RMetS Royal Meteorological Society
Geoscience Data Journal
Open Access

Surface ocean CO_2 in 1990–2011 modelled using a feed-forward neural network

Jiye Zeng*, Yukihiro Nojiri, Shin-ichiro Nakaoka, Hideaki Nakajima and Tomoko Shirai

Centre for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan

**Correspondence: Jiye Zeng, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, 305-8506, Japan, E-mail: zeng@nies.go.jp*

Funding information: None.

This dataset includes the monthly distributions of CO_2 fugacity in the world surface oceans reconstructed using a feed-forward neural network model and the CO_2 measurements of the Surface Ocean CO_2 Atlas version 2.0. It has a spatial resolution of $1 \times 1°$ and spans a period of 22 years, from January 1990 to December 2011. The dataset also includes necessary parameters for the reconstruction and an estimate of the CO_2 fluxes between the ocean and the atmosphere. The aim of this work is to provide a dataset for estimating the oceans' contribution to the global carbon budget.

Geosci. Data J. ***2****: 47–51 (2015), doi: 10.1002/gdj3.26*

Received: 2 September 2014, revised: 8 June 2015, accepted: 10 June 2015

Key words: ocean, CO_2, neural network, model

Dataset

Identifier: doi: http://doi.pangaea.de/10.1594/PANGAEA.834398
Creator: Jiye Zeng
Title: Surface ocean CO_2 in 1990–2011 reconstructed using a feed-forward neural network
Publisher: NIES
Publication year: 2014
Resource Type: NetCDF
Version: 3.6.1

Introduction

Understanding the global distribution of the surface ocean CO_2 (SOC) fugacity plays an important role in accurately estimating the oceans' contribution to the global carbon budget, as indicated in Le Quéré *et al.* (2013) and Wanninkhof *et al.* (2013). However, available measurements are insufficient in most parts of the oceans for direct estimates of the oceans' contribution. Although the composite map of SOC measurements from 1990 to 2011, shown in Zeng *et al.* (2014), indicates that about 60% of the oceanic areas were sampled, the area ratio is only about 7–25% when divided by the same months of all years (Figure 1(a)) and is even smaller, between 0% and 4% (Figure 1(b)), when calculated for individual months. Insufficient measurements demand using models to estimate the global SOC in multiple years. However, the relatively large temporal and spatial variations in SOC raise the difficulties in SOC modelling. Recent studies showed that the amplitude of the seasonal SOC changes can be 100 μatm or more (Wanninkhof *et al.*, 2013), which is about 10 times of what has been observed for atmospheric CO_2 (e.g. Bacastow *et al.*, 1985); and that the spatial decorrelation length scales are on the order of 100 km (Li *et al.*, 2005) to 400 km (Jones *et al.*, 2012), which is smaller than that of the marine atmosphere. As a result, many works on modelling SOC focused on a mesoscale (e.g. Zeng *et al.*, 2002; Lèfevre *et al.*, 2005; Sarma *et al.*, 2006; Jamet *et al.*, 2007; Friedrich and Oschlies, 2009; Telszewski *et al.*, 2009; Takamura *et al.*, 2010; Landschützer *et al.*, 2013; Nakaoka *et al.*, 2013; Schuster *et al.*, 2013); and on the global scale, mapping SOC was confined to the climatology in a given year (e.g. Takahashi *et al.*, 2009; Zeng *et al.*, 2014).

This work extends the method of Zeng *et al.* (2014) to reconstruct the global monthly SOC in 1990–2011. The resulting dataset not only has a finer spatial resolution ($1° \times 1°$) compared to the

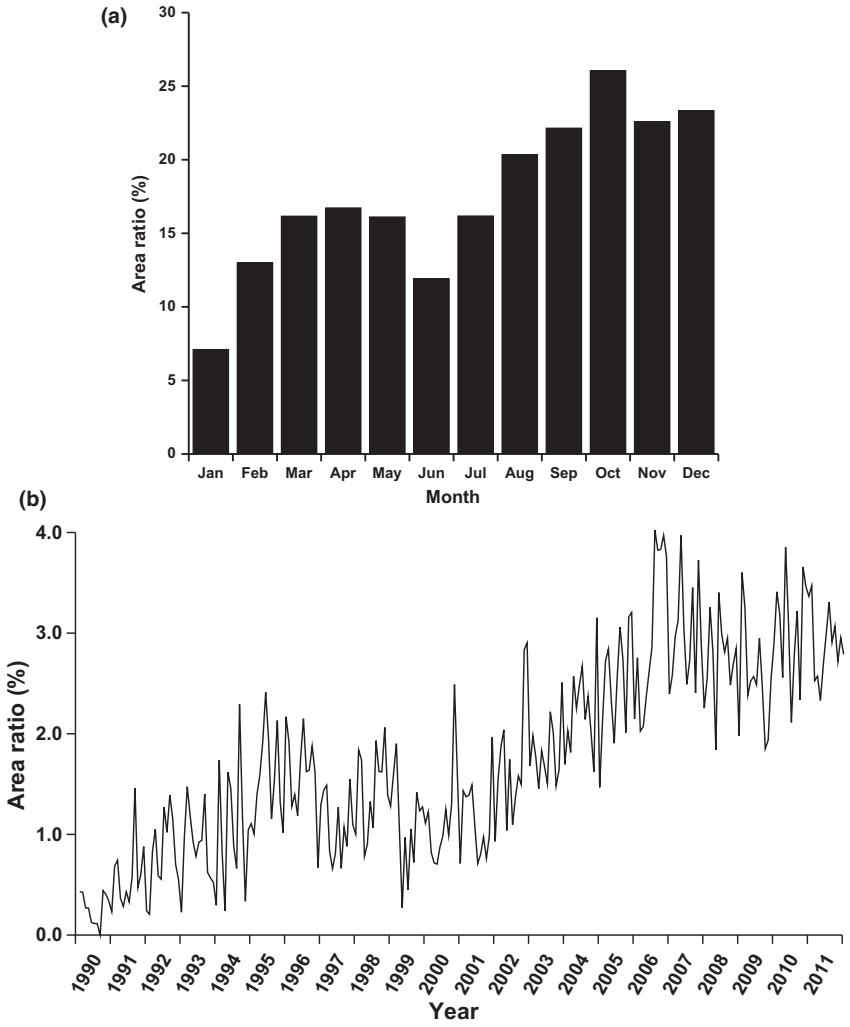

Figure 1. The ratios of the sum of areas with measurement data in 1 × 1° grid cells to the global open ocean area. (a) The sum of ratios in the same months of 1990–2011. (b) The ratios in individual months in 1990–2011.

climatology of Takahashi *et al.* (2009) (4° × 5°), which is currently the most frequently used product, but also provides global SOC maps in multiple years. While the SOC climatology of Zeng *et al.* (2014) does not cover areas where chlorophyll data were not available, this dataset filled those gaps.

1. Methods

The basic model equation of Zeng *et al.* (2014) defines the climatology of CO_2 fugacity (fCO_2) as a nonlinear function of month (MON), latitude (LAT), longitude (LON), sea surface temperature (SST), sea surface salinity (SSS), and chlorophyll-a concentration (CHL), i.e.

$$fCO_2 = F(\text{MON, LAT, LON, SST, SSS, CHL}). \quad (1)$$

We recognize that in areas where large spatial gaps exist the model has a potential to over-interpolating due to its nonlinear characteristics. To minimize such a problem, we implemented a three-stage modelling approach. In the first stage, we excluded the LAT and LON variables from the model; therefore the model equation becomes

$$fCO_2 = F(\text{MON, SST, SSS, CHL}). \quad (2)$$

We used model Equation (2) for spatial gap filling, i.e. filling grid cells with modelled fCO_2 if there is no measurement surrounding those grid cells within 10°. In the second stage, we used the measurements and gap-filled data with equation (1) for areas where CHL data are available. To estimate fCO_2 in CHL-missing areas, we excluded CHL from equation (1) so the model equation becomes:

$$fCO_2 = F(\text{MON, LAT, LON, SST, SSS}). \quad (3)$$

Therefore, we have two sets of model results in this stage: One includes the dependency of fCO_2 on CHL and another excludes that dependency. Finally in the third stage, we merged the two results to obtain the final product. The merging process first fills grid cells

with the model results of equation (1), and then with the results of equation (3).

It should be noted that the criteria for identifying the open ocean grid cells in this work are slightly different from those of Zeng *et al.* (2014). The criteria are elevation smaller than −500 m, SST larger than −10°C, SSS larger than 25, and ice cover smaller than 50%.

We used the climatology of all variables in training. The climatology of fCO_2 was obtained by normalizing fCO_2 measurements to the reference year of 2000 assuming a mean global increase rate of 1.5 μatm year^{-1} (Zeng *et al.*, 2014). For the fCO_2 reconstruction, the SST climatology was replaced by the monthly SST. It would be ideal if we could use also the monthly SSS and CHL, but their data were not available for the whole period. Finally, the model outputs of fCO_2 were adjusted by the same rate to yield the monthly fCO_2 in each year. The adjustment is necessary as the model output is the detrended fCO_2. Figure 2 shows modelled fCO_2 and observed climatology in January and June, 2000. We also estimated the global CO_2 fluxes (Figure 3) using the same parameters of Zeng *et al.* (2014).

2. Dataset location and format

The dataset in NetCDF format is archived at http://doi.pangaea.de/10.1594/PANGAEA.834398. The data file JTECH-D-13-00137.extended.nc contains the following variables.

Area – The grid cell area in 1 × 1 °.

Elevation – The grid mean of land topography and ocean bathymetry derived from the 1-arc-minute data of NOAA (http://www.ngdc.noaa.gov/mgg/global/global.html).

CHL_MM – The monthly climatology of chlorophyll concentration derived from the MODIS's AQUA and Terra datasets (http://oceancolor.gsfc.nasa.gov/cgi/l3).

SST_YYMM – The monthly SST extracted from the NOAA Optimum Interpolation (OI) V2 product (Reynolds *et al.*, 2002; http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html).

SST_MM – The monthly SST climatology derived from SST_YYMM.

SSS_MM – The monthly SSS extracted from the World Ocean Atlas 2009 (WOA09) product (Antonov *et al.*, 2010; http://www.nodc.noaa.gov/OC5/WOA09/netcdf_data.html).

Ice_MM – The ice cover in the reference year of 2000 derived from ECMWF's Interim data (http://apps.ecmwf.int/datasets/data/interim_full_moda/).

Ps_MM – The surface pressure in the reference year of 2000 derived from the ECMWF's Interim data.

WND_YYMM – The monthly wind speed extracted from the ECMWF's Interim data.

fCO_2_Obs2.0 – The monthly grid means of fCO_2 measurements extracted from the Surface Ocean CO_2

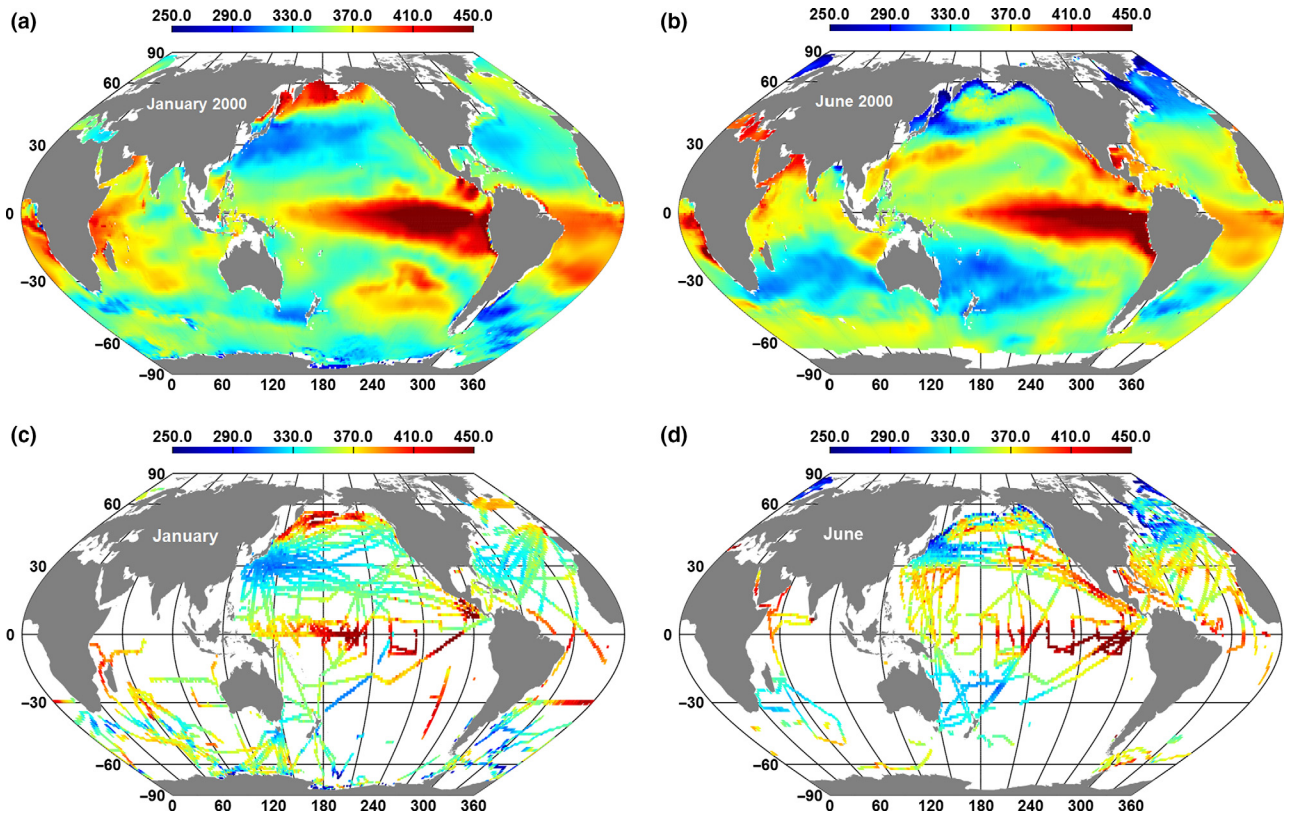

Figure 2. Modelled fCO_2 (a) and (b) and observed climatology (c) and (d) in January and June, 2000.

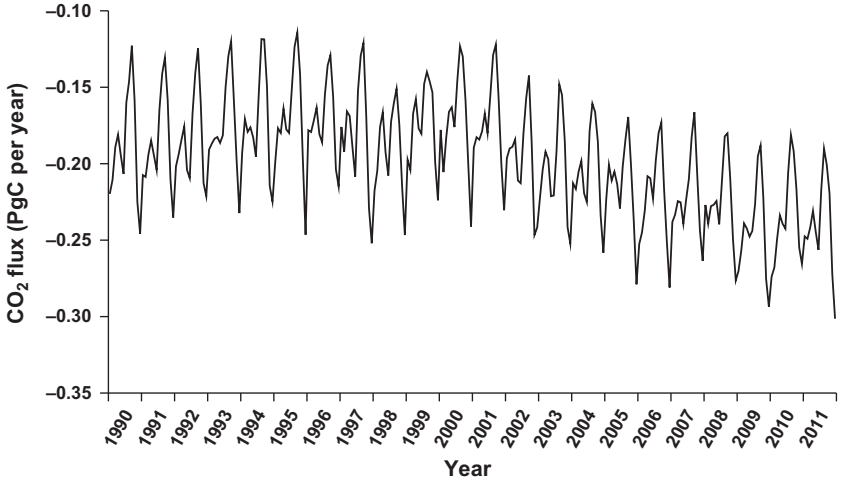

Figure 3. The CO_2 flux between the ocean and the atmosphere. A negative value indicates oceanic uptake.

Atlas (SOCAT) version 2.0 (Pfeil *et al.*, 2013; Sabine *et al.*, 2013; Bakker *et al.*, 2014; http://www.socat.info/).

fCO_2_Clm2.0 – The *f*CO_2 climatology with the reference year of 2000 derived from *f*CO_2_Obs2.0 assuming the increase rate of 1.5 μatm year^{-1} (Zeng *et al.*, 2014).

fCO_2_YYMM – The modelled monthly *f*CO_2.

gvCO_2_YYMM – The atmospheric xCO_2 (mixing ratio) extracted from the GlobalView product (ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/).

Flux_YYMM – The monthly CO_2 fluxes calculated by the method of Zeng *et al.* (2014).

Flux_YY – The sum of Flux_YYMM over the global oceans.

3. Dataset use

The dataset includes all necessary variables for reconstructing the monthly SOC and calculating CO_2 fluxes. Here, lists their uses.

1. Training of the neural network: CO_2_MM, SST_MM, SSS_MM, and CHL_MM.
2. Reconstruction of monthly *f*CO_2: SST_YYMM, SSS_MM, and CHL_MM.
3. Flux calculation: *f*CO_2_YYMM, gvCO_2_YYMM, SST_YYMM, Wnd_YYMM, and Area.

References

Antonov JI, Seidov D, Boyer TP, Locarnini RA, Mishonov AV, Garcia HE. 2010. *Salinity*. Vol. 2, World Ocean Atlas 2009, NOAA Atlas NESDIS 69, 184.

Bacastow RB, Keeling CD, Whorf TP. 1985. Seasonal amplitude increase in atmospheric CO_2 concentration at Mauna Loa, Hawaii, 1959–1982. *Journal of Geophysical Research* **90**(D6): 10529–10540.

Bakker DCE, Pfeil B, Smith K, Hankin S, Olsen A, Alin SR, Cosca C, Harasawa S, Kozyr A, Nojiri Y, O'Brien KM, Schuster U, Telszewski M, Tilbrook B, Wada C, Akl J, Barbero L, Bates NR, Boutin J, Bozec Y, Cai W-J, Castle RD, Chavez FP, Chen L, Chierici M, Currie K, de Baar HJW, Evans W, Feely RA, Fransson A, Gao Z, Hales B, Hardman-Mountford NJ, Hoppema M, Huang W-J, Hunt CW, Huss B, Ichikawa T, Johannessen T, Jones EM, Jones SD, Jutterström S, Kitidis V, Körtzinger A, Landschützer P, Lauvset SK, Lefèvre N, Manke AB, Mathis JT, Merlivat L, Metzl N, Murata A, Newberger T, Omar AM, Ono T, Park G-H, Paterson K, Pierrot D, Ríos AF, Sabine CL, Saito S, Salisbury J, Sarma VVSS, Schlitzer R, Sieger R, Skjelvan I, Steinhoff T, Sullivan KF, Sun H, Sutton AJ, Suzuki T, Sweeney C, Takahashi T, Tjiputra J, Tsurushima N, van Heuven SMAC, Vandemark D, Vlahos P, Wallace DWR, Wanninkhof R, Watson AJ. 2014. An update to the Surface Ocean CO_2 Atlas (SOCAT version 2). *Earth System Science Data* **6**: 69–90, doi:10.5194/essd-6-69-2014.

Friedrich T, Oschlies A. 2009. Neural network-based estimates of North Atlantic surface pCO_2 from satellite data: a methodological study. *Journal of Geophysical Research* **114**: C03020, doi:10.1029/2007JC004646.

Jamet C, Moulin C, Lefèvre N. 2007. Estimation of the oceanic pCO_2 in the North Atlantic from VOS lines in-situ measurements: parameters needed to generate seasonally mean maps. *Annales Geophysicae* **25**: 2247–2257, doi:10.5194/angeo-25-2247-2007.

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

Landschützer P, Gruber N, Bakker DCE, Schuster U, Nakaoka S, Payne MR, Sasse T, Zeng J. 2013. A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences* **10**: 7793–7815, doi:10.5194/bg-10-7793-2013.

Le Quéré C, Peters GP, Andres RJ, Andrew RM, Boden TA, Ciais P, Friedlingstein P, Houghton RA, Marland G, Moriarty R, Sitch S, Tans P, Arneth A, Arvanitis A, Bakker DCE, Bopp L, Canadell JG, Chini LP, Doney SC, Harper A, Harris I, House JI, Jain AK, Jones SD, Kato E, Keeling RF, Goldewijk KK, Körtzinger A, Koven C, Lefèvre N, Maignan F, Omar A, Ono T, Park GH, Pfeil B, Poulter B, Raupach MR, Regnier P, Rödenbeck C, Saito S, Schwinger J, Segschneider J, Stocker BD, Takahashi T, Til-

brook B, van Heuven S, Viovy N, Wanninkhof R, Wiltshire A, Zaehle S. 2013. Global carbon budget 2013. *Earth System Science Data* **6**: 235–263, doi:10.5194/essd-6-235-2014.

Lèfevre N, Watson AJ, Watson AR. 2005. A comparison of multiple regression and neural network techniques for mapping *in situ* pCO_2 data. *Tellus* **57B**: 375–384, doi:10.1111/j.1600-0889.2005.00164.x.

Li Z, Adamec D, Takahashi T, Sutherland SC. 2005. Global autocorrelation scales of the partial pressure of oceanic CO_2. *Journal of Geophysical Research* **110**: C08002, doi:10.1029/2004JC002723.

Nakaoka S, Telszewski M, Nojiri Y, Yasunaka S, Miyazaki C, Mukai H, Usui N. 2013. Estimating temporal and spatial variation of ocean surface pCO_2 in the North Pacific using a self-organizing map neural network technique. *Biogeosciences* **10**: 6093–6106, doi:10.5194/bg-10-6093-2013.

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

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W. 2002. An improved *in situ* and satellite SST analysis for climate. *Journal of Climate* **15**: 1609–1625, doi:10.1175/1520-442(2002) 015,1609:AIISAS.2.0.CO;2.

Sabine CL, Hankin S, Koyuk H, Bakker DCE, Pfeil B, Olsen A, Metzl N, Kozyr A, Fassbender A, Manke A, Malczyk J, Akl J, Alin SR, Bellerby RGJ, Borges A, Boutin J, Brown PJ, Cai W-J, Chavez FP, Chen A, Cosca C, Feely RA, González-Dávila M, Goyet C, Hardman-Mountford N, Heinze C, Hoppema M, Hunt CW, Hydes D, Ishii M, Johannessen T, Key RM, Körtzinger A, Landschützer P, Lauvset SK, Lefèvre N, Lenton A, Lourantou A, Merlivat L, Midorikawa T, Mintrop L, Miyazaki C, Murata A, Nakadate A, Nakano Y, Nakaoka S, Nojiri Y, Omar AM, Padin XA, Park G-H, Paterson K, Perez FF, Pierrot D, Poisson A, Ríos AF, Salisbury J, Santana-Casiano JM, Sarma VVSS, Schlitzer R, Schneider B, Schuster U, Sieger R, Skjelvan I, Steinhoff T, Suzuki T, Takahashi T, Tedesco K, Telszewski M, Thomas H, Tilbrook B, Vandemark D, Veness T, Watson AJ, Weiss R, Wong CS, and Yoshikawa-Inoue H. 2013. Surface ocean CO2 atlas (SOCAT) gridded data products. *Earth System Science Data* **5**: 145–153, doi:10.5194/essd-5-145-2013.

Sarma VVSS, Saino T, Sasaoka K, Nojiri Y, Ono T, Ishii M, Inoue HY, Matsumoto K. 2006. Basin-scale pCO_2 distribution using satellite sea surface temperature, Chla, and climatological salinity in the North Pacific in spring and summer. *Global Biogeochemical Cycles* **20**: GB3005, doi: 10.1029/2005GB002594.

Schuster U, McKinley GA, Bates N, Chevalier F, Doney SC, Fay AR, Gonzalez-Davila M, Gruber N, Jones S, Krijnen J, Landschützer P, Lefèvre N, Manizza M, Mathis JT, Metzl N, Olsen A, Rios AF, Rödenbeck C, Santana-Casiano JM, Takahashi T, Wanninkhof R, Watson AJ. 2013. An assessment of the Atlantic and Arctic sea-air CO_2 fluxes, 1990–2009. *Biogeosciences* **10**: 607–627, doi:10.5194/bg-10-607-2013.

Takahashi T, Sutherland SC, Wanninkhof R, Sweeney C, Feely RA, Chipman DW, Hales B, Friederich G, Chavez F, Sabine C, Watson A, Bakker DCE, Schuster U, Metzl N, Inoue HY, Ishii M, Midorikawa T, Nojiri Y, Koertzinger A, Steinhoff T, Hoppema M, Olafsson J, Arnarson TS, Tilbrook B, Johannessen T, Olsen A, Bellerby R, Wong CS, Delille B, Bates NR, de Baar HJW. 2009. Climatological mean and decadal change in surface ocean pCO2, and net sea-air CO2 flux over the global oceans, *Deep-Sea Research II* **56**: 554–577, doi:10.1016/j.dsr2.2008.12.009.

Takamura TR, Inoue HY, Midorikawa T, Ishii M, Nojiri Y. 2010. Seasonal and inter-annual variations in pCO_2 and air-sea CO_2 fluxes in mid-latitudes of the western and eastern North Pacific during 1999–2006: recent results utilizing voluntary observation ships. *Journal of Meteorological Society of Japan* **88**: 883–898, doi:10.2151/jmsj.2010-602.

Telszewski M, Chazottes A, Schuster U, Watson AJ, Moulin C, Bakker DCE, Gonzalez-Davila M, Johannessen T, Kortzinger A, Luger H, Olsen A, Omar A, Padin XA, Rios AF, Steinhoff T, Santana-Casiano M, Wallace DWR, Wanninkhof R. 2009. Estimating the monthly pCO_2 distribution in the North Atlantic using a self-organizing neural network. *Biogeosciences* **6**: 1405–1421, doi:10.5194/bg-6-1405-2009.

Wanninkhof R, Park GH, Takahashi T, Sweeney C, Feely R, Nojiri Y, Gruber N, Doney SC, McKinley GA, Lenton A, Le Quéré C, Heinze C, Schwinger J, Graven H, Khatiwala S. 2013. Global ocean carbon uptake: magnitude, variability and trends. *Biogeosciences* **10**: 1983–2000, doi:10.5194/bg-10-1983-2013.

Zeng J. 2014. Surface ocean CO2 in 1990–2011 reconstructed using a feed-forward neural network. NIES. doi:http://doi.pangaea.de/10.1594/PANGAEA.834398.

Zeng J, Nojiri Y, Murphy PP, Wong CS, Fujinuma Y. 2002. A comparison of Delta pCO_2 distributions in the northern North Pacific using results from a commercial vessel in 1995–1999. *Deep-Sea Research II* **49**: 5303–5315, doi:10.1016/S0967-0645(02)00192-3.

Zeng J, Nojiri Y, Landschützer P, Telszewski M, Nakaoka S. 2014. A global surface ocean fCO_2 climatology based on a feed-forward neural network. *Journal of Atmospheric and Oceanic Technology* **31**: 1838–1849, doi:10.1175/JTECH-D-13-00137.1.