Project:
GEOCARBON

Project full title:
Operational Global Carbon Observing System

European Commission - FP7
Collaborative Project (large scale integrating project) - for specific cooperation actions (SICA) dedicated to international cooperation partner countries
Grant agreement no.: **283080**

Del. no: 4.2
Deliverable name: Annual updates of global air-sea CO2 fluxes
Version: V1
WP no: 4
Lead beneficiary: ETHZ
Delivery date from Annex I (project month): 12
Actual delivery date (project month): 14
Jointly written by Peter Landschützer (UEA) and Nicolas Gruber (ETHZ), November 30, 2012.

---

1. Introduction

1.1 Short summary

This is the first release of a new monthly-resolved global air-sea CO_2 flux product at 1 degree resolution covering the period from 1998-2008 and developed on the basis of the SOCAT fCO_2 data base. The observed fCO_2 data were first analysed and related to a small number of globally observed “drivers”, such as sea-surface temperature, sea-surface salinity, mixed layer depth, surface chlorophyll, and atmospheric CO_2. These relationships were established within a number of discrete biogeochemical provinces and using a non-linear approach on the basis of a neural network method (feed-forward network). In the second step, a global surface ocean fCO_2 field was generated on a monthly basis using these non-linear relationships and the globally available “drivers”. In the final step, a flux product was generated assuming a quadratic relationship of the gas transfer velocity with windspeed and using the atmospheric pCO_2 computed from atmospheric xCO_2 from the marine boundary layer reference data set of NOAA ESRL. The errors associated with the flux product are generally within ±20%, but vary substantially depending on the region and the variability within each of the provinces.

The global distribution of the long-term mean air-sea CO_2 fluxes compare well with previous estimates, particularly the Takahashi pCO_2 dataset based estimates, but reveals some considerable differences, especially with regard to the seasonality. It does resolve, for the first time, the interannal variability of the air-sea CO_2 fluxes at high resolution over an extended period, revealing the impact of ENSO and other climate modes on the exchange of CO_2 across the air-sea interface. Future releases will include an

update and extension using the new release of SOCAT (Vs 2.x) as well as alternative means to estimate these fluxes.

1.2 Rationale for this deliverable

The ocean is a very strong sink for atmospheric CO_2, having removed more than one third of the total anthropogenic CO_2 emissions since pre-industrial times (Sabine et al., 2004). Although much progress has been made with regard to our understanding of the long-term mean annual fluxes (e.g., Gruber et al., 2009; Takahashi et al., 2009), little is known on a global basis about how this exchange of CO_2 between the ocean and atmosphere is varying on a month-by-month basis. This product closes this gap and provides opportunities to better understand and quantify this oceanic sink and how it responds to natural and anthropogenic perturbations. At the same time, these data also provide one of the most important boundary conditions for atmospheric CO_2 data assimilation systems, improving their accuracy and decreasing the uncertainty.

1.3Problems encountered and envisaged solutions

No fundamental problems have been encountered in the generation of this dataset.

2 Full description

The air-sea CO_2 flux product was produced in four discrete steps. First, the data were clustered together with a set of "drivers" into a set of 16 biogeochemical regimes in order to the reduce the variance within each regime. Second, within each of the 16 regimes, a non-linear regression method based on a feed-forward network approach was employed to establish a relationship between the drivers and the observed fCO_2. Third, these relationships were then used to produce a global continuous fCO2 field at 1x1 deg spatial and at 1 month temporal resolution over the entire 1998 through 2007 period. In the fourth and final step, the oceanic fCO2 data were then combined with atmospheric data and a windspeed based estimate of the gas-transfer coefficient to estimate the air-sea CO_2 flux at the same resolution and covering the same period.

2.1Data

The flux product is based on the gridded fCO_2 data product provided by the Surface Ocean CO_2 Atlas (SOCAT) Version 1.5 (Sabine et al., 2012). The SOCAT project assembled measured fCO_2 data across all ocean basins, and extending back in time as far as reliable measurements exist. Version 1.5 employed here covers the period up until December 2007. The SOCAT data have undergone an extensive secondary quality control, which ensures that the data are homogeneous and internally consistent (See Deliverable 4.1 for details).

As drivers (input data) we used the NOAA Optimum Interpolation (IO) sea surface temperature (SST) v.2 (Reynolds et al., 2002), SeaWiFS mapped chlorophyll (CHL) (SeaWifsProject, http://oceancolor.gsfc.nasa.gov/cgi/l3), ECCO2 mixed layer depth data (MLD) (Menemenlis et al., 2008), SODA sea surface salinity (SSS) data (Carton and Giese, 2008) and atmospheric CO_2 (xCO_2,atm) from GLOBALVIEW-CO2 (2011). Furthermore, to divide the ocean into biogeochemical regimes the monthly pCO_2 climatology (pCO_2(Takahashi)) of Takahashi et al. (2009) was used as an additional input parameter. The period of analysis covers Jan 1998 until Dec 2007 with the start date being determined by the start of the chlorophyll coverage by SeaWiFS, and the end date being determined by the last data entry in SOCAT Version 1.5.

Data with an original resolution higher than the required 1x1° were binned onto the desired grid (by averaging over every data point within the new bin), whereas input data with a coarser resolution were interpolated using a 2-dimensional linear interpolation algorithm. We further computed monthly averages of all inputs with a higher temporal resolution. After co-locating the network inputs to the same points in time and space, they were organized as *r* input vectors and the corresponding observations were organized as the target vector. Every vector consists therefore of data points at the same geographical location on a 1x1° grid at the same point in time.

2.2 Clustering

To reduce the variance of the data and also the error of the fitted fields, a set of biogeochemical provinces were defined first. To this end, we used a self-organizing map (SOM) as presented in Kohonen (1987, 2001), which consists of a set of neurons organized on a multidimensional grid. The algorithm learns to recognise groups of similar input vectors during the training process and physically responds to similar inputs presented to the SOM by moving the gridpoints of the map space towards the input parameters. For our study we chose a map with 16 neurons, referring to 16 regimes we want to separate the global ocean, organized on a 2 dimensional 4x4 point hexagonal grid. We chose to perform a batch training (see e.g. Vesanto et al., 2000) of the SOM using SST, log(MLD), SSS and pCO_2(Takahashi) as elements of the input parameters. We further chose a normalization that leads to a biased clustering towards the pCO_2 climatology. The consequence of this clustering approach is that the 16 regions are determined to have small internal pCO_2 or fCO_2 variability and that they each follow the seasonality of the pCO_2 climatology. Therefore every region is strongly geographically intra-annually shifting, with little inter-annual shifts, as we used the climatology as SOM input for every year of our study period.

2.3 Non-linear regression

In the second step we divided the training dataset into the 16 ocean regimes and process each of them separately by performing a non-linear regression between the input vectors and corresponding SOCAT gridded fCO_2 data. We chose to use the absolute SST, log(MLD), log(CHL), SSS and xCO_2^{atm} as elements of the input vector. We further use de-seasonalized sets of our input parameters to their long term mean seasonal cycle in this subset. Our feed-forward network uses 2 layers of neurons, i.e., 1 hidden layer of neurons using a sigmoid transfer function and 1 linear output layer. This type of network is capable of approximating any function with a finite number of discontinuities (Demuth et al., 2008).

Since the number of inputs and targets varies per regime, it is not possible to provide one best number of neurons to use for all 16 regimes. We therefore performed a number of pre-trainings, increasing the number of hidden neurons paraboloidal starting from 2 neurons up to a number where the ratio between of the training sample size to the number of weights does exceed 30 to avoid overfitting (Amari et al., 1997). During every pre-training process the training vectors and the corresponding gridded SOCAT fCO2 targets are introduced to the network and the weights and biases (regression coefficients) were iteratively updated in the direction where the performance function decreases most rapidly (this performance function was defined here as the mean squared error between network outputs fCO2;est and fCO2 targets).

Our feed-forward network used the Levenberg-Marquardt (Marquardt, 1963) algorithm to update weights and biases in every iteration step to reduce the error between outputs and targets. After every iteration of each pre-training, the network was validated using a subset of data that were not used for training (early stopping approach, see e.g. Demuth et al., 2008).

During the actual training process the number of neurons was adjusted according to the best pre-training performance (where the error between validation data and SOCAT fCO_2 data is a minimum) for every 16 regimes separately. We performed 10 trainings where we randomly pick subsets of validation data.

After every training we used the trained network to simulate fCO_2 from the input dataset and average the output of the 10 training cycles, to end up with 1 estimate for our time period between 1998 to 2007 for each regime. After 16 network runs the results of the 16 regions were combined to retrieve the fCO_2 estimates from 1998 through 2007 on a global 1x1° grid.

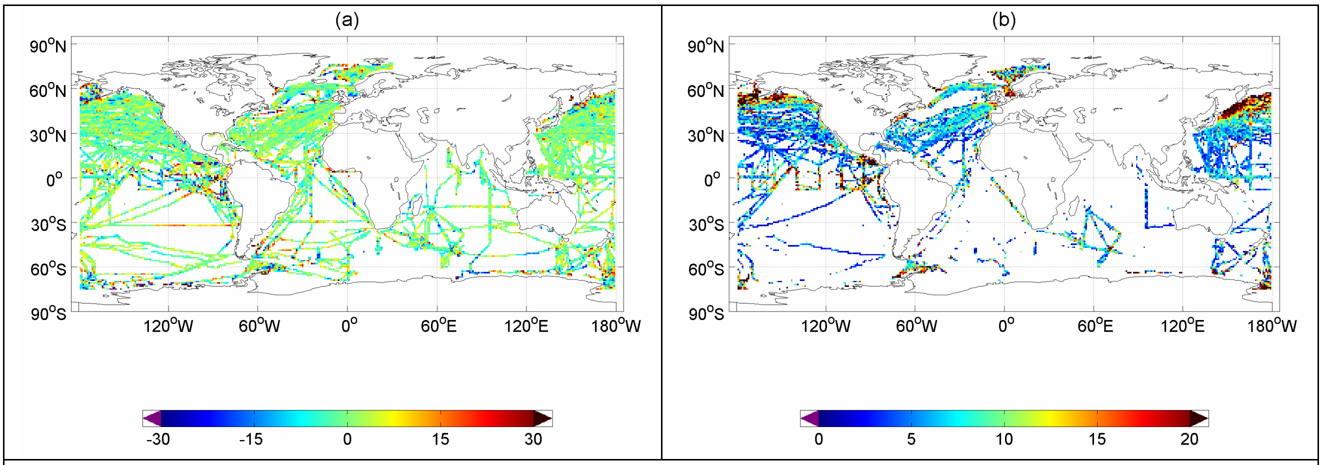

Figure 1: (a) mean and (b) standard error of fCO_2 (µatm) between neural network outputs and gridded SOCAT data.

2.4 Flux calculation

For the fourth and final step we estimated first the air-sea ΔpCO_2 by calculating the sea surface pCO_2 from the fCO_2 estimates using the formulation of Zeebe and Wolf-Gladrow (2001) and the atmospheric pCO_2 from GLOBALVIEW-CO2 (2011) atmospheric xCO_2, using the NCEP monthly mean sea level pressure (Kalnay et al., 1996) and taking into account the water vapour correction according to Dickson et al. (2007). The air-sea flux was then obtained by multiplying the air-sea ΔpCO_2 with a gas transfer coefficient, estimated on the basis of the windspeeds provived by NCEP and using the formulation of Wanninkhof (1992) with the gas transfer coefficient of Sweeney et al. (2007).

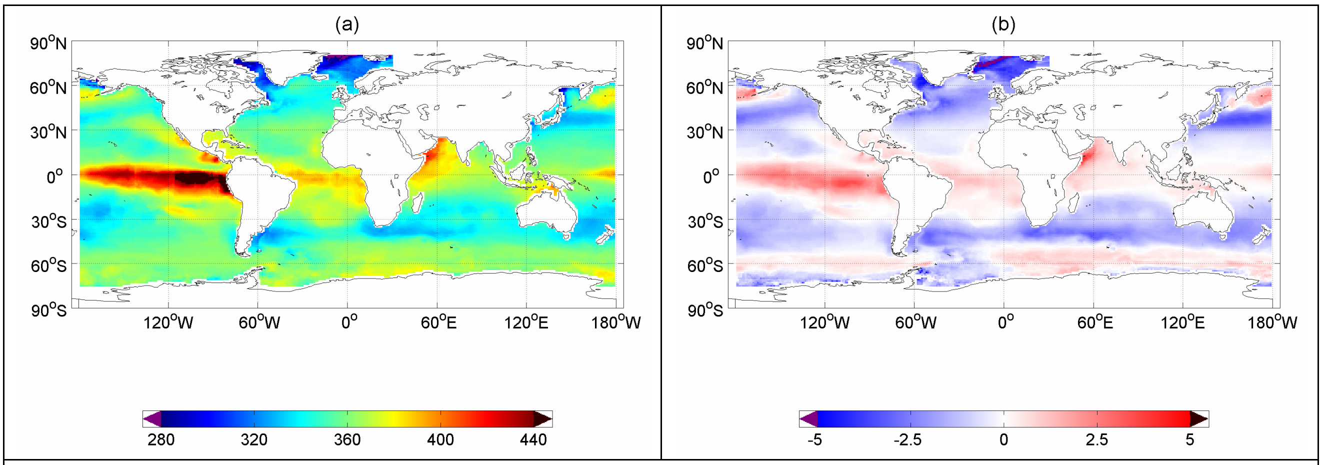

Figure 2: (a) Decadal mean neural network pCO_2 estimate in µatm and (b) decadal mean flux density in mol m^{-2} yr^{-1}

3 References

Amari, S., Murata, N., Muller, K.-R., Finke, M., and Yang, H. H.: Asymptotic Statistical Theory of

Overtraining and Cross-Validation, *Ieee Transactions On Neural Networks*, 8, 985–996, 1997.
Carton, J. A. and Giese, B. S.: A Reanalysis of Ocean Climate Using Simple Ocean Data Assimilation (SODA), *Monthly Weather Review*, 136, 2999–3017, doi:DOI:10.1175/2007MWR1978.1, 2008.
Demuth, H., Beale, M., and Hagan, M.: Neural Network Toolbox 6 Users Guide, The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA, 2008.
Dickson, A. G., Sabine, C. L., and Christian, J. R., eds.: *Guide to Best Practices for Ocean CO_2 Measurements*, PICES special Publication, IOCCP Report No . 8, 2007.
Gruber, N., Gloor, M., Mikaloff Fletcher, S. E., Doney, S. C., Dutkiewicz, S., Follows, M. J., Gerber, M., Jacobson, A. R., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A., Sarmiento, J. L., and Takahashi, T.: Oceanic sources, sinks, and transport of atmospheric CO_2, *Global Biogeochemical Cycles*, 23, GB1005, doi:10.1029/2008GB003349, 2009.
Kalnay et al., 1996 Kalnay et al.: The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteor. Soc.*, 77, 437–470, 1996.
Kohonen (1987, 2001) Kohonen, T.: *Self-Organization and Associative Memory*, Springer-Verlag, Berlin, 2nd edn., 1987.
Kohonen, T.: *Self-Organizing Maps*, Springer-Verlag, Berlin Heidelberg New York, 3rd edn., 2001.
Landschützer, P., U. Schuster, D. Bakker, and N. Gruber et al.: Seasonal to interannual variability of the Atlantic Ocean carbon sink. *Biogeosciences*, in prep.
Marquardt, D. W.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters, Journal of the Society for Industrial and Applied Mathematics, Vol. II, 431–441, 1963.
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.
Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., and Wang, W.: An improved in situ and satellite SST analysis for climate, *Journal of Climate*, 15, 1609–1625, 2002.
Sabine et al (2012) Surface Ocean CO_2 Atlas (SOCAT) gridded data products, *Earth Syst. Sci. Data Discuss*., 5, 781-804, doi:10.5194/essdd-5-781-2012, 2012.
Sabine. C.S., R. A. Feely, N. Gruber, R. M. Key, K. Lee, J. L. Bullister, R. Wanninkhof, C. S. Wong, D. W. R. Wallace, B. Tillbrook, F. J. Millero, T.-H. Peng, A. Kozyr, T. Ono, and A. F. Rios: The oceanic sink for anthropogenic CO_2, *Science*, 305, 367-371, 2004.
Sweeney et al. (2007). Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G., Sarmiento, J. L., and Wanninkhof, R.: Constraining global air-sea gas exchange for CO2 with recent bomb ^{14}C measurements, *Global Biogeochemical Cycles*, 21, GB2015, doi: 10.1029/2006GB002784, 2007.
Takahashi et al. : Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, *Deep-Sea Research II*, 56, 554–577, 2009.
Vesanto, J., Himberg, J., Alhoniemi, E., and Parhankangas, J.: *SOM Toolbox for Matlab 5*, Tech. rep., Helsinki University of Technology, 2000.
Wanninkhof, R.: Relation between wind speed and gas exchange over the ocean, *Journal of Geophysical Research*, 97, 7373–7383, 1992.
Zeebe, P. E. and Wolf-Gladrow, D.: *CO_2 in Seawater: Equilibrium, Kinetics, Isotopes*, Elsevier, 2001.