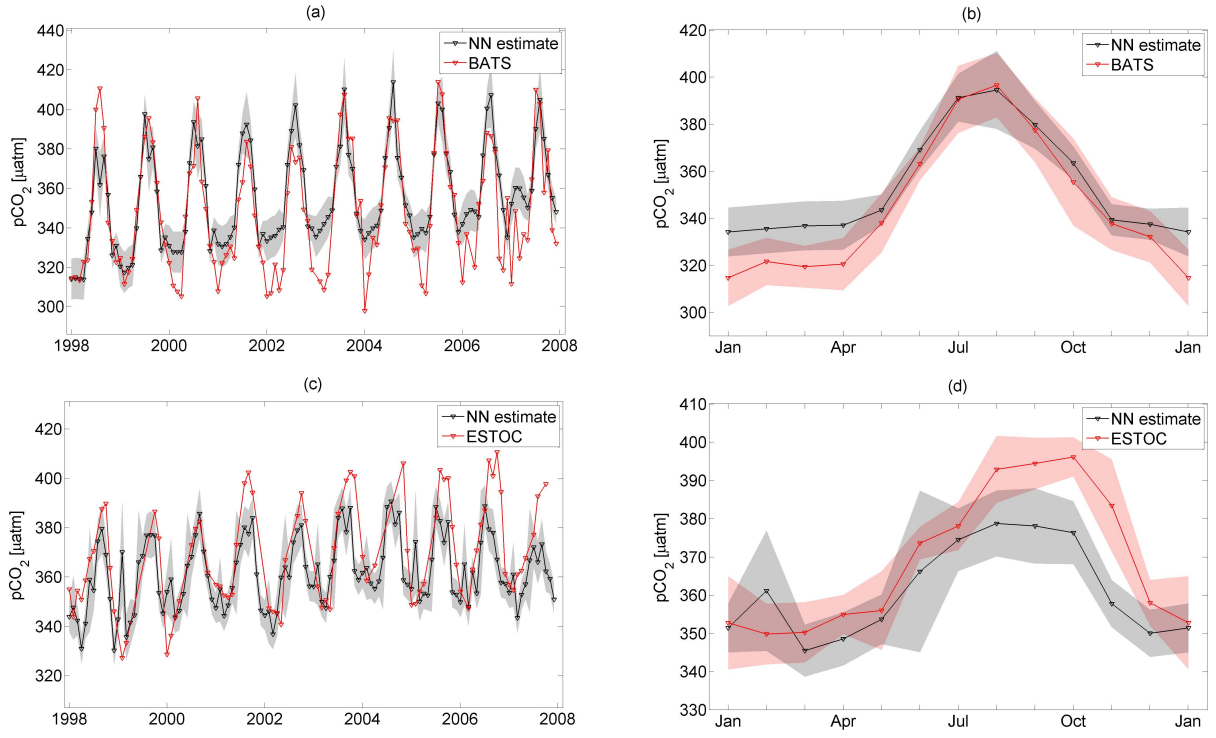

Fig. 4. Validation of neural network results with pCO_2 **(a, b)** from the BATS Hydrostation "S" (Bates, 2007; Gruber et al., 2002) and **(c, d)** from the ESTOC (González-Dávila et al., 2007) time-series stations. In **(a)** and **(c)** the actual time-series data are compared, while in **(b)** and **(d)**, the long-term mean seasonal cycle is evaluated. Grey shading shows the uncertainty of the neural network estimate based on its RMSE. Red shading shows the standard deviation of the mean seasonal cycle for each time-series station.

To estimate the discretization error associated with gridding for each RECCAP/ocean inversion region, we use the spatial decorrelation length scale of 400 km estimated by Jones et al. (2012) to compute the effective number of degrees of freedom. The uncertainty of the mean is then estimated by dividing the standard deviation by the square root of the effective degrees of freedom. This results in an uncertainty of between 1 and 2 µatm for the individual regions.

To estimate the spatial mean of the neural network error for each RECCAP/ocean inversion region, we estimate the spatial correlation by analysing the semi-variogram of the residuals (see e.g. Kalkhan, 2011) (for details see Appendix A4). While in some regions the spatial correlation of the residuals fall very quickly, the correlation of the residuals within one bin is substantial, yielding a substantial reduction in the effective degrees of freedom. The uncertainty between different regions ranges from 1 to nearly 4 µatm.

Adding the error from the gridding and the neural network together, and assuming a mean error of 0.2 µatm for the atmospheric pCO_2 (Takahashi et al., 2009), yields a total ΔpCO_2 error for the 4 regions of between 2 and 6 µatm. With a mean gas transfer rate in the Atlantic Ocean of 0.05 mol C m^{-2} yr^{-1} µatm^{-1} this results in a flux error between 0.03 and 0.06 Pg C yr^{-1} and an overall basin error of 0.07 Pg C yr^{-1} calculated by standard error propagation.

Furthermore, following Sweeney et al. (2007), we assume a random error of 30 % in the gas-transfer velocity. For our long-term mean estimate of the Atlantic Ocean (-0.45 Pg C yr^{-1}) the error due to the piston velocity uncertainty is 0.13 Pg C yr^{-1}. This results in a total uncertainty estimate for the Atlantic Ocean of ± 0.15 Pg C yr^{-1}, or roughly 33 %, with the largest contribution stemming from the uncertain gas transfer velocity.

3.4 Decadal mean pCO_2 and air–sea CO_2 flux

The lowest decadal mean sea surface pCO_2 values are found in the northern North Atlantic, especially the Labrador Sea, the Greenland Sea and the Norwegian Sea with pCO_2 below 320 µatm and in the midlatitudes, along the Gulf stream and North Atlantic Current and the South Atlantic south of 30° S (Fig. 6a). The highest pCO_2 values can be identified in the equatorial Atlantic, in the North Atlantic along the Caribbean Current and the tropical and subtropical South Atlantic northwards of 30° S. Further high values are estimated at 60° N around the Irminger basin and at 30° N in the subtropical North Atlantic along the Portugal Current, and in the eastern North Atlantic along the Canaries Current.

The decadal mean pCO_2 distribution from the neural network method is generally very similar to that estimated by Takahashi et al. (2009), with some important exceptions

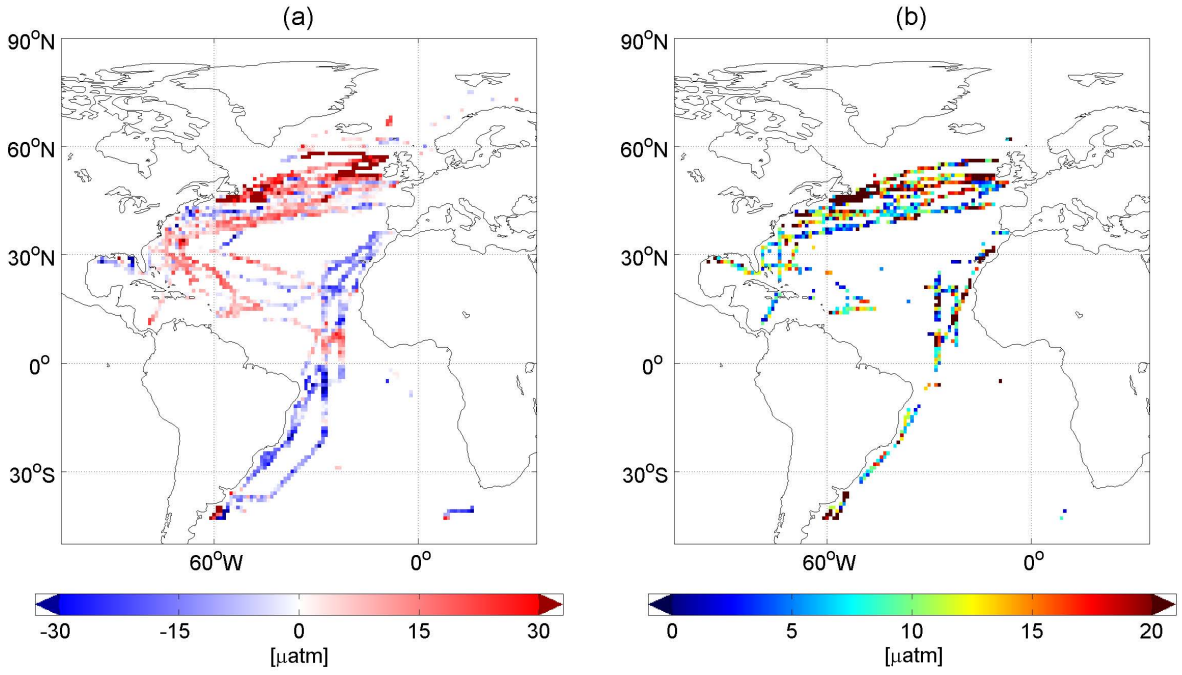

Fig. 5. (a) Temporal mean residuals and **(b)** standard deviation of the residuals in µatm between neural network estimates and independent data points obtained from the SOCAT v2 gridded observations (Bakker et al., 2013) for the period from 1998 to 2007. Pixels that have a value in **(a)** but not in **(b)** indicate where only 1 observation in time is available.

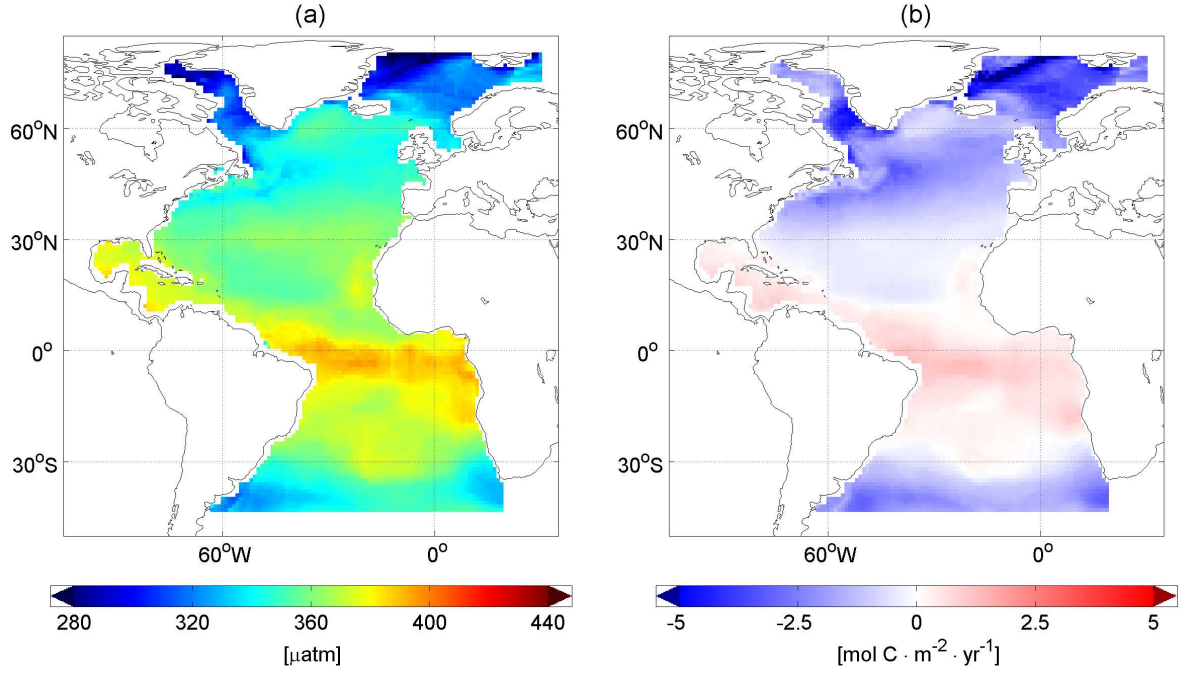

Fig. 6. (a) Decadal mean surface ocean pCO_2 and **(b)** CO_2 flux density in mol C m^{-2} yr^{-1} for the Atlantic Ocean. Negative flux densities indicate CO_2 uptake by the ocean.

(Fig. 7). To produce this comparison plot, we first binned our estimates to the same resolution (4° × 5°) as the original climatology of Takahashi et al. (2009). We then corrected our estimates to the year 2000 by subtracting 4.5 µatm on the basis of the assumption that the surface ocean follows the atmospheric trend of 1.5 µatm per year and the fact that our

estimate is centred around 2003. The strongest differences can be identified in the high latitudes of the North Atlantic within the Labrador Sea, the Greenland Sea and the Norwegian Sea. For the entire Atlantic we derive a mean difference of 0.38 µatm and a RMSE of 6.45 µatm.

Given the overall small bias and the low RMSE between the two very different methods to interpolate the data, it appears that the long-term mean surface ocean pCO_2 can be very robustly estimated from the available observations.

The CO_2 flux density (Fig. 6b) largely follows the pCO_2 pattern, although with some notable differences. Overall, the North Atlantic is a strong sink for atmospheric CO_2 in the mid and high latitudes, whereas the lower latitudes act as a source for atmospheric CO_2. The strongest CO_2 uptake in the North Atlantic occurs along the Gulf Stream and the North Atlantic Current, as well as in the Labrador, Norwegian, and Greenland seas, and in the South Atlantic south of 30° S.

We estimate a decadal mean flux of $-0.44 \pm 0.15\,\mathrm{Pg\,C\,yr^{-1}}$ for the Atlantic Ocean from 44° S to 76° N and 100° W to 19° E ($-0.45 \pm 0.15\,\mathrm{Pg\,C\,yr^{-1}}$ from 44° S to 79° N and west of 30° E). This is in good agreement with Schuster et al. (2013), who provided a "best" estimate of $-0.49 \pm 0.11\,\mathrm{Pg\,C\,yr^{-1}}$ (derived from the mean fluxes of the pCO_2 climatology and the ocean inversion fluxes within the RECCAP project).

Table 3 lists the long-term mean fluxes for the Atlantic Ocean as well as the four individual ocean RECCAP/ocean inversion regions considered by Schuster et al. (2013). While the basin average flux is well within the uncertainty range of the best estimate from Schuster et al. (2013), the subtropical North Atlantic (18–49° N) mean flux is just outside the uncertainty range of the RECCAP best estimate. In general, the neural network fluxes are closest to those of the pCO_2 climatology of Takahashi et al. (2009) with the exception of the subtropical South Atlantic (44–18° S) where the long-term mean flux is closest to the results of the ocean inversion and the ocean biogeochemical models. We estimate the main carbon sink region to be the high latitude North Atlantic with strong uptake throughout the year and a decadal average uptake of $-0.20 \pm 0.07\,\mathrm{Pg\,C\,yr^{-1}}$.

Comparing the decadal mean flux of -0.45±0.15 Pg C yr^{-1} to the results derived from the sensitivity runs SR1–4 reveals that the choice of products does not significantly influence the long-term mean result. The decadal mean fluxes from the sensitivity runs range from -0.41±0.14 (SR2) up to -0.48±0.16 Pg C yr^{-1} (SR4) and are therefore well within the estimated uncertainty range.

3.5 Seasonality

The 10 yr mean sea surface pCO_2 seasonal cycle exhibits strong latitudinal differences (Fig. 8). While we find the weakest seasonal signals north of 60° N, south of 40° S and near the Equator from 10° S to 10° N, the temperate North Atlantic (40–60° N) has a distinct seasonal cycle with high

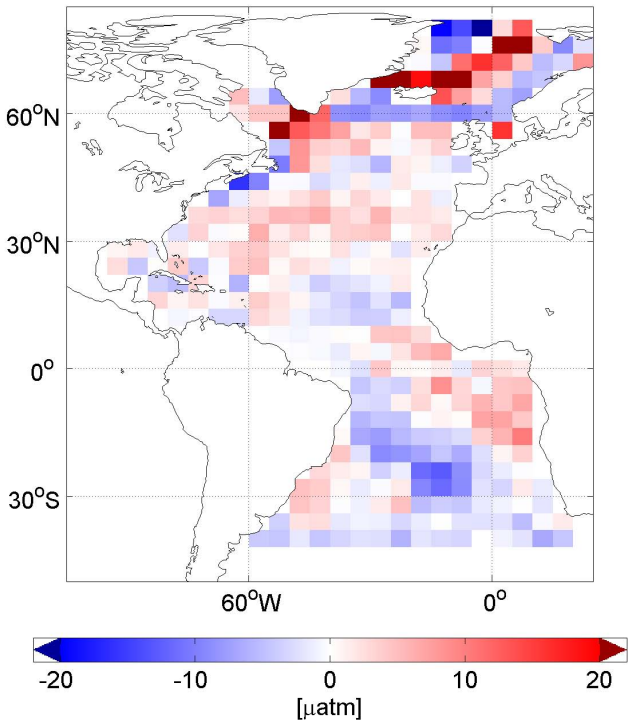

Fig. 7. Difference in the surface ocean pCO_2 in µatm between the decadal mean neural network estimates (this study), corrected to the year 2000, and the estimates from the climatology of Takahashi et al. (2009). Positive differences indicate higher pCO_2 for the neural network estimates. The neural network estimates have been interpolated to the (4° × 5°) grid of the Takahashi climatology.

pCO_2 from October to April and low values from May to September. The opposite cycle occurs in the subtropical North Atlantic between 10–40° N and 10–40° S with low partial pressures in winter and a seasonal maximum in the warmer summer months.

The mean seasonal cycle of our neural network-based estimates of pCO_2 agrees relatively well with the seasonal cycle estimated by Takahashi et al. (2009), but substantial differences exist at the regional level (Fig. 9). As above, we corrected our estimate to the year 2000 in order to have better comparable estimates. The strongest difference can be identified in the high latitude North Atlantic, where we estimate a stronger seasonal cycle with higher values in the winter and lower values in the summer, with differences of more than 10 µatm. In comparison, the differences throughout the Atlantic Ocean are mostly within the calculated RMSE of our method.

To determine the drivers behind the seasonal cycles, we split the long-term mean seasonal cycle at each grid cell into a thermal and into a non-thermal component (Sarmiento and Gruber, 2006; Takahashi et al., 2002, Eqs. 1, 2). The latter is driven by the seasonal changes in temperature and is computed on the basis of the well known temperature sensitivity

Table 3. Comparison of regional and basin-wide decadal mean CO_2 fluxes in Pg C yr^{-1} from the neural network-based method (column VIII) compared to a range of other methods. This includes (I) the pCO_2 climatology of Takahashi et al. (2009) and the tier 1 methodologies described in Schuster et al. (2013) which include (II) ocean inversion (Gruber et al., 2009), (III) atmospheric inversion (Peylin et al., 2013), (IV) ocean biogeochemical models as well as observation based results including (V) a SOCAT v1.5 based multi-parameter regression (Schuster et al., 2013) and (VI) an estimate based on the pCO_2 database of Takahashi et al. (2009) updated by McKinley et al. (2011). (VII) lists the best estimate from combining the different RECCAP methodologies as described in Schuster et al. (2013)

Region	(I) pCO_2 climatology [Pg C yr^{-1}]	(II) Ocean inversion [Pg C yr^{-1}]	(III) Atmospheric inversion [Pg C yr^{-1}]	(IV) OBGC models [Pg C yr^{-1}]	(V) SOCAT MPR [Pg C yr^{-1}]	(VI) pCO_2 database [Pg C yr^{-1}]	(VII) RECCAP best estimate [Pg C yr^{-1}]	(VIII) Neural network [Pg C yr^{-1}]
49–76° N	−0.23	−0.19	−0.28	−0.17	−0.07	−0.30	−0.21	−0.20
(West of 19° E)	± 0.12	± 0.06	± 0.03	± 0.02	± 0.04	± 0.13	± 0.06	± 0.07
18–49° N	−0.19	−0.34	−0.31	−0.13	−0.18	−0.24	−0.26	−0.19
	± 0.09	± 0.08	± 0.03	± 0.03	± 0.09	± 0.16	± 0.06	± 0.07
18° S–18° N	0.11	0.13	0.12	0.15	0.10	0.12	0.12	0.11
	± 0.05	± 0.06	± 0.05	± 0.06	± 0.05	± 0.14	± 0.04	± 0.07
44–18° S	−0.10	−0.17	−0.13	−0.17	−0.25	−0.21	−0.14	−0.16
	± 0.05	± 0.05	±0.02	± 0.01	± 0.12	± 0.23	± 0.04	± 0.06
Atlantic Ocean	**−0.42**	**−0.56**	**−0.60**	**−0.32**	**−0.40**	**-0.63**	**−0.49**	**−0.44**
	± 0.17	**± 0.13**	**± 0.07**	**± 0.07**	**± 0.16**	**± 0.34**	**± 0.11**	**± 0.15**

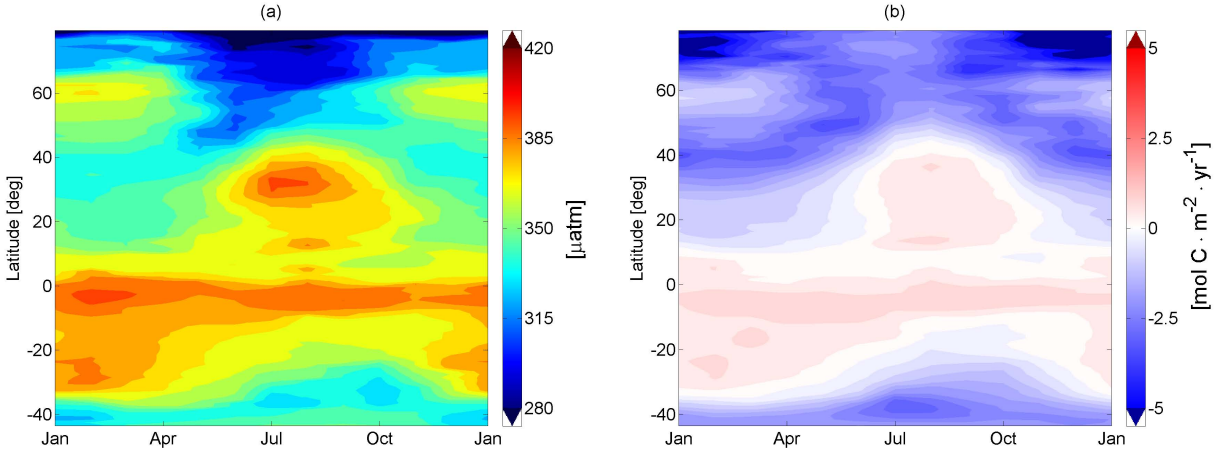

Fig. 8. Hovmöller plot of the zonally averaged long-term mean seasonal cycle of **(a)** pCO_2 in µatm and **(b)** the CO_2 flux density in mol C m^{-2} yr^{-1}.

of pCO_2 (Takahashi et al., 1993, Eq. 2) and the non-thermal component is computed by the difference.

The seasonal cycles of the thermally and non-thermally driven partial pressures tend to cancel each other (Fig. 10), consistent with previous analyses (Sarmiento and Gruber, 2006; Takahashi et al., 2002), i.e. assuming a 4 % change in pCO_2 per unit change in SST, and employing the same SST product used for the network training. In detail, we find in both hemispheres the non-thermally driven pCO_2 starts to decrease due to increasing biological production and reduced vertical mixing resulting in increased stratification in the warmer months. The thermally driven seasonal cycle, however, follows the increase in sea surface temperature and causes an increase in the sea surface pCO_2 due to a reduced solubility. Comparing Fig. 10 with Fig. 8 reveals that the non-thermal seasonal cycle of the sea surface pCO_2 dominates over the thermally driven seasonal cycle polewards of 40° N. In contrast, the seasonal cycle in the subtropical North and South Atlantic is driven by the thermal pCO_2 component. The thermal and non-thermal driven seasonal cycle of the equatorial band and in the South Atlantic south of 35° S compensate each other, resulting in little variability within each band.

The seasonal cycle of the CO_2 flux is largely driven by the seasonal cycle of the sea surface pCO_2 with only modest modifications by the seasonal cycles of wind and solubility

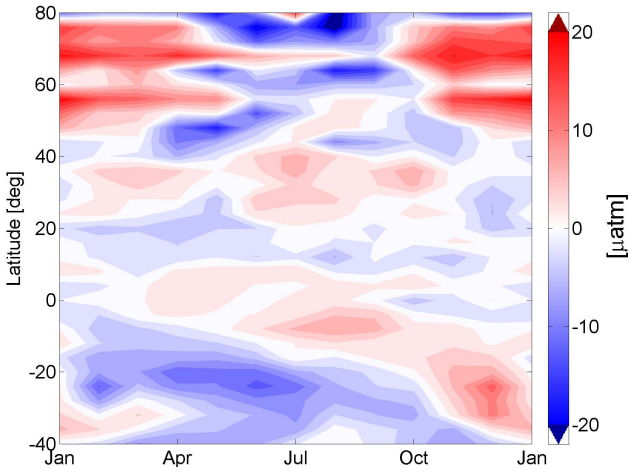

Fig. 9. Zonally averaged difference in the long-term mean seasonal cycle of the surface ocean pCO_2 between our neural network estimates (this study), corrected to the year 2000 and those from the climatology of Takahashi et al. (2009). Positive differences indicate higher partial pressures of CO_2 in the neural network-based estimates.

(Fig. 8b). The temperature driven solubility and the wind lead to a larger outgassing of CO_2 in large areas in the high latitudes in winter, where the sea surface pCO_2 is supersaturated, and to an increasing uptake of CO_2 in the subtropics and low latitudes and vice versa in summer. Throughout the entire Atlantic, the flux densities are considerably larger during the Northern Hemisphere winter season, compared to the summer. The neural network estimates show a strong seasonal CO_2 outgassing in summer in the northern subtropics, driven by the increasing pCO_2, with a 6 month difference in the Southern Hemisphere.

3.6 CO_2 trends and inter-annual variability

The main driving variable for trends in the sea surface pCO_2 is the atmospheric CO_2, but within our study period the neural network estimates show that these trends are not in parallel. Across large areas of the Atlantic, the 10 yr trend of surface ocean pCO_2 is estimated to be lower than that of atmospheric pCO_2 (Fig. 11a), but there are notable exceptions. In this plot, the atmospheric trend has been subtracted from the long-term mean trends for each $1° \times 1°$ pixel, so that positive values indicate a rate of increase faster than of the atmosphere and vice versa for negative values. Table 2 shows that the bias between estimates and observations is fairly constant at each year individually, suggesting that trends are captured well where observations exist.

The strongest increase in ocean surface pCO_2 relative to that in the atmosphere is found in the North Atlantic poleward of 40° N along the Gulf Stream and North Atlantic Current. Here, the neural network suggests an increase in sea surface pCO_2 of more than twice the atmospheric increase. Metzl et al. (2010) investigated the sink trend over a similar time period (2001–2008) in the North Atlantic subpolar gyre (53–57.5° N, 45–35° W and 57.5–62° N, 40–25° W). These authors found a particularly strong increase in the winter sea surface fCO_2 of 5.8 ± 1.1 and 7.2 ± 1.3 µatm yr^{-1}. While this is much stronger than suggested here (see Fig. 11) both studies agree on the North Atlantic subpolar gyre having a trend towards a stronger increase of the sea surface pCO_2 over the $\sim$2000 to $\sim$2007 period. McKinley et al. (2011) found a similar trend towards a weaker undersaturation in their subtropical seasonally stratified region around 40° N for the period 1993 until 2005, but did not identify a significant trend in the subpolar gyre over the same period. This may reflect differences in the time period, as their analyses with an earlier start, i.e. pre-1990, suggest also a trend toward a weaker undersaturation.

Decadal trends in surface ocean pCO_2 in the Labrador Sea and some parts of the Icelandic Seas were much smaller than that in the atmosphere, leading to an overall relatively small trend for the entire region north of 40° N. As the low latitudes of the North Atlantic (0–40° N) have close to zero trend relative to that of the atmosphere, the entire North Atlantic pCO_2 trend is also very close to that of the atmosphere, i.e. 1.80 ± 0.77 µatm yr^{-1} versus 1.90 ± 0.34 µatm yr^{-1} for the atmosphere.

As expected from the uptake of anthropogenic CO_2 by the surface ocean, the majority of the ocean pCO_2 trend stems from the non-thermal part, i.e. the increase in surface ocean DIC (dissolved inorganic carbon) (1.46 ± 1.75 µatm yr^{-1}). However, the overall warming trend in the North Atlantic over this decade further enhances the ocean pCO_2 trend quite considerably (with on average stronger than atmospheric increases of 2.25 ± 0.77 µatm yr^{-1} from 40 to 60° N and east of 50° W and 4.03 ± 3.18 µatm yr^{-1} at 60–70° N and 0–10° W). Splitting the trend into thermal and non-thermal component shows on average a linear trend of 1.46 ± 1.75 µatm yr^{-1} for the non-thermal components, while the thermally driven trend is on average 0.37 ± 1.47 µatm yr^{-1}. However, given their uncertainty, we cannot statistically distinguish both trends from zero. The small difference in the increase in surface ocean pCO_2 relative to the atmosphere results in an almost steady strength of the Atlantic carbon sink in time north of the Equator (-0.01 ± 0.02 Pg C yr^{-1} decade^{-1}).

Trends for the South Atlantic show a weaker increase in the sea surface pCO_2 relative to that in the atmosphere with the exception of the eastern South Atlantic and parts along the South American coast. On average, surface ocean pCO_2 increased only by 0.98 ± 0.97 µatm yr^{-1} over the 1998 through 2007 period, resulting in a carbon sink increase of -0.14 ± 0.02 Pg C yr^{-1} decade^{-1}. Similar to the North Atlantic, the non-thermal component of the pCO_2 with an average trend of 0.76 ± 1.30 µatm yr^{-1} appears to be stronger compared to 0.19 ± 0.79 µatm yr^{-1} of the thermal component, but given their uncertainty, both trends are again indistinguishable from zero.

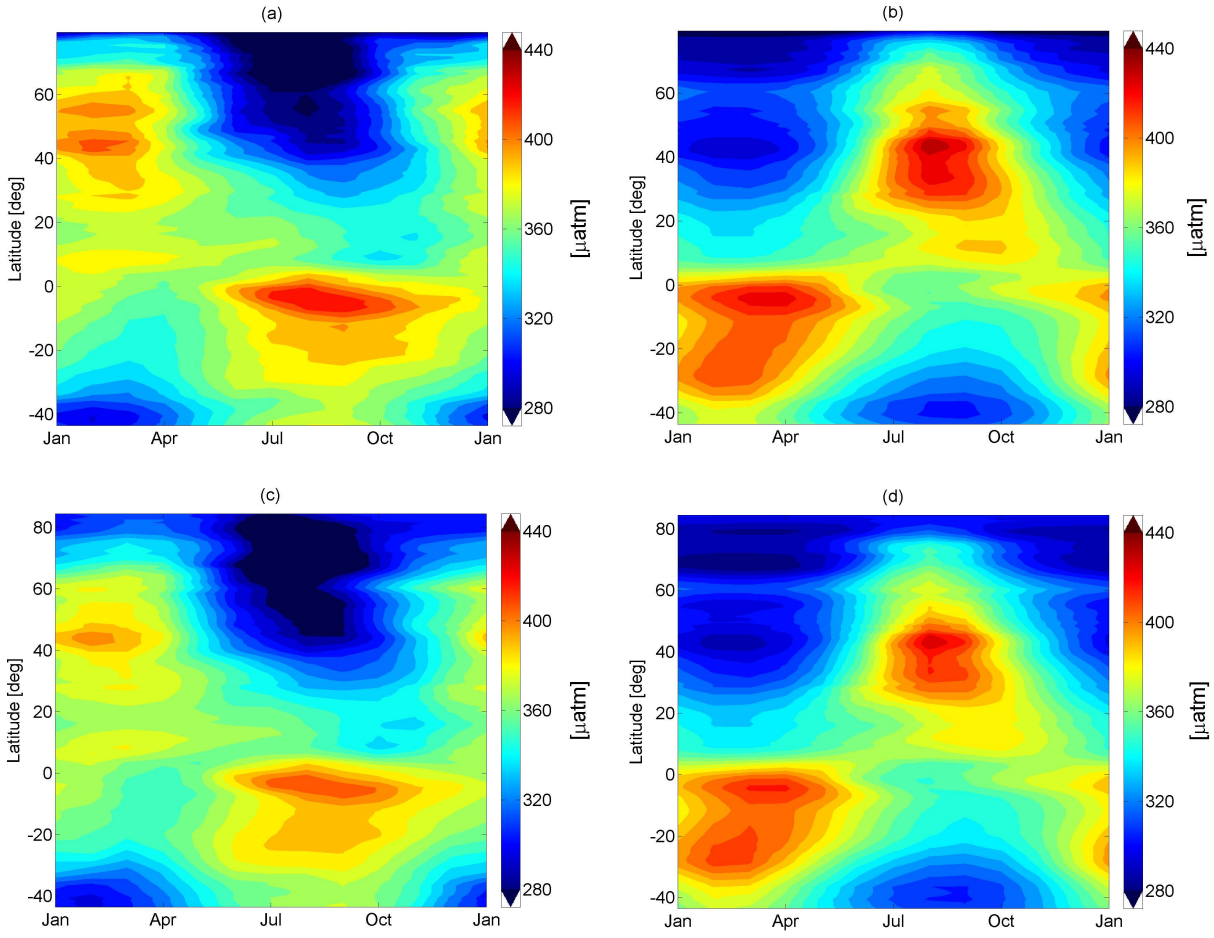

Fig. 10. Mean seasonal cycle of the zonally averaged sea surface pCO_2 driven by **(a)** the non-thermal component, such as changes in circulation, mixing, and biology, and **(b)** the thermal component, i.e. the changes in temperature. **(c)** and **(d)** show the non-thermal and thermal components respectively for the Takahashi et al. (2009) climatology. The decadal mean pCO_2 has been added to both components.

Taking the North and South Atlantic together, the trend over the entire study region is one toward a stronger sink over the 10 yr period with an overall mean trend of $1.46 \pm 0.76\,\mu\text{atm}\,\text{yr}^{-1}$ and a flux trend of $-0.15 \pm 0.04\,\text{Pg C}\,\text{yr}^{-1}\,\text{decade}^{-1}$.

The sensitivity runs reveal that trend estimates are barely influenced by the choice of the input data product, with the exception of SR2, i.e. where we had replaced the ECCO2 MLD product with the de Boyer Montegut (de Boyer Montegut et al., 2004) MLD climatology. While pCO_2 trends are statistically indistinguishable between our neural network estimate ($1.46 \pm 0.76\,\mu\text{atm}\,\text{yr}^{-1}$) and SR1–4 ($1.42 \pm 0.59$, 1.25 ± 0.48, 1.48 ± 0.77 and $1.37 \pm 0.73\,\mu\text{atm}\,\text{yr}^{-1}$ respectively), this is not always true for the fluxes. Here, SR2 reveals a flux trend ($-0.26 \pm 0.03\,\text{Pg C}\,\text{yr}^{-1}\,\text{decade}^{-1}$) outside the uncertainty interval of our neural network estimate ($-0.15 \pm 0.04\,\text{Pg C}\,\text{yr}^{-1}\,\text{decade}^{-1}$).

It is not possible to conclude from our data whether the 10 yr trends identified here are part of truly long-term trends (30 yr or longer), or whether they are part of decadal timescale fluctuations (Thomas et al., 2008; Gruber, 2009; McKinley et al., 2011). The most recent studies by McKinley et al. (2011) and Fay and McKinley (2013) suggest the latter to be the case. The authors show that short-term trends on timescales similar to this study are strongly influenced by the chosen start and end year and strongly reflect climate mode signals such as ENSO and NAO. However, reported 50 yr trends in heat storage (Levitus et al., 2012) and interior ocean oxygen changes in the North Atlantic (Stendardo and Gruber, 2012) indicate that this region has been subject to multi-decadal changes, particularly in the subpolar gyre. It is also tempting to point out that the resulting pattern of a decreasing sink in large areas of the North Atlantic and increasing sink in the South Atlantic appears to be mirrored in the observation of a faster rate of accumulation in anthropogenic CO_2 in the South compared to the North Atlantic (Wanninkhof et al., 2010). One needs to be careful, though, as the surface ocean trends are for the sum of natural and anthropogenic CO_2, while the ocean interior trends are for anthropogenic CO_2 only.

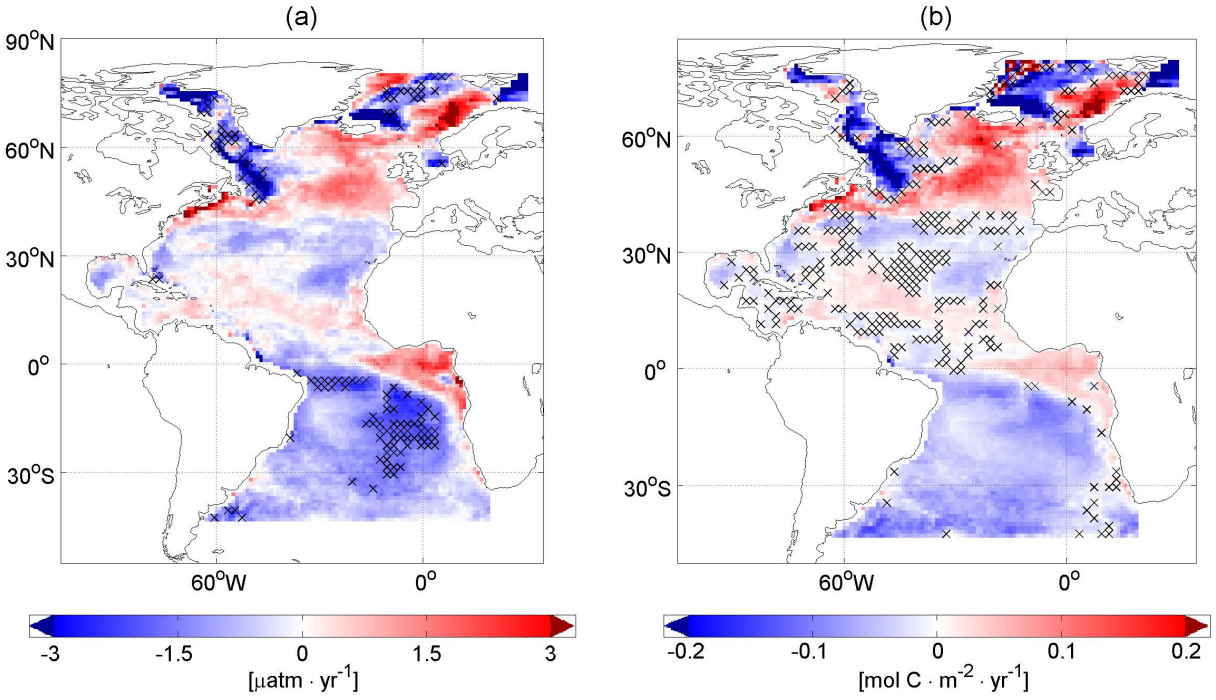

Fig. 11. Linear trends **(a)** in sea surface pCO_2 relative to that in the atmosphere and **(b)** in the CO_2 flux over the period 1998–2007. The relative trend in sea surface pCO_2 was computed by subtracting the atmospheric mean trend. Areas with cross-hatch indicate where the trend is outside the 95 % confidence level ($p \geq 0.05$). Trends were derived by first applying a 12 month running mean to each pixel to deseasonalize the data and then calculating the slope of a linear fit to these deseasonalized data.

The largest year-to-year variability in sea surface pCO_2 is found within the North Atlantic north of 40° N and in the eastern equatorial and South Atlantic. In contrast, the subtropics in both hemispheres show much less year-to-year variability.

Integrating our monthly air–sea CO_2 flux estimates for each year over the Atlantic Ocean reveals the largest annual mean flux differences during the second half of our study period (Fig. 12a), where annual mean fluxes range from $-0.39 \pm 0.13\,\mathrm{Pg\,C\,yr^{-1}}$ in 2001 up to $-0.56 \pm 0.18\,\mathrm{Pg\,C\,yr^{-1}}$ in 2006. Figure 12b illustrates the inter-annual variabilities (IAV) for both hemispheres and the entire Atlantic Ocean. The IAV, calculated as a 12 month running average, is fairly constant from 1998 to 2004 with a weak flux decrease in the Northern Hemisphere counterbalanced by a weak increase in the Southern Hemisphere. After 2004 the Atlantic Ocean sink increases mainly due to increases in the Southern Hemisphere. The standard deviations of the IAV (calculated as a 12 month running average and further detrended) for the Atlantic Ocean north of the Equator, south of the Equator and the entire basin are $\pm 0.02\,\mathrm{Pg\,C\,yr^{-1}}$, $\pm 0.02\,\mathrm{Pg\,C\,yr^{-1}}$ and $\pm 0.04\,\mathrm{Pg\,C\,yr^{-1}}$ respectively, which indicates only limited inter-annual variability in both hemispheres on a basin scale. This Atlantic Ocean low variability is further confirmed by the sensitivity runs, ranging from $\pm 0.03\,\mathrm{Pg\,C\,yr^{-1}}$ (SR1, SR2) to

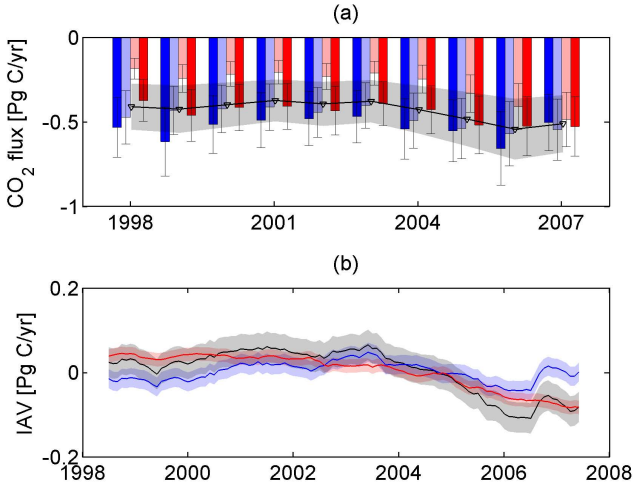

Fig. 12. (a) Seasonal and annual mean fluxes from 1998 to 2007 in the Atlantic Ocean (44° S–79° N and west of 30° E). Dark blue shows the results for the Northern Hemisphere winter months (DJF), light blue the spring months (MAM), light red the summer months (JJA), dark red the autumn months (SON). The annual mean flux is plotted as a black line on top. **(b)** Inter-annual variability (calculated using a 12 month running mean) for the Northern Hemisphere (blue line), the Southern Hemisphere (red line) and the entire Atlantic Ocean (black line).

$\pm 0.04\,\mathrm{Pg\,C\,yr^{-1}}$ (SR3, SR4), indicating that our result is not sensitive with regards to the data choice. This shows that the main findings are statistically indistinguishable from those derived without chlorophyll *a* (SR4), indicating the possibility to expand the analysis period back in time in future studies.

Inter-annual variability of the sea surface pCO_2 in the North Atlantic has previously been linked to variations in the NAO (e.g. Gruber et al., 2002; Schuster and Watson, 2007; Thomas et al., 2008). The NAO is the dominant large-scale climate mode in the Atlantic Ocean (e.g. Hurrell, 1995) and impacts sea surface pCO_2 via changes in the driving parameters. During positive NAO phases, sea surface temperature shows a tripole pattern with cold anomalies in the subpolar region and warm anomalies in the midlatitudes and corresponding changes in vertical mixing and nutrient supply (Marshall et al., 2001). Corbière et al. (2007) found an increase in sea surface pCO_2 in the subpolar gyre between the mid-1990s and mid-2000s linked to an increase in SST due to a shift from positive to negative NAO. We investigate the effect of the NAO by focusing on two $10° \times 10°$ boxes, one located in the subtropical North Atlantic (20–30° N and 40–50° W) and the other in the subpolar North Atlantic (50–60° N and 30–40° W). Figure 13 illustrates the pCO_2 estimate and their anomalies for each box together with the NAO index. We find a weak but significant ($p \leq 0.05$) positive correlation in the subtropics ($R = 0.32$) and negative correlation in the subpolar box ($R = -0.31$). This pattern is consistent with that identified by Thomas et al. (2008) on the basis of a modelling study (see also summary by Gruber (2009) and recent multi-model analysis by Keller et al. (2012), and the available time-series analyses (e.g Gruber et al., 2002; Bates, 2007) from the BATS site). The correlation patterns are derived from the neural network estimates, hence the NAO signal stems from the signal of the input data. Clearly, an important driver are the NAO-associated SST anomalies, but strongly modified by the various physical and biogeochemical changes that are driven by the NAO-induced changes in heat fluxes and wind stress (see e.g. Keller et al., 2012).

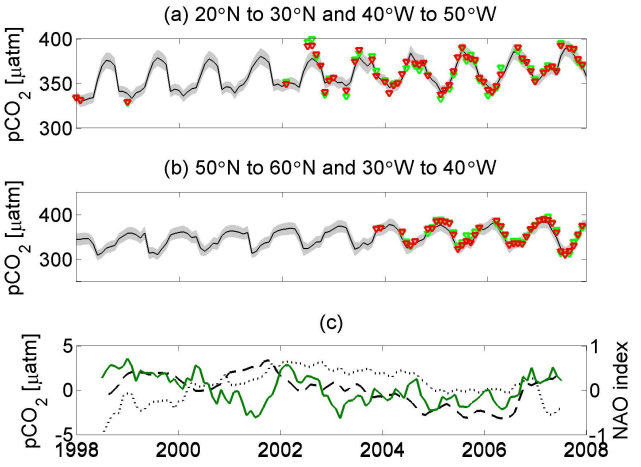

Fig. 13. (a) Temporal evolution of the pCO_2 (in µatm) in the subtropical (20–30° N, 40–50° W) and **(b)** in the subpolar box (50–60° N, 30–40° W). The black line shows the spatial average pCO_2 within each $10° \times 10°$ box. Red triangles illustrate the average sea surface measured pCO_2 within each box where observations are available and the green triangles represent the average of the neural network pCO_2 of those $1° \times 1°$ pixels which have co-located pCO_2 observations in SOCAT v1.5. **(c)** pCO_2 anomalies (left axis in µatm – detrended and smoothed using a 12 month running average filter) compared to the NAO index (right axis – smoothed using a 12 month running average filter). The dashed line shows the anomaly for the subtropical box, the dotted line for the subpolar box and the green line illustrates the NAO index.

4 Summary and conclusions

We have developed a novel two-step neural network approach to diagnose monthly ocean surface pCO_2 fields from 1998 through 2007 in the Atlantic Ocean using the SOCAT v1.5 pCO_2 measurement network. Independent testing indicates our estimates are accurate to within 22.8 µatm for the entire Atlantic Ocean, which decreases to 16.2 µatm for data south of 40° N. Our study suggests a decadal mean CO_2 flux from 1998 through 2007 of $-0.45 \pm 0.15\,\mathrm{Pg\,C\,yr^{-1}}$ for the Atlantic Ocean from 44° S to 79° N and west of 30° E. This result is in good accordance with the recent assessment from the RECCAP project (Schuster et al., 2013). We find the strongest seasonal variability in our predicted sea surface pCO_2 and air–sea fluxes within the subtropics in the Northern and Southern hemispheres, i.e. the zones where the temperature effect dominates the seasonal cycle of sea surface pCO_2. Trends in sea surface pCO_2 suggest that in large areas poleward of 40° N the rate of increase in oceanic pCO_2 was faster than in the atmosphere, leading to a regional weakening of the carbon sink strength. However, this is counterbalanced on the basin scale by weaker surface ocean pCO_2 trends elsewhere in the North Atlantic. The South Atlantic, in contrast, shows an increasing carbon sink strength throughout the study period. In total, the Atlantic Ocean carbon sink increased by $-0.15 \pm 0.04\,\mathrm{Pg\,C\,yr^{-1}\,decade^{-1}}$. The standard deviation of the inter-annual variability of the spatially integrated CO_2 flux within the study period in the Atlantic Ocean was $\pm 0.04\,\mathrm{Pg\,C\,yr^{-1}}$ with both low inter-annual variability in the Southern Hemisphere ($\pm 0.02\,\mathrm{Pg\,C\,yr^{-1}}$) and in the Northern Hemisphere ($\pm 0.02\,\mathrm{Pg\,C\,yr^{-1}}$).

It would be beneficial to extend the study period to further investigate responses to climate modes such as the NAO and to investigate multi-decadal variabilities. The absence of chlorophyll *a* permits us to prolong our analysis period and this option will be explored in subsequent work. However, chlorophyll *a* is a simple, but important proxy representing the relation between biology and pCO_2 and our results provide no evidence that chlorophyll can be neglected

when considering longer timescales. Chlorophyll a is available from models before the launch of satellite observations, but the products to-date have not achieved sufficient reliability as yet.

Our product shows that the data collection and synthesis effort of the marine carbon community makes it possible to investigate the seasonal to inter-annual variability of the ocean carbon sink on a basin-scale based on observations. Future measurements are expected to increase the accuracy of such observation based estimates, and further improvements of the methods used to model the observations will result in providing better historical estimates and more accurate products for these important fluxes.

Appendix A

A1 Dividing the global ocean into biogeochemical provinces using a self-organizing map

A map with 16 neurons was chosen, organized on a 2-dimensional 4×4 point hexagonal grid. This means that the input data are clustered into 16 neurons, which represent the 16 biogeochemical provinces. The term neuron refers to a processing unit, which consist of a weight vector, where each element of the weight vector corresponds to one input parameter. In our case each weight vector consists of 4 elements (SST, log(MLD), SSS, $pCO_{2,\mathrm{Takahashi}}$), representing its coordinates and the distance between 2 neurons is calculated via a distance function. These processing units are initially spread over a 2-dimensional field, in our case in a hexagonal formation, forming a single layer of neurons. Our experience has shown, however, that the choice of neuron topology does not have a significant effect on the final province distribution. The use of neurons, their initialization and their distance relation describes the biggest difference towards other clustering algorithms, e.g. k means clustering. For our study, the Euclidean distance between a neurons weight vector and the input vectors of the SINP data set was used for the distance function. The weight matrix ($\mathbf{W}_{m=16,n=4}$), which is formed by the 16 neurons with their 4 vector elements, was randomly initialized.

After initialization, the training vectors are introduced to the SOM with each training parameter as 1 element of the vector. For the jth input vector $\boldsymbol{p}_n^j$ with the length n the Euclidean distances to each of the $i = 1, \ldots, 16$ neurons represented by each row i of the weight matrix are calculated:

$$d_i^j = \sqrt{\sum_{l=1}^{n} (\mathbf{W}_{i,l} - \boldsymbol{p}_l^j)^2}, \tag{A1}$$

where $\boldsymbol{d}_m^j$ comprises a vector containing the Euclidean distances to each neuron i of the input vector $\boldsymbol{p}_n^j$. The smallest element of the distance vector, i.e. the shortest distance element, marks the distance towards the closest neuron, called the winning neuron. The neuron i, gets updated by moving towards the average position of all the training vectors j it was identified as a winner, or a close neighbour of the winner. This sort of training is called "batch training". This is done by adjusting the ith row of the initial weights matrix after the iterative step q following Vesanto et al. (2000):

$$\mathbf{W}_{i,n}(q+1) = \frac{\sum_{j=1}^{r} \mathbf{S}(d_{\mathrm{neighbour}} - d_i^j) \cdot \boldsymbol{p}_n^j}{\sum_{j=1}^{r} \mathbf{S}(d_{\mathrm{neighbour}} - d_i^j)}, \tag{A2}$$

where $\mathbf{W}_{(i,n)}$ is the ith row of the input weight matrix, d_i^j is the Euclidean distance between the neuron i and the presented jth input vector and $d_{\mathrm{neighbour}}$ is the neighbourhood radius. $\mathbf{S}$ describes the step function. Neighbouring neurons will only be updated if $\mathbf{S}(d_{\mathrm{neighbour}} - d_i^j) > 0$. During the training of our SOM we decrease $d_{\mathrm{neighbour}}$ in 2 steps from an initial coarse training phase where we update the surrounding nearest neighbours that lie within a neighbourhood radius of 3, which lasts for the first 100 iterations, to a final fine training phase where only the winning neuron is updated.

After $q = 200$ iterations (presenting all our input vectors to the SOM 200 times) we stop the training and the input vectors are re-introduced, without updating the SOM. In return every vector receives the neuron number i of the winning neuron, where $d_i^j = \min(\boldsymbol{d}_m^j)$, until every training vector is labelled with a number between 1 and 16, representing the province the vector belongs to. Since every training vector has a geographical location we can now divide the global ocean into these 16 provinces, i.e. the 16 biogeochemical provinces.

We forced the relative weights of the input data toward the climatological pCO_2 data, in order to minimize the variance of pCO_2 within each biogeochemical province. To do so, we did not normalize our input data, with the exception of MLD, which we log-transformed (Table 1). As a consequence, the range between the lowest and highest value of pCO_2 is one order of magnitude larger than that for SST, and about another order of magnitude larger than that for the remaining input parameters (log(MLD), SSS). This was done to reduce the biases in the second stage of the fitting, i.e. in the feed-forward network. As a consequence, the biogeochemical provinces follow the seasonal pattern of the pCO_2 climatology, meaning that the seasonality of pCO_2 at any given location will be mostly determined by the seasonal changes of the biogeochemical provinces and to a lesser degree by the seasonal cycle of the input data in the second state. In addition, owing to the climatological nature of the used pCO_2 data, there are little inter-annual shifts in the distribution of the biogeochemical provinces.

A2 Reconstructing the sea surface pCO_2 using a feed-forward network

Our feed-forward network uses 2 layers of neurons, 1 hidden layer of neurons using a sigmoid transfer function and 1 linear output layer. The hidden layer response to the input vector $\boldsymbol{p}_n^j$ can be written as follows:

$$a_i = \frac{2}{1+\exp(-2\cdot(\mathbf{W}_{i,n}\cdot\boldsymbol{p}_n^j + b_i))} - 1, \quad \text{(A3)}$$

where a_i is the hidden layer response of ith neuron. The response of all neurons m then forms the output vector of the hidden layer $\boldsymbol{a}_m$. This vector serves as input for the second layer of neurons, the linear output layer.

$$p\mathrm{CO}_{2,\mathrm{est}}^j = \mathbf{W}_m\cdot\boldsymbol{a}_m + b \quad \text{(A4)}$$

Eqs. (A3) and (A4) show how the network calculates the scalar output $p\mathrm{CO}_{2,\mathrm{est}}^j$ for the jth input vector $\boldsymbol{p}_n^j$ in 2 steps, each step referring to 1 layer of the network. In the hidden layer the input vectors are multiplied with the weight matrix of the hidden layer $\mathbf{W}_{m,n}$ and added to the layer bias vector $\boldsymbol{b}_m$. The output vector of the hidden layer $\boldsymbol{a}_m$ is created using a tangent sigmoid transfer function (Eq. A3), that computes elements of $\boldsymbol{a}_m$ in the range from -1 to $+1$. Similar to the SOM, the size of $\mathbf{W}_{m,n}$ is determined by the size n of the input vector and the number m of neurons in the hidden layer. The length of $\boldsymbol{b}_m$ and $\boldsymbol{a}_m$ is as well determined by the number of neurons m. In the linear output layer $\boldsymbol{a}_m$ is processed the same way as $\boldsymbol{p}_n^j$ in the hidden layer, with the exception that the output layer only consists of 1 neuron to produce one scalar $p\mathrm{CO}_2$ estimate ($p\mathrm{CO}_{2,\mathrm{est}}^j$) for every input vector. Furthermore, the linear output layer allows $p\mathrm{CO}_{2,\mathrm{est}}^j$ to have any value between $-$*infinity* and $+$*infinity*. During the training the weights and biases of each layer get iteratively adjusted to minimize the error between the network output $p\mathrm{CO}_{2,\mathrm{est}}^j$ and the scalar target element t^j from the SOCAT v1.5 database that corresponds to $\boldsymbol{p}_n^j$. Therefore, the network can only be trained by those input vectors which do have co-located observations.

Before the training starts, the training vectors with corresponding observations from our FINP^k data set are provided to initialize the network layer size. The size of the 2 network layers is determined by the number of neurons and the number of input vector elements n. We randomly generate the networks initial weights and biases. An important parameter that has to be provided before training starts is the number of neurons. Too few neurons are not able to reproduce realistic results, whereas too many neurons decrease the computational performance and cause over-fitting and therefore the network is not able to generalize (Demuth et al., 2008). Since the number of inputs and targets varies per province, we cannot provide one best number of neurons to use for all 16 provinces. We therefore perform a pre-training, increasing the number of neurons in the first or hidden layer parabolically starting from 2 neurons up to a number where the ratio between of the training sample size to the number of weights does exceed 30. Amari et al. (1997) proposed this ratio to prevent artificial neural networks from over-fitting.

During every pre-training process the FINP^k set is divided into 2 separate sub-sets. The first (FITR^k) is used to train the network and the second (FIVAL^k) is used for validation within the method. Amari et al. (1997) suggested an optimal split (r_{opt}) between training and validation data as a function of the modifiable parameters m:

$$r_{\mathrm{opt}} = \frac{1}{\sqrt{2\cdot m}}. \quad \text{(A5)}$$

Modifiable parameter refers to the weights and biases of the network. During every pre-training process the FITR^k training vectors and the corresponding FITR^k targets are introduced to the network and the weights and biases are iteratively updated in the direction where the performance function, which is the mean squared error between network outputs $p\mathrm{CO}_{2,\mathrm{est}}^j$ and FITR^k targets t^j, decreases most rapidly. Our feed-forward network uses the Levenberg–Marquardt (Marquardt, 1963) algorithm to update weights and biases in every iteration step to reduce the mean squared error between outputs and targets. The application of the algorithm in neural networks is described in more detail in Hagan and Menhaj (1994); Hagan et al. (1996).

After every iteration of each pre-training the network is validated by using the FIVAL^k sub-set. The updated weights and biases are used to simulate outputs from the FIVAL^k inputs and the mean squared error between these outputs and the FIVAL^k targets is calculated. Every pre-training of the network stops automatically when 6 consecutive iterations do not reduce the network's error on the FIVAL^k targets to prevent the network from over-fitting. After the pre-trainings with increasing number of neurons we select the one where the mean squared error of the validation data set FIVAL is a minimum and receive the optimal number of neurons for the actual training process.

During the actual training process the number of neurons is adjusted according to the best pre-training performance for each of the 16 provinces separately. We perform 10 trainings where we randomly pick validation data according to Eq. (A5) out of the entire pool of observations to avoid over fitting of the network output. After every training we use the trained network to simulate $p\mathrm{CO}_{2,\mathrm{est}}^j$ from the FINP^k data set and average the output of the 10 training cycles, to end up with 1 estimate for our time period from 1998 through 2007 for each province. After 16 FFN runs we can now combine our results of the 16 provinces to retrieve our $p\mathrm{CO}_2$ estimates from 1998 to 2007 on a global $1^\circ \times 1^\circ$ grid.

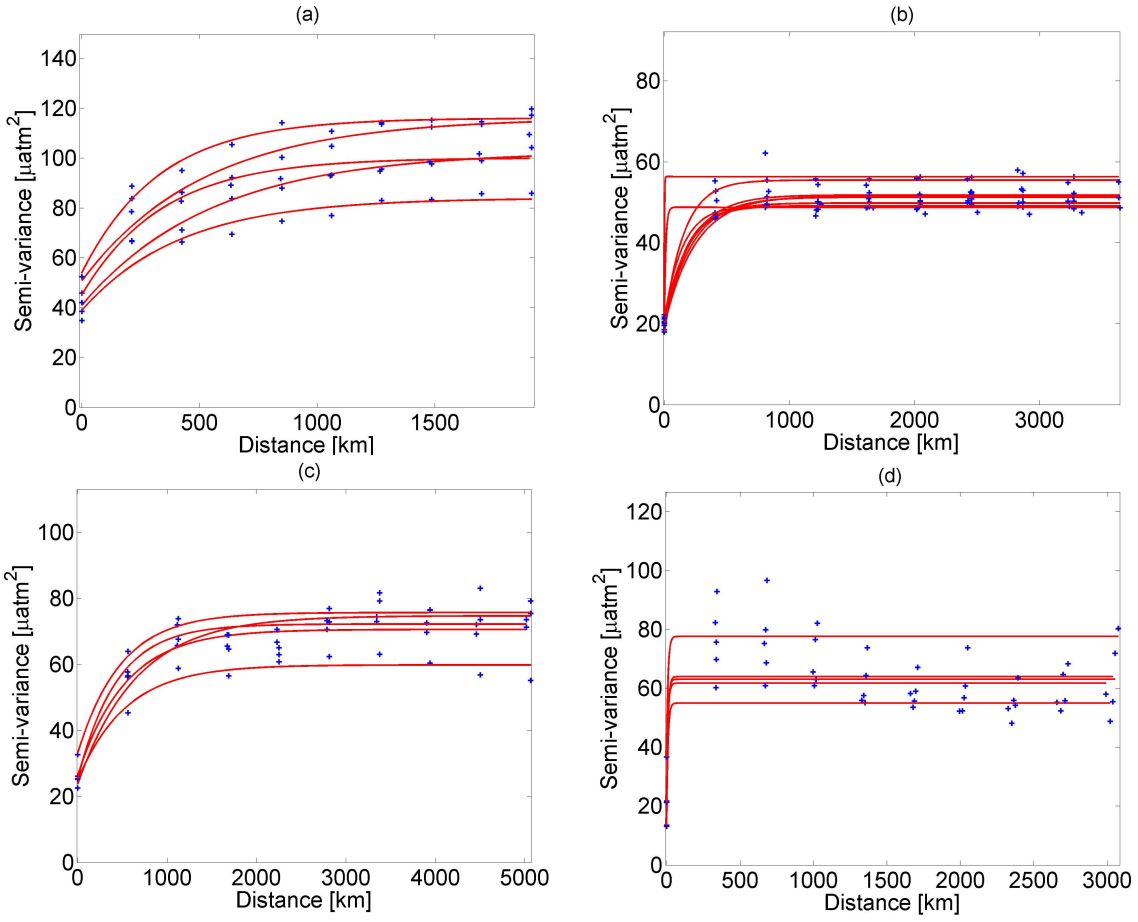

Fig. A1. Empirical semi-variograms of randomly chosen ensembles of the residuals as a function of distance in **(a)** the subpolar North Atlantic with a median decorrelation length of 430 km, **(b)** subtropical North Atlantic with a median decorrelation length of 166 km, **(c)** the equatorial Atlantic with a median decorrelation length of 530 km and **(d)** the South Atlantic with a median decorrelation length of 9 km. Region borders are listed in Table 3.

A3 Air–sea CO_2 flux calculation

We calculate the air–sea flux density in mol C m^{-2} yr^{-1} for each month and 1° × 1° pixel from

$$F_{CO_2} = -k_w \cdot S_{CO_2} \cdot (1 - f_{ice}) \cdot (pCO_{2,atm,wet} - pCO_2), \quad \text{(A6)}$$

where S_{CO_2} is the mainly temperature driven solubility of CO_2 (calculated in mol C m^{-3} µatm^{-1}) and k_w is the gas transfer velocity (calculated in m yr^{-1}) and where the flux is defined positive upward, i.e. outgassing is positive, and uptake is negative. f_{ice} refers to the percent of ice cover within a region derived from Rayner et al. (2003). For the gas transfer velocity (here calculated in cm h^{-1}) we decided to use the formulation of Wanninkhof (1992) with the scaling factor of Sweeney et al. (2007), i.e.

$$k_w = 0.27 \cdot (Sc/660)^{-\frac{1}{2}} \cdot u^2, \quad \text{(A7)}$$

where Sc the dimensionless Schmidt number and u the monthly mean CCMP wind speed (Atlas et al., 2011) at a height of 10 m above the sea surface.

The solubility of CO_2 is calculated according to Weiss (1974) and the Schmidt number according to Wanninkhof (1992) using the same SST and SSS data we used for the training of our network.

The partial pressure of atmospheric CO_2, i.e. $pCO_{2,atm,wet}$, was computed from the dry air mixing ratio xCO_2 of GLOBALVIEW-CO2 (2011), taking into account the water vapour correction according to Dickson et al. (2007):

$$pCO_{2,atm,wet} = xCO_{2,atm} \cdot [P_{atm,surf} - P_{H_2O}], \quad \text{(A8)}$$

where $P_{atm,surf}$ is the sea-level pressure from NCEP (Kalnay et al., 1996), and P_{H_2O} describes the water vapour pressure.

A4 Accounting for spatial autocorrelation in flux uncertainty analysis

For each RECCAP/ocean inversion region (see Table 3 for region borders), we first divide the residuals into 5 randomly chosen mutually exclusive ensembles, with the exception of the subtropical North Atlantic, where we use 10 ensembles, due to the larger amount of data. For each ensemble, we compute the semi-variance of the residuals and their point-to-point Haversine distance matrix, and then fit an exponential function of the form

$$a + b \cdot \exp\left(\frac{-x}{c}\right) \tag{A9}$$

to the semi-variogram in order to estimate the correlation length (parameter c) between the residuals. We find that the semi-variograms are very sensitive to extreme values of the residuals, forcing us to use Chauvenet's criterion to reject them prior to the computation and the fit. By using several different ensembles per region, we will account for the potential biasing effect of their removal.

Figure A1 shows the semi-variograms of all ensembles in the Atlantic Ocean. Correlation lengths of the residuals vary between 9 km, where the ensembles are well below the distance between 2 neighbouring grid boxes, and 532 km. However, in all cases the semi-variogram shows a large lag 0 correlation, (semi-variance at 0 distance varies between 20 and 60 μatm^2 within the different ensembles in the different regions) indicating the residuals within one grid cell are correlated with each other, leading to a substantial reduction of the degrees of freedom.

Acknowledgements. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7 2007-2013) under grant agreement no. 238366. Support for N. Gruber, U. Schuster and D. C. E. Bakker has been provided by EU grant 264879 (CARBOCHANGE). Support for N. Gruber and U. Schuster has been provided by EU grant 283080 (GEO-CARBON). We thank the captains, officers, and crew of all ships on which measurements that contributed to this study were made. We thank all scientists and institutes involved in creating the SOCAT database.

Edited by: L. Bopp

References

Amari, S., Murata, N., Müller, K.-R., Finke, M., and Yang, H. H.: Asymptotic Statistical Theory of Overtraining and Cross-Validation, IEEE T. Neural Networ., 8, 985–996, 1997.

Antonov, J. I., Seidov, D., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., and Garcia, H. E.: World Ocean Atlas 2009, Volume 2: Salinity, NOAA Atlas NESDIS 69, U.S. Government Printing Office, Washington, DC, 2010.

Atlas, R., Hoffman, R. N., Ardizzone, J., Leidner, S. M., Jusem, J. C., Smith, D. K., and Gombos, D.: A cross-calibrated multi-platform ocean surface wind velocity product for meteorological and oceanographic applications, B. Am. Meteorol. Soc., 92, 157–174, 2011.

Bakker, D. C. E., Pfeil, B., Smith, K., Hankin, S., Olsen, A., Alin, S. R., Cosca, C., Harasawa, S., Kozyr, A., Nojiri, Y., O'Brien, K. M., Schuster, U., Telszewski, M., Tilbrook, B., Wada, C., Akl, J., Barbero, L., Bates, N., Boutin, J., Cai, W.-J., Castle, R. D., Chavez, F. P., Chen, L., Chierici, M., Currie, K., de Baar, H. J. W., Evans, W., Feely, R. A., Fransson, A., Gao, Z., Hales, B., Hardman-Mountford, N., Hoppema, M., Huang, W.-J., Hunt, C. W., Huss, B., Ichikawa, T., Johannessen, T., Jones, E. M., Jones, S. D., Jutterström, S., Kitidis, V., Körtzinger, A., Landschtzer, P., Lauvset, S. K., Lefèvre, N., Manke, A. B., Mathis, J. T., Merlivat, L., Metzl, N., Murata, A., Newberger, T., Ono, T., Park, G.-H., Paterson, K., Pierrot, D., Ríos, A. F., Sabine, C. L., Saito, S., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Sieger, R., Skjelvan, I., Steinhoff, T., Sullivan, K., Sun, H., Sutton, A. J., Suzuki, T., Sweeney, C., Takahashi, T., Tjiputra, J., Tsurushima, N., van Heuven, S. M. A. C., Vandemark, D., Vlahos, P., Wallace, D. W. R., Wanninkhof, R., and Watson, A. J.: An update to the Surface Ocean CO_2 Atlas (SOCAT version 2), Earth Syst. Sci. Data Discuss., 6, 465–512, doi:10.5194/essdd-6-465-2013, 2013.

Bates, N. R.: Interannual variability of the oceanic CO_2 sink in the subtropical gyre of the North Atlantic Ocean over the last 2 decades, J. Geophys. Res., 112, C09013, doi:10.1029/2006JC003759, 2007.

Bennington, V., McKinley, G. A., Dutkiewicz, S., and Ullman, D.: What does chlorophyll variability tell us about export and air-sea CO_2 flux variability in the North Atlantic?, Global Biogeochem. Cy., 23, GB3002, doi:10.1029/2008GB003241, 2009.

Carton, J. A. and Giese, B. S.: A Reanalysis of Ocean Climate Using Simple Ocean Data Assimilation (SODA), Mon. Weather Rev., 136, 2999–3017, doi:10.1175/2007MWR1978.1, 2008.

Chierici, M., Olsen, A., Johannessen, T., Trinañes, J., and Wanninkhof, R.: Algorithms to estimate the carbon dioxide uptake in the northern North Atlantic using ship board observations, satellite and ocean analysis data, Deep-Sea Res. Pt. II, 56, 630–639, 2009.

Corbière, A., Metzl, N., Reverdin, G., Brunet, C., and Takahashi, T.: Interannual and decadal variability of the oceanic carbon sink in the North Atlantic subpolar gyre, Tellus, 59B, 168–178, 2007.

de Boyer Montegut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, J. Geophys. Res., 109, C12003, doi:10.1029/2004JC002378, 2004.

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

Dickson, A. G., Sabine, C. L., and Christian, J. R., (Eds.): Guide to Best Practices for Ocean CO_2 Measurements, PICES Special Publication, IOCCP Report No . 8, 2007.

Doney, S., Fabry, V., Feely, R. A., and Kleypas, J.: Ocean Acidification: The Other CO_2 Problem, Annu. Rev. Mar. Sci., 1, 169–192, doi:10.1146/annurev.marine.010908.163834, 2009.

Fay, A. R. and McKinley, G. A.: Global trends in surface ocean pCO_2 from in situ data, Global Biogeochem. Cy., 27, 1–17, doi:10.1002/gbc.20051, 2013.

Friedrich, T. and Oschlies, A.: Neural network-based estimates of North Atlantic surface pCO_2 from satellite data:

A methodological study, J. Geophys. Res., 114, C03020, doi:10.1029/2007JC004646, 2009.

GLOBALVIEW-CO2: Cooperative Atmospheric Data Integration Project – Carbon Dioxide, CD-ROM, NOAA ESRL, Boulder, Colorado [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW], 2011.

González-Dávila, M., Santana-Casiano, J. M., and González-Dávila, E. F.: Interannual variability of the upper ocean carbon cycle in the Northeast Atlantic Ocean, Geophys. Res. Lett., 34, L07608, doi:10.1029/2006GL028145, 2007.

Graven, H. D., Guilderson, T., and Keeling, R.: Observations of radiocarbon in CO_2 at seven global sampling sites in the Scripps flask network: Analysis of spatial gradients and seasonal cycles, J. Geophys. Res., 117, D02303, doi:10.1029/2011JD016533, 2012.

Gruber, N.: Fickle trends in the ocean, Nature, 458, 155–156, 2009.

Gruber, N., Keeling, C. D., and Bates, N. R.: Interannual Variability in the North Atlantic Ocean Carbon Sink, Science, 298, 2374–2378, doi:10.1126/science.1077077, 2002.

Gruber, N., Gloor, M., Mikaloff Fletcher, S. E., Doney, S. C., Dutkiewicz, S., Follows, M. J., Gerber, M., Jacobson, A. R., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A., Sarmiento, J. L., and Takahashi, T.: Oceanic sources, sinks, and transport of atmospheric CO_2, Global Biogeochem. Cy., 23, GB1005, doi:10.1029/2008GB003349, 2009.

Gurney, K., Baker, D., Rayner, P., and Denning, S.: Interannual variations in continental-scale net carbon exchange and sensitivity to observing networks estimated from atmospheric CO_2 inversions for the period 1980 to 2005, Global Biogeochem. Cy., 22, GB3025, doi:10.1029/2007GB003082, 2008.

Hagan, M., Demuth, H. B., and Beale, M. H.: Neural Network Design, PWS Publishing, Boston, MA, 1996.

Hagan, M. T. and Menhaj, M.: Training feed-forward networks with the Marquardt algorithm, IEEE T. Neural Networ., 5, 989–993, 1994.

Hurrell, J. W.: Decadal trends in the North-Atlantic Oscillation: regional temperatures and precipitation, Science, 269, 676–679, 1995.

Jones, S. D., Le Quéré, C., and Rödenbeck, C.: Autocorrelation characteristics of surface ocean pCO_2 and air-sea CO_2 fluxes, Global Biogeochem. Cy., 26, GB2042, doi:10.1029/2010GB004017, 2012.

Kalkhan, M. A.: Spatial Statistics: GeoSpatial Information Modelling and Thematic Mapping, Taylor & Francis Group, 2011.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, B. Am. Meteorol. Soc., 77, 437–470, 1996.

Keller, K., Joos, F., Raible, C. C., Cocco, V., Frölicher, T., Dunne, J. P., Gehlen, M., Bopp, L., Orr, J. C., Tjiputra, J., Heinze, C., Segscheider, J., Roy, T., and Metzl, N.: Variability of the ocean carbon cycle in response to the North Atlantic Oscillation, Tellus, 64, 18738, doi:10.3402/tellusb.v64i0.18738, 2012.

Kohonen, T.: Self-Organization and Associative Memory, Springer-Verlag, Berlin, 2nd Edn., 1987.

Kohonen, T.: Self-Organizing Maps, Springer-Verlag, Berlin Heidelberg New York, 3rd Edn., 2001.

Körtzinger, A.: Methods of Seawater Analysis, chap. Determination of carbon dioxide partial pressure (pCO_2), 149–158, Verlag Chemie, 1999.

Le Quéré, C., Rödenbeck, C., Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A., Labuschagne, C., Ramonet, M., Nakazawa, T., Metzl, N., Gillett, N., and Heimann, M.: Saturation of the Southern Ocean CO_2 sink due to recent climate change, Science, 316, 1735–1738, doi:10.1126/science.1136188, 2007.

Lefèvre, N., Watson, A. J., Olsen, A., Rios, A. F., Pérez, F. F., and Johannessen, T.: A decrease in the sink for atmospheric CO_2 in the North Atlantic, Geophys. Res. Lett., 31, L07306, doi:10.1029/2003GL018957, 2004.

Lefèvre, N., Watson, A. J., and Watson, A. R.: A comparison of multiple regression and neural network techniques for mapping in situ pCO_2 data, Tellus, 57B, 375–384, 2005.

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., Mishonov, A. V., Reagan, J. R., Seidov, D., Yarosh, E. S., and Zweng, M. M.: World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010, Geophys. Res. Lett., 39, L10603, doi:10.1029/2012GL051106, 2012.

Lüger, H., Wanninkhof, R., Wallace, D. W. R., and Körtzinger, A.: CO_2 fluxes in the subtropical and subarctic North Atlantic based on measurements from a volunteer observing ship, J. Geophys. Res., 111, C06024, doi:10.1029/2005JC003101, 2006.

Marquardt, D. W.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters, J. Soc. Ind. Appl. Math., II, 431–441, 1963.

Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., Hurrel, J., McCartney, M., Saravanan, R., and Visbeck, M.: North Atlantic Climate Variability: Phenomena, Impacts and Mechanisms, Int. J. Climatol., 21, 1863–1898, doi:10.1002/joc.693, 2001.

McKinley, G. A., Fay, A. R., Takahashi, T., and Metzl, N.: Convergence of atmospheric and North Atlantic carbon dioxide trends on multidecadal timescales, Nat. Geosci., 4, 606–610, doi:10.1038/NGEO1193, 2011.

Menemenlis, D., Campin, J., Heimbach, P., Hill, C., Lee, T., Nguyen, A., Schodlok, M., and Zhang, H.: ECCO2: High resolution global ocean and sea ice data synthesis, Mercator Ocean Quarterly Newsletter, 31, 13–21, 2008.

Metzl, N., Corbière, A., Reverdin, G., Lenton, A., Takahashi, T., Olsen, A., Johannessen, T., Pierrot, D., Wanninkhof, R., Ólafsdóttir, S. R., Olafsson, J., and Ramonet, M.: Recent acceleration of the sea surface fCO_2 growth rate in the North Atlantic subpolar gyre (1993–2008) revealed by winter observations, Global Biogeochem. Cy., 24, GB4004, doi:10.1029/2009gb003658, 2010.

Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., Follows, M., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A., and Sarmiento, J. L.: Inverse estimates of anthropogenic CO_2 uptake, transport, and storage by the ocean, Global Biogeochem. Cy., 20, GB2002, doi:10.1029/2005GB002530, 2006.

Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Gloor, M., Doney, S. C., Dutkiewicz, S., Gerber, M., Follows, M., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A., and Sarmiento, J.: Inverse estimates of the oceanic sources and sinks of natural CO_2 and the implied oceanic carbon transport, Global

Biogeochem. Cy., 21, GB1010, doi:10.1029/2006GB002751, 2007.

Olsen, A., Omar, A. M., Bellerby, R. G. J., Johannessen, T., Ninnemann, U. S., Brown, K. R., Olsson, K. A., Olafsson, J., Nondal, G., Kivimäe, C., Kringstad, S., Neill, C., and Olafsdottir, S.: Magnitude and origin of the anthropogenic CO_2 increase and ^{13}C Suess effect in the Nordic seas since 1981, Global Biogeochem. Cy., 20, GB3027, doi:10.1029/2005GB002669, 2006.

Peng, T.-H. and Wanninkhof, R.: Increase in anthropogenic CO_2 in the Atlantic Ocean in the last two decades, Deep-Sea Res. Pt. I, 57, 755–770, 2010.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699–6720, doi:10.5194/bg-10-6699-2013, 2013.

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., 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., 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, 125–143, doi:10.5194/essd-5-125-2013, 2013.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108, 4407, doi:10.1029/2002JD002670, 2003.

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. Climate, 15, 1609–1625, 2002.

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T.-H., Kozyr, A., Ono, T., and Rios, A. F.: The oceanic sink for anthropogenic CO_2, Science (New York, N.Y.), 305, 367–71, doi:10.1126/science.1097403, 2004.

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., Sarma, V. V. S. 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, Earth Syst. Sci. Data, 5, 145–153, doi:10.5194/essd-5-145-2013, 2013.

Sarmiento, J. and Gruber, N.: Ocean Biogeochemical Dynamics, Princeton University Press, 2006.

Sasse, T. P., McNeil, B. I., and Abramowitz, G.: A new constraint on global air-sea CO_2 fluxes using bottle carbon data, Geophys. Res. Lett., 40, 1594–1599, doi:10.1002/grl.50342, 2013.

Schuster, U. and Watson, A. J.: A variable and decreasing sink for atrmospheric CO_2 in the North Atlantic, J. Geophys. Res., 112, C11006, doi:10.1029/2006JC003941, 2007.

Schuster, U., Watson, A. J., Bates, N. R., Corbière, A., González-Dávila, M., Metzl, N., Pierrot, D., and Santana-Casiano, M.: Trends in North Atlantic sea-surface fCO_2 from 1990 to 2006, Deep-Sea Res. Pt. II, 56, 620–629, 2009.

Schuster, U., McKinley, G. A., Bates, N., Chevallier, F., Doney, S. C., Fay, A. R., González-Dávila, M., Gruber, N., Jones, S., Krijnen, J., Landschützer, P., Lefèvre, N., Manizza, M., Mathis, J., Metzl, N., Olsen, A., Rios, A. F., Rödenbeck, C., Santana-Casiano, J. M., Takahashi, T., Wanninkhof, R., and Watson, A. J.: An assessment of the Atlantic and Arctic sea–air CO_2 fluxes, 1990–2009, Biogeosciences, 10, 607–627, doi:10.5194/bg-10-607-2013, 2013.

Stendardo, I. and Gruber, N.: Oxygen trends over five decades in the North Atlantic, J. Geophys. Res., 117, C11004, doi:10.1029/2012JC007909, 2012.

Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G., Sarmiento, J. L., and Wanninkhof, R.: Constraining global air-sea gas exchange for CO_2 with recent bomb ^{14}C measurements, Global Biogeochem. Cy., 21, GB2015, doi:10.1029/2006GB002784, 2007.

Takahashi, T., Olafsson, J., Goddard, J., Chipman, D., and Sutherland, S.: Seasonal variation of CO_2 and nutrients in the high-latitude surface oceans: A comparative study, Global Biogeochem. Cy., 7, 843–878, doi:10.1029/93GB02263, 1993.

Takahashi, T., Wanninkhof, R., Feely, R., Weiss, R., Chipman, D., Bates, N., Olafsson, J., Sabine, C., and Sutherland, S.: Net sea-air CO_2 flux over the global oceans: An improved estimate based on the sea-air pCO_2 difference, in: Proceedings of the Second-Internat ional Symposium, CO_2 in the Oceans, edited by: Nojiri, Y., Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan, 9–14, 1999.

Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., Bates, N., Wanninkhof, R., Feely, R. F., Sabine, C., Olafsson, J., and Nojiri, Y.: Global sea-air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, Deep-Sea Res. Pt. II, 49, 1601–1622, 2002.

Takahashi, T., Weiss, R., Wanninkhof, R., Chipman, D., and Feely, R.: Global air-sea flux of CO_2: an estimate based on measurements of sea-air pCO_2 difference, P. Natl. Acad. Sci., 94, 8292–8299, 2003.

Takahashi, T., Sutherland, S., Wanninkhof, R., Sweeney, C., Feely, R., Chipman, D., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafson, J., Arnarson, T., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C., Delille, B., Bates, N., and de Baar, H.: Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, Deep-Sea Res. Pt. II, 56, 554–577, 2009.

Telszewski, M., Chazottes, A., Schuster, U., Watson, A. J., Moulin, C., Bakker, D. C. E., González-Dávila, M., Johannessen, T., Körtzinger, A., Lüger, H., Olsen, A., Omar, A., Padin, X. A., Ríos, A. F., Steinhoff, T., Santana-Casiano, M., Wallace, D. W. R., and Wanninkhof, R.: Estimating the monthly pCO_2 distribution in the North Atlantic using a self-organizing neural network, Biogeosciences, 6, 1405–1421, doi:10.5194/bg-6-1405-2009, 2009.

Thomas, H., Prowe, F. A. E., Lima, I. D., Doney, S. C., Wanninkhof, R., Greatbatch, R. J., Schuster, U., and Corbière, A.: Changes in the North Atlantic Oscillation influence CO_2 uptake in the North Atlantic over the past 2 decades, Global Biogeochem. Cy., 22, GB4027, doi:10.1029/2007GB003167, 2008.

Ullman, D. J., McKinley, G. A., Bennington, V., and Dutkiewicz, S.: Trends in the North Atlantic carbon sink: 1992–2006, Global Biogeochem. Cy., 23, GB4011, doi:10.1029/2008GB003383, 2009.

Vesanto, J., Himberg, J., Alhoniemi, E., and Parhankangas, J.: SOM Toolbox for Matlab 5, Tech. rep., Helsinki University of Technology, 2000.

Wanninkhof, R.: Relation between wind speed and gas exchange over the ocean, J. Geophys. Res., 97, 7373–7383, 1992.

Wanninkhof, R., Doney, S., Bullister, J. L., Levine, N. M., Warner, M., and Gruber, N.: Detecting anthropogenic CO_2 changes in the interior Atlantic Ocean between 1989 and 2005, J. Geophys. Res., 115, C11028, doi:10.1029/2010JC006251, 2010.

Weiss, R. F.: Carbon Dioxide in Water and Seawater: The Solubility of a Non-Ideal Gas, Mar. Chem., 2, 203–215, 1974.