Global carbon fluxes using multi-output Gaussian processes regression and MODIS products

Journal:	*Journal of Selected Topics in Applied Earth Observations and Remote Sensing*
Manuscript ID	JSTARS-2024-01285
Manuscript type:	Regular
Date Submitted by the Author:	18-Apr-2024
Complete List of Authors:	Campos-Taberner, Manuel; Universitat de Valencia, Earth Physics and Thermodynamics GILABERT, MARIA AMPARO; UNIVERSITAT DE VALÈNCIA, FÍSICA DE LA TERRA I TERMODINÀMICA; Sánchez-Ruiz, Sergio; Universitat de Valencia, FÍSICA DE LA TERRA I TERMODINÀMICA MARTINEZ, BEATRIZ; UNIVERSITY OF VALENCIA, THERMODINAMICS; Jiménez-Guisado, Adrián; Universitat de Valencia, Earth Physics and Thermodynamics García-Haro, F. Javier; University of Valencia, Thermodynamics Guanter, Luis; Universitat Politecnica de Valencia, Applied Physics
Keywords:	Gaussian processes, Remote sensing, Vegetation mapping
Subject Category Please select the subject category that best reflects the scope of your manuscript:	Methodologies and Applications to: Vegetation and Land Surface

Global carbon fluxes using multi-output Gaussian processes regression and MODIS products

Manuel Campos-Taberner, María Amparo Gilabert, Sergio Sánchez-Ruiz, Beatriz Martínez, Adrián Jiménez-Guisado, Francisco Javier García-Haro, and Luis Guanter

***Abstract*—The quantification of carbon fluxes is crucial due to their role in the global carbon cycle having a direct impact on Earth's climate. In the last years, considerable efforts have been done to scale carbon fluxes from eddy covariance (EC) data to the globe. In this work, a data-driven approach that exploits a multi-output Gaussian processes regression algorithm ($\mathcal{G}$-model) is proposed to jointly estimate gross primary production (GPP), terrestrial ecosystem respiration (TER), and net ecosystem exchange (NEE) at global scale. The $\mathcal{G}$-model not only provides an estimate of the carbon fluxes but also an uncertainty. Moreover, it derives the three fluxes jointly preserving their physical relationship. The predictors are selected from a set of the MODerate-resolution Imaging Spectroradiometer (MODIS) products available on Google Earth Engine. The performance of the model revealed high accuracies (R^2 reaching 0.82, 0.69, and 0.80 in the case of GPP, NEE, and TER, respectively), and low root mean square errors (1.55 g m^{-2} d^{-1} in the case of GPP, 1.09 g m^{-2} d^{-1} for the NEE, and 1.14 g m^{-2} d^{-1} for TER) over the FLUXNET2015 data set at 8-day time scale. The GPP estimates provided by $\mathcal{G}$-model outperformed the MOD17A2 product, and a state-of-the-art GPP product (PML_V2) without using meteorological forcing data sets. The results reported mean annual amounts of 133.7 Pg yr^{-1}, 114.8 Pg yr^{-1}, and 18.9 Pg yr^{-1} for GPP, TER, and NEE, respectively during the 2002–2023 period. The proposed approach paves the way for the development of multi-output strategies that preserve the physical relationships among carbon fluxes in upscaling processes.**

***Index Terms*—Gross primary production (GPP), Terrestrial ecosystem respiration (TER), Net ecosystem exchange (NEE), machine learning; data-driven; Multi-output Gaussian process regression**

I. INTRODUCTION

CARBON fluxes (CFs) are essential for understanding the Earth's carbon cycle and its influence on the environment. The estimation of CFs is a measure of how much carbon is being exchanged between different components of Earth's climate system such as soils, vegetation, oceans and atmosphere. These exchanges occur via natural processes such as photosynthesis, respiration, and decomposition. The three main carbon fluxes are the gross primary production (GPP), terrestrial ecosystem respiration (TER), and net ecosystem exchange (NEE). GPP stands for the total amount of carbon that is fixed by plants through photosynthesis, whereas TER refers to the total amount of carbon that is released into the atmosphere by respiration of plants and other organisms present in the ecosystem. NEE is the difference between carbon uptake and carbon release by an ecosystem.

In-situ carbon fluxes are obtained from different networks of micrometeorological towers that use eddy covariance (EC) technique to estimate GPP from NEE and TER. Towers operate in many countries, across America, Europe and Asia, through the networks AmeriFlux (https://ameriflux.lbl.gov), ICOS (Integrated Carbon Observation System) (https://www.icos-cp.eu/), AsiaFlux (http://www.asiaflux.net), or the FLUXNET global network (https://fluxnet.org/). These measurements are representative of the corresponding EC tower footprint, which may vary from hectometric to kilometric scales depending on the site characteristics. At regional and global scales CFs can be estimated using models that simulate the carbon cycle adding the influence of several factors such as vegetation types, soil properties and climate conditions [1]–[3]. In addition, CFs can also be estimated from polar and geostationary remote sensing (RS) data such as the the MODerate resolution Imaging

This study was supported by Grant PID2020-118036RB-I00 funded by MCIN/AEI/ 10.13039/501100011033, and by LSA-SAF (EUMETSAT).
Corresponding author: Manuel Campos-Taberner.

Manuel Campos-Taberner is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: manuel.campos@uv.es).

María Amparo Gilabert is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: m.amparo.gilabert@uv.es).

Sergio Sánchez-Ruiz is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: sergio.sanchez@uv.es).

Beatriz Martínez is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: beatriz.martinez@uv.es).

Adrián Jiménez-Guisado is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: adrian.jimenez@uv.es).

Francisco Javier García-Haro is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: j.garcia.haro@uv.es).

Luis Guanter is with the Land and Atmosphere Remote Sensing group (LARS), Universitat Politècnica de València, Valencia, Spain. (e-mail: lguanter@fis.upv.es).

Supplementary material and color versions of the figures in this article are available online at http://ieeexplore.ieee.org

Spectroradiometer (MODIS) [4] or the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) on board the Meteosat Second Generation (MSG) [5], which are based on Production Efficiency Models (PEMs). A simple yet effective method to estimate GPP from RS data is to employ vegetation indices or spectral bands combinations as proxies of GPP [6]–[7]. Similarly, Solar-induced fluorescence (SIF) can be used for inferring GPP [8]–[9]. Other approaches link RS and meteorological data with EC measures [10]–[15]. These data-driven techniques rely on the relationship among vegetation properties and carbon fluxes to build a single model (one per flux to be retrieved) with machine learning (ML) methods. The use of a single ML model for every predictive variable can lead to good independent performances but does not guarantee that the relationship among variables is respected, leading to non-robust approaches. Some studies have shown that the use of a unique multi-output model could improve the results when estimating biophysical parameters at the same time [16]–[17]. However, this technique has not been tested on upscaling global carbon fluxes exchange yet, where the balance between carbon uptake through photosynthesis and carbon release through respiration leads to net ecosystem exchange GPP-TER=NEE.

In this framework, the objectives of this paper are (i) to develop a data-driven approach to estimate CFs at global scale blending MODIS products and EC data, (ii) to build a single multi-output model able to jointly estimate GPP, NEE, and TER in a robust way, and (iii) to validate the estimates against global in-situ data, and compare with two reference products (MOD17A2H and PML_V2). The approach relies on the use of multitemporal MODIS products from Google Earth engine (GEE) as inputs in a multi-output Gaussian Processes Regression algorithm ($\mathcal{G}$-model), which exploits the covariance among CFs. The remainder of the manuscript is structured as follows. Section II describes in-situ and remote sensing products, and the proposed multi-output algorithm. Section III shows the model performance, the explanatory power of the inputs, the derived global carbon fluxes maps, and a comparison with other products. Section IV discusses the results, and section V highlights the main conclusions.

II. MATERIALS AND METHODS

A. Tower data

The approach proposed in this study exploits data coming from the FLUXNET network. The FLUXNET2015 dataset was downloaded (https://fluxnet.org/data/fluxnet2015-dataset/), which allowed the use of data coming from 211 sites globally distributed (see Supplementary Fig. SI). In particular, we used the dataset update dating from February 6, 2020. The distribution of the flux sites according to the International Geosphere-Biosphere Programme (IGBP) can be found as Supplementary Table SI. Daily GPP, NEE, and TER data were computed for every site from the average of the night- and day-time partitioning methods present in the FLUXNET2015 dataset [18]. The sign convention used in this study was such that CO_2 flux to the surface was positive so that NEE=GPP-TER, therefore positive values of NEE stand for vegetation carbon uptake, and negatives for release. The NEE quality flag was used to ensure the consistency between GPP, TER and NEE, and similarly to other studies we excluded daily data when more than 20% of the half-hourly data were based on gap-filling with low confidence [11],[19]. Finally, daily data were temporally aggregated (8-day) to match the RS products.

B. Remote sensing products

The biosphere-atmosphere interactions related to photosynthesis are mainly driven by Photosynthetically Active Radiation (PAR), as well as environmental and physiological variables, such as water availability, temperature, chlorophyll content, and leaf area index (LAI) [20]. On this basis, five MODIS-based RS products were used to extract the predictors for the regression method. Overall, 8 predictors were selected, namely: PAR from the MCD18C2 product; day- and night-time land surface temperature (LST_D, LST_N) from the MOD11A2 product; Evapotranspiration (ET) and Potential Evapotranspiration (PET) from the MOD16A2 product, as well as the water stress factor (C_{ws}) computed as $C_{ws}=\frac{ET}{PET}$; the kernel version of the normalized difference vegetation index (kNDVI) [21] computed from the MCD43A4 red and NIR bands; and LAI from the MCD15A2H product. The GEE platform was used to download the selected MODIS products over the flux sites from 2002 (when starts the availability of all considered MODIS products) to 2014 (when ends the availability of the FLUXNET2015 data), thus fitting the coincident temporal range between the FLUXNET2015 dataset and the RS observations. The final value of the predictors that was assigned to every tower was the mean of the pixel values contained in a 500-m buffer weighted by their spatial extension. The quality flag of every predictor was also used to discard bad-quality pixels. In the case of the MCD18C2 product, the values were integrated to obtain a daily-integrated PAR, and aggregated to 8-day for temporal consistency among the rest of the predictors and EC data. Eventually, two reference GPP products (MOD17A2H, and PML_V2) were downloaded from GEE with the same aforementioned processing to compare the GPP retrievals. Table I highlights the main features of the products.

TABLE I
MAIN FEATURES OF THE SELECTED MODIS-BASED PRODUCTS

	GEE name	Spatial resolution	Temporal frequency	Reference
Predictors	MCD18C2.061	0.05º	3 hours	[22]
	MCD43A4.061	500 m	daily	[23]
	MOD11A2.061	1 km	8 days	[24]
	MOD16A2.061	500 m	8 days	[25]
	MCD15A2H.061	500 m	8 days	[26]
For comparison	MOD17A2H.061	500 m	8 days	[4]
	PML_V2 0.1.7	500 m	8 days	[13]

C. Multi-output Gaussian processes regression ($\mathcal{G}$-model)

Gaussian processes (GPs) [27] are statistical techniques that offer a probabilistic approach for kernel regression methods in which the likelihood of unobserved latent functions is supposed to be modeled by a multivariate Gaussian prior. The approach makes use of a latent function of the inputs $f(\mathbf{x})$ (the RS predictors) plus constant Gaussian noise to estimate the outputs (the carbon flux of interest) as $y=f(\mathbf{x})+\varepsilon_n$, $\varepsilon_n \sim N(0,\sigma_n^2)$. The latent function is chosen to be a zero mean Gaussian process prior, or $(x)\sim GP(0,k_\theta)$, where k_θ is a covariance function parametrized by θ. The noise power is given by the hyperparameter σ_n^2, and it is assumed that the noise ε_n follows a prior Gaussian distribution. Given the GP priors, samples drawn from $f(x)$ at $\mathbf{x}_i=\{x_i^1,\ldots,x_i^B\}_{i=1}^N$ (N is the number of training samples, and B the number of predictors) follow a joint multivariate zero mean Gaussian distribution defined by the covariance matrix $\mathbf{K}$, where $[\mathbf{K}]_{ij}=k_\theta(x_i,x_j)$.

The output of the model is computed from the predictive distribution described by the equations:

$$p(y_*|\mathbf{x}_*,D)=N(y_*|\mu_{GPR*},\sigma^2_{GPR*}), \quad (1)$$

$$\mu_{GPR*}=\mathbf{k}_*^\top(\mathbf{K}+\sigma_n^2\mathbf{I}_n)^{-1}\mathbf{y}=\mathbf{k}_*^\top\alpha, \quad (2)$$

$$\sigma^2_{GPR*}=\sigma^2+k_{**}-\mathbf{k}_*^\top(\mathbf{K}+\sigma_n^2\mathbf{I}_n)^{-1}\mathbf{k}_*, \quad (3)$$

where D stands for the training dataset (being $\mathbf{x}_*$ and y_* the input and output), $\mathbf{k}_*=[k(x_*,x_1),\ldots,k(x_*,x_N)]$ is an $N\times 1$ vector and $k_{**}=k(x_*,x_*)$. The model provides a full posterior probability from which predictions (μ_{GPR*}) and confidence estimates (σ^2_{GPR*}) can be computed. In the single output approach $\mathbf{y}\in\mathbb{R}^{N\times 1}$, the output of interest (e.g., GPP) is computed from the inputs (8 RS predictors, $\mathbf{x}_i\in\mathbb{R}^{N\times 8}$) as:

$$\hat{y}=f(\mathbf{x})=\sum_{i=1}^{N}\alpha_i k_{\boldsymbol{\theta}}(\mathbf{x}_i,\mathbf{x})+\alpha_0, \quad (4)$$

where α_i is the weight assigned to each predictor $\mathbf{x}_i$, α_0 is the bias term, and $k_{\boldsymbol{\theta}}$ is the covariance function evaluating the similarity between $\mathbf{x}$ and all the N training spectra. For this task the automatic relevance determination kernel was selected

$$K(\mathbf{x}_i,\mathbf{x}_j)=\nu\exp\left(-\sum_{b=1}^{B}\left(x_i^b-x_j^b\right)^2/2\sigma_b^2\right)+\sigma_n^2\delta_{ij}, \quad (5)$$

where ν is a scaling factor, σ_n accounts for the standard deviation of the noise, B is the number of predictors (in our case, B=8), and σ_b can be related to the spread of each predictor b. The model hyperparameters $\boldsymbol{\theta}=[\nu,\sigma_n,\sigma_1,\ldots,\sigma_b]$ can be estimated by maximizing the marginal log-likelihood [27]:

$$\log p(\mathbf{y}|\mathbf{x}_i,\theta)=-\frac{1}{2}\mathbf{y}^\top(\mathbf{K}+\sigma_n^2\mathbf{I}_n)^{-1}\mathbf{y}-\frac{1}{2}\log|\mathbf{K}+\sigma_n^2\mathbf{I}_n|-\frac{N}{2}\log(2\pi). \quad (6)$$

The model can be adapted to deal with multi-output regression problems ($\mathcal{G}$-model) adapting the kernel hyperparameters for a unique kernel which is able to deal with all the outputs. In the $\mathcal{G}$-model we need to optimize the hyperparameters considering that $\mathbf{y}\in\mathbb{R}^{N\times D}$ instead $\mathbf{y}\in\mathbb{R}^{N\times 1}$ where D is the total number of outputs (in our case, D=3). Therefore, a global cost function summarizing all the cost functions (one per output) into a single one must be defined. Here we use a global cost function C just as the squared sum of the standard model cost function of each output [17]:

$$C=\sum_{d=1}^{D}\log p(\mathbf{y}_d|\mathbf{x})^2=\sum_{d=1}^{D}\left(-\frac{1}{2}\mathbf{y}_d^\top\boldsymbol{\alpha}-\sum_{j=1}^{N}\log L_{jj}-\frac{N}{2}\log(2\pi)\right)^2, \quad (7)$$

where L_{jj} is the Cholesky factorization of the covariance matrix for every output.

D. Algorithm calibration and performance

The $\mathcal{G}$-model was evaluated with 14562 samples obtained matching the EC data and the RS products. The inputs underwent min-max scaling to transform them into the 0–1 range, as the predictors exhibited a substantial range of variation. Employing this technique is highly advisable to mitigate the risk of suboptimal performance by machine learning algorithms. The 70% of the samples were randomly selected for training and we assessed the results in the remaining 30% (test set unused during model training). Different random initializations were conducted and no significant changes in performance were obtained (not shown for brevity). Root mean square error (RMSE), mean error (ME), mean absolute error (MAE), and the coefficient of determination (R^2) were computed over the test set to evaluate the accuracy, bias, and the goodness-of-fit of the model.

In addition, the model behavior was tested by identifying the most relevant predictors for every estimated flux. This was achieved by means of a sensitivity analysis carried out through an added-noise permutation approach [28]. The inputs were perturbed with Gaussian white noise $N(0,\sigma^2)$, being σ^2 the 3% of the perturbed signal amplitude. The noise was added to a single predictor remaining the rest of the predictors unperturbed, and repeating this process for every predictor. The predictor relevance for every flux was computed as the difference between the accuracy obtained with no perturbation, and the obtained when the perturbation was applied. Finally, the results were normalized with respect to the most relevant predictor.

III. Results

A. Assessment of the model performance over tower data

Fig. 1 shows the direct validation obtained over the test set. The accuracy metrics revealed good correspondence between estimations and in-situ measurements. It is worth mentioning the high R^2 values, which reach 0.82, 0.69, and 0.80 in the case of GPP, NEE, and TER respectively. RMSE was 1.55 g m^{-2} d^{-1} for GPP, 1.09 g m^{-2} d^{-1} for NEE, and 1.14 g m^{-2} d^{-1} for TER. The NEE and TER predictions revealed low and very similar MAE and ME (see Fig. 1) whereas GPP exhibit slightly higher MAE (1.04 g m^{-2} d^{-1}) and lower ME (0.01 g m^{-2} d^{-1}). In addition, Fig. 2(d-e) shows the performance of MOD17A2 and PML_V2 products over the same test set. Note that these products do not provide NEE nor TER. The RMSEs and MAEs of MOD17A2 and PML_V2 are 2.44 g m^{-2} d^{-1} and 2.11 g m^{-2} d^{-1}, and 1.66 g m^{-2} d^{-1} and 1.48 g m^{-2} d^{-1}, respectively. The reported biases regarding EC data are -0.50 g m^{-2} d^{-1} and -0.31 g m^{-2} d^{-1} in the case of MOD17A2 and PML_V2, respectively.

TABLE II
ACCURACY ASSESSMENT PER BIOME IN THE EC VALIDATION SET

	GPP				NEE				TER			
	ME	RMSE	MAE	R^2	ME	RMSE	MAE	R^2	ME	RMSE	MAE	R^2
CRO	0.62	2.56	1.93	0.57	0.37	1.26	0.91	0.77	0.32	1.26	0.96	0.58
CSH	0.61	1.18	0.92	0.69	0.14	0.64	0.53	0.31	0.14	0.62	0.51	0.73
DBF	0.23	1.77	1.26	0.81	0.29	1.23	0.94	**0.80**	0.20	1.13	0.89	0.60
DNF(*)	0.61	0.85	0.64	0.78	1.19	1.37	1.19	0.74	1.25	1.39	1.25	0.42
EBF	0.15	1.33	0.93	0.88	**0.01**	1.02	0.75	0.41	0.06	1.17	0.84	**0.88**
ENF	0.28	1.41	1.02	0.79	0.00	1.27	0.84	0.51	0.07	1.40	0.88	0.73
GRA	0.17	1.31	0.87	0.88	0.28	1.15	0.75	0.46	0.28	1.22	0.79	0.87
MF	0.25	1.52	1.18	0.74	0.02	0.94	0.75	0.75	0.17	1.03	0.84	0.52
OSH	0.07	0.73	0.53	0.74	0.07	**0.48**	**0.34**	0.23	0.16	0.52	0.37	0.72
SAV	0.15	0.94	0.63	0.83	0.04	0.57	0.43	0.76	**0.05**	0.57	0.44	0.87
WET	0.16	2.11	1.47	0.70	0.20	1.14	0.87	0.79	0.14	1.10	0.80	0.77
WSA	**0.01**	**0.58**	**0.45**	**0.91**	0.09	0.51	0.36	0.71	**0.05**	**0.45**	**0.33**	0.86

Best results highlighted in bold. GPP, NEE; TER, ME, RMSE and MAE in g m^{-2} d^{-1}. Codes: croplands (CRO), closed shrublands (CSH), deciduous broadleaf forest (DBF), deciduous needleleaf forests (DNF), evergreen broadleaf forests (EBF), evergreen needleleaf forests (ENF), grasslands (GRA), mixed forests (MF), open shrublands (OSH), savannas (SA), wetlands (WET), and woody savannas (WSA). (*) Only one site.

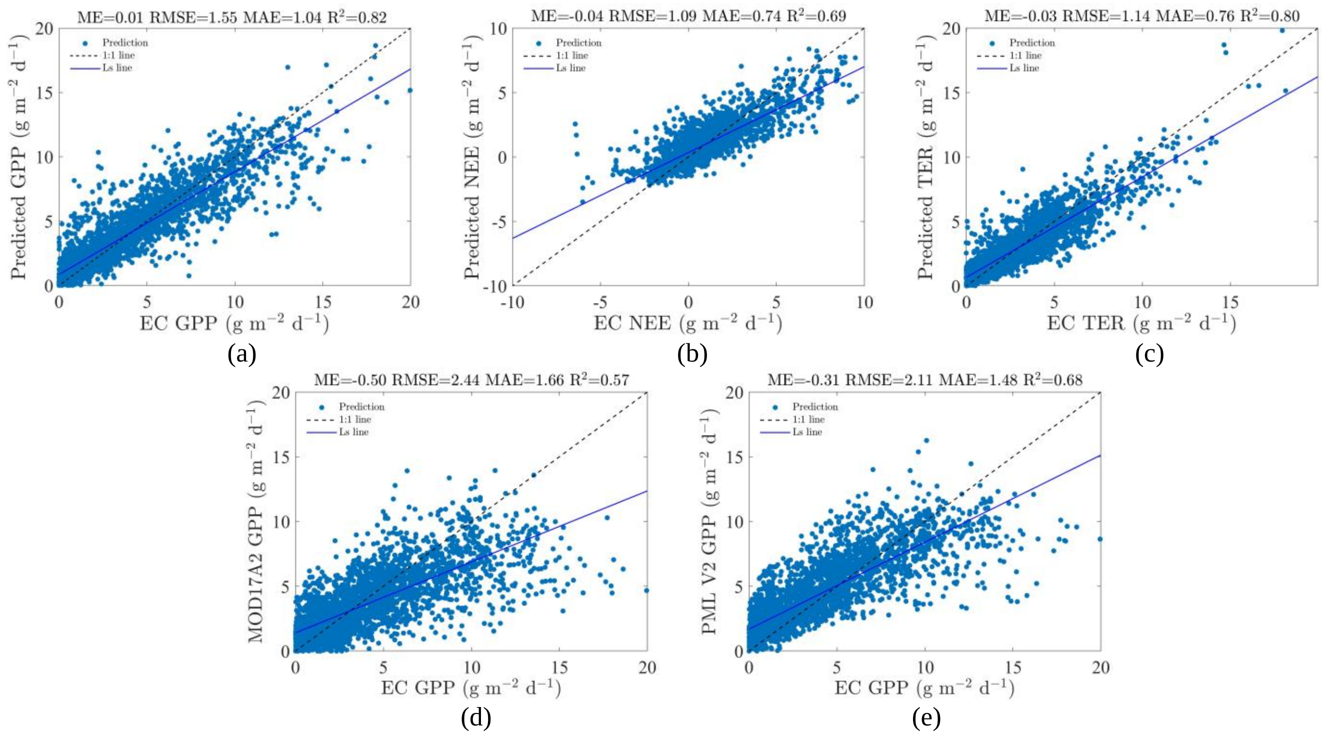
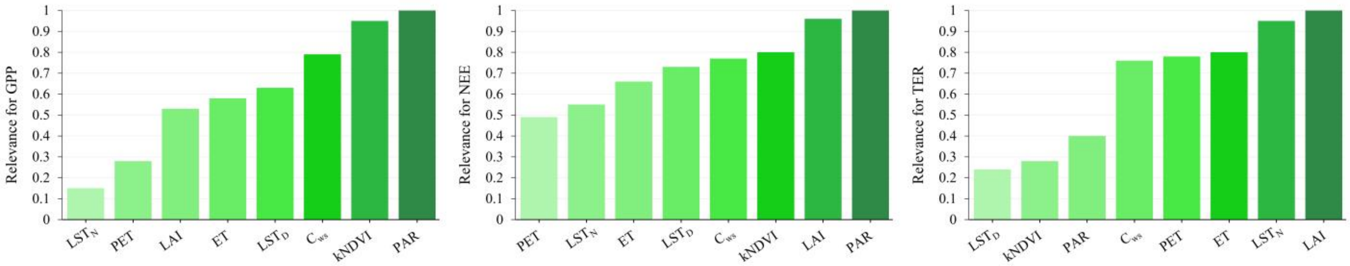


Fig. 1. Carbon fluxes, (a) GPP, (b) NEE, and (c) TER obtained with the $\mathcal{G}$-model over the EC validation set. ME, RMSE and MAE units are g m^{-2} d^{-1}. GPP estimates of (d) MOD15A2 and (e) PML_V2 GPP obtained over the EC validation set. ME, RMSE and MAE units are g m^{-2} d^{-1}.

The $\mathcal{G}$-model accuracy metrics were disaggregated per biome type as shown in Table II. Note that the results over deciduous needleleaf forests are obtained only over 1 available site and may not be representative of this biome. The best results in all metrics for GPP were obtained over woody savannas. In the case of NEE the best accuracy metrics were observed over evergreen broadleaf forests and deciduous needleleaf forests (for ME and R^2), and open shrublands (for RMSE and MAE). The lowest errors for TER were reported over woody savannas, whereas the highest correlation was found over evergreen broadleaf forests.

B. Sensitivity analysis

Fig. 2 shows the relevance of every predictor for GPP, NEE and TER retrievals. The relevance of every predictor was normalized with the value of the maximum relevance predictor in every case. The $\mathcal{G}$-model sensitivity analysis revealed the daily integrated PAR from the MCD18C2 product as the most influential input for both GPP and NEE estimates. LAI from the MOD15A2 product proved as the most relevant predictor for TER retrievals.



(a) (b) (c)

Fig. 2. $\mathcal{G}$-model predictor importance for (a) GPP, (b) NEE, and (c) TER.

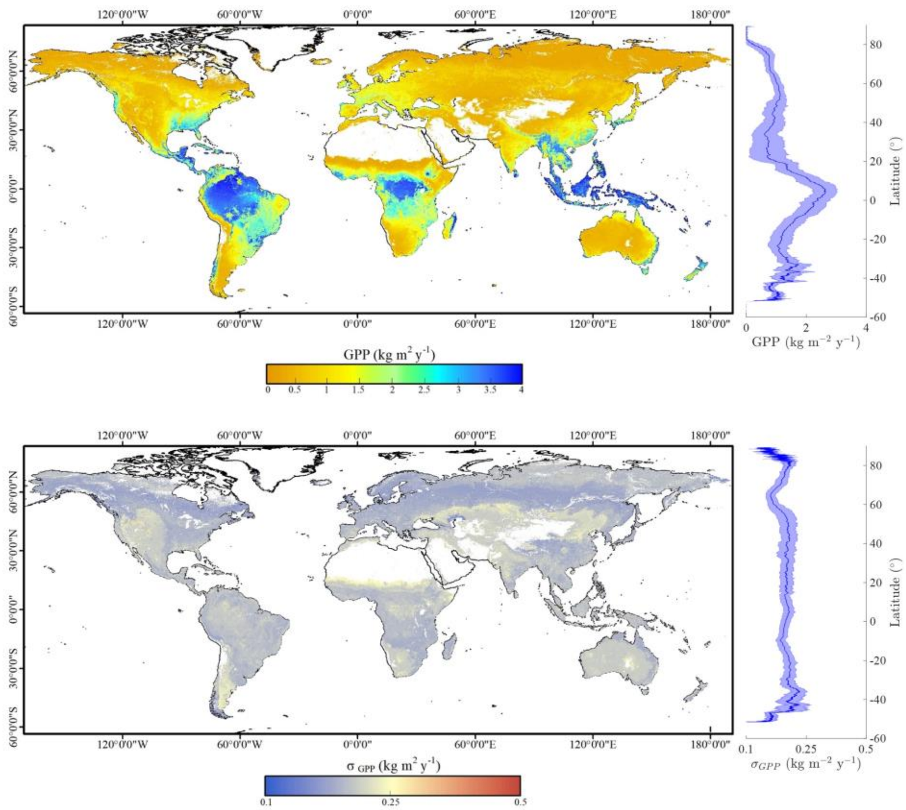


Fig. 3. Global average maps for annual GPP (top), and model uncertainty σ_{GPP} (bottom) computed for the 2002–2023 period.

C. Mean annual fluxes

After the completion of the $\mathcal{G}$-model training phase, the model was effectively applied to derive global GPP, NEE and TER maps. Figs. 3–5 show the mean annual predictions and uncertainties of every flux computed for the 2002–2023 period along with their latitudinal profiles. The spatial pattern of the GPP (Fig. 3, top) and TER (Fig. 4, top) shows that tropics, encompassing regions like Southeast Asia, central Africa, and South America, exhibit a notable high mean annual GPP. Conversely, high latitudes, western North America, and both central Asia and Australia, are characterized by considerably lower mean annual GPP and TER. NEE map (Fig. 5, top) presents a similar spatial pattern. It is worth mentioning that the spatial distribution of the mean annual uncertainty of every flux is slightly different. The σ_{GPP} (Fig. 3, bottom) exhibits a relatively consistent latitudinal profile, hovering around a value of 0.2 kg m^{-2} yr^{-1}. However, there is a zone between 62–65º N where the values are lower (see blueish colors in the north of the σ_{GPP} map). In the case of σ_{TER} and σ_{NEE} their spatial patterns are similar (Figs. 4–5, bottom). Their latitudinal uncertainty values are broadly constant around 0.15–0.2 kg m^{-2} yr^{-1} with no major variations, however σ_{NEE} presents lower spatial variability (Fig. 5, bottom).

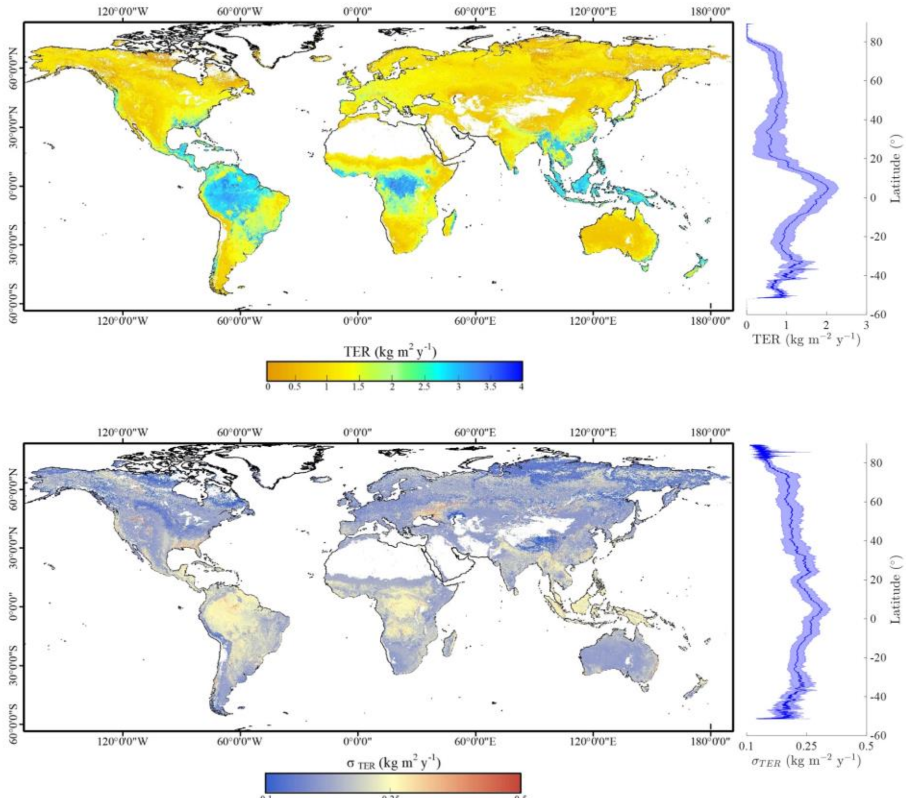


Fig. 4. Global average maps for annual TER (top), and model uncertainty σ_{TER} (bottom) computed for the 2002–2023 period.

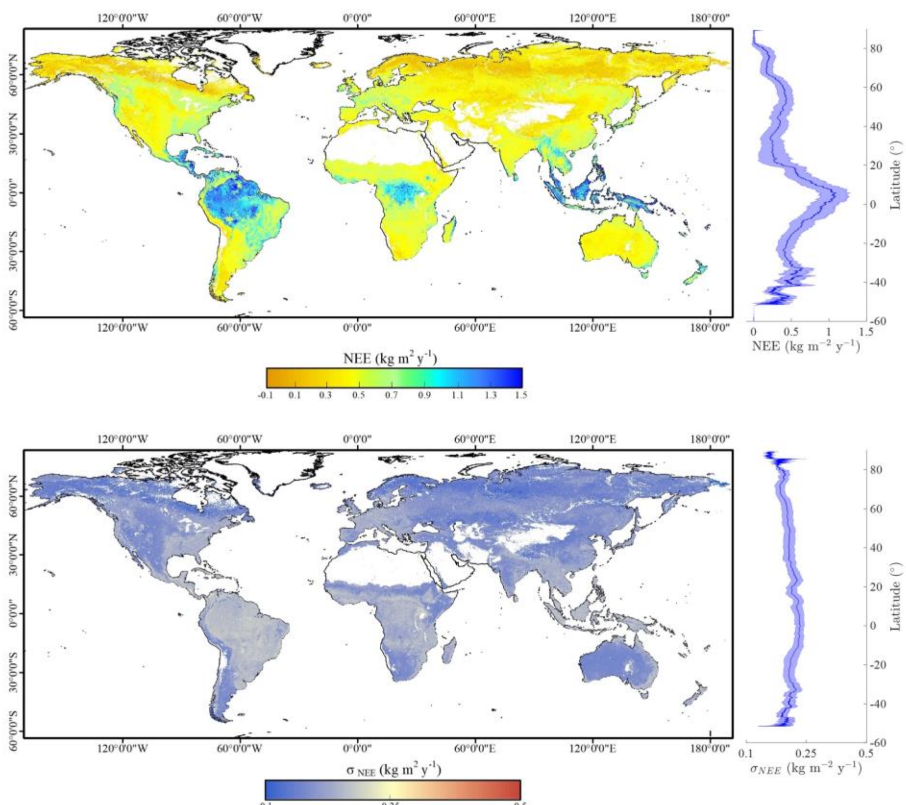
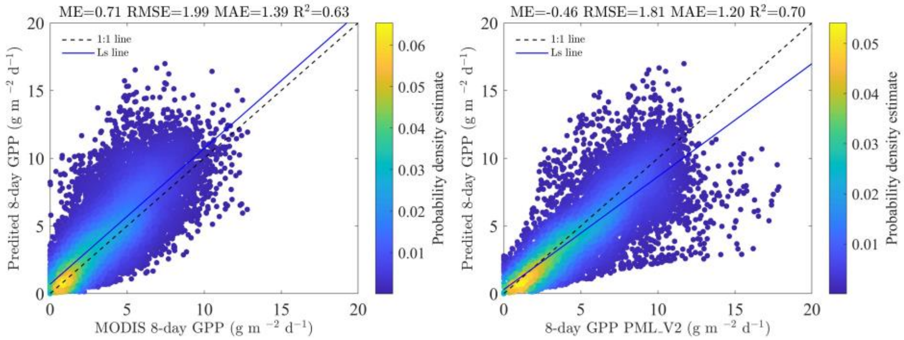


Fig. 5. Global average maps for annual NEE (top), and model uncertainty σ_{NEE} (bottom) computed for the 2002–2023 period.



(a) (b)

Fig. 6. GPP comparison between the proposed approach and (a) the MOD17A2, and (b) the PML_V2 products, respectively. ME, RMSE and MAE units are g m^{-2} d^{-1}.

TABLE III

GPP ACCURACY METRICS PER BIOME WITH REGARD THE MOD17A2 AND PML_V2 PRODUCTS

	Vs MOD17A2				Vs PML_V2			
	ME	RMSE	MAE	R^2	ME	RMSE	MAE	R^2
CRO	0.84	2.51	1.91	0.35	-1.23	3.26	2.48	0.26
CSH	0.82	1.48	1.05	0.26	-0.84	1.03	0.87	**0.83**
DBF	0.20	1.90	1.39	0.67	-0.70	1.72	1.21	0.79
DNF(*)	-0.80	1.38	1.50	0.81	-0.98	1.16	0.98	0.88
EBF	0.21	2.11	1.60	0.64	-0.17	1.55	1.17	0.77
ENF	0.44	1.98	1.46	0.49	-0.40	1.71	1.17	0.61
GRA	0.28	1.88	1.31	0.67	-0.29	1.53	1.03	0.78
MF	0.36	2.17	1.55	0.51	-0.56	2.29	1.57	0.52
OSH	-0.72	**1.22**	**0.90**	**0.72**	-0.68	1.07	0.78	0.80
SAV	0.21	1.24	0.92	0.56	**0.07**	1.00	0.70	0.68
WET	0.38	2.18	1.89	0.49	-0.48	2.17	1.59	0.54
WSA	-0.22	1.27	0.93	0.59	-0.21	**0.78**	**0.57**	0.81

Best results highlighted in bold. GPP, NEE; TER, ME, RMSE and MAE in g m^{-2} d^{-1}. Codes: croplands (CRO), closed shrublands (CSH), deciduous broadleaf forest (DBF), deciduous needleleaf forests (DNF), evergreen broadleaf forests (EBF), evergreen needleleaf forests (ENF), grasslands (GRA), mixed forests (MF), open shrublands (OSH), savannas (SA), wetlands (WET), and woody savannas (WSA). (*) Only one site.

The results reported global amounts of 133.7 Pg yr^{-1}, 114.8 Pg yr^{-1}, and 18.9 Pg yr^{-1} for GPP, TER, and NEE, respectively in the 2002–2023 period. Note that those values preserve the relationship among carbon fluxes (NEE=GPP-TER). In the case of MOD17A2 and PML_V2 products we found a GPP global mean of 118.3 Pg yr^{-1} and 143.6 Pg yr^{-1}, respectively. Note that in the case of the PML_V2 the period was 2002–2020 due to data availability.

D. Comparison with MOD17A2 and PML_V2 products

The GPP estimates were compared with the provided ones by the MOD17A2, and the PML_V2 products available on GEE. The comparison was carried out at site level (Fig. 6). The R^2 values regarding the MOD17A2 and PML_V2 were 0.63 and 0.70, respectively. The $\mathcal{G}$-model predictions present a positive bias (0.71 g m^{-2} d^{-1}) regarding the MOD17A2, and a negative one with respect to the PML_V2 estimates (-0.46 g m^{-2} d^{-1}). Table III shows the comparison per biome type. The best concordances regarding the MOD17A2 product have been found over open shrublands in terms of RMSE, MAE, and R^2, and over crops in terms of ME. Over closed shrublands has been found the highest agreement regarding the PML_V2 estimates, but the lowest R^2 has been observed regarding the MOD17A2. The lowest correspondences regarding the two reference datasets have been reported over crops.

Fig. 7 shows difference maps of the mean annual GPP regarding the reference datasets (Δ_{GPP}=GPP$_{\mathcal{G}\mathrm{model}}$-reference). In general, the $\mathcal{G}$-model agrees well with the PML_V2, however a slight overestimation of the GPP estimates provided by the PML_V2 regarding the $\mathcal{G}$-model is reported at high latitudes (see Fig. 7, top). Fig. 7 (bottom) shows the MOD17A2 estimates systematically underestimates GPP regarding the $\mathcal{G}$-model (mainly in the tropics) except over some locations.

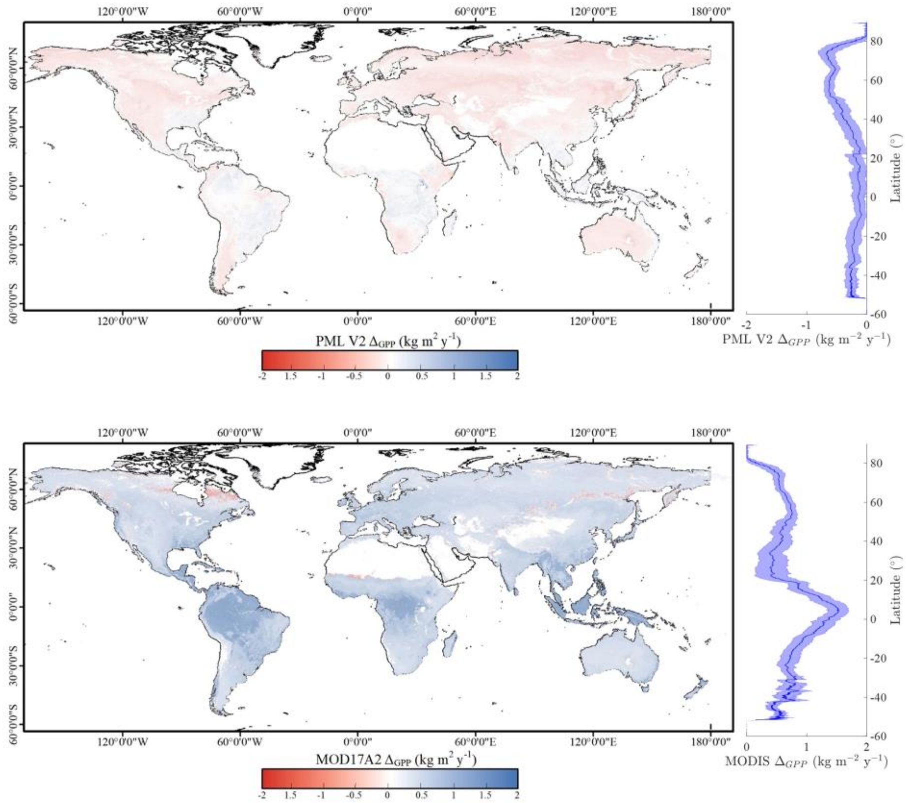


Fig. 7. Global difference maps for mean annual GPP regarding (top) PML_V2 product, and (bottom) MOD17A2 for the 2002–2020 period.

Finally, Fig. 8 shows the global GPP computed for every year during the 2002–2023 period in the case of the $\mathcal{G}$-model, the MOD17A2, and the PML_V2 product (2002–2020 period). Global GPP provided by the $\mathcal{G}$-model varies from 131.8 Pg C yr^{-1} to 137.6 Pg yr^{-1}, whereas the MOD17A2 reported lower amounts in the same period (118.1 Pg yr^{-1} to 119.9 Pg C yr^{-1}). The PML_V2 product showed a range of 128.5 Pg yr^{-1} to 133.9 Pg yr^{-1} (from 2002 to 2020). The corresponding linear trends in global GPP reported by the $\mathcal{G}$-model, the PML_V2, and the MOD17A2 are 0.28±0.02 Pg yr^{-1} yr^{-1}, 0.21±0.03 Pg yr^{-1} yr^{-1}, and 0.09±0.02 Pg yr^{-1} yr^{-1}, respectively.

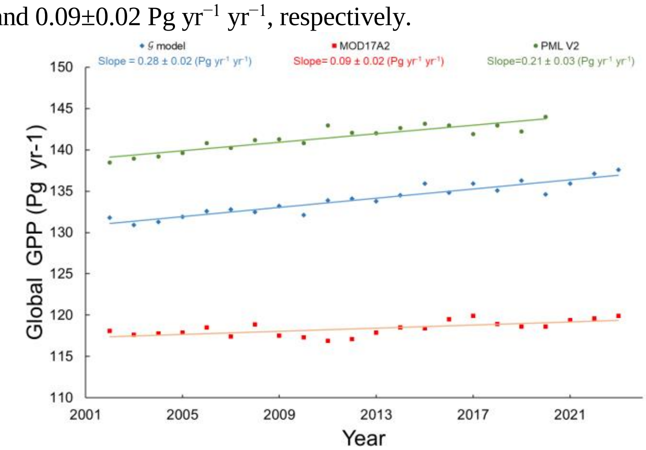


Fig. 8. Trends for mean annual GPP obtained by the G-model and the MOD17A2 for the 2002–2023 period, and the PML_V2 for 2002 to 2020.

IV. Discussion

A. Model performance and explanatory variables

The state-of-the-art machine learning approaches that provide the highest accuracies are based on a combination of RS observations and meteorological variables as inputs [14]–[16]. Nevertheless, this study exclusively utilized a set of RS predictors derived from MODIS data into a machine learning regression method, resulting in a remarkable level of concordance between GPP, NEE, and TER retrievals and in-situ data. The interactions between CO_2 fluxes and RS/meteo variables depend significantly on ecosystem types and state [29]. A main advantage of the proposed approach is its potential to obtain globally those fluxes without a priori knowledge of the pixel characteristics. The results of the accuracy assessment show an outperformance of the $\mathcal{G}$-model regarding both MOD17A2, and PML_V2 GPP products. The highest correlations for GPP and TER were obtained over woody savannas, savannas, grasslands, (in line with the results of Joiner *et al.* [19]) and evergreen broadleaf forests. In the case of NEE the best correlation was obtained over deciduous broadleaf forest, wetlands and crops. Crops is the biome where the worst results were reported in the case of GPP and TER ($R^2 \leq 0.58$), however the model performance in terms of correlation for NEE over this biome was good (R^2 = 0.77). Closed and open shrublands are the biomes where the lowest correlations for NEE ($R^2 \leq 0.31$) have been found.

The $\mathcal{G}$-model consistently predicted GPP, NEE, and TER while effectively maintaining their inherent relationship. This is an important feature of this work since the usual procedure in machine learning approaches does not exploit the covariance of the carbon fluxes when upscaling in-situ data to global scale. Tramontana *et al*. [30] developed an approach for partitioning carbon fluxes preserving the relationship among them using neural networks physically constrained, but only applied with in-situ data.

The selection of the RS predictive variables is based on the main drivers of photosynthesis, transpiration and respiration processes. In this regard, MODIS products that are related with water, radiation, temperature, and evapotranspiration were selected. The ranking of the explanatory variables highlighted PAR as the most influential predictor of GPP estimates as expected according to Monteith logic [31]. However, the vegetation index kNDVI has also a great influence, which suggests that it might be connected with the light use efficiency (LUE). In fact, Wu *et al*. [32] proposed a GPP approach incorporating vegetation indices for LUE and the fraction of absorbed photosynthetically active radiation (FAPAR) estimation. In literature, different authors have shown the usefulness of vegetation indices as proxies of both LUE and FAPAR [33]–[34].

The relevance analysis for TER estimates revealed LAI as the most important variable. LST_N, ET, PET, and C_{ws} also revealed as important factors for TER. In the case of NEE, the main explanatory variables were the two most relevant for GPP and TER, i.e., PAR and LAI. Globally, temperature is an influencing factor on the three fluxes: the diurnal temperature slightly affects GPP (i.e., photosynthesis), probably through the LUE in case of a thermal stress (similar to the role of evapotranspiration in case of water stress), and night temperature affects respiration process and, to a lesser extent, the NEE. Respiration provides the energy for a plant to acquire nutrients and to produce and maintain biomass. The total plant respiration can be split into three functional components: growth respiration, maintenance respiration, and the respiratory cost of ion uptake [35]. The maintenance respiration is expected to be greater in productive ecosystems and then to be positively correlated with LAI, with a rather high relevance in respiration estimation as shown in Fig. 3. Maintenance respiration depends also on environmental stresses: for a given ecosystem it increases with temperature and drought. However, plants from hot environments have lower respirations rates at a given temperature than plants from cold places [35]. In any case, temperature and water balance are both rather relevant variables in respiration processes.

B. Global estimates

GPP estimates at site level provided by the $\mathcal{G}$-model are highly correlated with both MOD17A2 and PML_V2 estimates (R^2 =0.63 and R^2=0.71). One factor that explains the high correlation is the use of similar inputs in the approaches. The per biome type comparison found the best agreements over open shrublands in the case of MOD17A2, and over closed shrublands in the case of PML_V2. It is worth mentioning that the worst consistencies between products were obtained over crops in both cases. The spatial distribution of the GPP is consistent with the MOD17A2, and PML_V2 products. The major concentrations of GPP and TER are located in the tropics and decreases to high and low latitudes.

The $\mathcal{G}$-model provided a global GPP (133.7 Pg yr^{-1}) similar to the one reported by other studies. In instance, Zhang *et al*. [37] reported an average global annual GPP around 125 Pg yr^{-1} for the VPM and FluxCom products. Joiner *et al*. [19] revealed an annual GPP value around 140 Pg yr^{-1}. Jung *et al*. [14] found mean global GPP of 111 Pg yr^{-1} for FluxCom-RS, 120 Pg yr^{-1} for FluxCOM-RS+METEO product, and a range of 83–172 Pg yr^{-1} for TRENDY v7 models. Dong *et al*. [37] carried out an inter-comparison of global GPP annual mean provided by PEM models, which reported a range of variation of 125–165 Pg yr^{-1}. In addition, Badgley *et al*. [6] reported a range of 131–163 Pg yr^{-1} using the NIRv as proxy for GPP. Recently, Li *et al*. [38] reported 125.74 Pg yr^{-1} as a mean value for the 1982–2019 period, and Tagesson *et al*., [39] found 121.8 Pg yr^{-1} over the 1982–2015 period. The TER and NEE budgets found in this study (114.8 Pg yr^{-1}, and 18.9 Pg yr^{-1}, respectively) are closer to the ones reported by Zeng *et al*. [40] (TER ranging from115 to 121 Pg yr^{-1}, and NEE ranging from 20 to 22 Pg yr^{-1}). In addition, Li *et al*. [40] found 109.3 Pg yr^{-1} for TER, and 16.28 Pg yr^{-1} for NEE, and Tagesson *et al*. [41] reported 105.6 Pg yr^{-1} for TER in the 1982–2015 period. These values are lower than the found in our study. The range of variation of the values reported in the literature indicate a significant level of uncertainty persists across different approaches.

The 2002–2023 temporal analysis provided a GPP trend (0.28±0.02 Pg yr^{-1} yr^{-1}) similar to other studies. Tagesson *et al*., [39] found a trend of 0.27±0.02 Pg yr^{-1} yr^{-1} in the 1982–2015 period, Guo *et al*. [15] and Zeng *et al*. [40] reported 0,21 Pg yr^{-1} yr^{-1} and 0.49 Pg yr^{-1} yr^{-1} both computed during the 1999–2019 period. The inter-comparison conducted by Dong *et al*. [37] reported a range of global GPP trends of different products, which varies from −0.22 to 0.51 Pg yr^{-1} yr^{-1} over different periods.

C. Model Uncertainty

Although the results show a good level of agreement regarding in-situ data and other products, there exist sources of uncertainty that propagates into the model. The spatial mismatch between the footprint of the EC towers and the RS data spatial resolution is source of uncertainty in CFs retrievals. Uncertainties in the EC covariance data, and in the RS-based products can propagate into carbon fluxes retrieval [42]–[44]. In addition, the different spatial resolutions of the model inputs may also affect the stability of the predictions [45].

The $\mathcal{G}$-model provides not only a joint estimate of GPP, NEE, and TER but also an uncertainty (or confidence interval) for the retrievals. Note that the uncertainty of the predictive mean estimate does not depend on the value of the predictive mean but accounts for the similarity of the new inputs ($\mathbf{x}_*$) regarding the inputs trained in the model ($\mathbf{x}_i$). Therefore, high uncertainties would be obtained in scenarios where the behavior of the inputs are not accounted during model training. The latitudinal profiles of the uncertainties are quite constant and do not follow the latitudinal variation of the GPP, NEE, and TER. This decorrelation between predictions and uncertainties suggests that the uncertainties are not influenced by the value of the prediction.

V. Conclusions

This study developed a data-driven approach using a multi-output regression method that is able to predict GPP, TER, and NEE at the same time. The $\mathcal{G}$-model estimates the three fluxes jointly also conserving their physical relationship, which is one of the drawbacks of the classic machine learning algorithms. The predictors of the model came from five MODIS products, which resulted in 8 explanatory variables. Among them, the PAR and kNDVI were identified as the most relevant inputs for GPP, whereas the LAI and LST_N were the most influential for TER retrievals. PAR and LAI were also the most important for the NEE. The model was well calibrated at tower level, and successfully upscaled to globe. At tower level the best results were obtained for GPP (R^2=0.82), followed by TER (R^2=0.80), and NEE (R^2=0.69). The RMSE was 1.55 g m^{-2} d^{-1} in the case of GPP, 1.09 g m^{-2} d^{-1} for NEE, and 1.14 g m^{-2} d^{-1} for TER. These results, obtained from only MODIS-based inputs, outperformed MOD17A2H and PML_V2 products. At global scale the spatial distribution of the derived maps agreed with MOD17A2H and PML_V2. The results provide a mean annual GPP value of 133.7 Pg yr^{-1} and a global trend of 0.28 Pg yr^{-1} yr^{-1} for the 2002–2023 period that fall within the range reported by other works. However, further studies are needed to better understand the spatial and temporal patterns of global carbon fluxes that help to reduce uncertainties among approaches.

References

[1] M. A. White, P. E. Thornton, S. W. Running, and R. R. Nemani, "Parameterization and sensitivity analysis of the BIOME–BGC terrestrial ecosystem model: Net primary production controls" *Earth Interact.*, vol. 4, no 3, pp.1–85, 2000, doi: 10.1175/1087-3562(2000)004%3C0003:PASAOT%3E2.0.CO;2

[2] S. Sitch, *et al*., "Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs)", *Glob. Change Biol.*, vol. 14, no 9, pp. 2015–2039, 2008, doi: 10.1111/j.1365-2486.2008.01626.x

[3] S. Sitch, *et al*., "Recent trends and drivers of regional sources and sinks of carbon dioxide", *Biogeosciences*, vol. 12, no 3, pp. 653–679, 2015, doi: 10.5194/bg-12-653-2015

[4] S. Running, Q. Mu, and M. Zhao, "MOD17A3H MODIS/Terra Net Primary Production Yearly L4 Global 500m SIN Grid V006," NASA EOSDIS Land Processes DAAC, 2015, doi: 10.5067/MODIS/MOD17A3H.006.

[5] B. Martínez, *et al*., "Evaluation of the LSA-SAF gross primary production product derived from SEVIRI/MSG data (MGPP)", *ISPRS J. Photogramm.*, vol. 159, pp. 220–236, 2020, doi: 10.1016/j.isprsjprs.2019.11.010

[6] G. Badgley, L. D. L. Anderegg, J. A. Berry, and C. B. Field, "Terrestrial gross primary production: Using NIRV to scale from site to globe," *Global Change Biol.*, vol. 25, no. 11, pp. 3731–3740, 2019, doi: 10.1111/gcb.14729

[7] R. Wang, *et al*., "Detecting intra-and inter-annual variability in gross primary productivity of a North American grassland using MODIS MAIAC data", *Agr. Forest Meteorol*, vol. 281, Art. no. 107859, 2020, doi: 10.1016/j.agrformet.2019.107859

[8] Y. Sun, *et al*., "OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence", *Science*, vol. 358; Art. no. eaam5747, 2017, doi: 10.1126/science.aam5747

[9] D. Hao, *et al*., "Adjusting solar-induced fluorescence to nadir-viewing provides a better proxy for GPP", *ISPRS J. Photogramm.*, vol. 186, pp. 157–169, 2022, doi: 10.1016/j.isprsjprs.2022.01.016

[10] J. Xiao, *et al*., "Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data", *Agr. Forest Meteorol*., vol. 148, pp. 1827–184, 2008, doi: 10.1016/j.agrformet.2008.06.015

[11] G. Tramontana, *et al*., "Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms", *Biogeosciences*, vol. 13, pp. 4291–4313, 2016, doi: 10.5194/bg-13-4291-2016

[12] K. Ichii, *et al*., "New data-driven estimation of terrestrial CO2 fluxes in Asia using a standardized database of eddy covariance measurements, remote sensing data, and support vector regression", *J. Geophys. Res-Biogeo*., vol. 122, pp. 767–795, 2017, doi: 10.1002/2016JG003640

[13] Y. Zhang, *et al*., "Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017", *Remote Sens. Environ.*, vol. 222, pp. 165–182, 2019, doi: 10.1016/j.rse.2018.12.031

[14] M. Jung, *et al*., "Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach", *Biogeosciences*, vol. 17, pp. 1343–1365, 2020, doi: 10.5194/bg-17-1343-2020

[15] R. Guo, *et al*., "Estimating Global GPP from the Plant Functional Type Perspective Using a Machine Learning Approach", *J. Geophys. Res-Biogeo*., vol. 128, no 4, Art. no. e2022JG007100, 2023, doi: 10.1029/2022JG007100

[16] D. Tuia, J. Verrelst, L. Alonso, F. Perez-Cruz, and G. Camps-Valls, "Multioutput support vector regression for remote sensing biophysical parameter estimation," *IEEE Geosci. Remote Sens. Lett*., vol. 8, no. 4, pp. 804–808, Jul. 2011, doi: 10.1109/LGRS.2011.2109934

[17] F. J. García-Haro *et al*., "Derivation of global vegetation biophysical parameters from EUMETSAT polar system," *ISPRS J. Photogramm*. Remote Sens., vol. 139, pp. 57–74, May 2018, doi: 10.1016/j.isprsjprs.2018.03.005.

[18] M. Reichstein, *et al*., "On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm", *Glob. Change Biol*., vol. 11, pp. 1424–1439, 2005, doi: 10.1111/j.1365-2486.2005.001002.x

[19] J. Joiner, Y. Yoshida, Y. Zhang, G. Duveiller, M. Jung, A. Lyapustin, Y. Wang, and C. Tucker, ''Estimation of terrestrial global gross primary production (GPP) with satellite data-driven models and eddy covariance flux data,'' *Remote Sens*., vol. 10, no. 9, Art. no. 1346, Aug. 2018, doi: 10.3390/rs10091346

[20] S. Bao *et al*., "Environment-sensitivity functions for gross primary productivity in light use efficiency models," *Agricult. Forest Meteorol*., vol. 312, Jan. 2022, Art. no. 108708, doi: 10.1016/j.agrformet.2021.108708

[21] G. Camps-Valls *et al*., "A unified vegetation index for quantifying the terrestrial biosphere," Sci. Adv., vol. 7, no. 9, Art. no. eabc7447, 2021, doi: 10.1126/sciadv.abc7447

[22] S. Liang, and D. Wang, "Moderate Resolution Imaging Spectroradiometer (MODIS) Downward Shortwave Radiation (MCD18A1) and Photosynthetically Active Radiation (MCD18A2) Algorithm Theoretical Basis Document", 2017, doi: 10.5067/MODIS/MCD18C2.061

[23] A. H. Strahler, *et al*., "MODIS BRDF/albedo product: algorithm theoretical basis document version 5.0", pp.1–53, 1999, doi: 10.5067/MODIS/MCD43A4.061 (1999).

[24] G. C. Hulley, S. J. Hook, and C. Hughes, "MODIS MOD21 land surface temperature and emissivity algorithm theoretical basis document. Pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration. 2012. [Online]. Available: https://emissivity.jpl.nasa.gov/downloads/examples/documents/MOD21_LSTE_ATBD_Hulley_v2.0_20121116_pmbv1.pdf

[25] Q. Mu, M. Zhao and S. W. Running, S. W. "MODIS Global Terrestrial Evapotranspiration (ET) Product (NASA MOD16A2/A3)". Algorithm Theoretical Basis Document. Collection 5. Numerical Terradynamic Simulation Group. College of Forestry and Conservation. The University of Montana, 2013, doi: 10.5067/MODIS/MOD16A2.006

[26] Y. Knyazikhin *et al*., "MODIS leaf area index (LAI) and fraction of photosynthetically active radiation absorbed by vegetation (FPAR) product (MOD15) algorithm theoretical basis document (Version 4.0)," NASA Goddard Space Flight Center (GSFC), Greenbelt, MD, USA, Tech. Rep., 1999. [Online]. Available: https://modis.gsfc.nasa.gov/data/atbd/atbd_mod15.pdf

[27] C. E. Rasmussen and K. I. Christopher, Gaussian Processes for Machine Learning. Massachusetts, MA, USA: MIT Press, 2005, pp. 7–31

[28] M. Campos-Taberner *et al*., "Understanding deep learning in land use classification based on sentinel-2 time series," *Sci. Rep*., vol. 10, no. 1, pp. 1–12, 2020, doi: 10.1038/s41598-020-74215-5

[29] Z. Man, *et al*. "Effect of climate change on CO2 flux in temperate grassland, subtropical artificial coniferous forest and tropical rain forest ecosystems", *Int. J. Environ. Res. Public Health*, vol. 18, pp. 13056, 2021, doi: 10.3390/ijerph182413056

[30] G. Tramontana *et al*., "Partitioning net carbon dioxide fluxes into photosynthesis and respiration using neural networks," *Global Change Biol*., vol. 26, no. 9, pp. 5235–5253, 2020, doi: 10.1111/gcb.15203

[31] J. L. Monteith, "Solar radiation and productivity in tropical ecosystems", *J. Appl. Ecol*., vol. 9, pp. 747-766, 1972. [Online]. Available: https://www.jstor.org/stable/2401901 (1972).

[32] C. Wu, Z. Niu, and S. Gao, "Gross primary production estimation from Modis data with vegetation index and photosynthetically active radiation in maize," *J. Geophys. Res. Atmospheres*, vol. 115, pp. 13–20, 2010, doi: 10.1029/2009JD013023

[33] Y. Inoue, J. Penuelas, A. Miyata, and M. Mano, "Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO2 flux measurements in rice," *Remote Sens. Environ*., vol. 112, no. 1, pp. 156–172, Jan. 2008, doi: 10.1016/j.rse.2007.04.011

[34] C. Y. Wu, J. W. Munger, Z. Niu, and D. Kuang, "Comparison of multiple models for estimating gross primary production using MODIS and eddy covariance data in Harvard forest," *Remote Sens. Environ*., vol. 114, pp. 2925–2939, Dec. 15, 2010, doi: 10.1016/j.rse.2010.07.012

[35] F. S. Chapin III, P. A. Matson, H. A. Mooney, and P. M. Vitousek, "Principles of terrestrial ecosystem ecology", Springer-Verlag, New York, 2002, 436 pp. doi: 10.1007/978-1-4419-9504-9

[36] Y. Zhang *et al*., "A global moderate resolution dataset of gross primary 700 production of vegetation for 2000–2016," *Sci. Data*, vol. 4, no. 1, 701 Dec. 2017, Art. no. 170165, doi: 10.1038/sdata.2017.165

[37] J. Dong, L. Li, Y. Li, and Q. Yu, "Inter-comparisons of mean, trend and interannual variability of global terrestrial gross primary production retrieved from remote sensing approach", *Sci. Total Environ*., vol. 822, Art. no. 153343, 2022, doi: 10.1016/j.scitotenv.2022.153343

[38] B. Li *et al*., "BESSv2. 0: A satellite-based and coupled-process model for quantifying long-term global land–atmosphere fluxes", *Remote Sens. of Environ.*, vol. 295, Art. no. 113696; 2023, doi: 10.1016/j.rse.2023.113696

[39] T. Tagesson *et al*. "A physiology-based Earth observation model indicates stagnation in the global gross primary production during recent decades", *Glob. Change Biol*., vol. 27, pp. 836–854, 2021, doi: 10.1111/gcb.15424

[40] J. Zeng *et al*., "Global terrestrial carbon fluxes of 1999-2019 estimated by upscaling eddy covariance data with a random forest," *Sci. Data*, vol. 7, no. 1, Sep. 2020, Art. no. 313, doi: 10.1038/s41597-020-00653-5

[41] T. Tagesson *et al*., "Increasing global ecosystem respiration between 1982 and 2015 from Earth observation-based modelling", *Glob. Ecol. Biogeogr*., vol. 33, pp. 116–130, 2012, doi: 10.1111/geb.13775

[42] F. A. Heinsch *et al*., "Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations", *IEEE Trans. Geosci*., vol. 44, pp. 1908–1925, 2006, doi: 10.1109/TGRS.2005.853936

[43] J. Kala *et al*., "Influence of leaf area index prescriptions on simulations of heat, moisture, and carbon fluxes", *J. Hydrometeorol*., vol. 15, pp. 489–503, 2014, doi: 10.1175/JHM-D-13-063.1

[44] Y. Liu *et al*., "Satellite-derived LAI products exhibit large discrepancies and can lead to substantial uncertainty in simulated carbon and water fluxes," *Remote Sens. Environ*., vol. 206, pp. 174–188, Mar. 2018, doi: 10.1016/j.rse.2017.12.024.

[45] G. Gelybó, Z. Barcza, A. Kern, and N. Kljun, ''Effect of spatial heterogeneity on the validation of remote sensing based GPP estimations,'' Agricult. Forest Meteorol., vols. 174–175, pp. 43–53, Jun. 2013, doi: 10.1016/j.agrformet.2013.02.003

SUPPLEMENTARY MATERIAL

Global carbon fluxes using multi-output Gaussian processes regression and MODIS products

Manuel Campos-Taberner, María Amparo Gilabert, Sergio Sánchez-Ruiz, Beatriz Martínez, Adrián Jiménez-Guisado, Francisco Javier García-Haro, and Luis Guanter

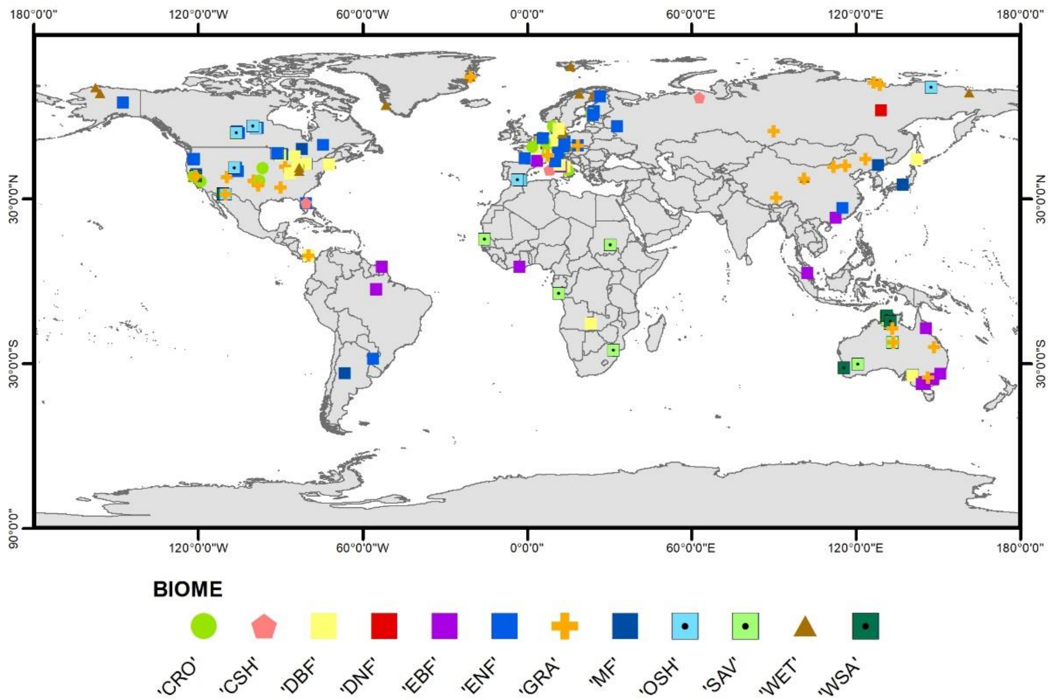


Fig. SI. Location of the FLUXNET2015 sites. The legend shows the biome of every flux site, namely: croplands (CRO), closed shrublands (CSH), deciduous broadleaf forest (DBF), deciduous needleleaf forests (DNF), evergreen broadleaf forests (EBF), evergreen needleleaf forests (ENF), grasslands (GRA), mixed forests (MF), open shrublands (OSH), savannas (SA), wetlands (WET), and woody savannas (WSA). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

TABLE SI

NUMBER OF FLUX SITES PER BIOME

Biome	CRO	CSH	DBF	DNF	EBF	ENF	GRA	MF	OSH	SAV	WET	WSA
Sites	20	3	26	1	15	49	39	9	13	9	21	6

Codes: croplands (CRO), closed shrublands (CSH), deciduous broadleaf forest (DBF), deciduous needleleaf forests (DNF), evergreen broadleaf forests (EBF), evergreen needleleaf forests (ENF), grasslands (GRA), mixed forests (MF), open shrublands (OSH), savannas (SA), wetlands (WET), and woody savannas (WSA).

Global carbon fluxes using multi-output Gaussian processes regression and MODIS products

Manuel Campos-Taberner, María Amparo Gilabert, Sergio Sánchez-Ruiz, Beatriz Martínez, Adrián Jiménez-Guisado, Francisco Javier García-Haro, and Luis Guanter

Abstract—**The quantification of carbon fluxes is crucial due to their role in the global carbon cycle having a direct impact on Earth's climate. In the last years, considerable efforts have been done to scale carbon fluxes from eddy covariance (EC) data to the globe. In this work, a data-driven approach that exploits a multi-output Gaussian processes regression algorithm ($\mathcal{G}$-model) is proposed to jointly estimate gross primary production (GPP), terrestrial ecosystem respiration (TER), and net ecosystem exchange (NEE) at global scale. The $\mathcal{G}$-model not only provides an estimate of the carbon fluxes but also an uncertainty. Moreover, it derives the three fluxes jointly preserving their physical relationship. The predictors are selected from a set of the MODerate-resolution Imaging Spectroradiometer (MODIS) products available on Google Earth Engine. The performance of the model revealed high accuracies (R^2 reaching 0.82, 0.69, and 0.80 in the case of GPP, NEE, and TER, respectively), and low root mean square errors (1.55 g m^{-2} d^{-1} in the case of GPP, 1.09 g m^{-2} d^{-1} for the NEE, and 1.14 g m^{-2} d^{-1} for TER) over the FLUXNET2015 data set at 8-day time scale. The GPP estimates provided by $\mathcal{G}$-model outperformed the MOD17A2 product, and a state-of-the-art GPP product (PML_V2) without using meteorological forcing data sets. The results reported mean annual amounts of 133.7 Pg yr^{-1}, 114.8 Pg yr^{-1}, and 18.9 Pg yr^{-1} for GPP, TER, and NEE, respectively during the 2002–2023 period. The proposed approach paves the way for the development of multi-output strategies that preserve the physical relationships among carbon fluxes in upscaling processes.**

This study was supported by Grant PID2020-118036RB-I00 funded by MCIN/AEI/ 10.13039/501100011033, and by LSA-SAF (EUMETSAT).
Corresponding author: Manuel Campos-Taberner.

Manuel Campos-Taberner is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: manuel.campos@uv.es).

María Amparo Gilabert is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: m.amparo.gilabert@uv.es).

Sergio Sánchez-Ruiz is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: sergio.sanchez@uv.es).

Beatriz Martínez is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: beatriz.martinez@uv.es).

Adrián Jiménez-Guisado is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: adrian.jimenez@uv.es).

Francisco Javier García-Haro is with the Environmental Remote Sensing group (UV-ERS), Universitat de València, Dr. Moliner, 46100 Burjassot, València, Spain. (e-mail: j.garcia.haro@uv.es).

Luis Guanter is with the Land and Atmosphere Remote Sensing group (LARS), Universitat Politècnica de València, Valencia, Spain. (e-mail: lguanter@fis.upv.es).

Supplementary material and color versions of the figures in this article are available online at http://ieeexplore.ieee.org

Index Terms—**Gross primary production (GPP), Terrestrial ecosystem respiration (TER), Net ecosystem exchange (NEE), machine learning; data-driven; Multi-output Gaussian process regression**

I. INTRODUCTION

CARBON fluxes (CFs) are essential for understanding the Earth's carbon cycle and its influence on the environment. The estimation of CFs is a measure of how much carbon is being exchanged between different components of Earth's climate system such as soils, vegetation, oceans and atmosphere. These exchanges occur via natural processes such as photosynthesis, respiration, and decomposition. The three main carbon fluxes are the gross primary production (GPP), terrestrial ecosystem respiration (TER), and net ecosystem exchange (NEE). GPP stands for the total amount of carbon that is fixed by plants through photosynthesis, whereas TER refers to the total amount of carbon that is released into the atmosphere by respiration of plants and other organisms present in the ecosystem. NEE is the difference between carbon uptake and carbon release by an ecosystem.

In-situ carbon fluxes are obtained from different networks of micrometeorological towers that use eddy covariance (EC) technique to estimate GPP from NEE and TER. Towers operate in many countries, across America, Europe and Asia, through the networks AmeriFlux (https://ameriflux.lbl.gov), ICOS (Integrated Carbon Observation System) (https://www.icos-cp.eu/), AsiaFlux (http://www.asiaflux.net), or the FLUXNET global network (https://fluxnet.org/). These measurements are representative of the corresponding EC tower footprint, which may vary from hectometric to kilometric scales depending on the site characteristics. At regional and global scales CFs can be estimated using models that simulate the carbon cycle adding the influence of several factors such as vegetation types, soil properties and climate conditions [1]–[3]. In addition, CFs can also be estimated from polar and geostationary remote sensing (RS) data such as the the MODerate resolution Imaging Spectroradiometer (MODIS) [4] or the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) on board the Meteosat Second Generation (MSG) [5], which are based on Production Efficiency Models (PEMs). A simple yet effective method to estimate GPP from RS data is to employ vegetation indices or spectral bands combinations as proxies of GPP [6]–[7]. Similarly, Solar-induced fluorescence (SIF) can be used for inferring GPP [8]–[9]. Other approaches link RS and meteorological data with EC measures [10]–[15]. These data-driven techniques rely on the relationship among vegetation properties and carbon fluxes to build a single model (one per flux to be retrieved) with machine learning (ML) methods. The use of a single ML model for every predictive variable can lead to good independent performances but does not guarantee that the relationship among variables is respected, leading to non-robust approaches. Some studies have shown that the use of a unique multi-output model could improve the results when estimating biophysical parameters at the same time [16]–[17]. However, this technique has not been tested on upscaling global carbon fluxes exchange yet, where the balance between carbon uptake through photosynthesis and carbon release through respiration leads to net ecosystem exchange GPP-TER=NEE.

In this framework, the objectives of this paper are (i) to develop a data-driven approach to estimate CFs at global scale blending MODIS products and EC data, (ii) to build a single multi-output model able to jointly estimate GPP, NEE, and TER in a robust way, and (iii) to validate the estimates against global in-situ data, and compare with two reference products (MOD17A2H and PML_V2). The approach relies on the use of multitemporal MODIS products from Google Earth engine (GEE) as inputs in a multi-output Gaussian Processes Regression algorithm ($\mathcal{G}$-model), which exploits the covariance among CFs. The remainder of the manuscript is structured as follows. Section II describes in-situ and remote sensing products, and the proposed multi-output algorithm. Section III shows the model performance, the explanatory power of the inputs, the derived global carbon fluxes maps, and a comparison with other products. Section IV discusses the results, and section V highlights the main conclusions.

II. Materials and methods

A. Tower data

The approach proposed in this study exploits data coming from the FLUXNET network. The FLUXNET2015 dataset was downloaded (https://fluxnet.org/data/fluxnet2015-dataset/), which allowed the use of data coming from 211 sites globally distributed (see Supplementary Fig. SI). In particular, we used the dataset update dating from February 6, 2020. The distribution of the flux sites according to the International Geosphere-Biosphere Programme (IGBP) can be found as Supplementary Table SI. Daily GPP, NEE, and TER data were computed for every site from the average of the night- and day-time partitioning methods present in the FLUXNET2015 dataset [18]. The sign convention used in this study was such that CO_2 flux to the surface was positive so that NEE=GPP-TER, therefore positive values of NEE stand for vegetation carbon uptake, and negatives for release. The NEE quality flag was used to ensure the consistency between GPP, TER and NEE, and similarly to other studies we excluded daily data when more than 20% of the half-hourly data were based on gap-filling with low confidence [11],[19]. Finally, daily data were temporally aggregated (8-day) to match the RS products.

B. Remote sensing products

The biosphere-atmosphere interactions related to photosynthesis are mainly driven by Photosynthetically Active Radiation (PAR), as well as environmental and physiological variables, such as water availability, temperature, chlorophyll content, and leaf area index (LAI) [20]. On this basis, five MODIS-based RS products were used to extract the predictors for the regression method. Overall, 8 predictors were selected, namely: PAR from the MCD18C2 product; day- and night-time land surface temperature (LST_D, LST_N) from the MOD11A2 product; Evapotranspiration (ET) and Potential Evapotranspiration (PET) from the MOD16A2 product, as well as the water stress factor (C_{ws}) computed as $C_{ws} = \frac{ET}{PET}$; the kernel version of the normalized difference vegetation index (kNDVI) [21] computed from the MCD43A4 red and NIR bands; and LAI from the MCD15A2H product. The GEE platform was used to download the selected MODIS products over the flux sites from 2002 (when starts the availability of all considered MODIS products) to 2014 (when ends the availability of the FLUXNET2015 data), thus fitting the coincident temporal range between the FLUXNET2015 dataset and the RS observations. The final value of the predictors that was assigned to every tower was the mean of the pixel values contained in a 500-m buffer weighted by their spatial extension. The quality flag of every predictor was also used to discard bad-quality pixels. In the case of the MCD18C2 product, the values were integrated to obtain a daily-integrated PAR, and aggregated to 8-day for temporal consistency among the rest of the predictors and EC data. Eventually, two reference GPP products (MOD17A2H, and PML_V2) were downloaded from GEE with the same aforementioned processing to compare the GPP retrievals. Table I highlights the main features of the products.

TABLE I
MAIN FEATURES OF THE SELECTED MODIS-BASED PRODUCTS

	GEE name	Spatial resolution	Temporal frequency	Reference
Predictors	MCD18C2.061	0.05°	3 hours	[22]
	MCD43A4.061	500 m	daily	[23]
	MOD11A2.061	1 km	8 days	[24]
	MOD16A2.061	500 m	8 days	[25]
	MCD15A2H.061	500 m	8 days	[26]
For comparison	MOD17A2H.061	500 m	8 days	[4]
	PML_V2 0.1.7	500 m	8 days	[13]

C. Multi-output Gaussian processes regression ($\mathcal{G}$-model)

Gaussian processes (GPs) [27] are statistical techniques that offer a probabilistic approach for kernel regression methods in which the likelihood of unobserved latent functions is supposed to be modeled by a multivariate Gaussian prior. The approach makes use of a latent function of the inputs $f(\mathbf{x})$ (the RS predictors) plus constant Gaussian noise to estimate the outputs (the carbon flux of interest) as $y = f(\mathbf{x}) + \varepsilon_n$, $\varepsilon_n \sim N(0,\sigma_n^2)$. The latent function is chosen to be a zero mean Gaussian process prior, or $(x) \sim GP(0,k_\theta)$, where k_θ is a covariance function parametrized by θ. The noise power is given by the hyperparameter σ_n^2, and it is assumed that the noise ε_n follows a prior Gaussian distribution. Given the GP priors, samples drawn from f(x) at $\mathbf{x}_i = \{x_i^1,\ldots,x_i^B\}_{i=1}^N$ (*N* is the number of training samples, and *B* the number of predictors) follow a joint multivariate zero mean Gaussian distribution defined by the covariance matrix $\mathbf{K}$, where $[\mathbf{K}]_{ij} = k_\theta(x_i,x_j)$.

The output of the model is computed from the predictive distribution described by the equations:

$$p(y_* | \mathbf{x}_*, D) = N(y_* | \mu_{GPR*}, \sigma_{GPR*}^2), \quad (1)$$

$$\mu_{GPR*} = \mathbf{k}_*^T (\mathbf{K} + \sigma_n^2 \mathbf{I}_n)^{-1} \mathbf{y} = \mathbf{k}_*^T \alpha, \quad (2)$$

$$\sigma_{GPR*}^2 = \sigma^2 + k_{**} - \mathbf{k}_*^T (\mathbf{K} + \sigma_n^2 \mathbf{I}_n)^{-1} \mathbf{k}_*, \quad (3)$$

where *D* stands for the training dataset (being $\mathbf{x}_*$ and y_* the input and output), $\mathbf{k}_* = [k(x_*,x_1),\ldots,k(x_*,x_N)]$ is an $N\times1$ vector and $k_{**} = k(x_*,x_*)$. The model provides a full posterior probability from which predictions (μ_{GPR*}) and confidence estimates (σ_{GPR*}^2) can be computed. In the single output approach $\mathbf{y} \in \mathbb{R}^{N\times1}$, the output of interest (e.g., GPP) is computed from the inputs (8 RS predictors, $\mathbf{x}_i \in \mathbb{R}^{N\times8}$) as:

$$\hat{y} = f(\mathbf{x}) = \sum_{i=1}^N \alpha_i k_\theta(\mathbf{x}_i,\mathbf{x}) + \alpha_0, \quad (4)$$

where α_i is the weight assigned to each predictor $\mathbf{x}_i$, α_0 is the bias term, and k_θ is the covariance function evaluating the similarity between $\mathbf{x}$ and all the *N* training spectra. For this task the automatic relevance determination kernel was selected

$$K(\mathbf{x}_i,\mathbf{x}_j) = v\exp\left(-\sum_{b=1}^B (\mathbf{x}_i^b - \mathbf{x}_j^b)^2 / 2\sigma_b^2\right) + \sigma_n^2 \delta_{ij}, \quad (5)$$

where *v* is a scaling factor, σ_n accounts for the standard deviation of the noise, *B* is the number of predictors (in our case, B=8), and σ_b can be related to the spread of each predictor *b*. The model hyperparameters $\boldsymbol{\theta} = [v,\sigma_n, \sigma_1, \ldots, \sigma_b]$ can be estimated by maximizing the marginal log-likelihood [27]:

$$\log p(\mathbf{y}|\mathbf{x}_i,\theta) = -\frac{1}{2}\mathbf{y}^T (\mathbf{K} + \sigma_n^2 \mathbf{I}_n)^{-1} \mathbf{y} - \frac{1}{2}\log|\mathbf{K} + \sigma_n^2 \mathbf{I}_n| - \frac{N}{2}\log(2\pi). \quad (6)$$

The model can be adapted to deal with multi-output regression problems ($\mathcal{G}$-model) adapting the kernel hyperparameters for a unique kernel which is able to deal with all the outputs. In the $\mathcal{G}$-model we need to optimize the hyperparameters considering that $\mathbf{y} \in \mathbb{R}^{N\times D}$ instead $\mathbf{y} \in \mathbb{R}^{N\times1}$ where *D* is the total number of outputs (in our case, *D*=3). Therefore, a global cost function summarizing all the cost functions (one per output) into a single one must be defined. Here we use a global cost function *C* just as the squared sum of the standard model cost function of each output [17]:

$$C = \sum_{d=1}^D \log p(\mathbf{y}_d|\mathbf{x})^2 = \sum_{d=1}^D \left(-\frac{1}{2}\mathbf{y}_d^T \boldsymbol{\alpha} - \sum_{j=1}^N \log L_{jj} - \frac{N}{2}\log(2\pi)\right)^2, \quad (7)$$

where L_{jj} is the Cholesky factorization of the covariance matrix for every output.

D. Algorithm calibration and performance

The $\mathcal{G}$-model was evaluated with 14562 samples obtained matching the EC data and the RS products. The inputs underwent min-max scaling to transform them into the 0–1 range, as the predictors exhibited a substantial range of variation. Employing this technique is highly advisable to mitigate the risk of suboptimal performance by machine learning algorithms. The 70% of the samples were randomly selected for training and we assessed the results in the remaining 30% (test set unused during model training). Different random initializations were conducted and no significant changes in performance were obtained (not shown for brevity). Root mean square error (RMSE), mean error (ME), mean absolute error (MAE), and the coefficient of determination (R^2) were computed over the test set to evaluate the accuracy, bias, and the goodness-of-fit of the model.

In addition, the model behavior was tested by identifying the most relevant predictors for every estimated flux. This was achieved by means of a sensitivity analysis carried out through an added-noise permutation approach [28]. The inputs were perturbed with Gaussian white noise $N(0,\sigma^2)$, being σ^2 the 3% of the perturbed signal amplitude. The noise was added to a single predictor remaining the rest of the predictors unperturbed, and repeating this process for every predictor. The predictor relevance for every flux was computed as the difference between the accuracy obtained with no perturbation, and the obtained when the perturbation was applied. Finally, the results were normalized with respect to the most relevant predictor.

III. Results

A. Assessment of the model performance over tower data

Fig. 1 shows the direct validation obtained over the test set. The accuracy metrics revealed good correspondence between estimations and in-situ measurements. It is worth mentioning the high R^2 values, which reach 0.82, 0.69, and 0.80 in the case of GPP, NEE, and TER respectively. RMSE was 1.55 g m^{-2} d^{-1} for GPP, 1.09 g m^{-2} d^{-1} for NEE, and 1.14 g m^{-2} d^{-1} for TER. The NEE and TER predictions revealed low and very similar MAE and ME (see Fig. 1) whereas GPP exhibit slightly higher MAE (1.04 g m^{-2} d^{-1}) and lower ME (0.01 g m^{-2} d^{-1}). In addition, Fig. 2(d-e) shows the performance of MOD17A2 and PML_V2 products over the same test set. Note that these products do not provide NEE nor TER. The RMSEs and MAEs of MOD17A2 and PML_V2 are 2.44 g m^{-2} d^{-1} and 2.11 g m^{-2} d^{-1}, and 1.66 g m^{-2} d^{-1} and 1.48 g m^{-2} d^{-1}, respectively. The reported biases regarding EC data are -0.50 g m^{-2} d^{-1} and -0.31 g m^{-2} d^{-1} in the case of MOD17A2 and PML_V2, respectively.

TABLE II

ACCURACY ASSESSMENT PER BIOME IN THE EC VALIDATION SET

	GPP				NEE				TER			
	ME	RMSE	MAE	R^2	ME	RMSE	MAE	R^2	ME	RMSE	MAE	R^2
CRO	0.62	2.56	1.93	0.57	0.37	1.26	0.91	0.77	0.32	1.26	0.96	0.58
CSH	0.61	1.18	0.92	0.69	0.14	0.64	0.53	0.31	0.14	0.62	0.51	0.73
DBF	0.23	1.77	1.26	0.81	0.29	1.23	0.94	**0.80**	0.20	1.13	0.89	0.60
DNF(*)	0.61	0.85	0.64	0.78	1.19	1.37	1.19	0.74	1.25	1.39	1.25	0.42
EBF	0.15	1.33	0.93	0.88	**0.01**	1.02	0.75	0.41	0.06	1.17	0.84	**0.88**
ENF	0.28	1.41	1.02	0.79	0.00	1.27	0.84	0.51	0.07	1.40	0.88	0.73
GRA	0.17	1.31	0.87	0.88	0.28	1.15	0.75	0.46	0.28	1.22	0.79	0.87
MF	0.25	1.52	1.18	0.74	0.02	0.94	0.75	0.75	0.17	1.03	0.84	0.52
OSH	0.07	0.73	0.53	0.74	0.07	**0.48**	**0.34**	0.23	0.16	0.52	0.37	0.72
SAV	0.15	0.94	0.63	0.83	0.04	0.57	0.43	0.76	**0.05**	0.57	0.44	0.87
WET	0.16	2.11	1.47	0.70	0.20	1.14	0.87	0.79	0.14	1.10	0.80	0.77
WSA	**0.01**	**0.58**	**0.45**	**0.91**	0.09	0.51	0.36	0.71	**0.05**	**0.45**	**0.33**	0.86

Best results highlighted in bold. GPP, NEE; TER, ME, RMSE and MAE in g m^{-2} d^{-1}. Codes: croplands (CRO), closed shrublands (CSH), deciduous broadleaf forest (DBF), deciduous needleleaf forests (DNF), evergreen broadleaf forests (EBF), evergreen needleleaf forests (ENF), grasslands (GRA), mixed forests (MF), open shrublands (OSH), savannas (SA), wetlands (WET), and woody savannas (WSA). (*) Only one site.

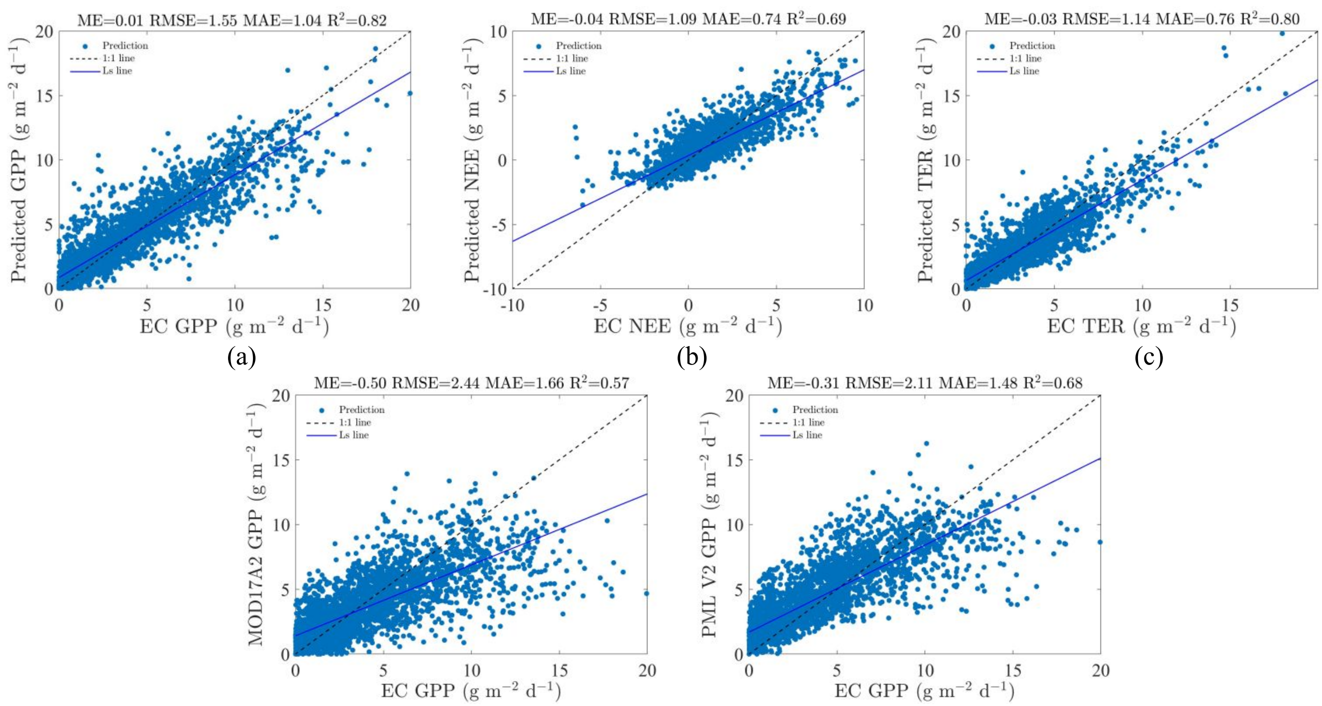
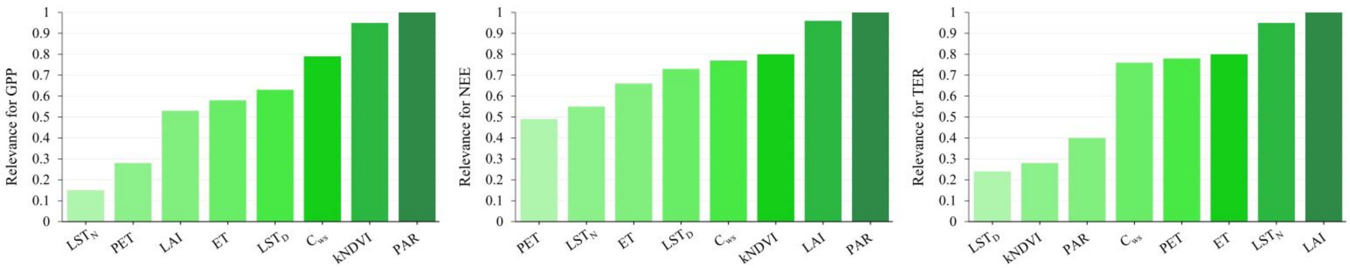


Fig. 1. Carbon fluxes, (a) GPP, (b) NEE, and (c) TER obtained with the $\mathcal{G}$-model over the EC validation set. ME, RMSE and MAE units are g m^{-2} d^{-1}. GPP estimates of (d) MOD15A2 and (e) PML_V2 GPP obtained over the EC validation set. ME, RMSE and MAE units are g m^{-2} d^{-1}.

The $\mathcal{G}$-model accuracy metrics were disaggregated per biome type as shown in Table II. Note that the results over deciduous needleleaf forests are obtained only over 1 available site and may not be representative of this biome. The best results in all metrics for GPP were obtained over woody savannas. In the case of NEE the best accuracy metrics were observed over evergreen broadleaf forests and deciduous needleleaf forests (for ME and R^2), and open shrublands (for RMSE and MAE). The lowest errors for TER were reported over woody savannas, whereas the highest correlation was found over evergreen broadleaf forests.

B. Sensitivity analysis

Fig. 2 shows the relevance of every predictor for GPP, NEE and TER retrievals. The relevance of every predictor was normalized with the value of the maximum relevance predictor in every case. The $\mathcal{G}$-model sensitivity analysis revealed the daily integrated PAR from the MCD18C2 product as the most influential input for both GPP and NEE estimates. LAI from the MOD15A2 product proved as the most relevant predictor for TER retrievals.



(a) (b) (c)

Fig. 2. $\mathcal{G}$-model predictor importance for (a) GPP, (b) NEE, and (c) TER.

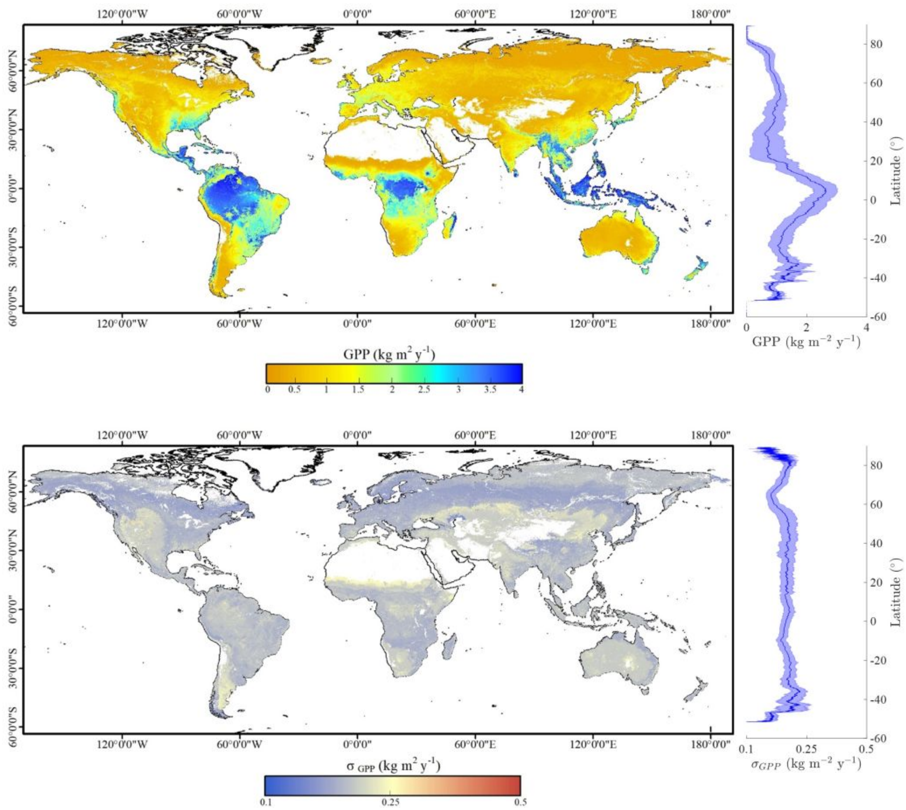


Fig. 3. Global average maps for annual GPP (top), and model uncertainty σ_{GPP} (bottom) computed for the 2002–2023 period.

C. Mean annual fluxes

After the completion of the $\mathcal{G}$-model training phase, the model was effectively applied to derive global GPP, NEE and TER maps. Figs. 3–5 show the mean annual predictions and uncertainties of every flux computed for the 2002–2023 period along with their latitudinal profiles. The spatial pattern of the GPP (Fig. 3, top) and TER (Fig. 4, top) shows that tropics, encompassing regions like Southeast Asia, central Africa, and South America, exhibit a notable high mean annual GPP. Conversely, high latitudes, western North America, and both central Asia and Australia, are characterized by considerably lower mean annual GPP and TER. NEE map (Fig. 5, top) presents a similar spatial pattern. It is worth mentioning that the spatial distribution of the mean annual uncertainty of every flux is slightly different. The σ_{GPP} (Fig. 3, bottom) exhibits a relatively consistent latitudinal profile, hovering around a value of 0.2 kg m^{-2} yr^{-1}. However, there is a zone between 62–65º N where the values are lower (see blueish colors in the north of the σ_{GPP} map). In the case of σ_{TER} and σ_{NEE} their spatial patterns are similar (Figs. 4–5, bottom). Their latitudinal uncertainty values are broadly constant around 0.15–0.2 kg m^{-2} yr^{-1} with no major variations, however σ_{NEE} presents lower spatial variability (Fig. 5, bottom).

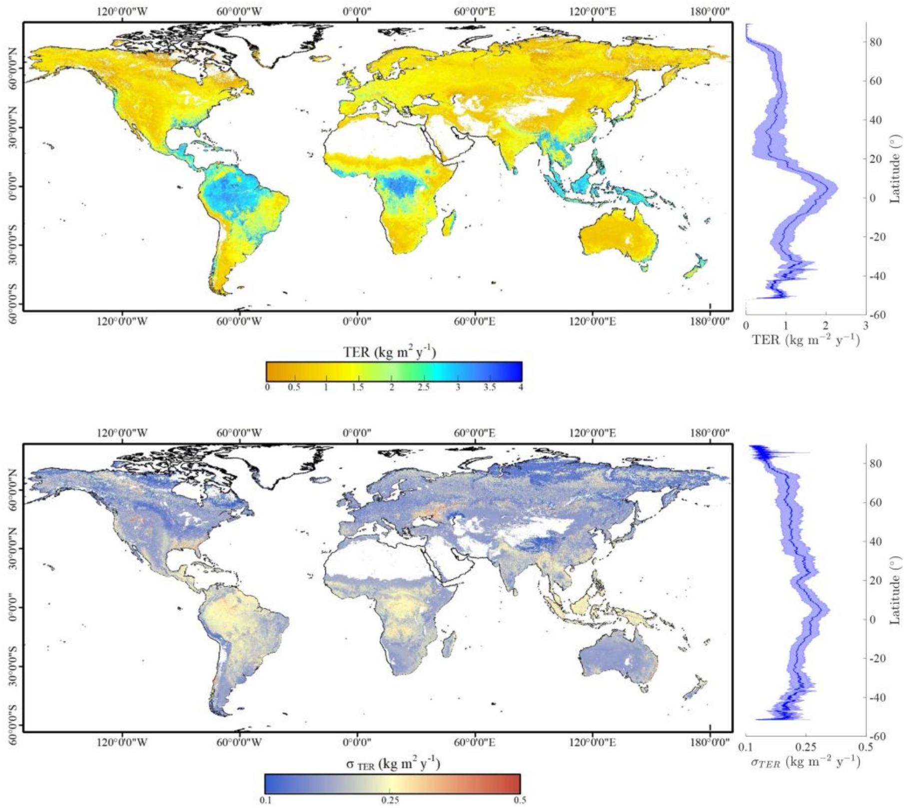


Fig. 4. Global average maps for annual TER (top), and model uncertainty σ_{TER} (bottom) computed for the 2002–2023 period.

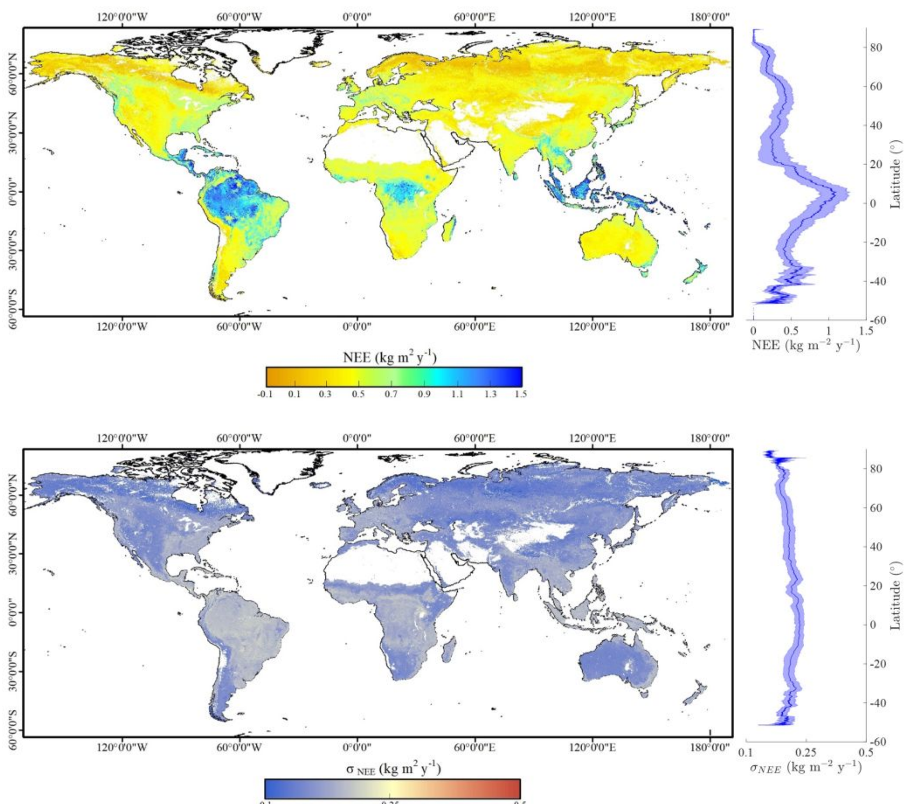
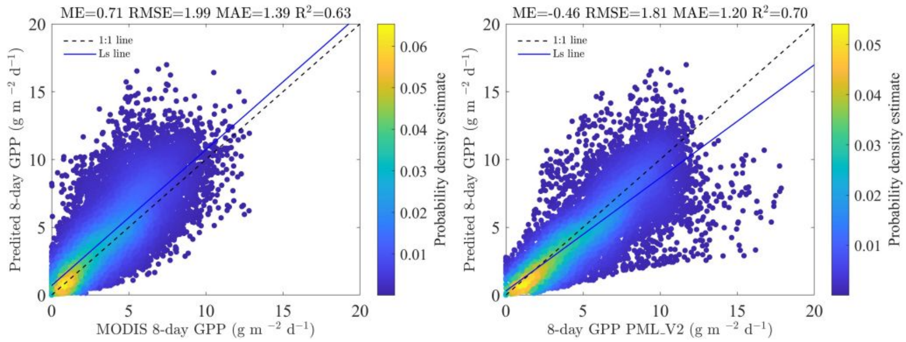


Fig. 5. Global average maps for annual NEE (top), and model uncertainty σ_{NEE} (bottom) computed for the 2002–2023 period.



(a) (b)

Fig. 6. GPP comparison between the proposed approach and (a) the MOD17A2, and (b) the PML_V2 products, respectively. ME, RMSE and MAE units are g m^{-2} d^{-1}.

TABLE III
GPP ACCURACY METRICS PER BIOME WITH REGARD THE MOD17A2 AND PML_V2 PRODUCTS

	Vs MOD17A2				Vs PML_V2			
	ME	RMSE	MAE	R^2	ME	RMSE	MAE	R^2
CRO	0.84	2.51	1.91	0.35	-1.23	3.26	2.48	0.26
CSH	0.82	1.48	1.05	0.26	-0.84	1.03	0.87	**0.83**
DBF	0.20	1.90	1.39	0.67	-0.70	1.72	1.21	0.79
DNF(*)	-0.80	1.38	1.50	0.81	-0.98	1.16	0.98	0.88
EBF	0.21	2.11	1.60	0.64	-0.17	1.55	1.17	0.77
ENF	0.44	1.98	1.46	0.49	-0.40	1.71	1.17	0.61
GRA	0.28	1.88	1.31	0.67	-0.29	1.53	1.03	0.78
MF	0.36	2.17	1.55	0.51	-0.56	2.29	1.57	0.52
OSH	-0.72	**1.22**	**0.90**	**0.72**	-0.68	1.07	0.78	0.80
SAV	0.21	1.24	0.92	0.56	**0.07**	1.00	0.70	0.68
WET	0.38	2.18	1.89	0.49	-0.48	2.17	1.59	0.54
WSA	-0.22	1.27	0.93	0.59	-0.21	**0.78**	**0.57**	0.81

Best results highlighted in bold. GPP, NEE; TER, ME, RMSE and MAE in g m^{-2} d^{-1}. Codes: croplands (CRO), closed shrublands (CSH), deciduous broadleaf forest (DBF), deciduous needleleaf forests (DNF), evergreen broadleaf forests (EBF), evergreen needleleaf forests (ENF), grasslands (GRA), mixed forests (MF), open shrublands (OSH), savannas (SA), wetlands (WET), and woody savannas (WSA). (*) Only one site.

The results reported global amounts of 133.7 Pg yr^{-1}, 114.8 Pg yr^{-1}, and 18.9 Pg yr^{-1} for GPP, TER, and NEE, respectively in the 2002–2023 period. Note that those values preserve the relationship among carbon fluxes (NEE = GPP - TER). In the case of MOD17A2 and PML_V2 products we found a GPP global mean of 118.3 Pg yr^{-1} and 143.6 Pg yr^{-1}, respectively. Note that in the case of the PML_V2 the period was 2002–2020 due to data availability.

D. Comparison with MOD17A2 and PML_V2 products

The GPP estimates were compared with the provided ones by the MOD17A2, and the PML_V2 products available on GEE. The comparison was carried out at site level (Fig. 6). The R^2 values regarding the MOD17A2 and PML_V2 were 0.63 and 0.70, respectively. The $\mathcal{G}$-model predictions present a positive bias (0.71 g m^{-2} d^{-1}) regarding the MOD17A2, and a negative one with respect to the PML_V2 estimates (-0.46 g m^{-2} d^{-1}). Table III shows the comparison per biome type. The best concordances regarding the MOD17A2 product have been found over open shrublands in terms of RMSE, MAE, and R^2, and over crops in terms of ME. Over closed shrublands has been found the highest agreement regarding the PML_V2 estimates, but the lowest R^2 has been observed regarding the MOD17A2. The lowest correspondences regarding the two reference datasets have been reported over crops.

Fig. 7 shows difference maps of the mean annual GPP regarding the reference datasets ($\Delta_{GPP} = GPP_{\mathcal{G}model}$ -reference). In general, the $\mathcal{G}$-model agrees well with the PML_V2, however a slight overestimation of the GPP estimates provided by the PML_V2 regarding the $\mathcal{G}$-model is reported at high latitudes (see Fig. 7, top). Fig. 7 (bottom) shows the MOD17A2 estimates systematically underestimates GPP regarding the $\mathcal{G}$-model (mainly in the tropics) except over some locations.

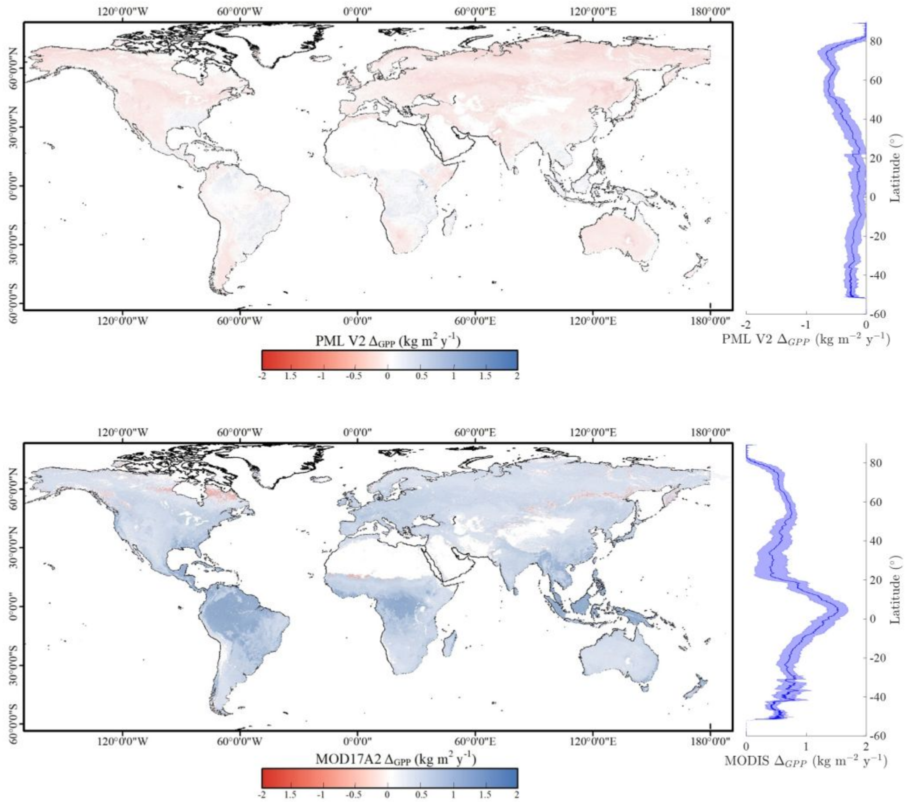


Fig. 7. Global difference maps for mean annual GPP regarding (top) PML_V2 product, and (bottom) MOD17A2 for the 2002–2020 period.

Finally, Fig. 8 shows the global GPP computed for every year during the 2002–2023 period in the case of the $\mathcal{G}$-model, the MOD17A2, and the PML_V2 product (2002–2020 period). Global GPP provided by the $\mathcal{G}$-model varies from 131.8 Pg C yr^{-1} to 137.6 Pg yr^{-1}, whereas the MOD17A2 reported lower amounts in the same period (118.1 Pg yr^{-1} to 119.9 Pg C yr^{-1}). The PML_V2 product showed a range of 128.5 Pg yr^{-1} to 133.9 Pg yr^{-1} (from 2002 to 2020). The corresponding linear trends in global GPP reported by the $\mathcal{G}$-model, the PML_V2, and the MOD17A2 are 0.28±0.02 Pg yr^{-1} yr^{-1}, 0.21±0.03 Pg yr^{-1} yr^{-1}, and 0.09±0.02 Pg yr^{-1} yr^{-1}, respectively.

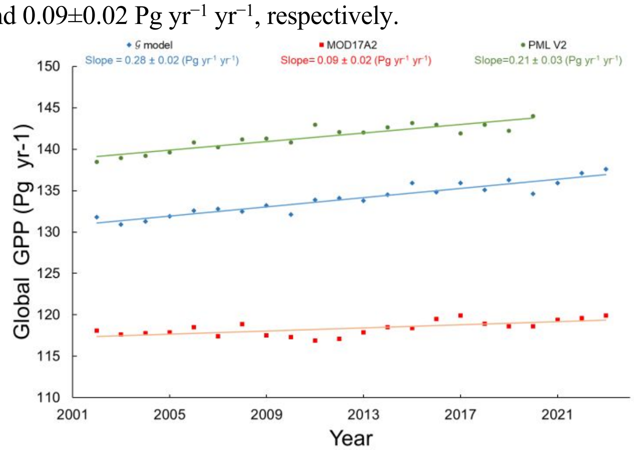


Fig. 8. Trends for mean annual GPP obtained by the G-model and the MOD17A2 for the 2002–2023 period, and the PML_V2 for 2002 to 2020.

IV. Discussion

A. Model performance and explanatory variables

The state-of-the-art machine learning approaches that provide the highest accuracies are based on a combination of RS observations and meteorological variables as inputs [14]–[16]. Nevertheless, this study exclusively utilized a set of RS predictors derived from MODIS data into a machine learning regression method, resulting in a remarkable level of concordance between GPP, NEE, and TER retrievals and in-situ data. The interactions between CO_2 fluxes and RS/meteo variables depend significantly on ecosystem types and state [29]. A main advantage of the proposed approach is its potential to obtain globally those fluxes without a priori knowledge of the pixel characteristics. The results of the accuracy assessment show an outperformance of the $\mathcal{G}$-model regarding both MOD17A2, and PML_V2 GPP products. The highest correlations for GPP and TER were obtained over woody savannas, savannas, grasslands, (in line with the results of Joiner *et al*. [19]) and evergreen broadleaf forests. In the case of NEE the best correlation was obtained over deciduous broadleaf forest, wetlands and crops. Crops is the biome where the worst results were reported in the case of GPP and TER ($R^2 \leq 0.58$), however the model performance in terms of correlation for NEE over this biome was good ($R^2 = 0.77$). Closed and open shrublands are the biomes where the lowest correlations for NEE ($R^2 \leq 0.31$) have been found.

The $\mathcal{G}$-model consistently predicted GPP, NEE, and TER while effectively maintaining their inherent relationship. This is an important feature of this work since the usual procedure in machine learning approaches does not exploit the covariance of the carbon fluxes when upscaling in-situ data to global scale. Tramontana *et al*. [30] developed an approach for partitioning carbon fluxes preserving the relationship among them using neural networks physically constrained, but only applied with in-situ data.

The selection of the RS predictive variables is based on the main drivers of photosynthesis, transpiration and respiration processes. In this regard, MODIS products that are related with water, radiation, temperature, and evapotranspiration were selected. The ranking of the explanatory variables highlighted PAR as the most influential predictor of GPP estimates as expected according to Monteith logic [31]. However, the vegetation index kNDVI has also a great influence, which suggests that it might be connected with the light use efficiency (LUE). In fact, Wu *et al*. [32] proposed a GPP approach incorporating vegetation indices for LUE and the fraction of absorbed photosynthetically active radiation (FAPAR) estimation. In literature, different authors have shown the usefulness of vegetation indices as proxies of both LUE and FAPAR [33]–[34].

The relevance analysis for TER estimates revealed LAI as the most important variable. LST_N, ET, PET, and C_{ws} also revealed as important factors for TER. In the case of NEE, the main explanatory variables were the two most relevant for GPP and TER, i.e., PAR and LAI. Globally, temperature is an influencing factor on the three fluxes: the diurnal temperature slightly affects GPP (i.e., photosynthesis), probably through the LUE in case of a thermal stress (similar to the role of evapotranspiration in case of water stress), and night temperature affects respiration process and, to a lesser extent, the NEE. Respiration provides the energy for a plant to acquire nutrients and to produce and maintain biomass. The total plant respiration can be split into three functional components: growth respiration, maintenance respiration, and the respiratory cost of ion uptake [35]. The maintenance respiration is expected to be greater in productive ecosystems and then to be positively correlated with LAI, with a rather high relevance in respiration estimation as shown in Fig. 3. Maintenance respiration depends also on environmental stresses: for a given ecosystem it increases with temperature and drought. However, plants from hot environments have lower respirations rates at a given temperature than plants from cold places [35]. In any case, temperature and water balance are both rather relevant variables in respiration processes.

B. Global estimates

GPP estimates at site level provided by the $\mathcal{G}$-model are highly correlated with both MOD17A2 and PML_V2 estimates (R^2 =0.63 and R^2=0.71). One factor that explains the high correlation is the use of similar inputs in the approaches. The per biome type comparison found the best agreements over open shrublands in the case of MOD17A2, and over closed shrublands in the case of PML_V2. It is worth mentioning that the worst consistencies between products were obtained over crops in both cases. The spatial distribution of the GPP is consistent with the MOD17A2, and PML_V2 products. The major concentrations of GPP and TER are located in the tropics and decreases to high and low latitudes.

The $\mathcal{G}$-model provided a global GPP (133.7 Pg yr^{-1}) similar to the one reported by other studies. In instance, Zhang *et al*. [37] reported an average global annual GPP around 125 Pg yr^{-1} for the VPM and FluxCom products. Joiner *et al*. [19] revealed an annual GPP value around 140 Pg yr^{-1}. Jung *et al*. [14] found mean global GPP of 111 Pg yr^{-1} for FluxCom-RS, 120 Pg yr^{-1} for FluxCOM-RS+METEO product, and a range of 83–172 Pg yr^{-1} for TRENDY v7 models. Dong *et al*. [37] carried out an inter-comparison of global GPP annual mean provided by PEM models, which reported a range of variation of 125–165 Pg yr^{-1}. In addition, Badgley *et al*. [6] reported a range of 131–163 Pg yr^{-1} using the NIRv as proxy for GPP. Recently, Li *et al*. [38] reported 125.74 Pg yr^{-1} as a mean value for the 1982–2019 period, and Tagesson *et al*., [39] found 121.8 Pg yr^{-1} over the 1982–2015 period. The TER and NEE budgets found in this study (114.8 Pg yr^{-1}, and 18.9 Pg yr^{-1}, respectively) are closer to the ones reported by Zeng *et al*. [40] (TER ranging from115 to 121 Pg yr^{-1}, and NEE ranging from 20 to 22 Pg yr^{-1}). In addition, Li *et al*. [40] found 109.3 Pg yr^{-1} for TER, and 16.28 Pg yr^{-1} for NEE, and Tagesson *et al*. [41] reported 105.6 Pg yr^{-1} for TER in the 1982–2015 period. These values are lower than the found in our study. The range of variation of the values reported in the literature indicate a significant level of uncertainty persists across different approaches.

The 2002–2023 temporal analysis provided a GPP trend (0.28±0.02 Pg yr^{-1} yr^{-1}) similar to other studies. Tagesson *et al*., [39] found a trend of 0.27±0.02 Pg yr^{-1} yr^{-1} in the 1982–2015 period, Guo *et al*. [15] and Zeng *et al*. [40] reported 0,21 Pg yr^{-1} yr^{-1} and 0.49 Pg yr^{-1} yr^{-1} both computed during the 1999–2019 period. The inter-comparison conducted by Dong *et al*. [37] reported a range of global GPP trends of different products, which varies from −0.22 to 0.51 Pg yr^{-1} yr^{-1} over different periods.

C. Model Uncertainty

Although the results show a good level of agreement regarding in-situ data and other products, there exist sources of uncertainty that propagates into the model. The spatial mismatch between the footprint of the EC towers and the RS data spatial resolution is source of uncertainty in CFs retrievals. Uncertainties in the EC covariance data, and in the RS-based products can propagate into carbon fluxes retrieval [42]–[44]. In addition, the different spatial resolutions of the model inputs may also affect the stability of the predictions [45].

The $\mathcal{G}$-model provides not only a joint estimate of GPP, NEE, and TER but also an uncertainty (or confidence interval) for the retrievals. Note that the uncertainty of the predictive mean estimate does not depend on the value of the predictive mean but accounts for the similarity of the new inputs ($\mathbf{x}_*$) regarding the inputs trained in the model ($\mathbf{x}_i$). Therefore, high uncertainties would be obtained in scenarios where the behavior of the inputs are not accounted during model training. The latitudinal profiles of the uncertainties are quite constant and do not follow the latitudinal variation of the GPP, NEE, and TER. This decorrelation between predictions and uncertainties suggests that the uncertainties are not influenced by the value of the prediction.

V. Conclusions

This study developed a data-driven approach using a multi-output regression method that is able to predict GPP, TER, and NEE at the same time. The $\mathcal{G}$-model estimates the three fluxes jointly also conserving their physical relationship, which is one of the drawbacks of the classic machine learning algorithms. The predictors of the model came from five MODIS products, which resulted in 8 explanatory variables. Among them, the PAR and kNDVI were identified as the most relevant inputs for GPP, whereas the LAI and LST_N were the most influential for TER retrievals. PAR and LAI were also the most important for the NEE. The model was well calibrated at tower level, and successfully upscaled to globe. At tower level the best results were obtained for GPP (R^2=0.82), followed by TER (R^2=0.80), and NEE (R^2=0.69). The RMSE was 1.55 g m^{-2} d^{-1} in the case of GPP, 1.09 g m^{-2} d^{-1} for NEE, and 1.14 g m^{-2} d^{-1} for TER. These results, obtained from only MODIS-based inputs, outperformed MOD17A2H and PML_V2 products. At global scale the spatial distribution of the derived maps agreed with MOD17A2H and PML_V2. The results provide a mean annual GPP value of 133.7 Pg yr^{-1} and a global trend of 0.28 Pg yr^{-1} yr^{-1} for the 2002–2023 period that fall within the range reported by other works. However, further studies are needed to better understand the spatial and temporal patterns of global carbon fluxes that help to reduce uncertainties among approaches.

References

[1] M. A. White, P. E. Thornton, S. W. Running, and R. R. Nemani, "Parameterization and sensitivity analysis of the BIOME–BGC terrestrial ecosystem model: Net primary production controls" *Earth Interact.,* vol. 4, no 3, pp.1–85, 2000, doi: 10.1175/1087-3562(2000)004%3C0003:PASAOT%3E2.0.CO;2

[2] S. Sitch, *et al*., "Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs)", *Glob. Change Biol.,* vol. 14, no 9, pp. 2015–2039, 2008, doi: 10.1111/j.1365-2486.2008.01626.x

[3] S. Sitch, *et al*., "Recent trends and drivers of regional sources and sinks of carbon dioxide"*, Biogeosciences*, vol. 12, no 3, pp. 653–679, 2015, doi: 10.5194/bg-12-653-2015

[4] S. Running, Q. Mu, and M. Zhao, "MOD17A3H MODIS/Terra Net Primary Production Yearly L4 Global 500m SIN Grid V006," NASA EOSDIS Land Processes DAAC, 2015, doi: 10.5067/MODIS/MOD17A3H.006.

[5] B. Martínez, *et al*., "Evaluation of the LSA-SAF gross primary production product derived from SEVIRI/MSG data (MGPP)", *ISPRS J. Photogramm.,* vol. 159, pp. 220–236, 2020, doi: 10.1016/j.isprsjprs.2019.11.010

[6] G. Badgley, L. D. L. Anderegg, J. A. Berry, and C. B. Field, "Terrestrial gross primary production: Using NIRV to scale from site to globe," *Global Change Biol*., vol. 25, no. 11, pp. 3731–3740, 2019, doi: 10.1111/gcb.14729

[7] R. Wang, *et al*., "Detecting intra-and inter-annual variability in gross primary productivity of a North American grassland using MODIS MAIAC data", *Agr. Forest Meteorol*, vol. 281, Art. no. 107859, 2020, doi: 10.1016/j.agrformet.2019.107859

[8] Y. Sun, *et al*., "OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence", *Science*, vol. 358; Art. no. eaam5747, 2017, doi: 10.1126/science.aam5747

[9] D. Hao, *et al*., "Adjusting solar-induced fluorescence to nadir-viewing provides a better proxy for GPP", *ISPRS J. Photogramm*., vol. 186, pp. 157–169, 2022, doi: 10.1016/j.isprsjprs.2022.01.016

[10] J. Xiao, *et al*., "Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data", *Agr. Forest Meteorol*., vol. 148, pp. 1827–184, 2008, doi: 10.1016/j.agrformet.2008.06.015

[11] G. Tramontana, *et al*., "Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms", *Biogeosciences,* vol. 13, pp. 4291–4313, 2016, doi: 10.5194/bg-13-4291-2016

[12] K. Ichii, *et al*., "New data-driven estimation of terrestrial CO2 fluxes in Asia using a standardized database of eddy covariance measurements, remote sensing data, and support vector regression", *J. Geophys. Res-Biogeo*., vol. 122, pp. 767–795, 2017, doi: 10.1002/2016JG003640

[13] Y. Zhang, *et al*., "Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017", *Remote Sens. Environ.,* vol. 222, pp. 165–182, 2019, doi: 10.1016/j.rse.2018.12.031

[14] M. Jung, *et al*., "Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach", *Biogeosciences,* vol. 17, pp. 1343–1365, 2020, doi: 10.5194/bg-17-1343-2020

[15] R. Guo, *et al*., "Estimating Global GPP from the Plant Functional Type Perspective Using a Machine Learning Approach", *J. Geophys. Res-Biogeo*., vol. 128, no 4, Art. no. e2022JG007100, 2023, doi: 10.1029/2022JG007100

[16] D. Tuia, J. Verrelst, L. Alonso, F. Perez-Cruz, and G. Camps-Valls, "Multioutput support vector regression for remote sensing biophysical parameter estimation," *IEEE Geosci. Remote Sens. Lett*., vol. 8, no. 4, pp. 804–808, Jul. 2011, doi: 10.1109/LGRS.2011.2109934

[17] F. J. García-Haro *et al*., "Derivation of global vegetation biophysical parameters from EUMETSAT polar system," *ISPRS J. Photogramm*. Remote Sens., vol. 139, pp. 57–74, May 2018, doi: 10.1016/j.isprsjprs.2018.03.005.

[18] M. Reichstein, *et al*., "On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm", *Glob. Change Biol*., vol. 11, pp. 1424–1439, 2005, doi: 10.1111/j.1365-2486.2005.001002.x

[19] J. Joiner, Y. Yoshida, Y. Zhang, G. Duveiller, M. Jung, A. Lyapustin, Y. Wang, and C. Tucker, ''Estimation of terrestrial global gross primary production (GPP) with satellite data-driven models and eddy covariance flux data,'' *Remote Sens*., vol. 10, no. 9, Art. no. 1346, Aug. 2018, doi: 10.3390/rs10091346

[20] S. Bao *et al*., "Environment-sensitivity functions for gross primary productivity in light use efficiency models," *Agricult. Forest Meteorol.*, vol. 312, Jan. 2022, Art. no. 108708, doi: 10.1016/j.agrformet.2021.108708

[21] G. Camps-Valls *et al*., "A unified vegetation index for quantifying the terrestrial biosphere," Sci. Adv., vol. 7, no. 9, Art. no. eabc7447, 2021, doi: 10.1126/sciadv.abc7447

[22] S. Liang, and D. Wang, "Moderate Resolution Imaging Spectroradiometer (MODIS) Downward Shortwave Radiation (MCD18A1) and Photosynthetically Active Radiation (MCD18A2) Algorithm Theoretical Basis Document", 2017, doi: 10.5067/MODIS/MCD18C2.061

[23] A. H. Strahler, *et al*., "MODIS BRDF/albedo product: algorithm theoretical basis document version 5.0", pp.1–53, 1999, doi: 10.5067/MODIS/MCD43A4.061 (1999).

[24] G. C. Hulley, S. J. Hook, and C. Hughes, "MODIS MOD21 land surface temperature and emissivity algorithm theoretical basis document. Pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration. 2012. [Online]. Available: https://emissivity.jpl.nasa.gov/downloads/examples/documents/MOD21_LSTE_ATBD_Hulley_v2.0_20121116_pmbv1.pdf

[25] Q. Mu, M. Zhao and S. W. Running, S. W. "MODIS Global Terrestrial Evapotranspiration (ET) Product (NASA MOD16A2/A3)". Algorithm Theoretical Basis Document. Collection 5. Numerical Terradynamic Simulation Group. College of Forestry and Conservation. The University of Montana, 2013, doi: 10.5067/MODIS/MOD16A2.006

[26] Y. Knyazikhin *et al*., "MODIS leaf area index (LAI) and fraction of photosynthetically active radiation absorbed by vegetation (FPAR) product (MOD15) algorithm theoretical basis document (Version 4.0)," NASA Goddard Space Flight Center (GSFC), Greenbelt, MD, USA, Tech. Rep., 1999. [Online]. Available: https://modis.gsfc.nasa.gov/data/atbd/atbd_mod15.pdf

[27] C. E. Rasmussen and K. I. Christopher, Gaussian Processes for Machine Learning. Massachusetts, MA, USA: MIT Press, 2005, pp. 7–31

[28] M. Campos-Taberner *et al*., "Understanding deep learning in land use classification based on sentinel-2 time series," *Sci. Rep*., vol. 10, no. 1, pp. 1–12, 2020, doi: 10.1038/s41598-020-74215-5

[29] Z. Man, *et al*. "Effect of climate change on CO2 flux in temperate grassland, subtropical artificial coniferous forest and tropical rain forest ecosystems", *Int. J. Environ. Res. Public Health*, vol. 18, pp. 13056, 2021, doi: 10.3390/ijerph182413056

[30] G. Tramontana *et al*., "Partitioning net carbon dioxide fluxes into photosynthesis and respiration using neural networks," *Global Change Biol.*, vol. 26, no. 9, pp. 5235–5253, 2020, doi: 10.1111/gcb.15203

[31] J. L. Monteith, "Solar radiation and productivity in tropical ecosystems", *J. Appl. Ecol*., vol. 9, pp. 747-766, 1972. [Online]. Available: https://www.jstor.org/stable/2401901 (1972).

[32] C. Wu, Z. Niu, and S. Gao, "Gross primary production estimation from Modis data with vegetation index and photosynthetically active radiation in maize," *J. Geophys. Res. Atmospheres*, vol. 115, pp. 13–20, 2010, doi: 10.1029/2009JD013023

[33] Y. Inoue, J. Penuelas, A. Miyata, and M. Mano, "Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO2 flux measurements in rice," *Remote Sens. Environ*., vol. 112, no. 1, pp. 156–172, Jan. 2008, doi: 10.1016/j.rse.2007.04.011

[34] C. Y. Wu, J. W. Munger, Z. Niu, and D. Kuang, "Comparison of multiple models for estimating gross primary production using MODIS and eddy covariance data in Harvard forest," *Remote Sens. Environ*., vol. 114, pp. 2925–2939, Dec. 15, 2010, doi: 10.1016/j.rse.2010.07.012

[35] F. S. Chapin III, P. A. Matson, H. A. Mooney, and P. M. Vitousek, "Principles of terrestrial ecosystem ecology", Springer-Verlag, New York, 2002, 436 pp. doi: 10.1007/978-1-4419-9504-9

[36] Y. Zhang *et al*., "A global moderate resolution dataset of gross primary 700 production of vegetation for 2000–2016," *Sci. Data*, vol. 4, no. 1, 701 Dec. 2017, Art. no. 170165, doi: 10.1038/sdata.2017.165

[37] J. Dong, L. Li, Y. Li, and Q. Yu, "Inter-comparisons of mean, trend and interannual variability of global terrestrial gross primary production retrieved from remote sensing approach", *Sci. Total Environ*., vol. 822, Art. no. 153343, 2022, doi: 10.1016/j.scitotenv.2022.153343

[38] B. Li *et al*., "BESSv2. 0: A satellite-based and coupled-process model for quantifying long-term global land–atmosphere fluxes", *Remote Sens. of Environ.,* vol. 295, Art. no. 113696; 2023, doi: 10.1016/j.rse.2023.113696

[39] T. Tagesson *et al*. "A physiology-based Earth observation model indicates stagnation in the global gross primary production during recent decades", *Glob. Change Biol*., vol. 27, pp. 836–854, 2021, doi: 10.1111/gcb.15424

[40] J. Zeng *et al*., "Global terrestrial carbon fluxes of 1999-2019 estimated by upscaling eddy covariance data with a random forest," *Sci. Data*, vol. 7, no. 1, Sep. 2020, Art. no. 313, doi: 10.1038/s41597-020-00653-5

[41] T. Tagesson *et al*., "Increasing global ecosystem respiration between 1982 and 2015 from Earth observation-based modelling", *Glob. Ecol. Biogeogr*., vol. 33, pp. 116–130, 2012, doi: 10.1111/geb.13775

[42] F. A. Heinsch *et al*., "Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations", *IEEE Trans. Geosci*., vol. 44, pp. 1908–1925, 2006, doi: 10.1109/TGRS.2005.853936

[43] J. Kala *et al*., "Influence of leaf area index prescriptions on simulations of heat, moisture, and carbon fluxes", *J. Hydrometeorol*., vol. 15, pp. 489–503, 2014, doi: 10.1175/JHM-D-13-063.1

[44] Y. Liu *et al*., "Satellite-derived LAI products exhibit large discrepancies and can lead to substantial uncertainty in simulated carbon and water fluxes," *Remote Sens. Environ*., vol. 206, pp. 174–188, Mar. 2018, doi: 10.1016/j.rse.2017.12.024.

[45] G. Gelybó, Z. Barcza, A. Kern, and N. Kljun, ''Effect of spatial heterogeneity on the validation of remote sensing based GPP estimations,'' Agricult. Forest Meteorol., vols. 174–175, pp. 43–53, Jun. 2013, doi: 10.1016/j.agrformet.2013.02.003