CO2 Human Emissions

Estimation of biogenic fluxes

Sophia Walther, Martin Jung

che-project.eu

Co-ordinated by ECMWF

D3.2 Net biospheric CO2 fluxes

Dissemination Level:	Public/
Author(s):	Sophia Walther, Martin Jung (Max-Planck-Institute for Biogeo-chemistry)
Date:	20/06/2019
Version:	1.0
Contractual Delivery Date:	30/06/2019
Work Package/ Task:	WP3/ T3.1
Document Owner:	MPI-BGC
Contributors:	MPI-BGC
Status:	Final

CHE: CO2 Human Emissions Project

Coordination and Support Action (CSA)
H2020-EO-3-2017 Preparation for a European capacity to monitor CO2 anthropogenic emissions

Project Coordinator: Dr Gianpaolo Balsamo (ECMWF)
Project Start Date: 01/10/2017
Project Duration: 39 months

Published by the CHE Consortium

Contact:
ECMWF, Shinfield Park, Reading, RG2 9AX,
gianpaolo.balsamo@ecmwf.int

The CHE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776186.

Table of Contents

Figures

Tables

1 Executive Summary

For the quantification and monitoring of anthropogenic contributions to the atmospheric CO2 concentration as targeted by CHE, estimates of the natural carbon fluxes between the land surface and the atmosphere need to be included in any analysis. We provide here as deliverable to WP3 ('Coordinating efforts on uncertainty trade-off for fossil fuel emissions') task 1 data-driven estimates of the net carbon exchange as well as of the gross photosynthetic carbon uptake of the terrestrial biosphere merging in-situ measurements, machine learning and satellite observations. This report illustrates the datasets (hourly, 0.5deg, global) and the associated uncertainties derived from a leave-one-site-out cross validation exercise. The flux data sets were uploaded to the ECMWF ftp project server. They can furthermore be accessed through direct contract to Sophia Walther (swalth@bgc-jena.mpg.de) or Martin Jung (mjung@bgc-jena.mpg.de).

2 Introduction

2.1 Background

The aim of the CO_2 Human Emissions (CHE) project is the development of a pre-operational system to accurately observe and quantify man-made CO_2 emissions. The results will have direct impact on European policy development. The approaches to quantify anthropogenic CO2 include bottom-up inventories and inverse transport modelling, partly within a coupled carbon cycle fossil fuel data assimilation system. Inverse models exploit as input atmospheric CO2 and its tracers, as well as prior estimates of fossil fuel emissions and of natural terrestrial carbon fluxes.

WP3 evaluates the current status and possible improvements from enhanced space-borne and in-situ observation scenarios for fossil CO_2 emissions quantification based on observing system simulation experiments (OSSEs) and quantitative network design (QND) studies using different approaches (high resolution inverse transport modelling of CO_2 and co-emitted species, advanced carbon cycle-fossil fuel data assimilation systems integrating atmospheric, terrestrial and socioeconomic datasets). The transport model inversions require - amongst others - high-resolution prior biogenic fluxes with quantified uncertainties. Based on these experiments WP3 will report on a set of inversion strategies blending bottom-up and top-down approaches for estimating fossil CO_2 emissions.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverable

With this deliverable, we provide a data-driven high-resolution product for biogenic CO2 fluxes with quantified uncertainties based on integrating the harmonized and standardized eddy covariance measurements from in-situ networks (FLUXNET), with satellite remote sensing (i.e. MODIS), meteorological observations (ERA5 reanalysis) and machine learning approaches.

2.2.2 Work performed in this deliverable

Hourly global estimates of net ecosystem exchange (NEE) and gross photosynthetic carbon uptake (GPP) with a resolution of 0.5deg have been produced and uploaded on the project ftp server for the years 2008 and 2015. A description of the product and associated uncertainties is provided in this report.

2.2.3 Deviations and counter measures

Contrary to what has been stated in the task description of the project, data on NEE and GPP are provided for the years 2008 and 2015. This has been decided during project meetings and in accordance with the work package leaders for reasons of compatibility with nature runs for 2015 at ECMWF and exercises of carbon-cycle-fossil-fuel-data-assimilation for 2008 at the University of Lund. However, the data set has been produced for the whole time period 1980-2018 and therefore, other years can be made available on request.

3 Biogenic flux estimates

3.1 The methodology

The data-driven NEE and GPP data delivered for CHE rely on the methodology of FLUXCOM (Tramontana et al. 2016, Bodesheim et al. 2018). This means that based on the eddy-covariance-derived net and gross carbon fluxes in the FLUXNET La Thuile data set, a random forest is trained using the following predictor variables:

Spatial or seasonal:

- mean seasonal cycle of the product of the enhanced vegetation index (EVI) and potential radiation (Rpot)
- mean seasonal cycle of the product of the fraction of absorbed photosynthetically active radiation and daytime land surface temperature (LST)
- minimum of mean seasonal cycle of the normalized difference water index (NDWI)
- amplitude of mean seasonal cycle of band 4 BRDF reflectance
- mean seasonal cycle of night-time LST
- amplitude of mean seasonal cycle of normalized difference vegetation index (NDVI)
- plant functional type
- amplitude of mean seasonal cycle of water availability index 2

Daily:

- Water availability index 2
- product of global radiation (total solar incoming at the surface) and the mean seasonal cycle of the normalized difference vegetation index (NDVI)
- air temperature

Hourly:

- potential radiation
- derivative of potential radiation
- air temperature
- vapour pressure deficit (VPD)
- global radiation (total solar incoming at the surface)

Remotely sensed data are based on a mean seasonal cycle of measurements by the MODIS instrument (Tramontana et al. 2016). One model is trained for all hours of the day. In the forward runs, the differentiation between the hours is achieved through hourly meteorological information, in particular potential radiation and its derivative. Daily and hourly meteorological information is derived from ECMWF ERA5 reanalysis data for the predictions. Potential radiation is calculated based on time of the year, day and location.

Compared to previous published work by Bodesheim et al. (2018), in this set-up hourly meteorological information on air temperature, vapour pressure deficit and incoming solar radiation at the surface are used as additional predictors compared to only daily ones. Moreover, half-hourly fluxes by Bodesheim et al. (2018) are based on CRUNCEP reanalysis data, whereas the hourly natural carbon fluxes delivered here use ERA5 meteorological information. This is expected to affect the results as the meteorological driving data set has been identified as one of the major factors determining magnitude and patterns of the estimated fluxes. The inclusion of sub-daily meteorological information also affects diurnal cycles compared to the typical shapes based on only daily meteorology (Fig.4).

3.2 Biogenic net and gross carbon fluxes

The data delivered represent hourly flux estimates in units of mumol m-2 s-1 at a spatial resolution of 0.5deg. Please note that the data uploaded on the CHE ftp server in the WP3 folder (upload October 2018) were based on the first ERA5 release from ECMWF at a native resolution at 0.3deg. Those were recalculated after the extended release of ERA5 extending back to 1979 at 0.25deg obtained from Copernicus in early 2019. With the delivery of this report those datasets on the ftp are updated, which can result in slight differences in simulated magnitudes. Also, now the first six hours of 2008 are available (in the first release, 2008 was the first available year and radiation as a forecast variable had not been available for hours 1-6 on January, 1st 2008).

In the following, some typical temporal and spatial patterns of the simulated fluxes are shown to introduce and illustrate the results.

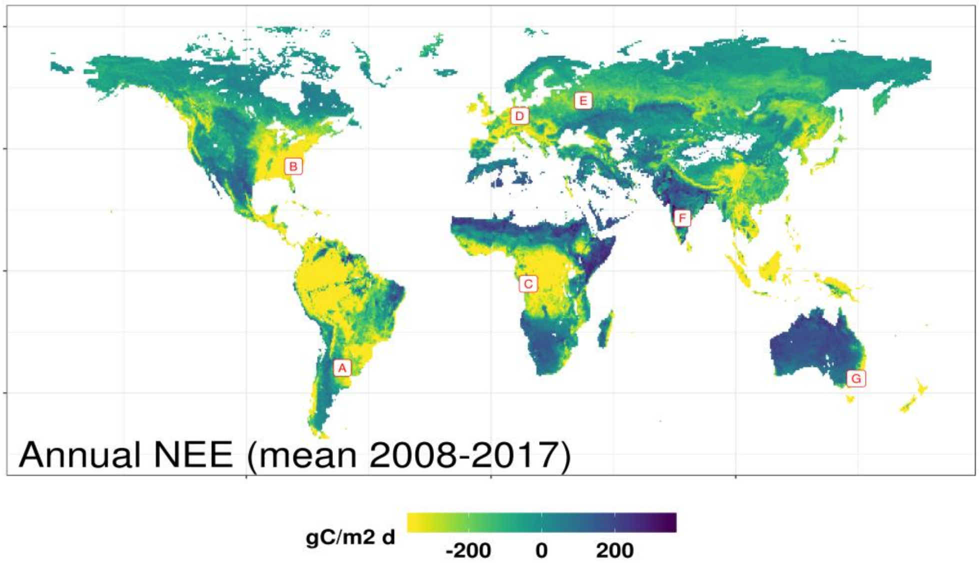

Figure 1: Total annual net ecosystem exchange averaged over 2008-2017. Spatial patterns are expected to be reliable. The tropics represent an unrealistically strong carbon sink.

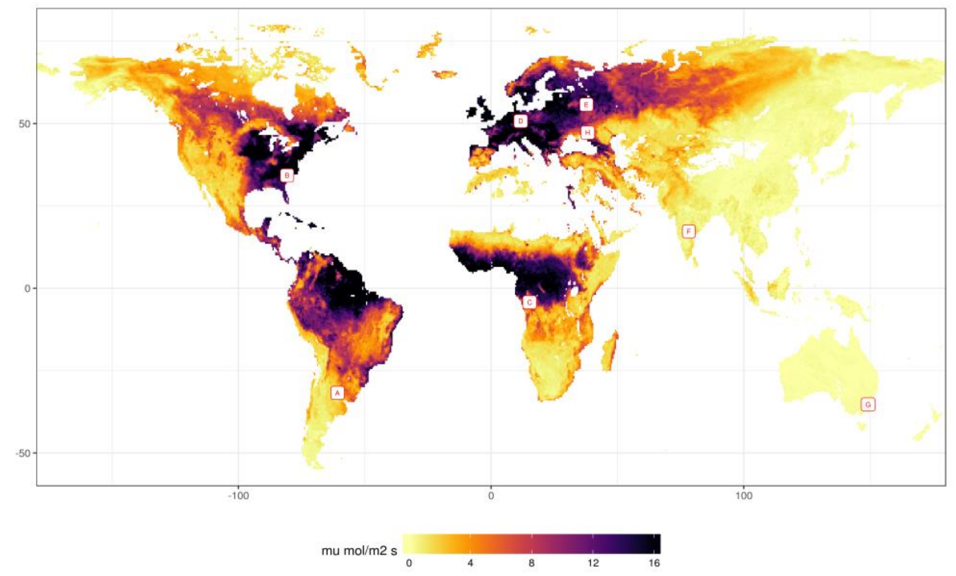

Figure 2: Snapshot of gross photosynthetic carbon uptake (GPP) predicted for 14UTC on DOY180 in 2008. The letters illustrate the places for which the diurnal and seasonal dynamics of GPP are shown in Fig.3.

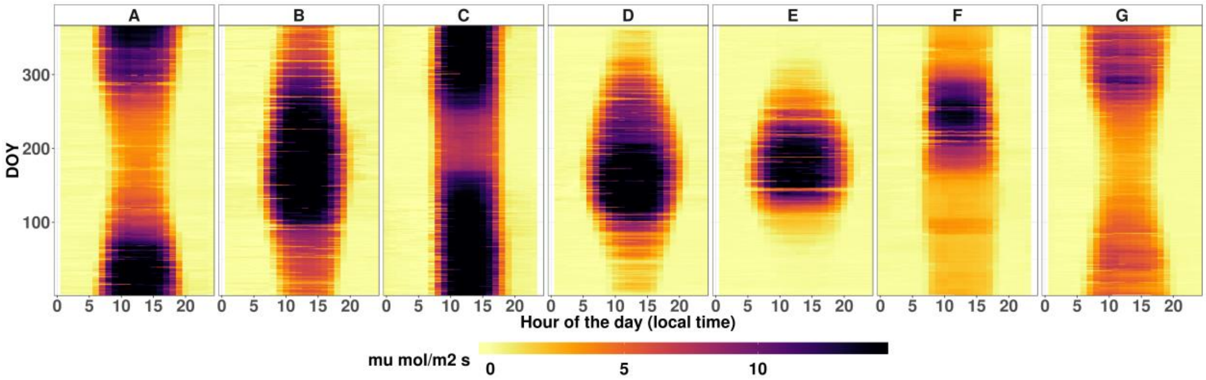

Figure 3: GPP in 2008 for the locations indicated in Figure 1 and Figure 2. Next to general differences in timing, duration and intensity of the growing season, these fingerprints illustrate short-term and sub-daily variability in productivity.

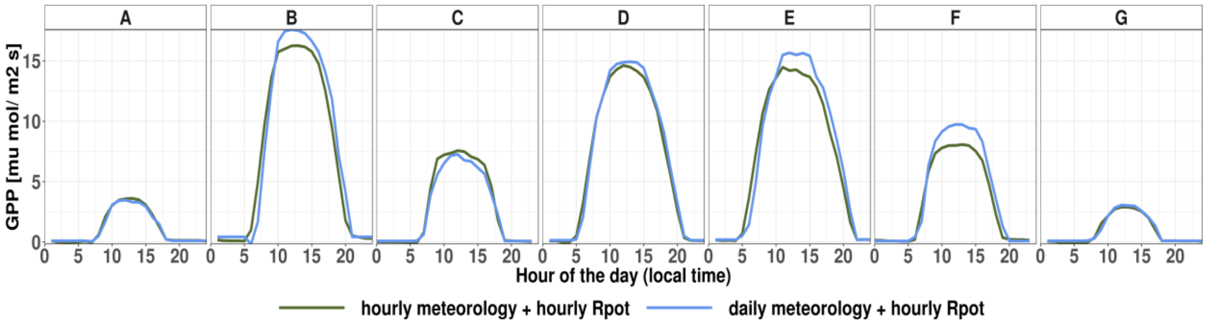

Figure 4: Average diurnal cycles of GPP in July 2008 for selected pixels. Comparison between the set-up in Bodesheim et al. 2018 using only daily meteorological information (CRUNCEP) and the setup of this deliverable using sub-daily meteorology from ERA5 illustrates differences in magnitude between a CRUNCEP and ERA5 forcing. It also shows a shift of the diurnal cycles towards morning when hourly meteorological information is available, probably because effects of atmospheric water demand can be better resolved using sub-daily VPD.

3.3 Uncertainty analysis

Evaluation of the model performance is only possible by comparison to net carbon flux measurements and estimates of the gross flux derived from them at eddy-covariance sites. These represent the best available direct measurements. We performed a consistent leave-one-site-out cross validation analysis, i.e. we repeated the training leaving out the data from one site and predicted the fluxes at this one site using meteorological information directly measured at the site. This has been done for each site in the training data set (191 sites with high quality fluxes and predictors available). The evaluation is based on half-hourly fluxes. In the following, the median across all sites of the individual coefficients of determination, the root-mean-squared errors as well as the Nash-Sutcliff modelling efficiencies (MEF, Nash and Sutcliff, 1970) are reported for GPP and NEE. The exact site-years used for validation vary between sites according to local data availability and range between 1991 and 2007.

Prediction of sub-daily variability (‘Mean monthly diurnal cycles’ in Table1) was assessed by aggregating to mean diurnal cycles per month (only if at least on ten days in a given month, year and site a valid value is available). Seasonality was evaluated by calculating mean seasonal cycles by aggregating first to daily means (if at least 24 half-hourly values in a given day, year and site were available), and subsequently averaging these daily averages over years per site. Finally, we assessed the ability of correctly predict deviations in the daily averages from the mean daily seasonality. Performance in the spatial domain is based on average fluxes per site.

In addition, we performed the cross-validation using meteorology extracted from ERA5 reanalysis data for a 0.25deg pixel closest to the respective site for the prediction. This approach is therefore closer to the method applied for the derivation of the global datasets. As hourly fluxes are simulated in the set-up with ERA5, we aggregate site-level observations to hourly first, and subsequently evaluate diurnal, seasonal, spatial and anomaly patterns.

Table 1: Overall leave-one-site-out cross validation results using either in-situ measured meteorological information or ERA5 reanalysis data. The performance is evaluated separately for NEE and GPP as well as for different temporal scales and in space. RMSE is given in units of mumol/ m² s.

	NEE (RMSE, r2, MEF)	GPP (RMSE, r2, MEF)
Tower meteo		
total	2.996 / 0.7 / 0.61	3.487 / 0.75 / 0.66
Mean daily seasonal cycles	1.351 / 0.65 / 0.47	1.718 / 0.82 / 0.69
Mean monthly diurnal cycles	1.705 / 0.88 / 0.76	1.977 / 0.91 / 0.80
Across-site-variability/ spatial	0.958 / 0.51 / 0.50	1.351 / 0.71 / 0.71
Deviations from mean daily seasonal cycles	0.458 / 0.44 / 0.46	0.531 / 0.40 / 0.43
ERA5 reanalysis meteo		
total	3.176 / 0.62 / 0.54	3.592 / 0.70 / 0.59
Mean daily seasonal cycles	1.514 / 0.59 / 0.39	1.986 / 0.78 / 0.60
Mean monthly diurnal cycles	1.864 / 0.84 / 0.70	2.213 / 0.88 / 0.73
Across-site-variability/ spatial	1.103 / 0.32 / 0.27	1.44 / 0.65 / 0.65

Deviations from mean daily seasonal cycles	0.554 / 0.34 / 0.35	0.641 / 0.30 / 0.30

Overall in the cross-validation using tower meteorology, the predictive power of our approach is slightly higher for GPP than for NEE judging from r2 and MEF (consistent with results in Tramontana et al., 2016), and diurnal patterns better than seasonal ones. NEE diurnal cycles are represented best, lower modelling efficiency occurs generally for seasonal patterns, particularly for shrublands and evergreen broadleaf forests (Fig.5). Those are often characterized by only few eddy-covariance sites and by gappy and contaminated remote sensing data due to clouds.

Similar conclusions can be drawn from the RMSE, r2 and MEF for the site-level predictions using ERA5 meteorology, despite slightly lower agreement of the simulated fluxes with in-situ observations. It is noteworthy, that despite high agreement in the diurnal and seasonal patterns, the performance got worse regarding spatial NEE patterns.

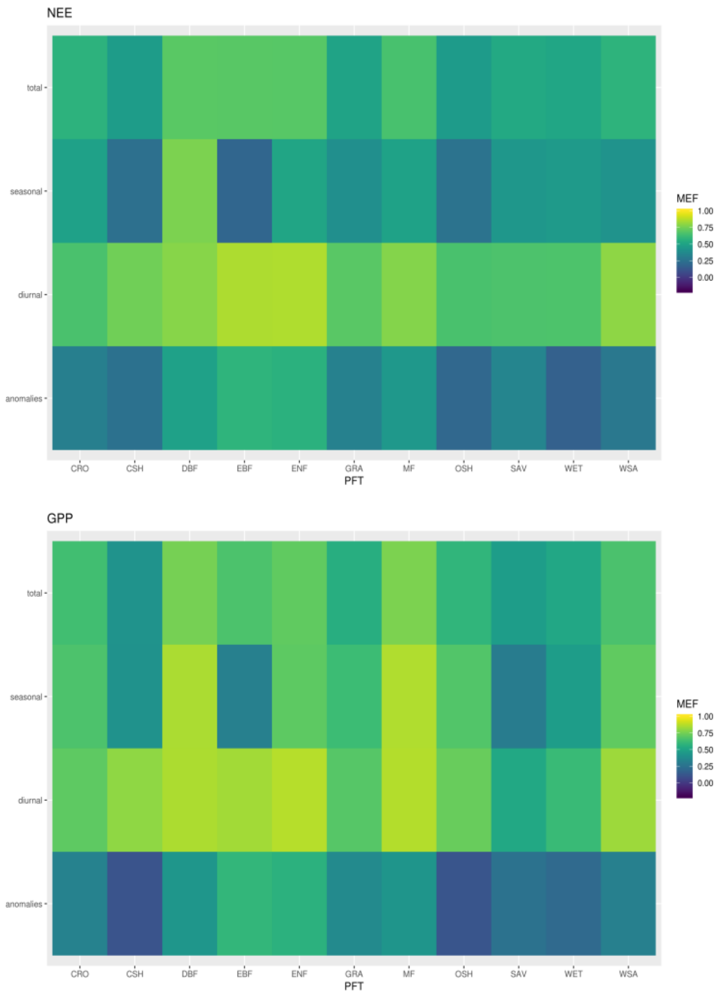

Figure 5: Average Nash-Sutcliff modelling efficiency for NEE (top) and GPP (bottom) across sites belonging to a certain plant functional type using in-situ measured meteorology as predictors.

Future efforts should focus on the uncertainty characterization in space and time (error-covariance-matrix). Furthermore, validation using independent flux estimates, e.g. from inversions or land surface models, would be valuable and desirable but bear the issue of scalemismatch between in-situ observations and grid cell sizes of typically 0.5deg/ 50km and are therefore barely representative.

4 Conclusion

We have built and delivered data-driven estimates of global terrestrial biogenic carbon fluxes (net and gross) at hourly resolution and half a degree spatial gridding as well as their associated uncertainties. This is part of WP3 (‘Coordinating efforts on uncertainty trade-off for fossil fuel emissions’) task 1. In CHE, these datasets can be used as priors in atmospheric inversions, as an independent data source to characterize uncertainties in a CCDAS framework or in an OSSE, or as an independent data set for cross-consistency checks, e.g. with carbon fluxes derived from land surface models. Furthermore, these datasets are of interest to the scientific community outside CHE for studies in the fields of ecology, hydrology, and meteorology and any studies evaluating the interactions between the land surface and the atmosphere.

5 References

Tramontana, G. et al. Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. Biogeosciences 13, 4291-4313 (2016)

Jung, M., et al. Compensatory water effects link yearly global land CO2 sink changes to temperature, Nature, 541(7638), 516-520 (2017)

Bodesheim, P., et al. Upscaled diurnal cycles of land–atmosphere fluxes: a new global half-hourly data product, Earth System Science Data, 10 (3), 1327-1365 (2018)

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., https://doi.org/10.1016/00221694(70)90255-6, 1970, 10, 282–290

6 List of abbreviations

- BRDF bidirectional-reflectance distribution function
- CHE CO2 human emissions
- DOY day of year
- ECMWF European Centre for Medium-Range Weather Forecast
- EVI enhanced vegetation index
- GPP Gross photosynthetic carbon uptake
- LST Land surface temperature
- MEF Modelling efficiency
- MODIS moderate resolution imaging spectroradiometer
- NDVI normalized difference vegetation index

- NDWI normalized difference water index
- NEE Net ecosystem exchange
- RMSE root mean squared error
- Rpot potential radiation (theoretical top of atmosphere incoming shortwave radiation)
- VPD vapour pressure deficit, measure of saturation of air with water vapour
- WP work package

Document History

Version	Author(s)	Date	Changes
0.1	Sophia Walther (Max-Planck-Institute for Biogeochemistry)	06/06/2019	Initial version of internal review
1.0	Sophia Walther (Max-Planck-Institute for Biogeochemistry)	20/06/2019	Minor adjustments based on reviewers' comments

Internal Review History

Internal Reviewers	Date	Comments
Thomas Kaminski (iLab)	13/06/2019	Approved with comments
Maximilian Harlander (ADSO)	11/06/2019	Approved with comments

Estimated Effort Contribution per Partner

Partner	Effort
Max-Planck-Institute for Biogeochemistry Jena	6 PM
Total	**6**