RINGO | Readiness of ICOS

Readiness of ICOS for Necessities of integrated Global Observatio

D1.6

Semi-automated near-real time (NRT) data pipeline for calculating atmosphere-ocean CO2 fluxes

RINGO (GA no 730944)

Public Document/define level

Deliverable:

Author(s): Jamie Shutler, Thomas Holding
Date: 14 June 2019
Activity: WP1
Lead Partner: NOC
Document Issue:
Dissemination Level:
Contact: j.d.shutler@exeter.ac.uk

	Name	Partner	Date
From	Jamie Shutler	University of Exeter	14 June 2019
Reviewed by	Andy Watson	University of Exeter	17 June 2019
Approved by	Evi-Carita Riikonen	ICOS ERIC	25 June 2019

Version	Date	Comments/Changes	Author/Partner
1.0	14 June 2019	Original	J. Shutler, T. Holding
1.1	17 June 2019	Original updated post internal review	J. Shutler

Deliverable Review Checklist

A list of checkpoints has been created to be ticked off by the Task Leader before finalizing the deliverable. These checkpoints are incorporated into the deliverable template where the Task Leader must tick off the list.

√

- Appearance is generally appealing and according to the RINGO template. Cover page has been updated according to the Deliverable details. √
- The executive summary is provided giving a short and to the point description of the deliverable. √
- All abbreviations are explained in a separate list. √
- All references are listed in a concise list. √
- The deliverable clearly identifies all contributions from partners and justifies the resources used. √
- A full spell check has been executed and is completed. √

DISCLAIMER

This document has been produced in the context of the *project* Readiness of ICOS for Necessities of integrated Global Observations (RINGO)

The Research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730944. All Information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors view.

Amendments, comments and suggestions should be sent to the authors.

ABSTRACT and Executive Summary

The flow (flux) of climate critical gases, such as carbon dioxide (CO_2), between the ocean and the atmosphere is a fundamental component of our climate and the biogeochemical development of the oceans. Therefore, the accurate calculation of these air-sea gas fluxes is critical if we are to monitor the health of our oceans and understand changes to our climate. The FluxEngine is an open source software toolbox that allows users to easily perform calculations of atmosphere-ocean gas fluxes from model, *in-situ* and Earth observation data. The original development and verification of the toolbox was described in Shutler et al., (2016) and the toolbox is now being used by scientists across multiple disciplines. The toolbox has now been extended to increase its flexibility and its application to in situ data and these extensions are described within Holding et al., (2019, in-review).

This technical report describes how the FluxEngine toolbox can now be used to provide routine in situ, regional, and global calculations of atmosphere-ocean gases fluxes and calculations of the integrated net CO_2 sink.

TABLE OF CONTENTS

1 INTRODUCTION

This report is written under the assumption that quality controlled collated oceanic gaseous CO_2 data (partial pressure of CO_2, pCO_2 or fugacity of CO_2, fCO_2) are being used as provided by the Surface Ocean CO_2 Atlas (SOCAT, Bakker et al., 2016). However the methods described are equally applicable to other collated pCO_2 datasets. This report is written under the assumption that a monthly 1º × 1º global analysis is being performed. However the methods are equally application to basin or regional calculations and all of the tools used are region and resolution independent.

The workflow for calculating atmosphere-ocean CO_2 fluxes is:

1. Download the latest SOCAT dataset in ASCII format (termed, *Downloading gaseous CO_2 data*)
2. Use the reanalysis script (which is part of the FluxEngine toolbox) to calculate pCO_2 and fCO_2 values at a consistent temperature and depth (termed, *The re-analysis of gaseous CO_2 data*).
3. Interpolation of these pCO_2 (or fCO_2) data (termed, *Interpolation of gaseous CO_2 data*).
4. Automated tools to download and process additional input data required for the flux calculation (termed, *Downloading and pre-processing other input datasets*).
5. Using the FluxEngine (termed, *CO_2 flux calculation*) to calculate:
    - atmosphere-ocean gas fluxes.
    - calculate integrated monthly and annual global net flux.

2. METHODS

The FluxEngine software can be downloaded from https://github.com/oceanflux-ghg/FluxEngine. Full instructions detailing how to install and verify the installation (on Windows, Mac and Linux operating systems) are included within the software download. Holding et al., (2019, in-review) gives a description of the verification process and detailed examples of how to use the toolbox.

2.1 Summary of directory structure and output files

The main 'drive script' for all the near-real time (NRT) approach for using the FluxEngine is:

- *run_analysis.sh* – The driver script used to download and process the required data sets, run FluxEngine, calculate global net integrated fluxes and create all output files.

The directory structure that is created when the FluxEngine is installed and as a result of using the methods described here is:

- *Root directory*
  - *configs* – Contains FluxEngine configuration files
  - *SOCAT_re-analysis* – Groups tools and data for the reanalysis step
  - *Processing <data name>* - contains the download and processed data for each input field. Also contains the scripts used by the driver script for downloading and processing each field.
  - *interpolated_fCO2_data* – default location for the interpolated fCO2 fields.
  - *fe_output* – contains the output from FluxEngine and the calculated global net fluxes.

2.2 Downloading gaseous CO_2 data (SOCAT)

SOCAT data can be downloaded in ASCII format from https://www.socat.info/

2.3 The re-analysis of gaseous CO_2 data

The reanalysis step uses the aqueous partial pressure of carbon CO_2 (pCO_2) and fugacity of CO_2 (fCO_2) data supplied in the SOCAT dataset and re-analyses them to be consistent and valid at a fixed depth sea surface temperature field. The full description of the methods and justification for their use can be found in (Goddijn-Murphy, et al., 2015). Mean monthly values are calculated from the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation Sea Surface Temperature (OISST) dataset (Reynolds et al., 2007) to provide the consistent temperature climatology.

The reanalysis step is performed by the *reanalyse_socat_driver.py* python script (part of the open source FluxEngine toolkit, Shutler et al., 2015), which is run as a command line tool requiring just the SOCAT data and a consistent sea surface temperature dataset. The SOCAT dataset must be downloaded manually but a tool is provided to download and resample the OISST temperature data into monthly means with a user-defined spatial resolution (here we use 1º × 1º). The bash script stored at *SOCAT_reanalysis/run_reanalyse_socat.sh* runs the reanalysis tool for SOCAT v5 and SOCAT v6, producing text formatted and netCDF output files containing the re-analysed data.

The re-analysed data are then concatenated with the original SOCAT data using a custom Python script (*combine_ascii.py*). Custom scripts are currently available for SOCAT v5 and SOCAT v6, and are located in the sub-directory for the reanalysis corresponding to each SOCAT version. The output from running this script is a single dataset containing the full SOCAT data augmented with additional columns containing the consistent sea surface temperature, re-analysed pCO_2 and re-analysed fCO_2. These additional columns are described in Table 1. Another custom Python script is supplied to calculate gridded (1º × 1º) versions of the re-analysed data in netCDF format (*combine_netcdf.py*), with full metadata provided in the output netCDF files.

Column name	**Description (units)**
Tcl_C	Consistent sea surface temperature data (ºC)
SST_C	Original SOCAT reported ocean temperature (ºC)
fCO2_Tym	Re-analysed fugacity of CO_2 (µatm)
fCO2_SST	Original SOCAT reported fugacity of CO_2 (µatm)
pCO2_Tym	Re-analysed partial pressure of CO_2 (µatm)
pCO2_SST	Original SOCAT reported partial pressure of CO_2 (µatm)

Table 1: Column names and descriptions in the re-analysed SOCAT dataset. Column names that are identical to those of the original SOCAT dataset are not shown.

2.4 Interpolation of gaseous CO_2 data

Spatially and temporally complete (or near-complete) monthly 1º × 1º pCO_2 (of fCO_2) data are required to calculate atmosphere-ocean gas fluxes. The intelligent interpolation of these data is not part of this deliverable but multiple published methods existing including those within Schuster et al., (2013) and Landschutzer et al., (2015).

2.5 Downloading and pre-processing other input datasets

A bash script is provided which carries out the downloading and pre-processing of input datasets, and then runs FluxEngine to calculate atmosphere-ocean CO_2 fluxes and the net integrated global CO_2 fluxes. This script, referred to as the driver script henceforth, is suitable for running on local machines or a remote server.

Six types of input data are required to perform the flux calculation (Shutler et al., 2015). These are listed, with a description of the datasets used, in Table 2. In addition, to calculate the net integrated global CO_2 fluxes, sea ice coverage is also required (Table 2). Configurable variables at the beginning of the driver script file allow the temporal range over which data will be downloaded and processed to be easily set. The spatial resolution of the gridded data can also be defined.

Each input data source is then downloaded into separate directories, e.g. *processing_ccmp_windu10/downloaded_file*s for wind speed, and processed into monthly mean gridded 1º × 1º netCDF files. The exception to this is the atmospheric CO_2 input data that must be downloaded manually. For this the NOAA Earth System Research Laboratory (ESRL) provide a web-form (link provided in Table 2) to download the Greenhouse Gas Marine Boundary Layer vCO_2 data product. The path to this downloaded data can be set in the configuration script, but must be downloaded prior to running the script so that it can process these data into the correct temporal and spatial resolution.

Input data	Source and citation / url
Aqueous CO_2	Interpolated, re-analysed SOCAT pCO_2 or fCO_2 data
Atmospheric CO_2	NOAA ESRL Greenhouse Gas Marine Boundary Layer Reference: https://www.esrl.noaa.gov/gmd/ccgg/mbl/mbl.html
Sea surface temperature	NOAA Optimum Interpolation Sea Surface Temperature (OISST) dataset (Reynolds et al., 2007)
Salinity	World Ocean Atlas (WOA) sea surface salinity (M. M. Zweng et al., 2018)
Air pressure	European Centre for Medium-Range Weather Forecasts (ECMWF), ERA-Interim air pressure: https://www.ecmwf.int/node/8174
Wind speed	Remote Sensing Systems (REMSS) Cross-Calibrated Multi-Platform (CCMP) wind speed data product, (Atlas et al., 2011): http://www.remss.com/measurements/ccmp/
Sea ice coverage	NOAA Optimum Interpolation Sea Surface Temperature (OISST) dataset, percentage sea ice coverage field (Reynolds et al., 2007)

Table 2: Datasets and data sources used in the calculation of atmosphere-ocean CO_2 fluxes. For this report all data are processed to 1º × 1º gridded monthly mean netCDF files. Sea ice coverage is used to calculate the global integrated net CO_2 flux.

2.6 CO_2 flux calculation

The flux calculation is carried out by FluxEngine and is driven by the same driver script that downloads and pre-processes the input datasets. The paths to the FluxEngine configuration files are defined at the beginning of the driver script, and the temporal range to run FluxEngine are also separately configurable for multiple linear regression interpolation and the neural network interpolation. Different configuration files are used to perform separate flux calculations dependent upon the method used to interpolate the gaseous CO2 data. E.g. for a multiple linear regression interpolated pCO_2 fields (*mlr_run.conf* in the *configs* subdirectory is used). The configuration files configure FluxEngine to use this equation for the flux calculation:

$$F = k(\alpha_W fCO_{2W} - \alpha_S fCO_{2A})$$

where F is the sea-to-air flux of CO_2, k is the gas transfer velocity, α_W is the solubility of CO_2 in the sub-skin layer of the ocean, fCO_{2W} is the fugacity of CO_2 in the sub-skin layer, α_W is the solubility of CO_2 at the ocean-atmosphere interface (skin layer), and fCO_{2A} is the fugacity of CO_2 in the atmosphere. This version of the flux equation takes into account the solubility difference between the sub-skin and (typically cooler) skin layer (Woolf, Land, et al., 2016). To calculate ocean skin temperature, the OISST sea surface temperature dataset is taken to be the sub-skin temperature, and the skin temperature is calculated by subtracting a constant 0.17K (see Donlon et al., 2002). To calculate gas transfer velocity, the wind speed based parameterisation provided by Nightingale et al. (2000) is used.

FluxEngine output from these runs contain atmosphere-to-ocean CO_2 flux fields, copies of the input data used in the calculation and intermediate calculation products (including gas transfer velocity and CO_2 concentration at the interface and sub-skin). Full details of each output variable is provided in the metadata of each netCDF file. These are located in *fe_output/mlr_fco2_runs* and *fe_output/nna_fco2_runs* for the multiple linear regression and neural network approaches, respectively.

Integrated net flux budgets are calculated monthly and annually using FluxEngine's *ofluxghg_flux_budgets.py* tool. This tool takes the gridded netCDF output from FluxEngine and calculated the global flux budgets accounting for the area and ice coverage of each grid cell. The tool is run separately for each interpolation method and writes output to the respective output directory (e.g. *fe_output/mlr_fco2_runs_global.txt* for the multiple learn regression approach). The driver script also processes these output to produce a comma separated value (.csv).

A summary of the output files produced by this analysis is given in Table 3.

Output	Location and format
FluxEngine output using fCO_2 field produced by the multiple linear regression interpolation method	fe_output/mlr_fco2_runs/ Output is organised into subdirectories by year and month. Each netCDF file contains monthly mean data on a 1º × 1º resolution grid for a single month.
FluxEngine output using fCO_2 field produced by the neural network interpolation method	fe_output/nna_fco2_runs/ Output is organised into subdirectories by year and month. Each netCDF file contains monthly mean data on a 1º × 1º resolution grid for a single month.
Global monthly and annual net fluxes from the fCO_2 field produced by the multiple linear regression interpolation method.	*fe_output/mlr_fco2_runs_global.txt* Comma separated value file.
Global monthly and annual net fluxes from the fCO_2 field produced by the neural network interpolation method.	*fe_output/nna_fco2_runs_global.txt* Comma separated value file.
Global annual net flux comparison table between interpolation methods	*fe_output/annual_net_flux.csv* Comma separated value file.

Table 3: A summary of the output data, file location(s) and format.

3. CONCLUSIONS

Shutler et al., (2016) and Holding et al., (2019, in-review) explain the development, verification, operation and use of the FluxEngine toolbox. This report explains how the toolbox can now be used to routinely calculate regional and global atmosphere-ocean gases fluxes and the resultant integrated net sink of CO_2 through exploiting satellite observations, model re-analysis data and collated *in situ* datasets.

4. REFERENCES

Atlas, R., Hoffman, R. N., Ardizzone, J., Leidner, S. M., Jusem, J. C., Smith, D. K., & Gombos, D. (2011). A Cross-calibrated, Multiplatform Ocean Surface Wind Velocity Product for Meteorological and Oceanographic Applications. *Bulletin of the American Meteorological Society*, *92*(2), 157–174. https://doi.org/10.1175/2010BAMS2946.1

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., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibánhez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., Van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., Xu, S. (2016). A multi-decade record of high quality fCO_2 data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT). Earth System Science Data , 8, 383-413. doi:10.5194/essd-8-383-2016.

Donlon, C. J., Minnett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., … Murray, M. J. (2002). Toward Improved Validation of Satellite Sea Surface Skin Temperature Measurements for Climate Research. *Journal of Climate*, *15*(4), 353–369. https://doi.org/10.1175/1520-0442(2002)015<0353:TIVOSS>2.0.CO;2

Goddijn-Murphy, L. M., Woolf, D. K., Land, P. E., Shutler, J. D., & Donlon, C. (2015). The OceanFlux Greenhouse Gases methodology for deriving a sea surface climatology of CO2 fugacity in support of air-sea gas flux studies. *Ocean Science*, *11*(4), 519–541. https://doi.org/10.5194/os-11-519-2015

Holding, T., Ashton, I. G., Shutler, J. D., Land, P. E., Nightingale, P. D., Rees, A. P., Brown, I., Piolle, J.-F., Kock, A., Bange, H. W., Woolf, D. K., Goddijn-Murphy, L., Pereira, R., Paul, F., Girand-Ardhuin, F., Chapron, B., Rehder, G., Ardhuin, F., and Donlon, C. J. (2019, in-review) The FluxEngine air-sea gas flux toolbox: simplified interface and extensions for *in situ* analyses and multiple sparingly soluble gases, *Ocean Sci. Discuss*., https://doi.org/10.5194/os-2019-45hol.

Landschützer, P., Gruber, N., Bakker, D. C. E., Haumann, F. A., Rödenbeck, C., van Heuven, S., Hoppema, M., Metzl, N., Sweeney, C., Takahashi, T., Tilbrook, B., and Wanninkhof, R. (2015). The reinvigoration of the Southern Ocean carbon sink. Science, 349, 1221-1224. doi:10.1126/science.aab2620.

M. M. Zweng, Reagan, J. R., Seidov, D., Boyer, T. P., Locarnini, R. A., Garcia, H. E., … Smolyar, I. (2018). *WORLD OCEAN ATLAS 2018 Volume 2: Salinity (pre-release)*. Retrieved from http://www.nodc.noaa.gov/

Nightingale, P. D., Malin, G., Law, C. S., Watson, A. J., Liss, P. S., Liddicoat, M. I., … Upstill-Goddard, R. C. (2000). In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers. *Global Biogeochemical Cycles*, *14*(1), 373–387. https://doi.org/10.1029/1999GB900091

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., Schlax, M. G., … Schlax, M. G. (2007). Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate*, *20*(22), 5473–5496. https://doi.org/10.1175/2007JCLI1824.1

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

Shutler, J. D., Land, P. E., Piolle, J.-F., Woolf, D. K., Goddijn-Murphy, L., Paul, F., … Donlon, C. J. (2015). FluxEngine: A Flexible Processing System for Calculating Atmosphere–Ocean Carbon Dioxide Gas Fluxes and Climatologies. *Journal of Atmospheric and Oceanic Technology*, *33*(4), 741–756. https://doi.org/10.1175/JTECH-D-14-00204.1

Woolf, D. K., Land, P. E., Shutler, J. D., Goddijn-Murphy, L. M., & Donlon, C. J. (2016). On the calculation of air-sea fluxes of CO_2 in the presence of temperature and salinity gradients. *Journal of Geophysical Research: Oceans*, *121*(2), 1229–1248. https://doi.org/10.1002/2015JC011427

5. ACRONYMS

ASCII, American Standard Code for Information Interchange (i.e. alpha numeric text format).

CO_2, Carbon dioxide

FluxEngine, The open-source software toolbox for calculating atmosphere-ocean gas fluxes

SOCAT, Surface Ocean CO_2 Atlas

5 ACRONYMS

This project has received funding from the European Union's Horizon 2020
research and innovation programme under grant agreement No 730944