



HAL
open science

LSCE-FFNN-v1: the reconstruction of surface ocean pCO₂

Anna Denvil-Sommer, Marion Gehlen, Mathieu Vrac, Carlos Mejia

► **To cite this version:**

Anna Denvil-Sommer, Marion Gehlen, Mathieu Vrac, Carlos Mejia. LSCE-FFNN-v1: the reconstruction of surface ocean pCO₂. *Geoscientific Model Development*, 2019, 12 (5), pp.2091-2105. 10.5194/gmd-12-2091-2019 . hal-02171978

HAL Id: hal-02171978

<https://hal.sorbonne-universite.fr/hal-02171978>

Submitted on 3 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



LSCE-FFNN-v1: a two-step neural network model for the reconstruction of surface ocean $p\text{CO}_2$ over the global ocean

Anna Denvil-Sommer¹, Marion Gehlen¹, Mathieu Vrac¹, and Carlos Mejia²

¹Laboratoire des Sciences du Climat et de l'Environnement (LSCE), Institut Pierre Simon Laplace (IPSL), CNRS/CEA/UVSQ/Univ. Paris-Saclay, Orme des Merisiers, Gif Sur Yvette, 91191, France

²Sorbonne Université, CNRS, IRD, MNHN, Institut Pierre Simon Laplace (IPSL), Paris, 75005, France

Correspondence: Anna Denvil-Sommer (anna.sommer.lab@gmail.com)

Received: 28 September 2018 – Discussion started: 7 November 2018

Revised: 17 April 2019 – Accepted: 23 April 2019 – Published: 29 May 2019

Abstract. A new feed-forward neural network (FFNN) model is presented to reconstruct surface ocean partial pressure of carbon dioxide ($p\text{CO}_2$) over the global ocean. The model consists of two steps: (1) the reconstruction of $p\text{CO}_2$ climatology, and (2) the reconstruction of $p\text{CO}_2$ anomalies with respect to the climatology. For the first step, a gridded climatology was used as the target, along with sea surface salinity (SSS), sea surface temperature (SST), sea surface height (SSH), chlorophyll a (Chl a), mixed layer depth (MLD), as well as latitude and longitude as predictors. For the second step, data from the Surface Ocean CO_2 Atlas (SOCAT) provided the target. The same set of predictors was used during step (2) augmented by their anomalies. During each step, the FFNN model reconstructs the nonlinear relationships between $p\text{CO}_2$ and the ocean predictors. It provides monthly surface ocean $p\text{CO}_2$ distributions on a $1^\circ \times 1^\circ$ grid for the period from 2001 to 2016. Global ocean $p\text{CO}_2$ was reconstructed with satisfying accuracy compared with independent observational data from SOCAT. However, errors were larger in regions with poor data coverage (e.g., the Indian Ocean, the Southern Ocean and the subpolar Pacific). The model captured the strong interannual variability of surface ocean $p\text{CO}_2$ with reasonable skill over the equatorial Pacific associated with ENSO (the El Niño–Southern Oscillation). Our model was compared to three $p\text{CO}_2$ mapping methods that participated in the Surface Ocean $p\text{CO}_2$ Mapping intercomparison (SOCOM) initiative. We found a good agreement in seasonal and interannual variability between the models over the global ocean. However, important differences still exist at the regional scale, especially in the

Southern Hemisphere and, in particular, in the southern Pacific and the Indian Ocean, as these regions suffer from poor data coverage. Large regional uncertainties in reconstructed surface ocean $p\text{CO}_2$ and sea–air CO_2 fluxes have a strong influence on global estimates of CO_2 fluxes and trends.

1 Introduction

The global ocean is a major sink of excess CO_2 that has been emitted to the atmosphere since the beginning of the industrial revolution. In 2011, the best estimate of the ocean inventory of anthropogenic carbon (C_{ant}) amounted to $155 \pm 30 \text{ PgC}$ or 28 % of the cumulated total CO_2 emissions attributed to human activities since 1750 (Ciais et al., 2013). Between 2000 and 2009, the yearly average ocean C_{ant} uptake was $2.3 \pm 0.7 \text{ PgC yr}^{-1}$ (Ciais et al., 2013). However, these global estimates hide substantial regional and interannual fluctuations (Rödenbeck et al., 2015), which need to be quantified in order to track the evolution of the Earth's carbon budget (e.g., Le Quéré et al., 2018).

Until recently, most estimates of interannual sea–air CO_2 flux variability were based on atmospheric inversions (Peylin et al., 2005, 2013; Rödenbeck et al., 2005) or global ocean circulation models (Orr et al., 2001; Aumont and Bopp, 2006; Le Quéré et al., 2010). However, models tend to underestimate the variability of sea–air CO_2 fluxes (Le Quéré et al., 2003), whereas atmospheric inversions suffer from a sparse network of atmospheric CO_2 measurements (Peylin et

al., 2013). These approaches are increasingly complemented by data-based techniques relying on in situ measurements of CO_2 fugacity or partial pressure (e.g., Takahashi et al., 2002, 2009; Nakaoka et al., 2013; Schuster et al., 2013; Landschützer et al., 2013, 2016; Rödenbeck et al., 2014, 2015; Bitting et al., 2018; Fay et al., 2014). These techniques rely on a variety of data-interpolation approaches developed to provide estimates in time and space of surface ocean $p\text{CO}_2$ (Rödenbeck et al., 2015) such as statistical interpolation, linear and nonlinear regressions, or model-based regressions or tuning (Rödenbeck et al., 2014, 2015). These methods, and their advantages and disadvantages, are compared and discussed in Rödenbeck et al. (2015). This intercomparison did not allow for the identification of a single optimal technique but rather pleaded in favor of exploiting the ensemble of methods.

Artificial neural networks (ANNs) have been widely used to reconstruct surface ocean $p\text{CO}_2$; Lefèvre et al. (2005), Friedrich and Oschlies (2009b), Telszewski et al. (2009), Landschützer et al. (2013) Nakaoka et al. (2013), Zeng et al. (2014) and Bitting et al. (2018) used ANNs to study the open ocean, whereas Laruelle et al. (2017) studied the coastal region using this method. ANNs fill the spatial and temporal gaps based on calibrated nonlinear statistical relationships between $p\text{CO}_2$ and its oceanic and atmospheric drivers. The existing products usually present monthly fields with a $1^\circ \times 1^\circ$ spatial resolution and capture a large part of the temporal–spatial variability. Methods based on ANNs are able to represent the relationships between $p\text{CO}_2$ and a variety of predictor combinations, but they are sensitive to the number of data used in the training algorithm and can generate artificial variability in regions with sparse data coverage (Bishop, 2006).

This study proposes an alternative implementation of a neural network applied to the reconstruction of surface ocean $p\text{CO}_2$ over the period from 2001 to 2016. It belongs to the category of feed-forward neural networks (FFNN) and consists of a two-step approach: (1) the reconstruction of monthly climatologies of global surface ocean $p\text{CO}_2$ based on data from Takahashi et al. (2009), and (2) the reconstruction of monthly anomalies (with respect to the monthly climatologies) on a $1^\circ \times 1^\circ$ grid exploiting the Surface Ocean CO_2 Atlas (SOCAT) (Bakker et al., 2016). The model is easily applied to the global ocean without any boundaries between the ocean basins or regions. However, as previously mentioned, it is still sensitive to the observational coverage. This limitation is partly overcome by the two-step approach, as the reconstruction of monthly climatologies draws on a global ocean gridded climatology (Takahashi et al., 2009); therefore, the FFNN output is kept close to realistic values. Furthermore, the reconstruction of monthly climatologies during the first step allows for a potential change in seasonal cycle in response to climate change to be taken into account when applied to time slices or to model output providing the drivers, but no carbon cycle variables.

The remainder of this paper is structured as follows: Sect. 2 introduces datasets used during this study and describes the neural network; Sect. 3 presents results for its validation and qualification, as well as its comparison to three mapping methods as part of the Surface Ocean $p\text{CO}_2$ Mapping intercomparison (SOCOM) exercise (Rödenbeck et al., 2015). Finally, results and perspectives are summarized in the last section.

2 Data and methods

2.1 Data

The standard set of variables known to represent the physical, chemical and biological drivers of surface ocean $p\text{CO}_2$ – mean state and variability – (Takahashi et al., 2009; Landschützer et al., 2013) were used as input variables (or predictors) for training the FFNN algorithm. These are sea surface salinity (SSS), sea surface temperature (SST), mixed layer depth (MLD), chlorophyll a concentration (Chl a) and the atmospheric CO_2 mole fraction ($x\text{CO}_{2,\text{atm}}$). Based on Rodgers et al. (2009) who reported a strong correlation between natural variations in dissolved inorganic carbon (DIC) and sea surface height (SSH), SSH was added as a new driver to this list. Initial tests suggested that the inclusion of the SSH did not significantly improve the accuracy of reconstructed $p\text{CO}_2$ at the global scale. At the basin and regional scale, however, adding the SSH improved the spatial pattern of reconstructed $p\text{CO}_2$ and the accuracy of our method.

For the first step, the reconstruction of monthly climatologies, the Takahashi et al. (2009) monthly $p\text{CO}_2$ gridded climatology ($1^\circ \times 1^\circ$) was used as the target. The original climatology was constructed by an advection-based interpolation method on a $4^\circ \times 5^\circ$ grid. It was interpolated on the $1^\circ \times 1^\circ$ SOCAT grid, which is also the resolution of the final output for the FFNN.

For the second step, the target was provided by the SOCAT v5 observational database (Bakker et al., 2016). We used a gridded version of this dataset that was derived by combining all SOCAT data collected within a $1^\circ \times 1^\circ$ box during a specific month. SOCAT v5 represents global observations of the sea surface fugacity of CO_2 ($f\text{CO}_2$) over the period from 1970 to 2016. It includes data from moorings, ships and drifters. These data are irregularly distributed over the global ocean with 188 274 gridded measurements over the Northern Hemisphere and 76 065 over the Southern Hemisphere. In order to ensure a satisfying spatial and temporal data coverage, we limited the reconstruction to the period from 2001 to 2016, which represents $\sim 77\%$ of the database (Fig. 1a).

The following formula is used to convert $f\text{CO}_2$ to $p\text{CO}_2$ (Körtzinger, 1999):

$$f\text{CO}_2 = p\text{CO}_2 \exp\left(p \frac{B + 2\delta}{RT}\right), \quad (1)$$

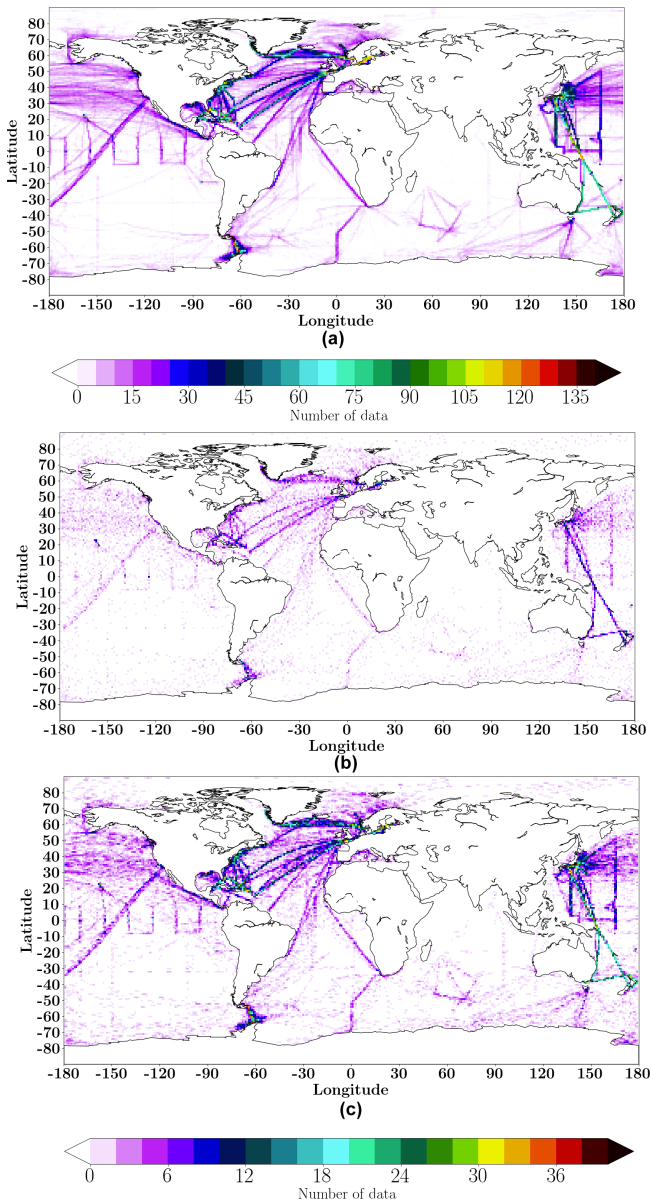


Figure 1. Spatial distribution of SOCAT data (number of measurements per grid point): (a) for the 2001–2016 period; (b) for all months of January for the 2001–2016 period; (c) for all months of December–January–February for the 2001–2016 period. Please note that the color bar under panel (c) refers to both (b) and (c).

where $f\text{CO}_2$ and $p\text{CO}_2$ are in micro-atmospheres (μatm), p is the total pressure (Pa), $R = 8.314 \text{ J K}^{-1}$ is the gas constant and T is the absolute temperature (K). Parameter B ($\text{m}^3 \text{ mol}^{-1}$) is estimated as $B = (-1636.75 + 12.040T - 3.27957 \times 10^{-2}T^2 + 3.16528 \times 10^{-5}T^3) \times 10^{-6}$. The parameter δ is the cross-virial coefficient ($\text{m}^3 \text{ mol}^{-1}$): $\delta = (57.7 - 0.118T) \times 10^{-6}$. The total pressure is from the Jena database (6 h, $5^\circ \times 5^\circ$; <http://www.bgc-jena.mpg.de/CarboScope/?ID=s>, last access: 13 September 2017) based

on the National Center for Environmental Prediction (NCEP) reanalysis (Kalnay et al., 1996).

Monthly global reprocessed products of physical variables from ARMOR3D L4 distributed through the Copernicus Marine Environment Monitoring Service (CMEMS; $0.25^\circ \times 0.25^\circ$; http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=MULTIOBS_GLO_PHY_REP_015_002, last access: 27 October 2017) were used for SSS, SST and SSH (Guinehut et al., 2012). The GlobColour project provided monthly chlorophyll (GlobColour “CHL” data) distributions at a $1^\circ \times 1^\circ$ resolution (http://www.globcolour.info/products_description.html, last access: 31 October 2017). For MLD, daily data from the “Estimating the Circulation and Climate of the Ocean” (ECCO2) project Phase II (Cube 92), at a $0.25^\circ \times 0.25^\circ$ resolution (Menemenlis et al., 2008) were used. For atmospheric $x\text{CO}_2$, the 6h data from Jena CO_2 inversion s76_v4.1 on a $5^\circ \times 5^\circ$ grid were selected (<http://www.bgc-jena.mpg.de/CarboScope/?ID=s>, last access: 13 September 2017). Finally, an ice mask based on daily “Operational Sea Surface Temperature and Sea Ice Analysis” (OSTIA) data with a gridded $0.05^\circ \times 0.05^\circ$ resolution (Donlon et al., 2011) was applied.

MLD and CHL data were log-transformed before their use in the FFNN algorithm because of their skewed distribution. In regions with no CHL data (high latitudes in winter) $\log(\text{CHL}) = 0$ was applied. It does not introduce discontinuities, as $\log(\text{CHL})$ is close to zero in the adjacent region.

All data were averaged or interpolated on a $1^\circ \times 1^\circ$ grid and, depending on the resolution of the dataset, averaged over the month. It is worth noting that all datasets have to be normalized (i.e., centered to zero-mean and reduced to unit standard deviation) before their use in the FFNN algorithm, for example,

$$\text{SSS}_n = \frac{\text{SSS} - \overline{\text{SSS}}}{\text{SD}(\text{SSS})}.$$

Normalization ensures that all predictors fall within a comparable range and avoids giving more weight to predictors with large variability ranges (Kallache et al., 2011).

As surface ocean $p\text{CO}_2$ also varies spatially, geographical positions, latitude (lat) and longitude (long), after conversion to radians were included as predictors. In order to normalize (lat, long) the following transformation is proposed:

$$\text{lat}_n = \sin(\text{lat} \cdot \pi/180^\circ)$$

$$\text{long}_{n,1} = \sin(\text{long} \cdot \pi/180^\circ)$$

$$\text{long}_{n,2} = \cos(\text{long} \cdot \pi/180^\circ)$$

Two functions *sin* and *cos* for longitudes are used to preserve its periodical 0 to 360° behavior and thus to consider the difference of positions before and after the 0° longitude. For step 2, data required for training were collocated at the SO-

CAT data positions that are used as a target for the FFNN model. Details are provided in the next section.

2.2 Methods

2.2.1 Network configuration and evaluation protocol

In this work, we use Keras, a high-level neural network Python library (“Keras: the Python Deep Learning library”, Chollet, 2015; <https://keras.io>) to build and train the FFNN models. The identification of an optimal configuration is the first step in the FFNN model building. This includes the choice of the number and size of hidden layers (i.e., intermediate layers between input and output layers), the connection type, the activation functions, the loss function and optimization algorithm, as well as the learning rate and other low-level parameters. Based on a series of tests and their statistical results (RMSE, correlation and bias) a hyperbolic tangent was chosen as an activation function for neurons in hidden layers, and a linear function was chosen for the output layer. As an optimization algorithm, the mini-batch gradient descent or “RMSprop” was used (adaptive learning rates for each weight, Chollet, 2015; Hinton et al., 2012). The number of layers and neurons utilized depends on the problem. For totally connected layers (i.e., a neuron in a hidden layer is connected to all neurons in the precedent layer and connects all neurons in the next one), which is the case here, it is enough to have only one single hidden layer, although two or more can help the approximation of complex functions (or complex relationships between the input and the output of the problem).

The number of FFNN layers and the number of neurons depends on the complexity of the problem: the more layers and neurons, the better the accuracy of the output. However, the size also depends on the number of patterns (data) used for training. The empirical rule recommends having a factor of 10 between the number of patterns (data) and the number of connections, or weights to adjust (in line with Amari et al., 1997, we use a factor of 10 that necessitates a cross-validation to avoid overfitting). This limits the size, the number of parameters and incidentally the number of neurons of the FFNN. This empirical rule was followed in this study.

1. *Step 1: reconstruction of monthly climatologies.* FFNN reconstructs a normalized monthly surface ocean $p\text{CO}_2$ climatology as a nonlinear function of normalized SSS, SST, SSH, Chl a , MLD climatologies and geographical position (lat and long):

$$p\text{CO}_{2,n} = (\text{SSS}_n, \text{SST}_n, \text{SSH}_n, \text{Chl}_n, \text{MLD}_n, \text{long}_n, \text{lat}_n). \quad (2)$$

Surface ocean $p\text{CO}_2$ from Takahashi et al. (2009) provided the target. The dataset was divided into 50 % for FFNN training and 25 % for its evaluation. This 25 % did not participate in the training. This set is used to

monitor the performance of the training process and to drive its convergence. The remaining 25 % (each fourth point) of the dataset was used after training for the FFNN model validation. More details about the FFNN training process can be found in Rumelhart et al. (1986) and Bishop (1995). Validation and evaluation datasets were chosen quasi-regularly in space and time to take all regions and seasonal variability into account. In order to improve the accuracy of the reconstruction, the model was applied separately for each month. We have developed a FFNN model with five layers (three hidden layers). Twelve models with a common architecture were trained. Tests with one model for 12 months showed a slight decrease in accuracy (not presented here). About 17 500 data were available for each month to train the model, resulting in monthly FFNN models with about 1856 parameters.

2. *Step 2: reconstruction of anomalies.* During the second step, normalized $p\text{CO}_2$ anomalies were reconstructed as a nonlinear function of normalized SSS, SST, SSH, Chl a , MLD, $x\text{CO}_2$ and their anomalies, as well as geographic position:

$$p\text{CO}_{2,\text{anom},n} = (\text{SSS}_n, \text{SST}_n, \text{SSH}_n, \text{Chl}_n, \text{MLD}_n, x\text{CO}_{2,n}, \text{SSS}_{\text{anom},n}, \text{SST}_{\text{anom},n}, \text{SSH}_{\text{anom},n}, \text{Chl}_{\text{anom},n}, \text{MLD}_{\text{anom},n}, x\text{CO}_{2,\text{anom},n}, \text{long}_{n,1}, \text{long}_{n,2}, \text{lat}_n) \quad (3)$$

Surface ocean $p\text{CO}_2$ anomalies were computed as the differences between colocated $p\text{CO}_2$ values based on SOCAT observations and monthly $p\text{CO}_2$ climatologies reconstructed during the first step that provided the targets:

$$p\text{CO}_{2,\text{anom}} = p\text{CO}_{2,\text{SOCAT}} - p\text{CO}_{2,\text{clim,FFNN}}. \quad (4)$$

The set of target data was again divided into 50 % for the training algorithm, 25 % for evaluation and 25 % for model validation. As in step (1) the model was trained separately for each climatological month. Thus, there were 12 models sharing a common architecture but trained on different data. In this step, in order to increase the amount of data during training and introduce information on the seasonal cycle, the model was trained using $p\text{CO}_2$ data from the month in question as a target as well as data from the previous and following month during the entire period from 2001 to 2016. Figure 1b and c show an example of the data distribution for the months of January over the period from 2001 to 2016 (Fig. 1b) and for the 3-month time window including December–January–February from 2001 to 2016 used in the training algorithm of the January FFNN model (Fig. 1c). In this particular example, the choice of 3 months provided a better cover of the region and doubled the number of data at high latitudes.

K -fold cross-validation was used for the evaluation and the validation of the FFNN architecture. Cross-validation relied on $K = 4$ different subsamples of the dataset to draw 25 % of independent data for validation (Fig. S1 in the Supplement). Each sampling fold was tested on five runs of the FFNN for each month. Each of these five runs was characterized by different initial values that were randomly chosen. From these five results, the best was chosen based on root-mean-square error (RMSE), the r^2 and the bias.

The final model architecture at step 2 had three layers (one hidden layer). About 10 000 samples were available for training for each month; thus, a model with 541 parameters was developed. Note that a higher number of parameters did not show a significant improvement in the accuracy.

2.2.2 Reconstruction of surface ocean $p\text{CO}_2$

The previous section presented the development of the “optimal” architecture of a FFNN model for the reconstruction of global surface ocean $p\text{CO}_2$ and the estimation of its accuracy. This FFNN model was used to provide the final product for scientific analysis and comparison with other mapping approaches. In order to provide the final output, the selected FFNN architecture was trained on all available data: 100 % of data for training, 100 % for evaluation and 100 % for validation. The network was executed five times (different initial values) and the best model was selected based on validation results considering the root-mean-square error (RMSE), the r^2 and the bias computed between the network output and the SOCAT-derived surface ocean $p\text{CO}_2$ data. The final model output is referred to as the LSCE-FFNN product.

2.3 Computation of sea–air CO_2 fluxes

The sea–air CO_2 flux, f , was calculated following Rödenbeck et al. (2015) as

$$f = k\rho L (p\text{CO}_2 - p\text{CO}_2^{\text{atm}}), \quad (5)$$

where k is the piston velocity estimated according to Wanninkhof (1992):

$$k = \Gamma u^2 \left(Sc^{\text{CO}_2} / Sc^{\text{Ref}} \right)^{-0.5}. \quad (6)$$

The global scaling factor Γ was chosen as in Rödenbeck et al. (2014) with the global mean CO_2 piston velocity equaling 16.5 cm h^{-1} . Sc corresponds to the Schmidt number estimated according to Wanninkhof (1992). The wind speed was computed from 6-hourly NCEP wind speed data (Kalnay et al., 1996). ρ is seawater density in Eq. (5) and L is the temperature-dependent solubility (Weiss, 1974). $p\text{CO}_2$ corresponds to the surface ocean $p\text{CO}_2$ output of the mapping method. $p\text{CO}_2^{\text{atm}}$ was derived from the atmospheric CO_2 mixing ratio fields provided by the Jena inversion s76_v4.1 (<http://www.bgc-jena.mpg.de/CarboScope/>).

3 Results

3.1 Validation

The subset of data used for network validation, which was 25 % of the total, represented independent observations, as they did not participate in training during model development (see Section 2.2.1). The skill of the FFNN to reconstruct monthly climatologies of surface ocean $p\text{CO}_2$ was assessed by comparing collocated reconstructed $p\text{CO}_2$ and corresponding values from Takahashi et al. (2009). The global climatology was reconstructed with satisfying accuracy during step 1 with a RMSE of $0.17 \mu\text{atm}$ and an r^2 of 0.93. The model output of step 2 was assessed using K -fold cross-validation as previously presented: $K = 4$ different subsets of independent data were drawn from the dataset and the network was run five times on each subset. From these 20 results the best one was chosen based on the RMSE, the r^2 and mean absolute error (MAE) (the bias is presented in Table S1 in the Supplement). A combination of the four best model output was used for the statistical analysis summarized in Table 1. Metrics were computed over the full period (2001–2016) and with reference to SOCAT observations (independent data only). At the global scale, the analysis yielded a RMSE of $\sim 17.97 \mu\text{atm}$, whereas the MAE was $11.52 \mu\text{atm}$ and the r^2 was 0.76. These results were comparable to those obtained by Landschützer et al. (2013) for the assessment of a surface ocean $p\text{CO}_2$ reconstruction based on an alternative neural network-based approach. The RMSE between the SOCAT data and the $p\text{CO}_2$ climatology from Takahashi et al. (2009) equalled $41.87 \mu\text{atm}$, which was larger than errors computed for the regional comparison between FFNN and SOCAT (Table 1). We also estimated the RMSE for the case where 100 % of the data were used for training, which equalled $14.8 \mu\text{atm}$ and confirmed the absence of overfitting.

Figure 2a shows the time mean difference between the estimated $p\text{CO}_2$ and $p\text{CO}_2$ from SOCAT v5 data used for validation: $\text{mean}_t (p\text{CO}_{2,i,j,\text{FFNN}} - p\text{CO}_{2,i,j,\text{SOCAT}})$. Large differences occurred at high latitudes, in equatorial regions, and along the Gulf Stream and Kuroshio currents – the regions with strong horizontal gradients of $p\text{CO}_2$. Moreover, the standard deviation of the residuals (Fig. 2b) in these regions was larger, indicating that the model fails to accurately reproduce the temporal variability. The reduced skill of the model in these regions reflects the poor data coverage along with a strong seasonal variability (e.g., the Southern Ocean) and/or high kinetic energy (e.g., the Southern Ocean and the Kuroshio and Gulf Stream currents) (Fig. 1a). At the scale of oceanic regions, (Table 1) the largest RMSE and MAE values were computed for the Pacific subpolar ocean (RMSE = $34.77 \mu\text{atm}$ and MAE = $23.12 \mu\text{atm}$), whereas the lowest correlation coefficient was obtained for the equatorial Atlantic Ocean ($r^2 = 0.57$). These low scores directly reflect low data density and are to be contrasted with those obtained over regions with better data

Table 1. Statistical validation of LSCE-FFNN. Comparison between reconstructed surface ocean $p\text{CO}_2$ and $p\text{CO}_2$ values from the SOCAT v5 database not used in the training algorithm for the period from 2001 to 2016 over the global ocean (except for regions with ice cover) and for large oceanographic regions. The number of measurements per region is given in parentheses in the first column.

Model	Latitude boundaries	RMSE (μatm)	r^2	MAE (μatm)
FFNN Global		17.97	0.76	11.52
Arctic (150)	76–90° N	22.05	0.54	17.1
Subpolar Atlantic (21903)	49–76° N	22.99	0.76	15.04
Subpolar Pacific (4529)	49–76° N	34.77	0.65	23.12
Subtropical Atlantic (41331)	18–49° N	17.28	0.69	11.27
Subtropical Pacific (41867)	18–49° N	15.86	0.77	9.9
Equatorial Atlantic (7300)	18° S–18° N	17.27	0.57	11.44
Equatorial Pacific (27092)	18° S–18° N	15.73	0.79	10.33
South Atlantic (3002)	44–18° S	17.81	0.63	12.28
South Pacific (12934)	44–18° S	13.52	0.63	9.36
Indian Ocean (2871)	44° S–30° N	17.25	0.62	11.6
Southern Ocean (16334)	90–44° S	17.4	0.58	11.92

coverage (e.g., the subtropical North Pacific with $\text{RMSE} = 15.86 \mu\text{atm}$, $\text{MAE} = 9.9 \mu\text{atm}$ and $r^2 = 0.77$, or the subpolar Atlantic with $\text{RMSE} = 22.99 \mu\text{atm}$, $\text{MAE} = 15.04 \mu\text{atm}$ and $r^2 = 0.76$). Despite large time mean differences computed over the eastern equatorial Pacific, scores are satisfying at the regional scale indicating error compensation by improved scores over the western basin ($\text{RMSE} = 15.73 \mu\text{atm}$, $\text{MAE} = 10.33 \mu\text{atm}$ and $r^2 = 0.79$). Scores are low in the Southern Hemisphere (Table 1) and time mean differences are large (Fig. 2a) reflecting sparse data coverage (Fig. 1a).

3.2 Qualification

This section presents the assessment of the final time series of reconstructed surface ocean $p\text{CO}_2$. The time series was computed using the best monthly models as described in Sect. 2.2, as well as 100 % of the data for learning, evaluation and validation.

Results of the LSCE-FFNN mapping model were compared to three published mapping methods which participated in the “Surface Ocean $p\text{CO}_2$ Mapping Intercomparison” (SOCOM) exercise presented in Rödenbeck et al. (2015) (<http://www.bgc-jena.mpg.de/SOCOM/>). These methods are as follows: (1) Jena-MLS oc_v1.5 (Rödenbeck et al., 2014), which is a statistical interpolation scheme (data-driven mixed-layer scheme; principal drivers used in parameterization: ocean-internal carbon sources/sinks, SST, wind speed, mixed-layer depth climatology and alkalinity climatology); (2) JMA-MLR (updated version up to 2016) (Iida et al., 2015), which is based on multi-linear regressions with SST, SSS and Chl a as independent variables; and (3) ETH-SOMFFN v2016 (Landschützer et al., 2014), which is a two-step neural network model with SST, SSS, MLD, Chl a and $x\text{CO}_2$ as drivers. The time series of $p\text{CO}_2$ and sea–air CO_2 flux (f) were assessed over 17 biomes defined by Fay and

McKinley (2014) (Fig. 3, Table 2). These biomes were derived based on coherence in SST, Chl a , the ice fraction, and the maximum MLD and represent regions of coherent biogeochemical dynamics.

We followed the protocol and diagnostics proposed in Rödenbeck et al. (2015) for the comparison of each of the respective mapping methods with one another, with respect to observations. The following diagnostics were computed: (1) the relative interannual variability (IAV) mismatch R^{iaiv} (in percent), and (2) the amplitude of interannual variations. The relative interannual variability (IAV) mismatch R^{iaiv} (in percent) is the ratio of the mismatch amplitude M^{iaiv} of the difference between the model output and the observations (its temporal standard deviation) and the mismatch amplitude $M^{\text{iaiv}}_{\text{benchmark}}$ of the “benchmark”. The latter was derived from the mean seasonal cycle of the corresponding model output where the trend of increasing yearly atmospheric $p\text{CO}_2$ was added (see details in Rödenbeck et al., 2015). It corresponds to a climatology corrected for increasing atmospheric CO_2 , but without interannual variability.

$$R^{\text{iaiv}} = \frac{M^{\text{iaiv}}}{M^{\text{iaiv}}_{\text{benchmark}}} \cdot 100\%, \quad (7)$$

where

$$M^{\text{iaiv}} = \text{SD}(\text{mean}(p\text{CO}_{2, \text{Model}} - p\text{CO}_{2, \text{SOCAT}})),$$

$$M^{\text{iaiv}}_{\text{benchmark}} = \text{SD}(\text{mean}(D_{\text{season}})).$$

Here “mean” is a mean over the region and year, and

$$D_{\text{season}} = (p\text{CO}_{2, \text{SS}} + \text{trend}(\text{CO}_{2, \text{atm}})) - p\text{CO}_{2, \text{SOCAT}},$$

$p\text{CO}_{2, \text{SS}}$ is the seasonal cycle of $p\text{CO}_2$ from the corresponding mapping method. $\text{CO}_{2, \text{atm}}$ estimates from $x\text{CO}_2$ Jena CO_2 inversion s76_v4.1 were used.

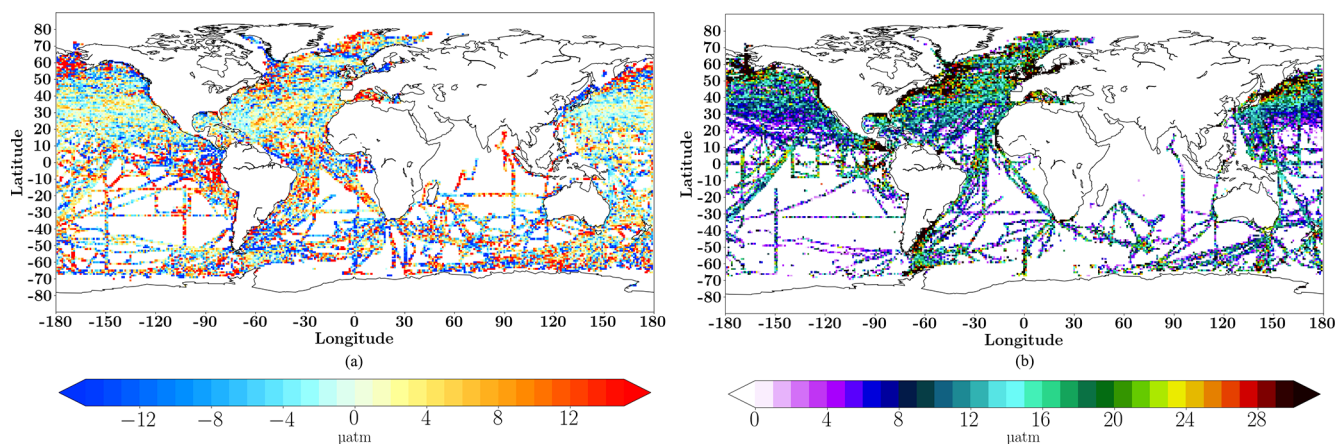


Figure 2. Time mean differences (μatm) between monthly LSCE-FFNN $p\text{CO}_2$ and SOCAT $p\text{CO}_2$ data used for evaluation of the model over the period from 2001 to 2016 (a) and its standard deviation (b).

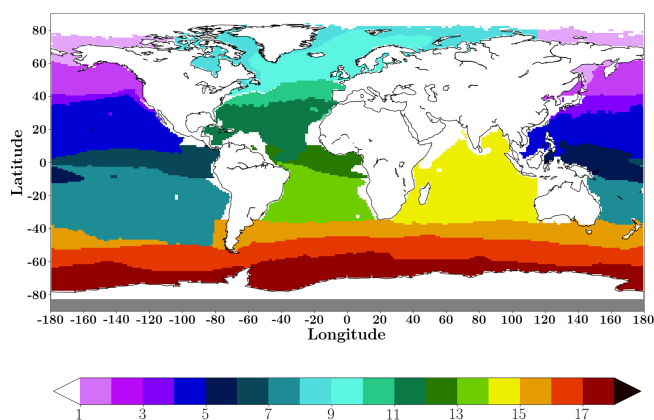


Figure 3. Map of biomes (following Rodenbeck et al., 2015; Fay and McKinley, 2014) used for comparison. See Table 2 for biome names.

R^{iaV} provides information on the capability of each method to reproduce the IAV compared to observations: a smaller R^{iaV} represents a better fit compared with the reference. The amplitude of the interannual variations (A^{iaV}) of sea–air flux of CO_2 (its 2-month running mean) is estimated as the temporal standard deviation over the period.

3.2.1 Interannual variability

The time series of globally averaged surface ocean $p\text{CO}_2$ over the period from 2001 to 2016 are presented in Fig. 4 for LSCE-FFNN and the three other models. Surface ocean $p\text{CO}_2$ (μatm) varied between the four mapping methods in the range of $\pm 7 \mu\text{atm}$ (Fig. 4a). Modeled $p\text{CO}_2$ values were at the lower end for ETH-SOMFFN and JMA-MLR, whereas LSCE-FFNN and Jena-MLS13 computed higher values. The same behavior was found for the 12-month running time series (Fig. 4b). Figure 4c shows the 12-month run-

Table 2. Biomes from Fay and McKinley (2014) used for time series comparison (Fig. 3).

Number	Name
1	(Omitted) North Pacific ice
2	Subpolar seasonally stratified North Pacific
3	Subtropical seasonally stratified North Pacific
4	Subtropical permanently stratified North Pacific
5	Equatorial West Pacific
6	Equatorial East Pacific
7	Subtropical permanently stratified South Pacific
8	(Omitted) North Atlantic ice
9	Subpolar seasonally stratified North Atlantic
10	Subtropical seasonally stratified North Atlantic
11	Subtropical permanently stratified North Atlantic
12	Equatorial Atlantic
13	Subtropical permanently stratified South Atlantic
14	Subtropical permanently stratified Indian Ocean
15	Subtropical seasonally stratified Southern Ocean
16	Subpolar seasonally stratified Southern Ocean
17	Southern Ocean ice

ning mean of the difference between computed $p\text{CO}_2$ and SOCAT data (model – SOCAT) over the globe. JMA-MLR mostly underestimated observed $p\text{CO}_2$ with a strong interannual variability of the misfit, especially at the end of the period, with up to $-5 \mu\text{atm}$. The difference between ETH-SOMFFN output and SOCAT data fluctuated in the range of $\pm 1 \mu\text{atm}$, with an increase in amplitude up to $-2 \mu\text{atm}$ from 2010 onward. Jena-MLS13 overestimated observations with the difference in the range of $0\text{--}1 \mu\text{atm}$. The difference between LSCE-FFNN and SOCAT varied around zero (between -0.7 and $1 \mu\text{atm}$).

The model was then assessed at biome scale. Results for all biomes are presented in the Supplement (Figs. S2, S3, S4). Two biomes with contrasting dynamics are discussed

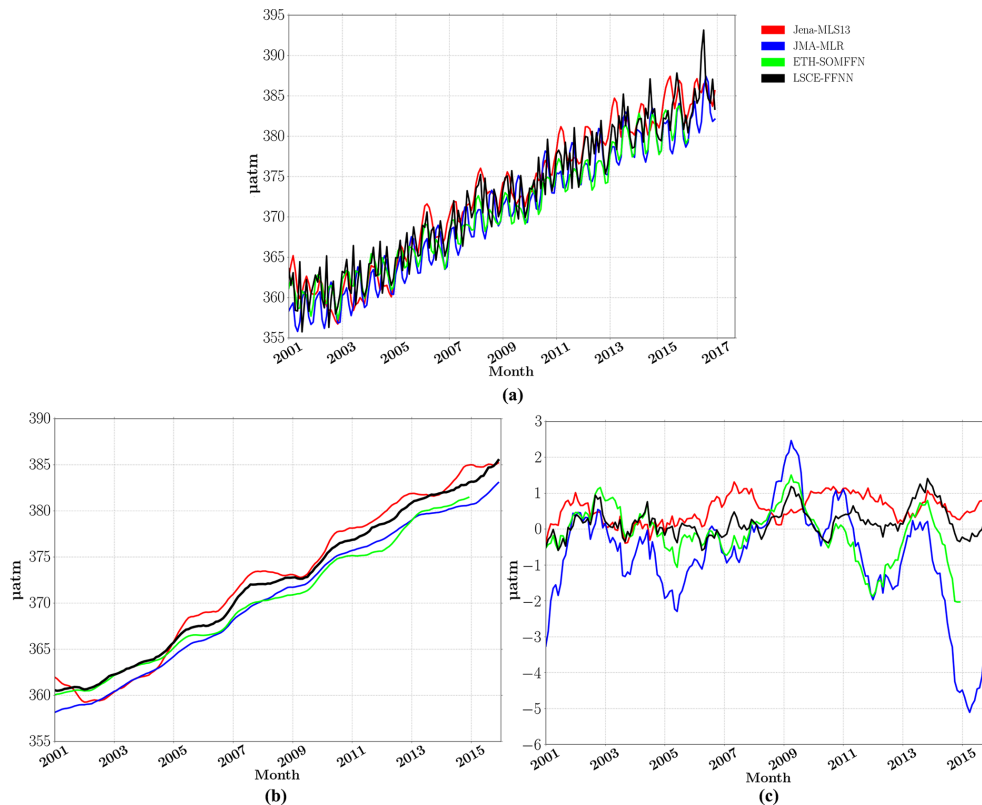


Figure 4. Global oceanic $p\text{CO}_2$: Jena (red), JMA (blue), ETH-SOMFFN (green) and LSCE-FFNN (black). (a) Global average monthly time series, (b) the global 12-month running mean average and (c) yearly $p\text{CO}_2$ mismatch (difference between mapping methods and SOCAT data).

hereafter in greater detail: (1) the equatorial East Pacific (biome 6) characterized by a strong IAV of surface ocean $p\text{CO}_2$ and sea–air CO_2 fluxes in response to ENSO, the El Niño–Southern Oscillation (Feely et al., 1999; Rödenbeck et al., 2015), and (2) the North Atlantic permanently stratified biome (biome 11) with a well-marked seasonal cycle, but little IAV (Schuster et al., 2013). Results for these biomes are presented in Fig. 5.

Biome 6 is relatively well-covered by observations and represents a key region for testing the skill of the model to reproduce the observed strong IAV linked to ENSO. El Niño events are characterized by positive SST anomalies, reduced upwelling and decreased surface ocean $p\text{CO}_2$ values. These episodes could be identified in all model time series (Fig. 5a) with reduced $p\text{CO}_2$ levels in 2004/05 and 2006/07 (weak El Niño), 2002/03 and 2009/10 (moderate El Niño), and 2015/16 (strong El Niño). JMA-MLR (blue curve in Fig. 5a) tended to underestimate $p\text{CO}_2$ during weak El Niño events; it was also underestimated during the La Niña 2011–2012 event by Jena-MLS13. LSCE-FFNN and ETH-SOMFFN, both based on a neural network approach, yielded similar results despite differences in network architecture and predictor datasets.

Data coverage is particularly high over Biome 11 (Fig. 5b, d, f). The seasonal cycle in this biome is dominantly driven by temperature. Modeled seasonal variability showed good agreement across the ensemble of methods (Fig. 5b) with an increase in spring–summer and a decrease in autumn–winter. However, the amplitude can be different by up to $10\ \mu\text{atm}$ between different models. The seasonal amplitude of $p\text{CO}_2$ computed by JMA-MLR increased from smaller values at the beginning of the time series to higher values in the middle of the 2005–2012 period. The variability of seasonal amplitude was the highest for Jena-MLS13 in line with the 12-month running mean time series (Fig. 5d). Again, similar seasonal amplitude and year-to-year variability of surface ocean $p\text{CO}_2$ were obtained with LSCE-FFNN and ETH-SOMFFN (Fig. 5b, d). The yearly $p\text{CO}_2$ mismatch (Fig. 5f) shows that observed surface ocean $p\text{CO}_2$ was underestimated by JMA-MLR at the beginning and at the end of the period by up to $-6\ \mu\text{atm}$, and overestimated during the 2007–2011 period by up to $8\ \mu\text{atm}$. Jena-MLS13 shows mostly positive differences in the range of 0 – $2\ \mu\text{atm}$ over the full period. LSCE-FFNN and ETH-SOMFFN vary around zero and between -2 and $2\ \mu\text{atm}$, respectively, and are close to each other.

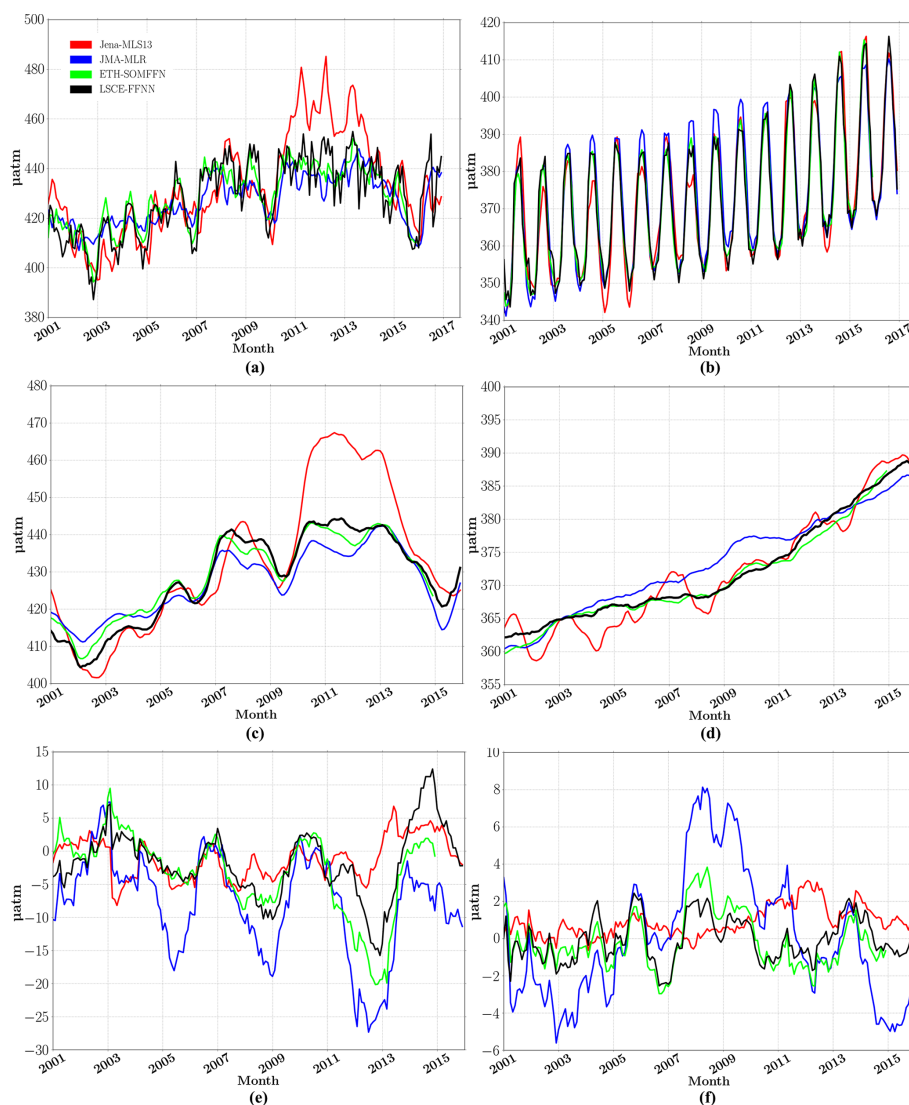


Figure 5. Surface ocean $p\text{CO}_2$ in the equatorial East Pacific (biome 6) (a, c, e) and the subtropical permanently stratified North Atlantic (biome 11) (b, d, f): FFNN (black), JMA (blue), Jena (red) and ETH-SOMFFN (green). (a, b) The monthly time series averaged over the biome. (c, d) The 12-month running mean averaged over the biome. (e, f) The yearly $p\text{CO}_2$ mismatch (difference between the mapping methods and the SOCAT data).

3.2.2 Sea–air CO_2 flux variability

Sea–air exchange of CO_2 was estimated using the same gas-exchange formulation (Eq. 5) and wind speed data (6-hourly NCEP wind speed) for each mapping data (Rödenbeck et al., 2005). It is worth noting that the sea–air flux is sensitive to the choice of the wind speed dataset (Roobaert et al., 2018).

Figure 6a presents the global 12-month running mean of the sea–air CO_2 flux for four mapping methods. All models showed an increase in CO_2 uptake in response to increasing atmospheric CO_2 levels, albeit with a strong between-model variability in multi-annual trends. There is less agreement between the methods compared to reconstructions of

surface ocean $p\text{CO}_2$ variability (Fig. 4b). This results from the contribution of uncertainties in sea–air CO_2 flux estimations over regions with poor data coverage (mostly in the South Hemisphere: the South Pacific, the South Atlantic, the Indian Ocean and the South Ocean; see Fig. S5). Nevertheless, the relative IAV mismatch was less than 30% for all methods (Fig. 6b), suggesting a reasonable fit to observational data. However, the relative IAV mismatch is a global score, and it is biased towards regions with good data coverage (Rödenbeck et al., 2015). The time series reconstructed in this study is too short to capture decadal variations and, in particular, the strengthening of the sink from 2000 onward (Landschützer et al., 2016). LSCE-FFNN computed a slowdown of ocean CO_2 uptake between 2010 and 2013 with

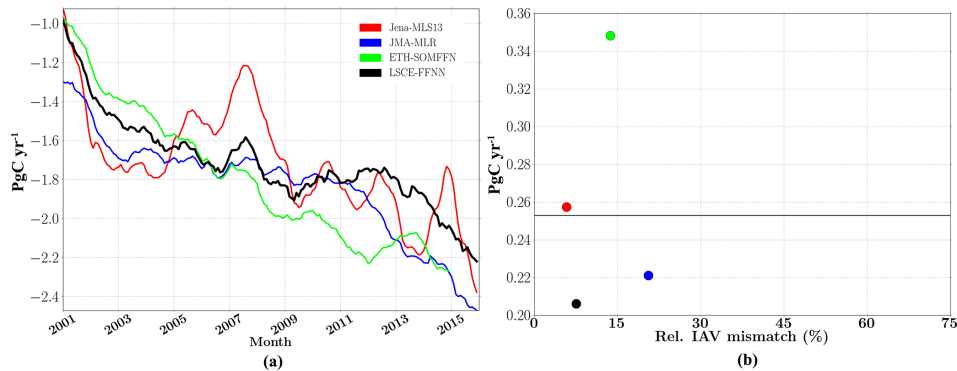


Figure 6. (a) Interannual global ocean sea–air CO_2 flux (12-month running mean); (b) the amplitude of interannual CO_2 flux plotted against the relative IAV mismatch amplitude. The weighted mean is given as a horizontal line.

a flux of $\sim -1.8 \text{ GtC yr}^{-1}$ compared with $\sim -2.2 \text{ GtC yr}^{-1}$ for ETH-SOMFFN. A leveling-off was also found for JMA-MLR, albeit shifted in time. In general, the amplitudes of reconstructed CO_2 fluxes across all four methods agreed within 0.2 to 0.36 PgC yr^{-1} . The weighted mean of IAV (horizontal line in Fig. 6b) computed from the four methods included here was 0.25 PgC yr^{-1} . This value is close to that reported by Rödenbeck et al. (2015) for the complete ensemble of SOCOM models (0.31 PgC yr^{-1}) estimated for the period from 1992 to 2009. The largest amplitude was obtained for ETH-SOMFFN ($\sim 0.35 \text{ PgC yr}^{-1}$). Conversely, LSCE-FFNN has the smallest amplitude with 0.21 PgC yr^{-1} . Jena-MLS13 and JMA-MLR lie very close to the weighted mean value with 0.26 and 0.22 PgC yr^{-1} , respectively. The weighted mean and the dispersion of individual models around it, reflect the period of analysis (2001–2015; ETH-SOMFFN output provided up to 2015) and the total number of models contributing to it (see Rödenbeck et al., 2015 for comparison). As such, it does not provide information on the skill of any particular model.

The interannual variability of reconstructed sea–air CO_2 fluxes (12-month running mean) showed good agreement for biome 6 (East Pacific equatorial, Fig. 7a). A small discrepancy was found at the beginning of the period. A strong increase was computed by Jena-MLS13 for 2010–2014 that was also identified in the $p\text{CO}_2$ variability (Fig. 5a). Despite this, Jena-MLS13 had a low relative R^{IAV} (26%), which confirmed a tendency mentioned in Rödenbeck et al. (2015): that mapping products with a small relative IAV mismatch show larger amplitude. LSCE-FFNN and ETH-SOMFFN yielded comparable results (Fig. 7a, c) with relative IAV mismatches of 46% and 53%, respectively, and with amplitudes of $\sim 0.03 \text{ PgC yr}^{-1}$. Interannual variability reproduced by JMA-MLR falls within the range of the other models (Fig. 7c), but with a R^{IAV} of $\sim 68\%$.

Reconstructed sea–air CO_2 fluxes over the North Atlantic subtropical permanently stratified region (biome 11) show large between model differences in amplitudes and variabil-

ity. The two models based on a neural network again show good agreement with an R^{IAV} of 17% for LSCE-FFNN and 20% for ETH-SOMFFN. Jena-MLS13 produced a strong seasonal variability (Fig. 7b) up to 0.06 PgC yr^{-1} , and a small R^{IAV} of $\sim 11\%$. Contrary to the other approaches, JMA-MLR did not reproduce a decrease in sea–air CO_2 in the middle of the period by up to 0.02 PgC yr^{-1} (Fig. 7b). The model is characterized by an R^{IAV} of 46% and an amplitude of $0.013 \text{ PgC yr}^{-1}$.

3.2.3 Sea–air CO_2 flux trend

The long-term trend of sea–air CO_2 fluxes is dominantly driven by the increase in atmospheric CO_2 (see Fig. S7). On shorter timescales, such as for the period from 2001 to 2016, the interannual variability at regional scales reflects natural modes of climate variability and local oceanographic dynamics (Heinze et al., 2015).

Figure 8 shows the significant linear trends ($p_{\text{val}} = 0.05$) of sea–air CO_2 fluxes for LSCE-FFNN (a), Jena-MLS13 (b), ETH-SOMFFN (c) and JMA-MLR (d). A total (averaged over the globe) negative trend was computed for all models, albeit with large regional contrasts, and LSCE-FFNN falls within this range: Jena-MLS13, $-0.0012 \text{ PgC yr}^{-1} \text{ yr}^{-1}$ ($-0.0028 \text{ PgC yr}^{-1} \text{ yr}^{-1}$, total value with no significant t test; Fig. S8); LSCE-FFNN, $-0.00087 \text{ PgC yr}^{-1} \text{ yr}^{-1}$ ($-0.0032 \text{ PgC yr}^{-1} \text{ yr}^{-1}$); JMA-MLR, $-0.0013 \text{ PgC yr}^{-1} \text{ yr}^{-1}$ ($-0.0037 \text{ PgC yr}^{-1} \text{ yr}^{-1}$); ETH-SOMFFN, $-0.0025 \text{ PgC yr}^{-1} \text{ yr}^{-1}$ ($-0.0059 \text{ PgC yr}^{-1} \text{ yr}^{-1}$). LSCE-FFNN computed negative trends over most of the Atlantic basin, the Indian Ocean and south of 40° S , which contrasts with decreasing fluxes over the Pacific and locally in the Antarctic Circumpolar Current. At first order, this broad regional pattern is found in all models. However, regional maxima and minima are more pronounced in Jena-MLS13 (Fig. 8b) and ETH-SOMFFN (Fig. 8c), whereas a patchy distribution at the sub-basin scale is diagnosed for JMA-MLR.

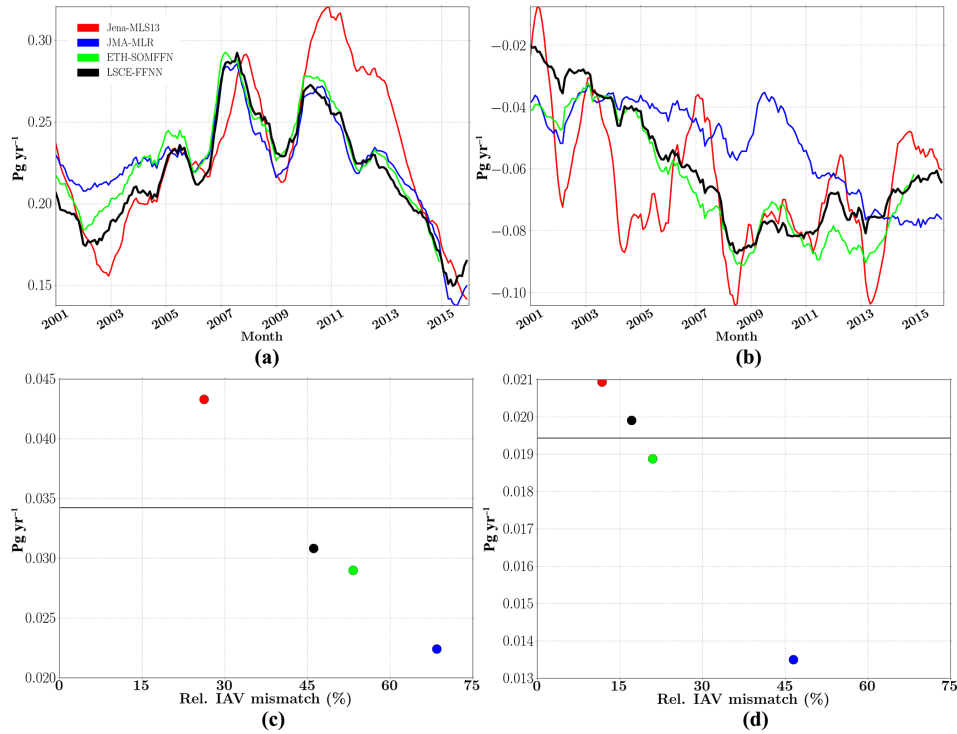


Figure 7. Global ocean interannual sea-air CO₂ flux (12-month running mean) in (a) the equatorial East Pacific (biome 6) and (b) the subtropical permanently stratified North Atlantic (biome 11). The amplitude of interannual CO₂ flux plotted against the relative IAV mismatch amplitude in (c) the equatorial East Pacific (biome 6) and (d) the subtropical permanently stratified North Atlantic (biome 11). The weighted mean is given as a horizontal line.

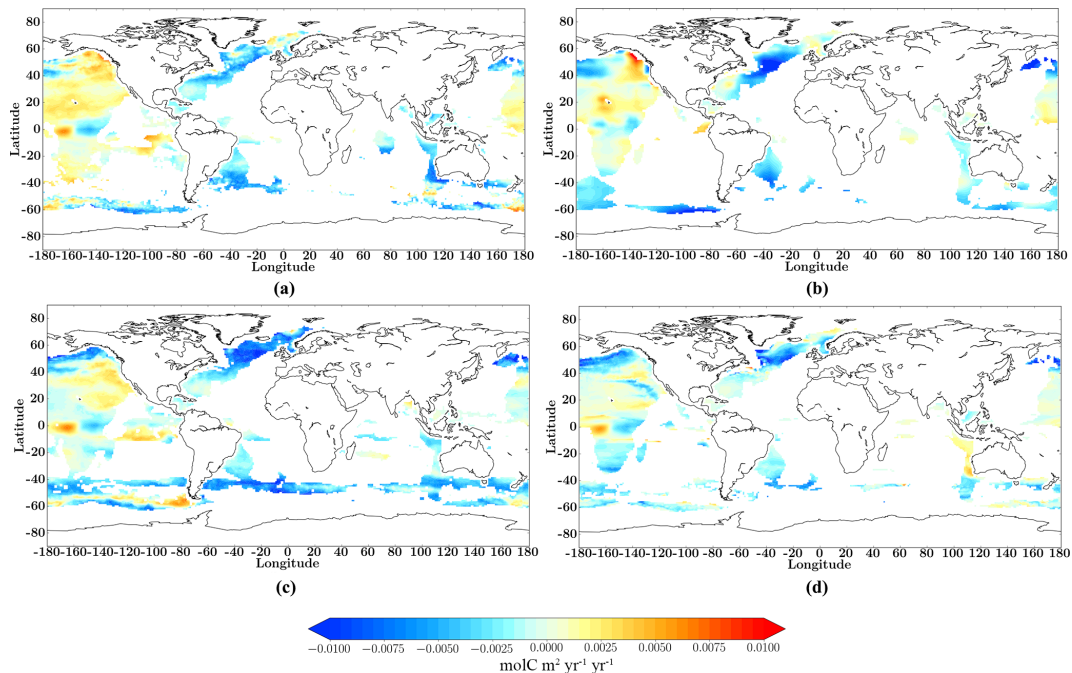


Figure 8. Significant ($p_{\text{val}} = 0.05$) linear trend of $f\text{CO}_2$ for the common period (2001–2015) for (a) LSCE-FFNN, (b) Jena-MLS13, (c) ETH-SOMFFN and (d) JMA-MLR.

Table 3. Mean of sea–air CO_2 flux (PgC yr^{-1}) over the global ocean and per region for the common period (2001–2015). Averages over the period from 2001 to 2009 are presented in parentheses. The last column presents a comparison with the best estimates from Schuster et al. (2013) for the Atlantic Ocean (1990–2009).

Region	Latitude boundaries	LSCE-FFNN	ETH-SOMFFN	Jena-MLS13	JMA-MLR	Schuster et al. (2013), 1990–2009
Global		−1.55 (−1.44)	−1.67 (−1.47)	−1.55 (−1.41)	−1.74 (−1.62)	–
Arctic	76–90° N	−0.001	−0.001	−0.001	−0.001	−0.12 ± 0.06
Subpolar Atlantic	49–76° N	−0.15 (−0.15)	−0.14 (−0.12)	−0.15 (−0.15)	−0.16 (−0.15)	−0.21 ± 0.06
Subpolar Pacific	49–76° N	−0.003 (−0.005)	−0.009 (−0.004)	−0.006 (−0.004)	−0.027 (−0.021)	–
Subtropical Atlantic	18–49° N	−0.21 (−0.19)	−0.21 (−0.19)	−0.2 (−0.18)	−0.21 (−0.2)	−0.26 ± 0.06
Subtropical Pacific	18–49° N	−0.45 (−0.46)	−0.49 (−0.48)	−0.47 (−0.46)	−0.49 (−0.47)	–
Equatorial Atlantic	18° S–18° N	0.085 (0.09)	0.085 (0.095)	0.08 (0.082)	0.1 (0.11)	0.12 ± 0.04
Equatorial Pacific	18° S–18° N	0.42 (0.41)	0.4 (0.4)	0.44 (0.42)	0.38 (0.37)	–
South Atlantic	44–18° S	−0.17 (−0.16)	−0.18 (−0.16)	−0.18 (−0.17)	−0.23 (−0.22)	−0.14 ± 0.04
South Pacific	44–18° S	−0.33 (−0.34)	−0.4 (−0.39)	−0.35 (−0.34)	−0.49 (−0.47)	–
Indian Ocean	44° S–30° N	−0.25 (−0.2)	−0.32 (−0.29)	−0.27 (−0.26)	−0.27 (−0.29)	–
Southern Ocean	90–44° S	−0.38	−0.29	−0.36	−0.26	–

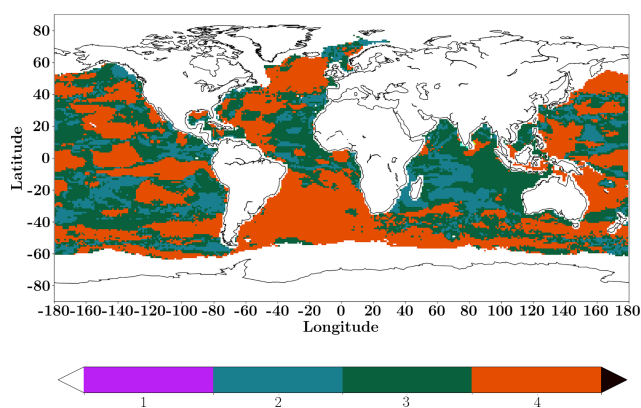


Figure 9. Agreement between the four mapping methods with respect to their linear trend of sea–air CO_2 flux. The color bar represents the number of products that have the same sign of linear trend.

The agreement in sign of computed linear trends from the four models is presented in Fig. 9 (total linear trends with no significant t test). Over most of the ocean, all four models show very close sea–air CO_2 tendency. In the Indian Ocean (biome 14), in comparison, a positive trend was computed for JMA-MLR ($0.0004 \text{ PgC yr}^{-1} \text{ yr}^{-1}$; $0.00006 \text{ PgC yr}^{-1} \text{ yr}^{-1}$ with a t test), whereas the other three models present a negative trend. The differences between models were also found in the Pacific Ocean, especially the southern Pacific. In the eastern equatorial Pacific region (biome 6) a total significant negative trend is presented by all models. All models reproduced a maximum in the southern part of biome 6, but they disagreed about its amplitude and spatial distribution. Almost everywhere over the Atlantic Ocean the mapping methods produced the same sign of linear trend (Fig. 9). However, in the eastern part of the subtropical North Atlantic, Jena-MLS13 gave a positive linear trend of $f\text{CO}_2$ (Fig. 8b).

According to LSCE-FFNN, the global ocean took up 1.55 PgC yr^{-1} on average between 2001 and 2015. This estimate is consistent with results from the other three models (Table 3; see Table S2 for estimations per biome). The spread between individual models falls in the range of the error reported in Landschützer et al. (2016), $\pm 0.4\text{--}0.6 \text{ PgC yr}^{-1}$. Per biome, estimates of CO_2 sea–air fluxes provided by LSCE-FFNN are also in good agreement with those derived from the other models.

4 Summary and conclusion

We proposed a new model for the reconstruction of monthly surface ocean $p\text{CO}_2$. The model is applied globally and allows a seamless reconstruction without introducing boundaries between the ocean basins or biomes. Our model relies on a two-step approach based on feed-forward neural networks (LSCE-FFNN). The first step corresponds to the reconstruction of a monthly $p\text{CO}_2$ climatology. It allows for the output of the FFNN to be kept close to the observed values in regions with poor data cover. In the second step, $p\text{CO}_2$ anomalies are reconstructed with respect to the climatology from the first step. The model was applied over the period from 2001 to 2016. Validation with independent data at the global scale indicated a RMSE of $17.57 \mu\text{atm}$, an r^2 of ~ 0.76 and an absolute bias of $11.52 \mu\text{atm}$. In order to assess the model further, it was compared to three different mapping models: ETH-SOMFFN (self-organizing maps + neural network), Jena-MLS13 (statistical interpolation) and JMA-MLR (linear regression) (Rödenbeck et al., 2015). Network qualification followed the protocol and diagnostics proposed in Rödenbeck et al. (2015).

Reconstructed surface ocean $p\text{CO}_2$ distributions were in good agreement with other models and observations. The seasonal variability was reproduced to a satisfying level by LSCE-FFNN, the yearly $p\text{CO}_2$ mismatch varied around zero

and the relative IAV mismatch was 7%. LSCE-FFNN proved skillful in reproducing the interannual variability of surface ocean $p\text{CO}_2$ over the eastern equatorial Pacific in response to ENSO. Reductions in surface ocean $p\text{CO}_2$ during El Niño events were well reproduced. The comparison between reconstructed and observed $p\text{CO}_2$ values yielded a RMSE of 15.73 μatm , an r^2 of 0.79 and an absolute bias of 10.33 μatm over the equatorial Pacific. The relative IAV misfit in this region was $\sim 17\%$. Despite the overall good agreement between models, important differences still exist at the regional scale, especially in the Southern Hemisphere and, in particular, in the southern Pacific and the Indian Ocean. These regions suffer from poor data coverage. Large regional uncertainties in reconstructed surface ocean $p\text{CO}_2$ and sea–air CO_2 fluxes have a strong influence on global estimates of CO_2 fluxes and trends.

Code and data availability. Python code for the $p\text{CO}_2$ climatology reconstruction (the first step of LSCE-FFNN model) and Python code for reconstruction of the $p\text{CO}_2$ anomalies (the second step of LSCE-FFNN model) are provided at the end of the Supplement.

Time series of reconstructed surface ocean $p\text{CO}_2$ and CO_2 fluxes are distributed through the Copernicus Marine Environment Monitoring Service (CMEMS), <http://marine.copernicus.eu/services-portfolio/access-to-products/>, search keyword: MULTI-OBS.

Supplement. The supplement related to this article is available online at: <https://doi.org/10.5194/gmd-12-2091-2019-supplement>.

Author contributions. ADS, MG, MV and CM contributed to the development of the methodology and designed the experiments, and ADS carried them out. ADS developed the model code and performed the simulations. ADS prepared the paper with contributions from all coauthors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors would like to thank the two referees, Christian Rödenbeck and Luke Gregor, for their helpful comments and questions, as well as Frederic Chevallier and Gilles Reverdin for their suggestions. MV acknowledges funding from the CoCliServ and EUPHEME projects (ERA4CS program).

Financial support. This study was funded by the AtlantOS project (EU Horizon 2020 research and innovation program, grant agreement no. 2014-633211) and the CoCliServ and EUPHEME projects.

Review statement. This paper was edited by James Annan and reviewed by Christian Rödenbeck and Luke Gregor.

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.
- Aumont, O. and Bopp, L.: Globalizing results from ocean in situ iron fertilization studies, *Global Biogeochem. Cy.*, 20, GB2017, <https://doi.org/10.1029/2005GB002591>, 2006.
- Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S.-I., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibáñez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., and Xu, S.: A multi-decade record of high-quality $f\text{CO}_2$ data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT), *Earth Syst. Sci. Data*, 8, 383–413, <https://doi.org/10.5194/essd-8-383-2016>, 2016.
- Bishop, C. M.: *Neural Networks for Pattern Recognition*, Oxford University Press, Cambridge, UK, 1995.
- Bishop, C. M.: *Pattern Recognition and Machine Learning*, Springer, Berlin, 2006.
- Bittig, H.C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N.L., Sauzède, R., Körtzinger, A., and Gattuso, J.-P.: An Alternative to Static Climatologies: Robust Estimation of Open Ocean CO_2 Variables and Nutrient Concentrations From T, S, and O_2 Data Using Bayesian Neural Networks, *Front. Mar. Sci.*, 5, 328, <https://doi.org/10.3389/fmars.2018.00328>, 2018.
- Chollet, F.: Keras, available at: <https://keras.io> (last access: 12 May 2019), 2015.
- Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and other biogeochemical cycles, in: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.
- Donlon, C. J., Martin, M., Stark, J. D., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: *The Operational Sea Surface Temperature and*

- Sea Ice analysis (OSTIA), *Remote Sens. Environ.*, 116, 140–158, <https://doi.org/10.1016/j.rse.2010.10.017>, 2011.
- Fay, A. R. and McKinley, G. A.: Global open-ocean biomes: mean and temporal variability, *Earth Syst. Sci. Data*, 6, 273–284, <https://doi.org/10.5194/essd-6-273-2014>, 2014.
- Fay, A. R., McKinley, G. A., and Lovenduski, N. S.: Southern Ocean carbon trends: Sensitivity to methods, *Geophys. Res. Lett.*, 41, 6833–6840, <https://doi.org/10.1002/2014GL061324>, 2014.
- Feely, R. A., Wanninkhof, R., Takahashi, T., and Tans, P.: Influence of El Niño on the equatorial Pacific contribution to atmospheric CO_2 accumulation, *Nature*, 398, 597–601, 1999.
- Friedrich, T. and Oschlies, A.: Neural network-based estimates of North Atlantic surface $p\text{CO}_2$ from satellite data: A methodological study, *J. Geophys. Res.*, 114, C03020, <https://doi.org/10.1029/2007JC004646>, 2009.
- Guinehut, S., Dhomps, A.-L., Larnicol, G., and Le Traon, P.-Y.: High resolution 3-D temperature and salinity fields derived from in situ and satellite observations, *Ocean Sci.*, 8, 845–857, <https://doi.org/10.5194/os-8-845-2012>, 2012.
- Heinze, C., Meyer, S., Goris, N., Anderson, L., Steinfeldt, R., Chang, N., Le Quéré, C., and Bakker, D. C. E.: The ocean carbon sink – impacts, vulnerabilities and challenges, *Earth Syst. Dynam.*, 6, 327–358, <https://doi.org/10.5194/esd-6-327-2015>, 2015.
- Hinton, G., Srivastava, N., and Swersky, K.: Lecture 6a: Overview of mini-batch gradient descent, Neural Networks for Machine Learning, Slides, available at: http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf (last access: 12 May 2019), 2012.
- Iida, Y., Kojima, A., Takatani, Y., Nakano, T., Midorikawa, T., and Ishii, M.: Trends in $p\text{CO}_2$ and sea-air CO_2 flux over the global open oceans for the last two decades, *J. Oceanogr.*, 71, 637–661, <https://doi.org/10.1007/s10872-015-0306-4>, 2015.
- Kallache, M., Vrac, M., Naveau, P., and Michelangeli, P.-A.: Non-stationary probabilistic downscaling of extreme precipitation, *J. Geophys. Res.*, 116, D05113, <https://doi.org/10.1029/2010JD014892>, 2011.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, *B. Am. Meteorol. Soc.*, 77, 437–471, 1996.
- Körtzinger, A.: Determination of carbon dioxide partial pressure ($p\text{CO}_2$), in: *Methods of Seawater Analysis*, 3rd edn., edited by: Grasshoff, K., Kremling, K., and Ehrhardt, M., Wiley-VCH Verlag GmbH, Weinheim, Germany, <https://doi.org/10.1002/9783527613984.ch9>, 1999.
- Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., Sasse, T. P., and Zeng, J.: A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink, *Biogeosciences*, 10, 7793–7815, <https://doi.org/10.5194/bg-10-7793-2013>, 2013.
- Landschützer, P., Gruber, N., Bakker, D. C. E., and Schuster, U.: Recent variability of the global ocean carbon sink, *Global Biogeochem. Cy.*, 28, 927–949, <https://doi.org/10.1002/2014GB004853>, 2014.
- Landschützer, P., Gruber, N., and Bakker, D. C. E.: Decadal variations and trends of the global ocean carbon sink, *Global Biogeochem. Cy.*, 30, 1396–1417, <https://doi.org/10.1002/2015GB005359>, 2016.
- Laruelle, G. G., Landschützer, P., Gruber, N., Tison, J.-L., Delille, B., and Regnier, P.: Global high-resolution monthly $p\text{CO}_2$ climatology for the coastal ocean derived from neural network interpolation, *Biogeosciences*, 14, 4545–4561, <https://doi.org/10.5194/bg-14-4545-2017>, 2017.
- Lefèvre, N., Watson, A. J., and Watson, A. R.: A comparison of multiple regression and neural network techniques for mapping *in situ* $p\text{CO}_2$ data, *Tellus B*, 57, 375–384, <https://doi.org/10.3402/tellusb.v57i5.16565>, 2005.
- Le Quéré, C., Aumont, O., Bopp, L., Bousquet, P., Ciais, P., Francey, R., Heimann, M., Keeling, C. D., Keeling, R. F., Khesghi, H., Peylin, P., Piper, S. C., Prentice, I. C., and Rayner, P. J.: Two decades of ocean CO_2 sink and variability, *Tellus B*, 55, 649–656, <https://doi.org/10.1034/j.1600-0889.2003.00043.x>, 2003.
- Le Quéré, C., Takahashi, T., Buitenhuis, E. T., Rödenbeck, C., and Sutherland, S. C.: Impact of climate change and variability on the global oceanic sink of CO_2 , *Glob. Biogeochem. Cy.*, 24, GB4007, <https://doi.org/10.1029/2009GB003599>, 2010.
- Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Nojiri, Y., Padin, X. A., Pregon, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S., and Zhu, D.: Global Carbon Budget 2017, *Earth Syst. Sci. Data*, 10, 405–448, <https://doi.org/10.5194/essd-10-405-2018>, 2018.
- 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.
- Nakaoka, S., Telszewski, M., Nojiri, Y., Yasunaka, S., Miyazaki, C., Mukai, H., and Usui, N.: Estimating temporal and spatial variation of ocean surface $p\text{CO}_2$ in the North Pacific using a self-organizing map neural network technique, *Biogeosciences*, 10, 6093–6106, <https://doi.org/10.5194/bg-10-6093-2013>, 2013.
- Orr, J. C., Monfray, P., Maier-Reimer, E., Mikolajewicz, U., Palmer, J., Taylor, N. K., Toggweiler, J. R., Sarmiento, J. L., Quere, C. L., Gruber, N., Sabine, C. L., Key, R. M., and Boutin, J.: Estimates of anthropogenic carbon uptake from four three-dimensional global ocean models, *Global Biogeochem. Cy.*, 15, 43–60, <https://doi.org/10.1029/2000GB001273>, 2001.
- Peylin, P., Bousquet, P., Le Quéré, C., Sitch, S., Friedlingstein, P., McKinley, G., Gruber, N., Rayner, P., and Ciais, P.: Multiple constraints on regional CO_2 flux variations over

- land and oceans, *Global Biogeochem. Cy.*, 19, GB1011, <https://doi.org/10.1029/2003GB002214>, 2005.
- 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, <https://doi.org/10.5194/bg-10-6699-2013>, 2013.
- Rödenbeck, C.: Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Technical Report 6, Max Planck Institute for Biogeochemistry, Jena, available at: http://www.bgc-jena.mpg.de/uploads/Publications/TechnicalReports/tech_report6.pdf (last access: 12 May 2019), 2005.
- Rödenbeck, C., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., Cassar, N., Reum, F., Keeling, R. F., and Heimann, M.: Interannual sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme, *Biogeosciences*, 11, 4599–4613, <https://doi.org/10.5194/bg-11-4599-2014>, 2014.
- Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., Landschützer, P., Metzl, N., Nakaoka, S., Olsen, A., Park, G.-H., Peylin, P., Rodgers, K. B., Sasse, T. P., Schuster, U., Shutler, J. D., Valsala, V., Wanninkhof, R., and Zeng, J.: Data-based estimates of the ocean carbon sink variability – first results of the Surface Ocean $p\text{CO}_2$ Mapping intercomparison (SOCOM), *Biogeosciences*, 12, 7251–7278, <https://doi.org/10.5194/bg-12-7251-2015>, 2015.
- Rodgers, K. B., Key, R. M., Gnanadesikan, A., Sarmiento, J. L., Aumont, O., Bopp, L., Doney, S. C., Dunne, J. P., Glover, D. M., Ishida, A., Ishii, M., Jacobson, A. R., Lo Monaco, C., Maier-Reimer, E., Mercier, H., Metzl, N., Pérez, F. F., Rios, A. F., Wanninkhof, R., Wetzel, P., Winn, C. D., and Yamanaka, Y.: Using altimetry to help explain patchy changes in hydrographic carbon measurements, *J. Geophys. Res.*, 114, C09013, <https://doi.org/10.1029/2008JC005183>, 2009.
- Roobaert, A., Laruelle, G. G., Landschützer, P., and Regnier, P.: Uncertainty in the global oceanic CO_2 uptake induced by wind forcing: quantification and spatial analysis, *Biogeosciences*, 15, 1701–1720, <https://doi.org/10.5194/bg-15-1701-2018>, 2018.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning internal representations by backpropagating errors, *Nature*, 323, 533–536, 1986.
- Sauzède, R., Claustre, H., Uitz, J., Jamet, C., Dall’Olmo, G., D’Ortenzio, F., Gentili, B., Poteau, A., and Schmechtig, C.: A neural network-based method for merging ocean color and Argo data to extend surface bio-optical properties to depth: Retrieval of the particulate backscattering coefficient, *J. Geophys. Res.-Oceans*, 121, 2552–2571, <https://doi.org/10.1002/2015JC011408>, 2016.
- 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, <https://doi.org/10.5194/bg-10-607-2013>, 2013.
- Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., Bates, N., Wanninkhof, R., Feely, R. A., Sabine, C., Olafsson, J., and Nojiri, Y.: Global sea–air CO_2 flux based on climatological surface ocean $p\text{CO}_2$, and seasonal biological and temperature effects, *Deep.-Sea Res. Pt. II*, 49, 1601–1622, [https://doi.org/10.1016/S0967-0645\(02\)00003-6](https://doi.org/10.1016/S0967-0645(02)00003-6), 2002.
- Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C. S., Delille, B., Bates, N. R., and de Baar, H. J. W.: Climatological mean and decadal change in surface ocean $p\text{CO}_2$, and net sea–air CO_2 flux over the global oceans, *Deep.-Sea Res. Pt. II*, 56, 554–577, <https://doi.org/10.1016/j.dsr2.2008.12.009>, 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 $p\text{CO}_2$ distribution in the North Atlantic using a self-organizing neural network, *Biogeosciences*, 6, 1405–1421, <https://doi.org/10.5194/bg-6-1405-2009>, 2009.
- Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean, *J. Geophys. Res.*, 97, 7373–7382, <https://doi.org/10.1029/92JC00188>, 1992.
- Weiss, R.: Carbon dioxide in water and seawater: the solubility of a non-ideal gas, *Mar. Chem.*, 2, 203–205, [https://doi.org/10.1016/0304-4203\(74\)90015-2](https://doi.org/10.1016/0304-4203(74)90015-2), 1974.
- Zeng, J., Nojiri, Y., Landschützer, P., Telszewski, M., and Nakaoka, S.: A global surface ocean $f\text{CO}_2$ climatology based on a feed-forward neural network, *J. Atmos. Ocean Technol.*, 31, 1838–1849, <https://doi.org/10.1175/JTECH-D-13-00137.1>, 2014.