LSCE-FFNN-v1: A two-step neural network model for the reconstruction of surface ocean pCO_2 over the Global Ocean.

Anna Denvil-Sommer[1], Marion Gehlen[1], Mathieu Vrac[1], Carlos Mejia[2]

[1]Laboratoire 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

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

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

Abstract.

A new Feed-Forward Neural Network (FFNN) model is presented to reconstruct surface ocean partial pressure of carbon dioxide (pCO_2) over the global ocean. The model consists of two steps: (1) reconstruction of pCO_2 climatology and (2) reconstruction of pCO_2 anomalies with respect to the climatology. For the first step, a gridded climatology was used as the target, along with sea surface salinity and temperature (SSS and SST), sea surface height (SSH), chlorophyll *a* (Chl), 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 non-linear relationships between pCO_2 and the ocean predictors. It provides monthly surface ocean pCO_2 distributions on a 1°x1° grid for the period 2001-2016. Global ocean pCO_2 was reconstructed with a satisfying accuracy compared to independent observational data from SOCAT. However, errors are larger in regions with poor data coverage (e.g. Indian Ocean, Southern Ocean, subpolar Pacific). The model captured the strong interannual variability of surface ocean pCO_2 with reasonable skills over the Equatorial Pacific associated with ENSO (El Niño Southern Oscillation). Our model was compared to three pCO_2 mapping methods that participated in the Surface Ocean pCO_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, the Southern Pacific and the Indian Ocean, as these regions suffer from poor data-coverage. Large regional uncertainties in reconstructed surface ocean pCO_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 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}) amounts to 155 ± 30 PgC or 28% of 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 ± 0.7 PgC yr^{-1} (Ciais et al., 2013). However, these global estimates hide substantial regional and inter-annual 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 inter-annual 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), while atmospheric inversions suffer from a still 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 pCO_2 (Rödenbeck et al., 2015) such as statistical interpolation, linear and non-linear regressions, or model-based regressions or tuning (Rödenbeck et al., 2014, 2015). These methods, their advantages and disadvantages are compared and discussed in Rödenbeck et al. (2015). This intercomparison did not allow identifying a single optimal technique but rather pleaded in favour of exploiting the ensemble of methods.

Artificial neural networks (ANN) have been widely used to reconstruct surface ocean pCO_2 (open ocean: Lefèvre et al., 2005; Friedrich and Oschlies, 2009b; Telszewski et al., 2009; Landschützer et al., 2013; Nakaoka et al., 2013; Zeng et al. 2014; Bitting et al., 2018; coastal region: Laruelle et al., 2017). ANN fill the spatial and temporal gaps based on calibrated non-linear statistical relationships between pCO_2 and its oceanic and atmospheric drivers. The existing products usually present monthly fields with a 1ºx1º spatial resolution and capture a large part of temporal-spatial variability. Methods based on ANN are able to represent the relationships between pCO2 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 pCO_2 over the period 2001-2016. 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 pCO_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°x1° 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 mentioned before, 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), thereby keeping FFNN output close to realistic values. Furthermore, the reconstruction of monthly climatologies during the first step allows taking into account a potential change in seasonal cycle in response to climate change 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: section 2 introduces datasets used during this study and describes the neural network; section 3 presents results for its validation and qualification, as well as a comparison to three mapping methods part of the Surface Ocean pCO_2 Mapping intercomparison (SOCOM) exercise (Rödenbeck et al., 2015). Results and perspectives are summarized in the last section.

2. Data and method.

2.1. Data.

The standard set of variables known to represent physical, chemical and biological drivers of surface ocean pCO_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), atmospheric CO_2 mole fraction ($xCO_{2,atm}$). 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. First tests suggested that the inclusion of SSH does not significantly improve the accuracy of reconstructed pCO_2 at global scale. At basin and regional scale, however, adding SSH improves the spatial pattern of reconstructed pCO_2 and the accuracy of our method.

For the first step, the reconstruction of monthly climatologies, the Takahashi et al. (2009) monthly pCO_2 gridded climatology (1°x1°) was used as the target. The original climatology was constructed by an advection-based interpolation method on a 4ºx5º grid. It was interpolated on the 1°x1° SOCAT grid which is also the resolution of the final output for the FFNN.

For the second step, the target is provided by the observational database SOCAT v5 (Bakker et al., 2016). We used a gridded version of this dataset that was derived by combining all SOCAT data collected within a 1ºx1º box during a specific month. SOCAT v5 represents global observations of sea surface fugacity of CO_2 (fCO_2) over the period 1970 to 2016. It includes data from moorings, ships and drifters. These data are

distributed irregularly over the global ocean with 188274 gridded measurements over the Northern Hemisphere and 76065 over the Southern Hemisphere. In order to ensure a satisfying spatial and temporal data coverage, we limited the reconstruction to the period 2001-2016, which represents ~77% of the database (Fig. 1(a)).

The following formula is used to convert fCO_2 to pCO_2 (Körtzinger et al., 1999):

$$fCO_2 = pCO_2 \exp\left(p\frac{B+2\delta}{RT}\right), \quad (1)$$

where fCO_2 and pCO_2 are in µatm, p is the total pressure (Pa), R=8.314 JK^{-1} is the gas constant, T is the absolute temperature (K). Parameter B (m^3mol^{-1}) is estimated as: B = (-1636.75 + 12.0408 T – 3.27957 * 10^{-2} T^2 + 3.16528 * 10^{-5} T^3) 10^{-6}. The parameter δ is the cross virial coefficient (m^3mol^{-1}): δ = (57.7 – 0.118T) 10^{-6}. The total pressure is from the Jena database (6h, 5ºx5º) based on the NCEP reanalysis (Kalnay et al., 1996) (http://www.bgc-jena.mpg.de/CarboScope/?ID=s).

Monthly global reprocessed products of physical variables from ARMOR3D L4 distributed through the Copernicus Marine Environment Monitoring Service (CMEMS) (0.25ºx0.25º) (http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=MULTIOBS_GLO_PHY_REP_015_002) were used for SSS, SST and SSH (Guinehut et al., 2012). The GlobColour project provided monthly CHL distributions at 1ºx1º resolution (http://www.globcolour.info/products_description.html). For MLD, daily data from the "Estimating the Circulation and Climate of the Ocean" (ECCO2) project Phase II (Cube 92), at 0.25ºx0.25º resolution (Menemenlis et al., 2008) were used. For xCO_2 atmospheric, the 6h data from Jena CO_2 inversion s76_v4.1 on a 5ºx5º grid were selected (http://www.bgc-jena.mpg.de/CarboScope/?ID=s). Finally, an ice mask based on daily "Operational Sea Surface Temperature and Sea Ice Analysis" (OSTIA) with a gridded 0.05ºx0.05º resolution (Donlon et al., 2011) was applied.

MLD and CHL 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(CHL) = 0 was applied. It does not introduce discontinuities since log(CHL) is close to zero in the adjacent region.

All data were averaged or interpolated on a 1ºx1º 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:

$$SSS_n = \frac{SSS - \overline{SSS}}{std(SSS)}.$$

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

As surface ocean pCO_2 also varies spatially, geographical positions (lat, lon) after conversion to radians

were included as predictors. In order to normalize (lat, lon) the following transformation is proposed:

$$lat_n = \sin\left(lat * \pi / 180^0\right)$$

$$lon_{n,1} = \sin\left(lon * \pi / 180^0\right)$$

$$lon_{n,2} = \cos\left(lon * \pi / 180^0\right)$$

Two functions *sin* and *cos* for longitudes are used to preserve its periodical 0 to 360 degrees behavior and thus to consider the difference of positions before and after the 0º longitude. For step 2, data required for training were co-located at the SOCAT data positions that are used as a target for the FFNN model. Details are provided in the next section.

2.2. Method.

a) 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 number and size of hidden layers (i.e., intermediate layers between input and output layers), connection type, activation functions, 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, bias) a hyperbolic tangent was chosen as an activation function for neurons in hidden layers, and a linear function for the output layer. As 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 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), that is the case here, it is enough to have only one single hidden layer but 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 the FFNN layers and the number of neurons depends on one side 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 advises to have 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 pCO_2 climatology as a nonlinear function of normalized SSS, SST, SSH, Chl, MLD climatologies and geographical position (longitude, latitude):

$$pCO_{2,n} = \left(SSS_n, SST_n, SSH_n, Chl_n, MLD_n, lon_n, lat_n\right) \quad (2)$$

Surface ocean pCO_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 4th point) of the dataset were 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 into account all regions and seasonal variability. In order to improve the accuracy of the reconstruction, the model was applied separately for each month. We have developed a FFNN model with 5 layers (3 hidden layers). 12 models with a common architecture were trained. Tests with one model for 12 months showed a slight decrease in accuracy (not presented here). About 17500 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 pCO_2 anomalies were reconstructed as a nonlinear function of normalized SSS, SST, SSH, Chl, MLD, xCO_2 and their anomalies, as well as geographic position:

$$pCO_{2,anom,n} = \left(SSS_n, SST_n, SSH_n, Chl_n, MLD_n, xCO_{2,n}, SSS_{anom,n}, SST_{anom,n}, SSH_{anom,n}, Chl_{anom,n}, MLD_{anom,n}, xCO_{2,anom,n}, lon_{n,1}, lon_{n,2}, lat_n\right) \quad (3)$$

Surface ocean pCO_2 anomalies computed as the differences between collocated pCO_2 values based on SOCAT observations and monthly pCO_2 climatologies reconstructed during the first step provided the targets:

$$pCO_{2,anom} = pCO_{2,SOCAT} - pCO_{2,clim,FFNN} \quad (4)$$

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. There were thus 12 models sharing a common architecture but trained on different data. At this step, in order to increase the amount of data during training and to introduce information on the seasonal cycle, the model was trained using as a target pCO_2 data from the month in question as well as those from the previous and following month during the entire period 2001-2016. Figures 1 (b) and 1 (c) show an example of data distribution for the sole months of January over the period 2001-2016 (Fig. 1 (b)) and for the three months time-window December-January-February 2001-2016 used in the training algorithm of the January FFNN model (Fig. 1 (c)). In this particular example, the choice of three 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 subsampling of the dataset to draw 25% of independent data for validation (Fig. S1). Each sampling fold was tested on 5 runs of the FFNN for each month. Each of these 5 runs is characterized by different initial values that are chosen randomly. From these 5 results, the best was chosen based on root-mean-square-error (RMSE), r^2 and bias.

The final model architecture at step 2 had 3 layers (1 hidden layer). About 10000 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 of accuracy.

b) Reconstruction of surface ocean pCO_2

The previous section presented the development of the "optimal" architecture of a FFNN model for the reconstruction of global surface ocean pCO_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 is trained on all available data: 100% of data for training, 100% for evaluation and 100% for validation. The network was executed 5 times (different initial values) and the best model was selected based on validation results considering root-mean-square-error (RMSE), r^2 and bias computed between network output and SOCAT derived surface ocean pCO_2 data. The final model output is referred to as the LSCE-FFNN product.

2.3. Computation of sea-air CO_2 fluxes.

Sea-air CO_2 flux *f* was calculated following Rödenbeck et al. (2015) as:

$$f = k\rho L\left(pCO_2 - pCO_2^{atm}\right) \quad (5)$$

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

$$k = \Gamma u^2 \left(Sc^{CO_2}/Sc^{Ref}\right)^{-0.5} . \quad (6)$$

The global scaling factor Γ was chosen as in Rödenbeck et al. (2014) with the global mean CO_2 piston velocity equaling to 16.5 cm/h. *Sc* corresponds to the Schmidt number estimated according to Wanninkhof (1992). The wind speed was computed from 6-hourly NCEP wind speed (Kalnay et al., 1996). ρ is seawater density in (5) and L is the temperature-dependent solubility (Weiss, 1974). pCO_2 corresponds to the surface ocean pCO_2 output of the mapping method. pCO_2^{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, that is 25% of the total, represents independent observations as they did not participate in training during model development (see 2.2a). The skill of the FFNN to reconstruct monthly climatologies of surface ocean pCO_2, was assessed by comparing collocated reconstructed pCO_2 and corresponding values from Takahashi et al. (2009). The global climatology was reconstructed with a satisfying accuracy during step 1 with a RMSE of 0.17 µatm and r^2 of 0.93. Model output of step 2 was assessed by K-fold cross-validation as presented before: K=4 different subsets of independent data were drawn from the dataset and the network was run 5 times on each subset. From these 20 results the best one was chosen based on RMSE, r^2 and mean absolute error (MAE) (the bias is presented in Table S1). The 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 ~17.97 µatm, while the MAE was 11.52 µatm and r^2 was 0.76. These results are comparable to those obtained by Landschützer et al. (2013) for the assessment of a surface ocean pCO_2 reconstruction based on an alternative neural network-based approach. The RMSE between SOCAT data and the climatology of pCO_2 from Takahashi et al. (2009) equals 41.87 µatm, larger than errors computed for the regional comparison between FFNN and SOCAT (Table 1). We also estimated the RMSE for the case of 100% data used for training. It equals 14.8 µatm and confirms the absence of overfitting.

Figure 2 (a) shows the time mean difference between the estimated pCO_2 and pCO_2 from SOCAT v5 data used for validation $mean_t\left(pCO_{2,i,j,FFNN} - pCO_{2,i,j,SOCAT}\right)$. Large differences occurred at high latitudes, in equatorial regions, along the Gulf Stream and Kuroshio currents – the regions with strong horizontal gradients of pCO_2. Moreover, the standard deviation of residuals (Figure 2 (b)) 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. Southern Ocean) and/or high kinetic energy (e.g. Southern Ocean, Kuroshio and Gulf Stream currents) (Fig. 1 (a)). At the scale of ocean regions, (Table 1) the largest RMSE and MAE were computed for the Pacific Subpolar ocean (RMSE = 34.77 µatm, MAE = 23.12 µatm), while 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 coverage (e.g. Subtropical North Pacific: RMSE = 15.86 µatm, MAE = 9.9 µatm, r^2 = 0.77 or Subpolar Atlantic: RMSE = 22.99 µatm, MAE = 15.04 µatm, 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 (RMSE = 15.73 µatm, MAE = 10.33 µatm, r^2 = 0.79). Scores are low in the Southern Hemisphere (Table 1) and time mean differences are large (Fig. 2 (a)) reflecting sparse data coverage (Fig. 1 (a)).

3.2. Qualification.

This section presents the assessment of the final time series of reconstructed surface ocean pCO_2. The time series was computed using the best monthly models as described in section 2.2, as well as 100% of 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 pCO2 Mapping Intercomparison" (SOCOM) exercise presented in Rödenbeck et al. (2015) (http://www.bgc-jena.mpg.de/SOCOM/). These methods are: (1) Jena-MLS oc_v1.5 (Rödenbeck et al., 2014), a statistical interpolation scheme (data-driven mixed-layer scheme; principal drivers used in parametrisation: ocean-internal carbon sources/sinks, SST, wind speed, mixed-layer depth climatology, alkalinity climatology); (2) JMA-MLR (updated version up to 2016) (Iida et al., 2015), based on multi-linear regressions with SST, SSS and Chl *a* as independent variables, and (3) ETH-SOMFFN v2016 (Landschützer et al., 2014), a two-step neural network model with SST, SSS, MLD, Chl *a*, xCO_2 as drivers. The time series of pCO_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*, ice fraction, 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 the mapping methods between each other, respectively to observations. The following diagnostics were computed: (1) the relative interannual variability (IAV) mismatch R^{iav} (in %) and (2) the amplitude of interannual variations. The relative interannual variability (IAV) mismatch R^{iav} (in %) is the ratio of the mismatch amplitude M^{iav} of the difference between the model output and observations (its temporal standard deviation) and the mismatch amplitude $M^{iav}_{benchmark}$ of the "benchmark". The latter was derived from the mean seasonal cycle of the corresponding model output where the trend of increasing yearly atmospheric pCO_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.

$$R^{iav} = \frac{M^{iav}}{M^{iav}_{benchmark}} * 100\% \quad , (6)$$

where

$$M^{iav} = std\left(mean\left(pCO_{2,Model} - pCO_{2,SOCAT}\right)\right) \quad ,$$

$$M^{iav}_{benchmark} = std\left(mean\left(D_{season}\right)\right) \quad ,$$

where “mean” is a mean over the region and year and

$$D_{season} = \left(pCO_{2,SS} + trend\left(CO_{2,atm} \right) \right) - pCO_{2,SOCAT} ,$$

$pCO_{2,SS}$ is the seasonal cycle of pCO_2 from the corresponding mapping method. $CO_{2,atm}$ estimates from *x*CO_2 Jena CO_2 inversion s76_v4.1 were used.

R^{iav} provides information on the capability of each method to reproduce the IAV compared to observations: a smaller R^{iav} stands for better fit compared to 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 pCO_2 over the period 2001-2016 are presented in Figure 4 for LSCE-FFNN and the three other models. Surface ocean pCO_2 (µatm) varied between the 4 mapping methods in the range of ±7 µatm (Fig. 4 (a)). Modeled pCO_2 values were at the lower end for ETH-SOMFFN and JMA-MLR, while LSCE-FFNN and Jena-MLS13 computed higher values. The same behavior was found for 12-month running mean time series (Fig. 4 (b)). Figure 4 (c) shows the 12-month running mean of the difference between computed pCO_2 and SOCAT data (model – SOCAT) over the globe. JMA-MLR mostly underestimated observed pCO_2 with a strong interannual variability of the misfit, especially at the end of the period with up to -5 µatm. The difference between ETH-SOMFFN output and SOCAT data fluctuated in the range of ±1 µatm, with an increase in amplitude up to -2 µatm from 2010 onward. Jena-MLS13 overestimated observations with the difference in the range of 0-1 µatm. The difference between LSCE-FFNN and SOCAT varies around zero between -0.7 and 1 µatm.

The model was assessed next at biome scale. Results for all biomes are presented in the supplementary material (Fig. S2, S3, S4). Two biomes with contrasting dynamics are discussed hereafter in greater detail: (1) the Equatorial East Pacific (biome 6) characterized by a strong IAV of surface ocean pCO_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 Figure 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 pCO_2 values. These episodes could be identified in all model time series (Fig. 5 (a)) with reduced pCO_2 levels in 2004/2005 and 2006/2007 (weak El Niño), 2002/2003 and 2009/2010 (moderate El Niño), and 2015/2016 (strong El Niño). JMA-MLR (blue curve) tended to underestimate pCO_2 during weak El Niño events. It was 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. 5 (b), (d), (f)). The seasonal cycle in this biome is dominantly driven by temperature. Modeled seasonal variability showed a good agreement across the ensemble of methods (Fig. 5(b)) with an increase in spring-summer and a decrease in autumn-winter. However, the amplitude can be different by up to 10 µatm between different models. The seasonal amplitude of pCO_2 computed by JMA-MLR increased from smaller values at the beginning of the time series to higher ones in the middle of the period 2005-2012. The variability of seasonal amplitude was the highest for Jena-MLS13 in line with the 12-month running mean time series (Fig. 5 (d)). Again, similar seasonal amplitude and year-to-year variability of surface ocean pCO_2 were obtained with LSCE-FFNN and ETH-SOMFFN (Fig. 5 (b), (d)). The yearly pCO_2 mismatch (Fig. 5 (f)) shows that observed surface ocean pCO_2 was underestimated by JMA-MLR at the beginning and at the end of the period by up to -6 µatm, and overestimated during 2007-2011 by up to 8 µatm. Jena-MLS13 shows mostly positive differences in the range 0-2 µatm over the full period. LSCE-FFNN and ETH-SOMFFN vary around zero and between -2 – 2 µatm, being close to each other.

3.2.2. Sea-air CO_2 flux variability.

Sea-air exchange of CO_2 was estimated using the same gas exchange formulation (4) and wind data speed (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 6 (a) 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 pCO_2 variability (Fig. 4 (b)). 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: South Pacific, South Atlantic, Indian Ocean, South Ocean; see Fig. S5). Nevertheless, the relative IAV mismatch was less than 30% for all methods (Fig. 6 (b)), suggesting a reasonable fit to observational data. The relative IAV mismatch is, however, 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 a flux of ~-1.8 GtC yr^{-1} compared to ~-2.2 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 0.2-0.36 PgC/yr. The weighted mean of IAV (horizontal line in Fig. 6 (b)) computed from the four methods included here was 0.25 PgC/yr. This value is close to the one of Rödenbeck et al. (2015) for the complete ensemble of SOCOM models (0.31 PgC/yr) estimated for the period 1992-2009. The largest amplitude was obtained for ETH-SOMFFN, ~0.35 PgC/yr. On the other hand, LSCE-FFNN has the smallest amplitude with 0.21 PgC/yr. Jena-MLS13 and JMA-MLR lie very close to the weighted mean value with 0.26 PgC/yr and 0.22 PgC/yr, 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 for comparison Rödenbeck et al., 2015). 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 a good agreement for biome 6 (East Pacific Equatorial, Fig. 7 (a)). 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 on pCO_2 variability (Fig. 5 (a)). Despite this, Jena-MLS13 has a low relative R^{IAV} (26%), which confirms 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. 7 (a), (c)) with relative IAV mismatches of 46% and 53%, respectively, and with amplitudes ~ 0.03 PgC/yr. Interannual variability reproduced by JMA-MLR falls within the range of the other models (Fig. 7 (c)), but with a R^{IAV} of ~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 variability. The two models based on a neural network show again a good agreement with R^{IAV} of 17% for LSCE-FFNN and 20% for ETH-SOMFFN. Jena-MLS13 produced a strong seasonal variability (Fig. 7 (b)) up to 0.06 PgC/yr, and small R^{IAV} of ~11%. Contrary to the other approaches, JMA-MLR did not reproduce a decrease in sea-air CO_2 at the middle of the period by up to 0.02 PgC/yr (Fig. 7 (b)). The model is characterized by a R^{IAV} of 46% and an amplitude of 0.013 PgC/yr.

3.3.3. 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 time scales, such as for the period 2001-2016, 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_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 the range: Jena-MLS13, -0.0012 PgC/yr/yr (-0.0028 PgC/yr/yr, total value without significant t-test, Fig. S8); LSCE-FFNN, -0.00087 PgC/yr/yr (-0.0032 PgC/yr/yr); JMA-MLR, -0.0013 PgC/yr/yr (-0.0037 PgC/yr/yr); ETH-SOMFFN, -0.0025 PgC/yr/yr (-0.0059 PgC/yr/yr). LSCE-FFNN computed negative trends over most of the Atlantic basin, 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. Regional maxima and minima are, however, more pronounced in Jena-MLS13 (Fig. 8 (b)) and ETH-SOMFFN (Fig. 8 (c)), while a patchy distribution at sub-basin scale is diagnosed for JMA-MLR.

The agreement in sign of computed linear trends from four models is presented in Fig. 9 (total linear trends without 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), on the other hand a positive trend was computed for JMA-MLR (0.0004 PgC/yr/yr, and with t-test: 0.00006 PgC/yr/yr) while the three other 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 disagree about its amplitude and spatial distribution. Almost everywhere over the Atlantic Ocean the mapping methods produced the same sign of linear trend (Fig. 9). Only in the eastern part of the subtropical North Atlantic Jena-MLS13 gave a positive linear trend of fCO_2 (Fig. 8 (b)).

According to LSCE-FFNN, the global ocean took up in average 1.55 PgC/yr between 2001-2015.This estimate is consistent with results from the other three models (Table 3) (see Table S2 for estimations per biomes). The spread between individual models falls in the range of the error reported in Landschützer et al. (2016), ±0.4-0.6 PgC/yr. Per biome, estimates of CO_2 sea-air fluxes provided by LSCE-FFNN are similarly 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 pCO_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 pCO_2 climatology. It allows to keep the output of the FFNN close to the observed values in regions with poor data cover. At the second step, pCO_2 anomalies are reconstructed with respect to the climatology from the first step. The model was applied over the period 2001-2016. Validation with independent data at global scale indicated a RMSE of 17.57 µatm, r^2 of ~0.76 and an absolute bias of 11.52 µatm. 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), JMA-MLR (linear regression) (Rödenbeck et al., 2015). Network qualification followed the protocol and diagnostics proposed in Rödenbeck et al. (2015).

Reconstructed surface ocean pCO_2 distributions were in good agreement with other models and observations. The seasonal variability was reproduced satisfyingly by LSCE-FFNN, the yearly pCO_2 mismatch varied around zero, and relative IAV mismatch was 7%. LSCE-FFNN proved skillful in reproducing the interannual variability of surface ocean pCO_2 over the Eastern Equatorial Pacific in response to ENSO. Reductions in surface ocean pCO_2 during El Niño events were well reproduced. The comparison between reconstructed and observed pCO_2 values yielded a RMSE of 15.73 µatm, 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 ~17%. Despite an overall good agreement between models, important differences still exist at the regional scale, especially in the Southern Hemisphere and in particular, the Southern Pacific and the Indian Ocean. These regions suffer from poor data-coverage. Large regional uncertainties in reconstructed surface ocean pCO_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 pCO_2 climatology reconstruction, 1st step of LSCE-FFNN model, python code for reconstruction of pCO_2 anomalies, 2nd step of LSCE-FFNN model, are provided at the end of supplementary material.

Time series of reconstructed surface ocean pCO_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: MULTIOBS.

Author contribution.

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

Acknowledgments.

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. This study was funded by the AtlantOS project (EU Horizon 2020 research and innovation program, grant

agreement no. 2014-633211). MV also acknowledges funding by the CoCliServ and EUPHEME projects (ERA4CS program).

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. Cycles, 20, GB2017, doi: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. et al.: A multi-decade record of high-quality fCO_2 data in version 5043 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 Univ. Press, Cambridge, U. K., 1995.

Bishop, C.M.: Pattern Recognition and Machine Learning, Springer, 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, doi: 10.3389/fmars.2018.00328, 2018.

Bricaud, A., Mejia, C., Blondeau-Patissier, D., Claustre, H., Crepon, M., and Thiria, S.: Retrieval of pigment concentrations and size structure of algal populations from the absorption spectra using multilayered perceptrons, Appl. Opt., 46, 8, 1251-1260, 2006.

Chollet, F., et al.: Keras. https://keras.io, 2015.

Ciais P., Sabine C., Bala G. et al.: 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[Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley (eds.)]. 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 Sensing of the Environment, doi: 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, doi: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, doi: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.: Basin-scale pCO_2 maps estimated from ARGO float data: A model study, J.

542Geophys. Res., 114, C10012, doi:10.1029/2009JC005322, 2009a.
543
544Friedrich, T., and Oschlies, A.: Neural network-based estimates of North Atlantic surface pCO_2 from
545satellite data: A methodological study, J. Geophys. Res., 114, C03020, doi: 10.1029/2007JC004646, 2009b.
546
547Gross, L., Thiria, S., Frouin, R., and Mitchell, B.G.: Artificial neural networks for modeling transfer
548function between marine reflectance and phytoplankton pigment concentration, J. Geophys. Res., 105, C2,
5493483-3949, doi: 10.1029/1999jc900278, 2000.
550
551Guinehut S., Dhomps, A.-L.,, Larnico, G.,l and Le Traon, P.-Y.:: High resolution 3D temperature and
552salinity fields derived from in situ and satellite observations, *Ocean Sci*., 8(5):845–857, 2012.
553
554Hinton, G., Srivastava, N., and Swersky, K.: Lecture 6a: Overview of mini-batch gradient descent. Neural
555Networks for Machine Learning. Slides:
556http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf, 2012.
557
558Iida, Y., Kojima, A., Takatani, Y., Nakano, T., Midorikawa, T., and Ishii, M.: Trends in pCO_2 and sea-air
559CO_2 flux over the global open oceans for the last two decades, J. Oceanogr., 71, 637–661,
560doi:10.1007/s10872-015-0306-4, 2015.
561
562Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White,
563G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C.,
564Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis
565project, B. Am. Meteorol. Soc., 77, 437–471, 1996.
566
567Kallache, M., Vrac, M., Naveau, P., Michelangeli, P.-A.: Non-stationary probabilistic downscaling of
568extreme precipitation, J. Geophys. Res. - Atmospheres, 116, D05113, doi:10.1029/2010JD014892, 2011.
569
570Körtzinger, A.: Methods of Seawater Analysis, chap. Determination of carbon dioxide partial pressure
571(pCO_2), 149–158, Verlag Chemie, 1999.
572
573Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., Sasse, T. P., and
574Zeng, J.: A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean
575carbon sink, Biogeosciences, 10, 7793–7815, https://doi.org/10.5194/bg-10-7793-2013, 2013.
576
577Landschützer, P., Gruber, N., Bakker, D. C. E., and Schuster, U.: Recent variability of the global ocean
578carbon sink, Global Biogeochem. Cy., 28, 927–949, https://doi.org/10.1002/2014GB004853, 2014.
579
580Landschützer, P., Gruber, N. & Bakker, D. C. E.: Decadal variations and trends of the global ocean carbon
581sink, Glob. Biogeochem. Cycles, 30, 1396–1417, https://doi.org/10.1002/2015GB005359, 2016.
582
583Laruelle, G. G., Landschützer, P., Gruber, N., Tison, J.-L., Delille, B., and Regnier, P.: Global high-
584resolution monthly pCO_2 climatology for the coastal ocean derived from neural network interpolation,
585Biogeosciences, 14, 4545-4561, https://doi.org/10.5194/bg-14-4545-2017, 2017.
586
587Lefèvre, N., Watson, A. J., and Watson, A. R.: A comparison of multiple regression and neural network
588techniques for mapping *in situ* pCO_2 data, Tellus, 57B, 375–384,
589https://doi.org/10.3402/tellusb.v57i5.16565, 2005.
590
591Le Quéré, C., and Coauthors: Two decades of ocean CO_2 sink and variability, Tellus, 55B, 649–656,
592https://doi.org/10.1034/j.1600-0889.2003.00043.x, 2003.
593
594Le Quéré, C., Takahashi, T., Buitenhuis, E. T., Rödenbeck, C., and Sutherland, S. C.: Impact of climate
595change and variability on the global oceanic sink of CO_2, Glob. Biogeochem. Cy., 24, GB4007,

596doi:10.1029/2009GB003599, 2010.
597
598Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., et al.: Global
599carbon budget 2017, Earth System Science Data, 10 (1), 405–448. https://doi.org/10.5194/essd-10-405-
6002018, 2018.
601
602Majkut, J. D., Carter, B.R., Frölicher, T.L., Dufour, C.O., Rodgers, K.B., and Sarmiento, J.L.: An observing
603system simulation for Southern Ocean carbon dioxide uptake, Philos. Trans. Roy. Soc. London, A372,
60420130046, doi:https://doi.org/10.1098/rsta.2013.0046, 2014.
605
606Mejia, C., Thiria, S., Tran, N., and Crepon, M.: Determination of the Geophysical Model Function of ERS1
607Scatterometeer by the use of Neural Networks, J. Geophys. Res., Vol. 103, No. C6, PP 12853-12868,
608https://doi.org/10.1029/97JC02178, 1998.
609
610Menemenlis, D., Campin, J., Heimbach, P., Hill, C., Lee, T., Nguyen, A., Schodlok, M., and Zhang, H.:
611ECCO2: High resolution global ocean and sea ice data synthesis, Mercator Ocean, Quarterly Newsletter,
61231, 13–21, 2008.
613
614Nakaoka, S., Telszewski, M., Nojiri, Y., Yasunaka, S., Miyazaki, C., Mukai, H., and Usui, N.: Estimating
615temporal and spatial variation of ocean surface pCO2 in the North Pacific using a self-organizing map
616neural network technique, Biogeosciences, 10, 6093–6106, doi:10.5194/bg-10-6093-2013, 2013.
617
618Organelli, E., Barbieux, M., Claustre, H., Schmechtig, C., Poteau, A., Bricaud, A., Boss, E., Briggs, N.,
619Dall'Olmo, G., D'Ortenzio, F., Leymarie, E., Mangin, A., Obolensky, G., Penkerc'h, C., Prieur, L., Roesler,
620C., Serra, R., Uitz, J., and Xing, X.: Two databases derived from BGC-Argo float measurements for marine
621biogeochemical and bio-optical applications, Earth Syst. Sci. Data, 9, 861-880,
622https://doi.org/10.5194/essd-9-861-2017, 2017.
623
624Orr, J. C., Monfray, P., Maier-Reimer, E., Mikolajewicz, U., Palmer, J., Taylor, N. K., Toggweiler, J. R.,
625Sarmiento, J. L., Quere, C. L., Gruber, N., Sabine, C. L., Key, R. M. and Boutin, J.: Estimates of
626anthropogenic carbon uptake from four three-dimensional global ocean models, Global Biogeochem. Cycl.,
62715, 43–60, https://doi.org/10.1029/2000GB001273, 2001.
628
629Peylin, P., Bousquet, P., Le Quéré, C., Sitch, S., Friedlingstein, P., McKinley, G., Gruber, N., Rayner, P.,
630and Ciais, P.: Multiple constraints on regional CO_2 flux variations over land and oceans, Glob.
631Biogeochem. Cycles, 19, GB1011, https://doi.org/10.1029/2003GB002214, 2005.
632
633Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K.,
634Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric
635carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699-6720,
636https://doi.org/10.5194/bg-10-6699-2013, 2013.
637
638Rödenbeck, C.: Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a
639global inversion of atmospheric transport, Technical Report 6, Max Planck Institute for Biogeochemistry,
640Jena, available at: http://www.bgc-jena.mpg.de/uploads/Publications/TechnicalReports/tech_report6.pdf,
6412005.
642
643Rödenbeck, C., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., Cassar, N., Reum, F., Keeling, R. F., and
644Heimann, M.: Interannual sea-air CO_2 flux variability from an observation-driven ocean mixed-layer
645scheme, Biogeosciences, 11, 4599–4613, doi:10.5194/bg-11-4599-2014, 2014.
646
647Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., Landschützer, P. et al.:
648Data-based estimates of the ocean carbon sink variability—first results of the Surface Ocean pCO_2

649Mapping intercomparison (SOCOM), Biogeosciences, 12, 7251–7278, https://doi.org/10.5194/bg-12-7251-
6502015, 2015.
651
652Rodgers, K. B., Key, R. M., Gnanadesikan, A., Sarmiento, J. L., Aumont, O., Bopp, L. et al.: Using
653altimetry to help explain patchy changes in hydrographic carbon measurements, J. Geophys. Res., 114,
654C09013, doi:10.1029/2008JC005183, 2009.
655
656Roobaert, A., Laruelle, G. G., Landschützer, P., and Regnier, P.: Uncertainty in the global oceanic CO2
657uptake induced by wind forcing: quantification and spatial analysis, Biogeosciences, 15, 1701-1720,
658https://doi.org/10.5194/bg-15-1701-2018, 2018.
659
660Rumelhart, D.E., Hinton, G.E., & Williams, R.J.: Learning internal representations by backpropagating
661errors, Nature, 323, 533–536, 1986.
662
663Sauzède, R., Claustre, H., Uitz, J., Jamet, C., Dall'Olmo, G., D'Ortenzio, F., Gentili, B., Poteau, A., and
664Schmechtig, C.: A neural network-based method for merging ocean color and Argo data to extend surface
665bio-optical properties to depth: Retrieval of the particulate backscattering coefficient, J. Geophys. Res.
666Oceans, 121, 2552–2571, doi:10.1002/2015JC011408, 2016.
667
668Schuster, U., McKinley, G. A., Bates, N., Chevallier, F., Doney, S. C., Fay, A. R., González-Dávila, M.,
669Gruber, N., Jones, S., Krijnen, J., Landschützer, P., Lefèvre, N., Manizza, M., Mathis, J., Metzl, N., Olsen,
670A., Rios, A. F., Rödenbeck, C., Santana-Casiano, J. M., Takahashi, T., Wanninkhof, R., and Watson, A. J.:
671An assessment of the Atlantic and Arctic sea–air CO2 fluxes, 1990–2009, Biogeosciences, 10, 607-627,
672https://doi.org/10.5194/bg-10-607-2013, 2013.
673
674Takahashi, T., Sutherland, S.C., Wanninkhof, R., Sweeney, C., Feely, R.A., Chipman, D.W., Hales, B.,
675Friederich, G., Chavez, F., Sabine, C., et al.: Climatological mean and decadal change in surface ocean
676pCO_2, and net sea-air CO_2 flux over the global oceans, Deep.-Sea Res. II, 56(8–10), 554–577,
677https://doi.org/10.1016/j.dsr2.2008.12.009, 2009.
678
679Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N. et al.: Global sea-air CO_2 flux based
680on climatological surface ocean pCO_2, and seasonal biological and temperature effects, Deep.-Sea Res. II,
68149, 1601–1622, https://doi.org/10.1016/S0967-0645(02)00003-6, 2002.
682
683Telszewski, M., Chazottes, A., Schuster, U., Watson, A. J., Moulin, C., Bakker, D. C. E., González-Dávila,
684M., Johannessen, T., Körtzinger, A., Lüger, H., Olsen, A., Omar, A., Padin, X. A., Ríos, A. F., Steinhoff, T.,
685Santana-Casiano, M., Wallace, D.W.R., and Wanninkhof, R.: Estimating the monthly pCO_2 distribution in
686the North Atlantic using a self-organizing neural network, Biogeosciences, 6, 1405–1421, doi:10.5194/bg-
6876-1405-2009, 2009.
688
689Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean, J. Geophys. Res.-
690Oceans, 97, 7373–7382, https://doi.org/10.1029/92JC00188, 1992.
691
692Weiss, R.: Carbon dioxide in water and seawater: the solubility of a non-ideal gas, Mar. Chem., 2, 203–205,
693https://doi.org/10.1016/0304-4203(74)90015-2, 1974.
694
695Williams, N.L., Juranek, L.W., Feely, R.A., Johnson, K.S., Sarmiento, J.L., Talley, L.D., Dickson, A.G.,
696Gray, A.R., Wanninkhof, R., Russell, J.L., Riser, S.C., Takeshita, Y.: Calculating surface ocean pCO_2 from
697biogeochemical Argo floats equipped with pH: An uncertainty analysis, Global Biogeochemical Cycles,
69831:591-604, https://doi.org/10.1002/2016GB005541, 2017.
699
700Zeng, J., Nojiri, Y., Landschützer, P., Telszewski, M., and Nakaoka, S.: A global surface ocean fCO_2
701climatology based on a feed-forward neural network, J. Atmos. Ocean Technol., 31, 1838–1849,
702https://doi.org/10.1175/JTECH-D-13-00137.1, 2014.

703

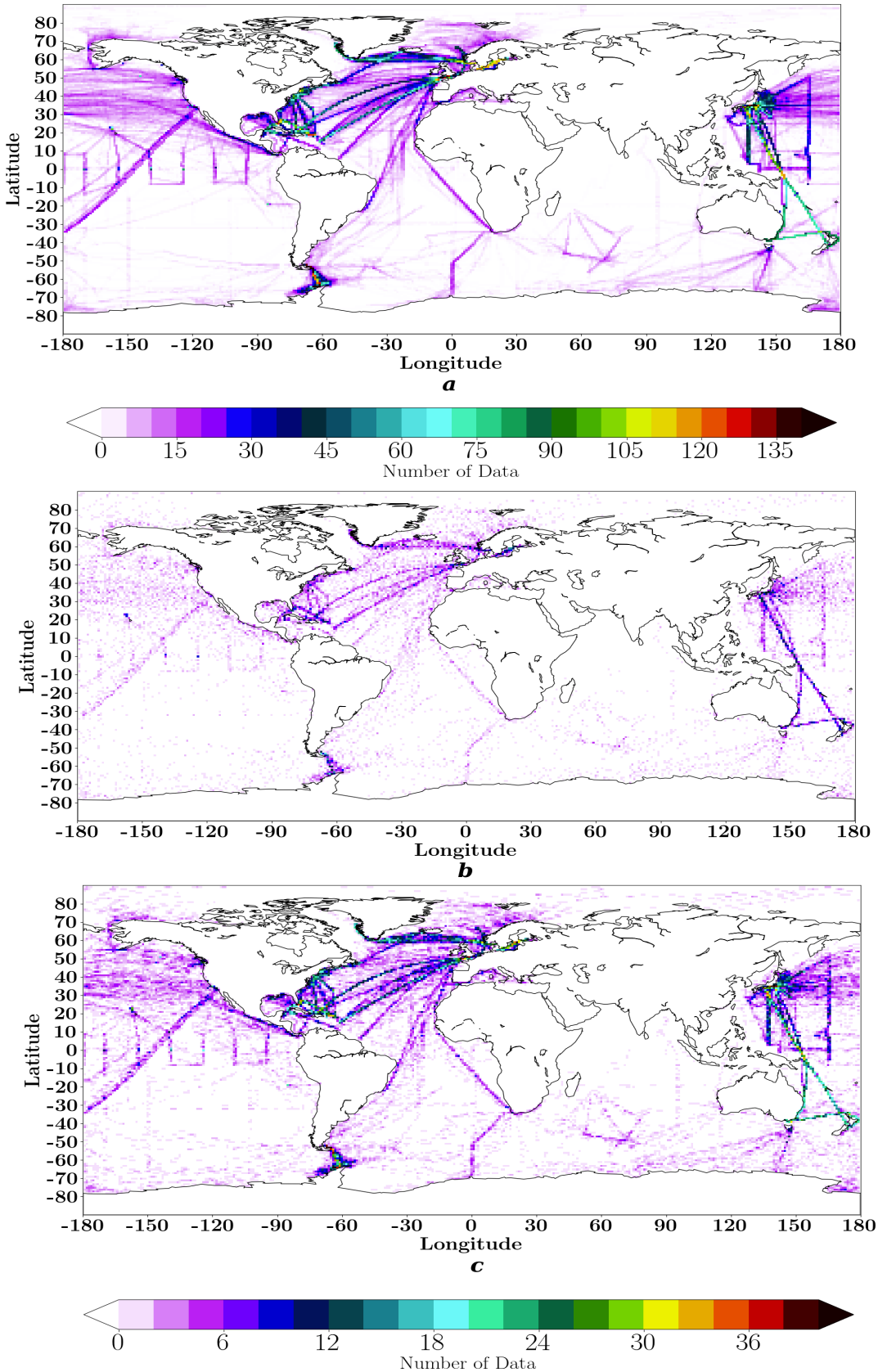

Figure 1: Spatial distribution of SOCAT data (number of measurements per grid point): (a) - period 2001-2016; (b) - all months of January for period 2001-2016; (c) - all months of December-January-February for period 2001-2016.

705
706
707
708

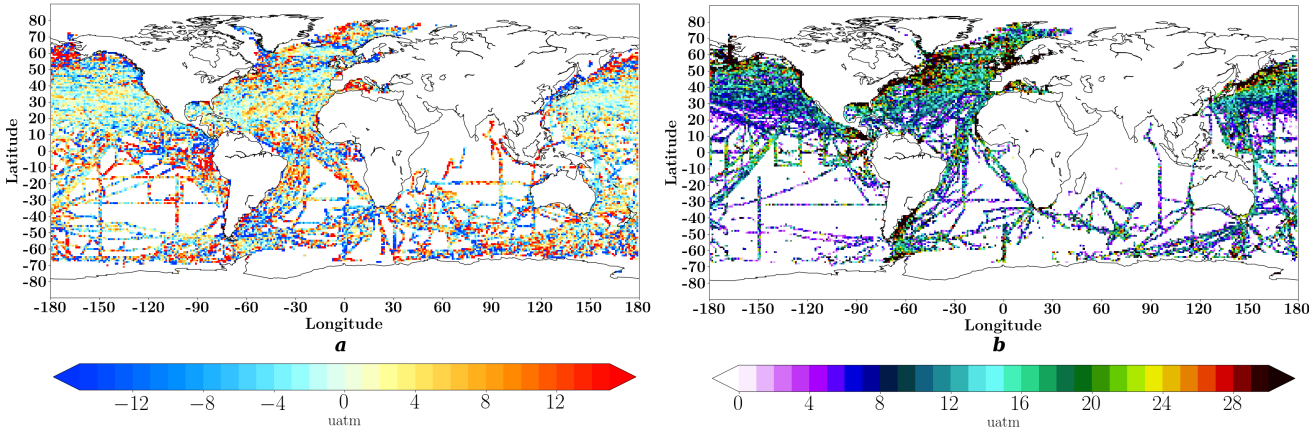

Figure 2: Time mean differences (µatm) between monthly LSCE-FFNN pCO_2 and SOCAT pCO_2 data used for evaluation of the model over the period 2001-2016 (a) and its std (b).

710

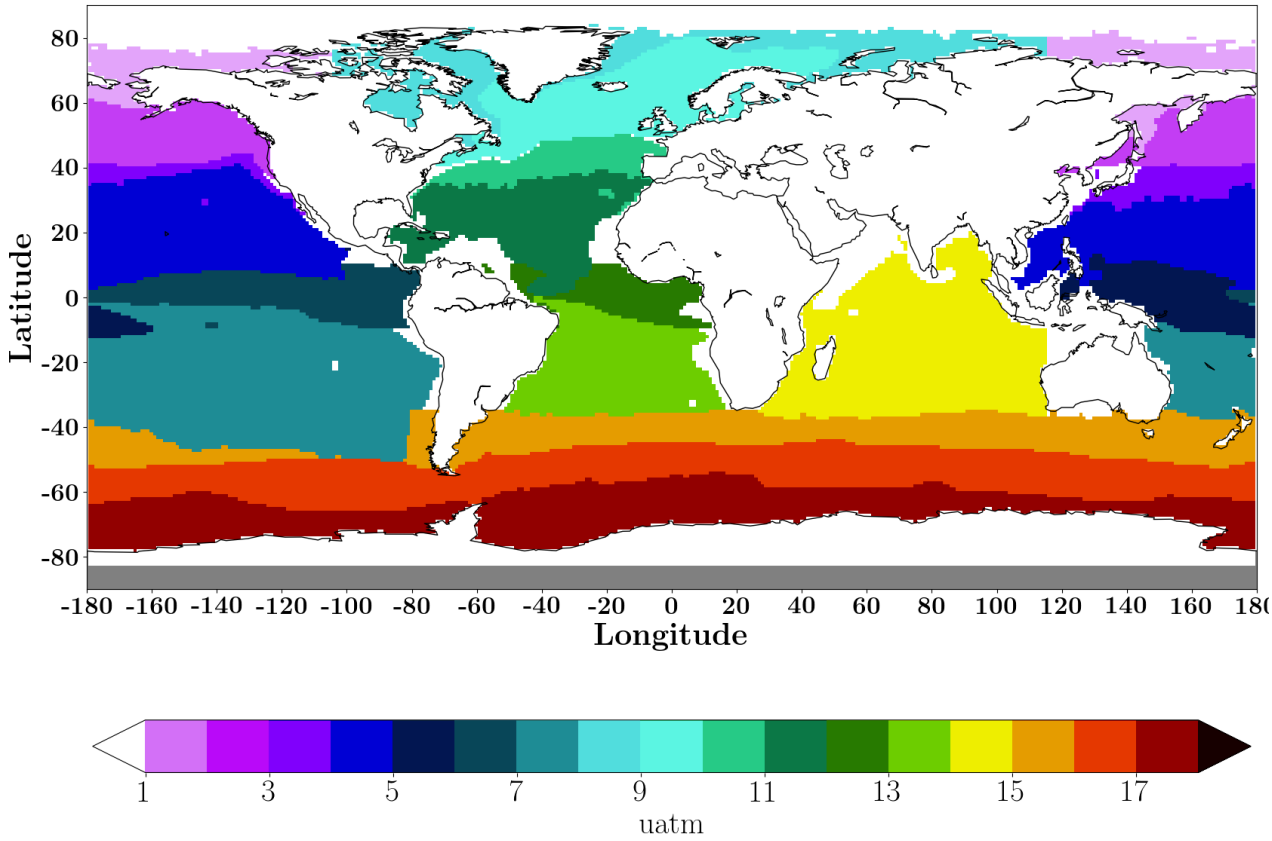

Figure 3: Map of biomes (after Rodenbeck et al. (2015); and Fay and McKinley (2014)) used for comparison. See table 2 for biome names.

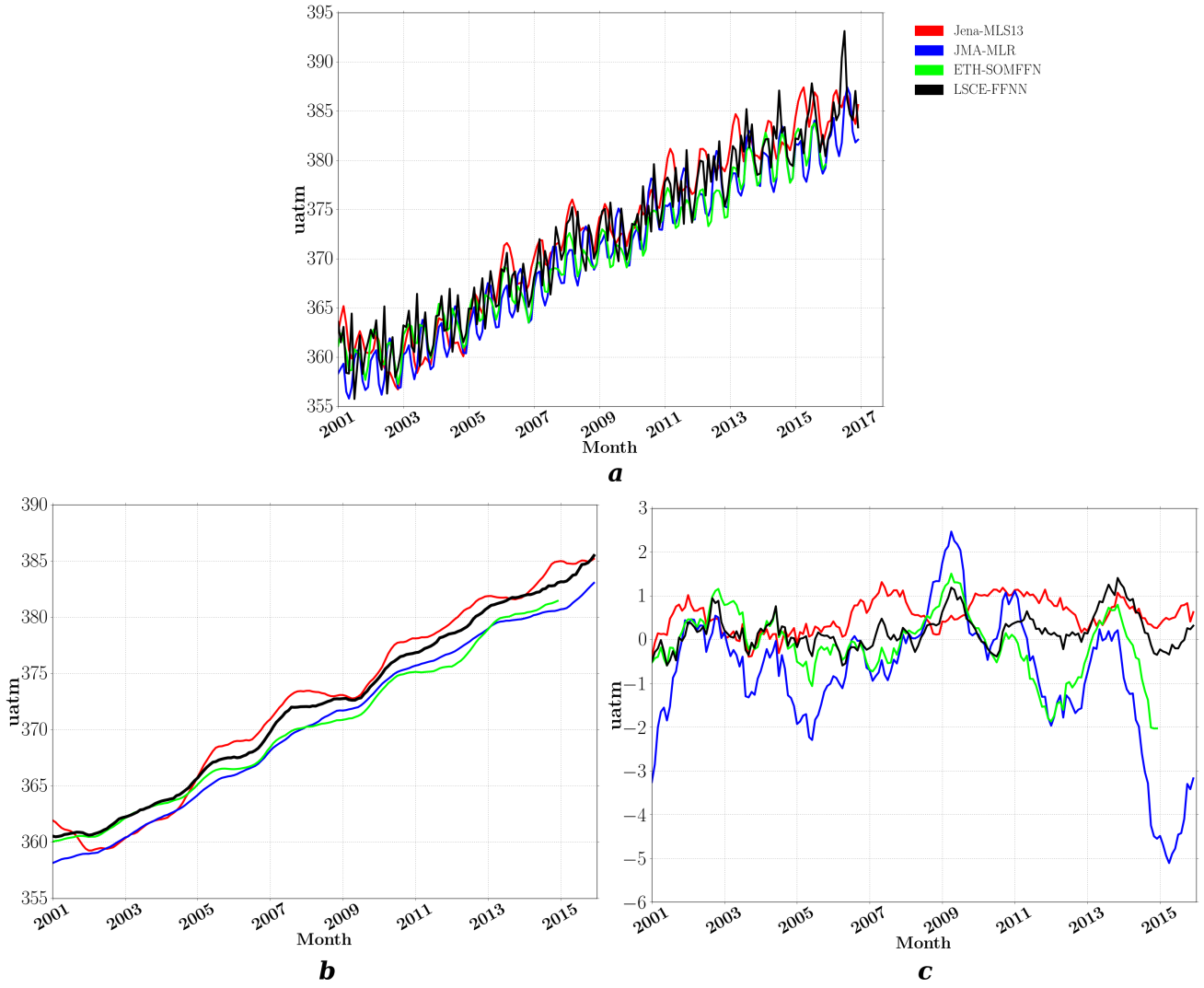

Figure 4: Global oceanic pCO_2: black - LSCE-FFNN, blue - JMA, red - Jena, green - ETH-SOMFFN; (a) – global average monthly time series, (b) – global 12-month running mean average, (c) - yearly pCO_2 mismatch (difference of mapping methods and SOCAT data).

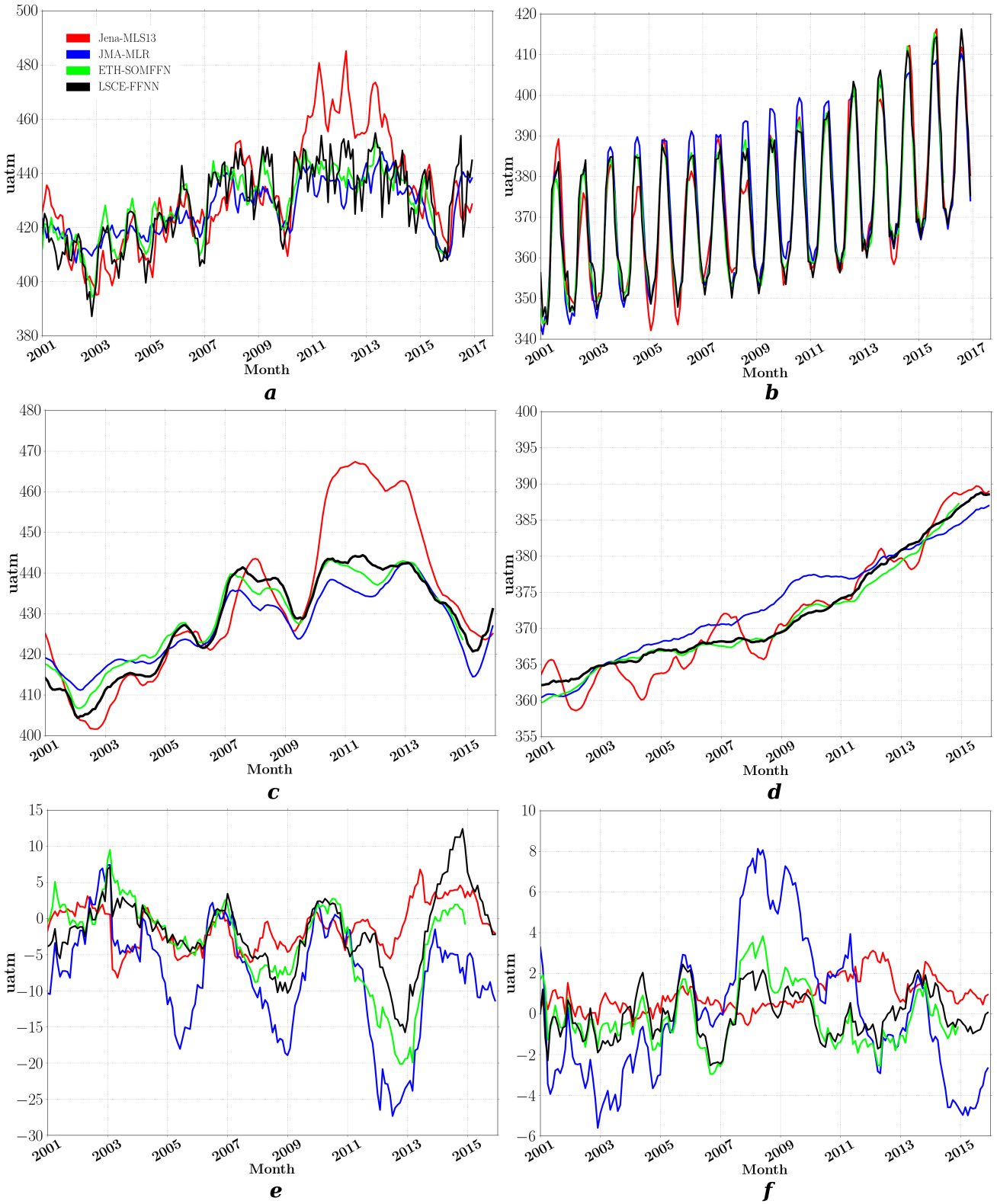

Figure 5: Surface ocean pCO_2: Equatorial East Pacific (biome 6) (left) and Subtropical Permanently Stratified North Atlantic (biome 11) (right): black – FFNN, blue – JMA, red – Jena, green – ETH-SOMFFN; (a), (b) – monthly time series averaged over biome; (c), (d) – 12-month running mean averaged over biome; (e), (f) – yearly pCO_2 mismatch (difference of mapping methods and SOCAT data).

715

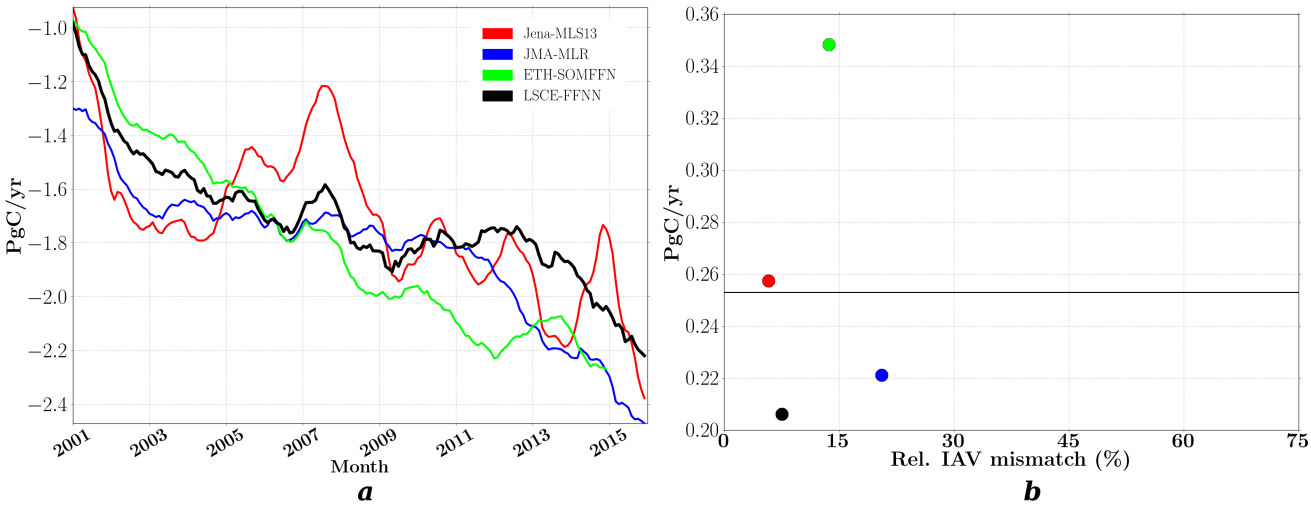

Figure 6: (a) – Interannual global ocean sea-air CO_2 flux (12-month running mean); (b) – amplitude of interannual CO_2 flux plotted against the relative IAV mismatch amplitude. The weighted mean is given as a horizontal line.

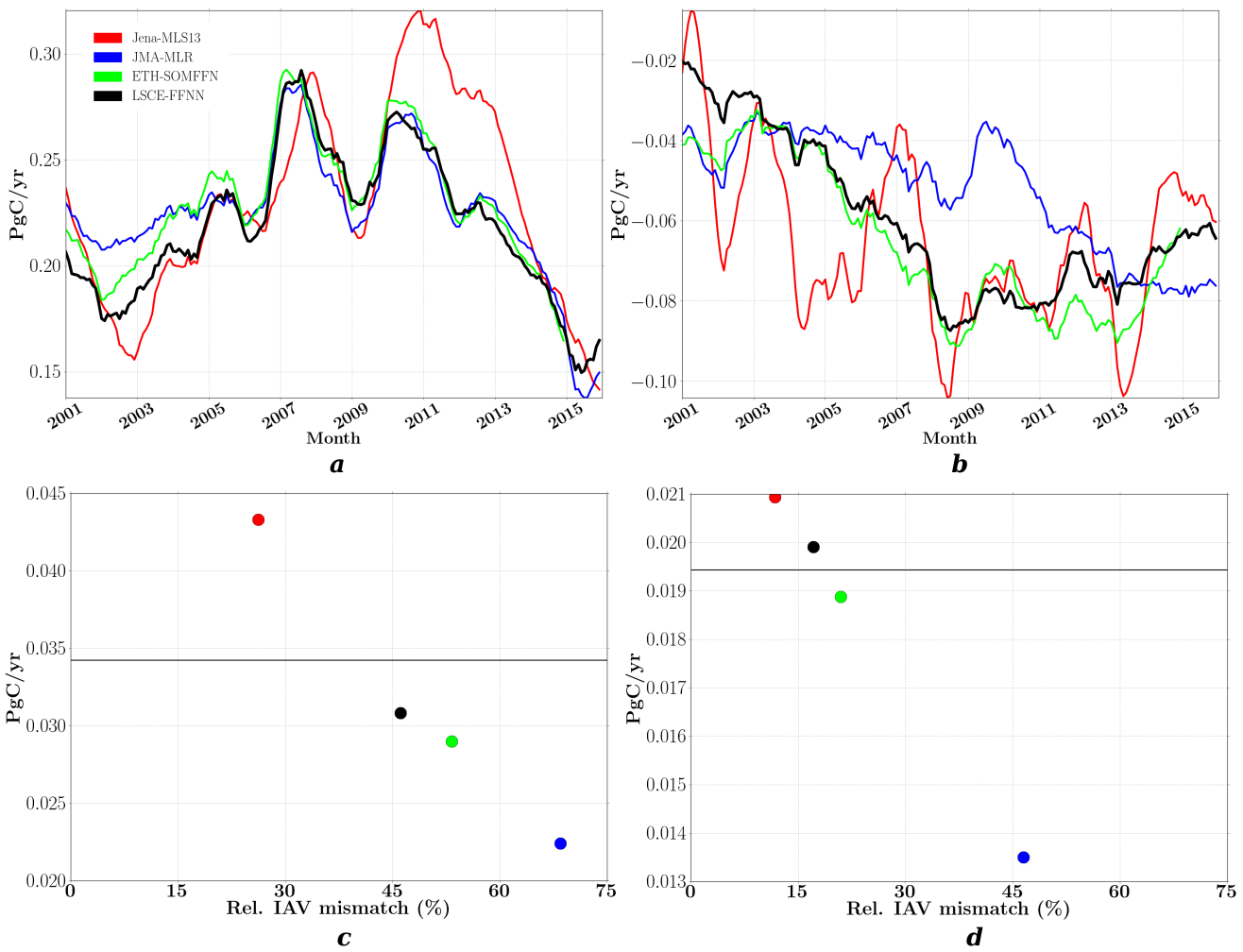

Figure 7: Global ocean interannual sea-air CO_2 flux (12-month running mean): (a) Equatorial East Pacific (biome 6) and (b) Subtropical Permanently Stratified North Atlantic (biome 11). Amplitude of interannual CO_2 flux plotted against the relative IAV mismatch amplitude: (c) Equatorial East Pacific (biome 6) (left) and (d) Subtropical Permanently Stratified North Atlantic (biome 11). The weighted mean is given as a horizontal line.

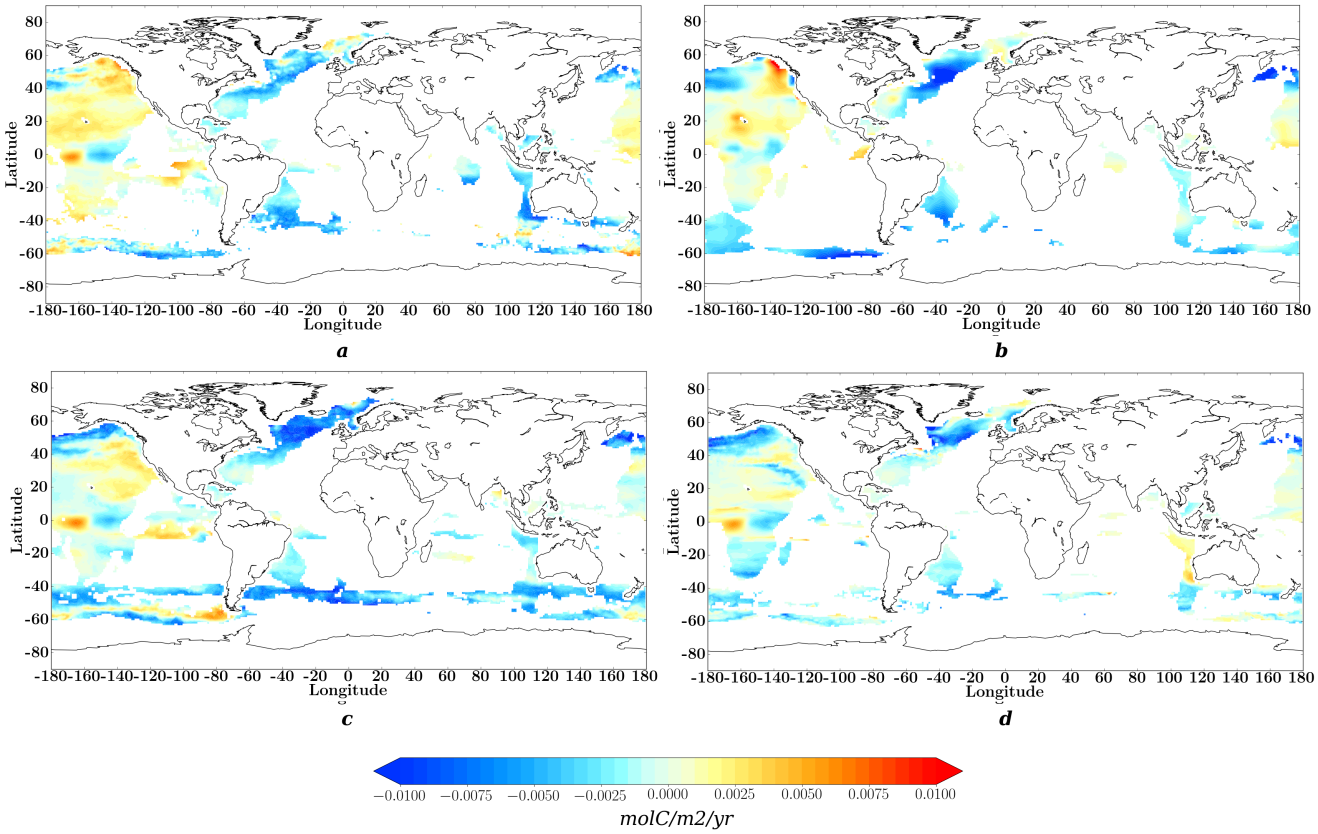

Figure 8: Significant (p_val = 0.05) linear trend of fCO_2 for common period 2001-2015: (a) – LSCE-FFNN; (b) – Jena-MLS13; (c) – ETH-SOMFFN; (d) – JMA-MLR.

721

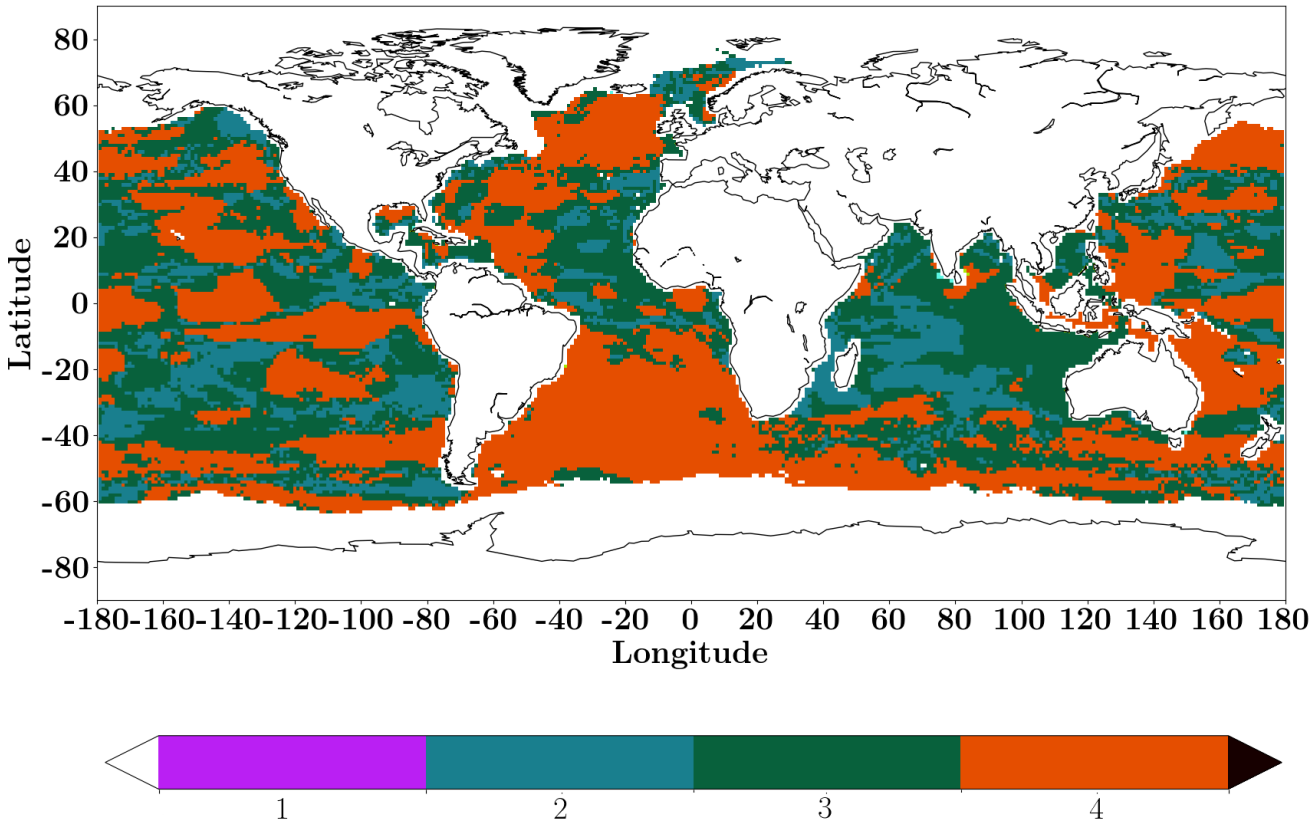

Figure 9: Agreement between four mapping methods in their linear trend of sea-air CO_2 flux. Color-bar represents the number of products that have the same sign of linear trend.

Table 1: Statistical validation of LSCE-FFNN. Comparison between reconstructed surface ocean pCO_2 and pCO_2 values from SOCAT v5 database not used in the training algorithm for the period 2001-2016 over the

745global ocean (except for regions with ice-cover) and for large oceanographic regions. In round brackets:
746number of measurements per region

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

747
748Table 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

Table 3: Mean of sea-air CO_2 flux (PgC/yr) over the Global Ocean and per regions for period in common (2001-2015). Averages over the period 2001-2009 are presented between brackets. The last column presents a comparison to 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	76ºN to 90ºN	-0.001	-0.001	-0.001	-0.001	-0.12±0.06
Subpolar Atlantic	49ºN to 76ºN	-0.15 (-0.15)	-0.14 (-0.12)	-0.15 (-0.15)	-0.16 (-0.15)	-0.21±0.06
Subpolar Pacific	49ºN to 76ºN	-0.003 (-0.005)	-0.009 (-0.004)	-0.006 (-0.004)	-0.027 (-0.021)	---
Subtropical Atlantic	18ºN to 49ºN	-0.21 (-0.19)	-0.21 (-0.19)	-0.2 (-0.18)	-0.21 (-0.2)	-0.26±0.06
Subtropical Pacific	18ºN to 49ºN	-0.45 (-0.46)	-0.49 (-0.48)	-0.47 (-0.46)	-0.49 (-0.47)	---
Equatorial Atlantic	18ºS to 18ºN	0.085 (0.09)	0.085 (0.095)	0.08 (0.082)	0.1 (0.11)	0.12±0.04
Equatorial Pacific	18ºS to 18ºN	0.42 (0.41)	0.4 (0.4)	0.44 (0.42)	0.38 (0.37)	---
South Atlantic	44ºS to 18ºS	-0.17 (-0.16)	-0.18 (-0.16)	-0.18 (-0.17)	-0.23 (-0.22)	-0.14±0.04
South Pacific	44ºS to 18ºS	-0.33 (-0.34)	-0.4 (-0.39)	-0.35 (-0.34)	-0.49 (-0.47)	---
Indian Ocean	44S to 30N	-0.25 (-0.2)	-0.32 (-0.29)	-0.27 (-0.26)	-0.27 (-0.29)	---
Southern Ocean	90ºS to 44ºS	-0.38	-0.29	-0.36	-0.26	---