https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Performance assessment of the global gross primary productivity datasets over India for climate and ecophysiological applications

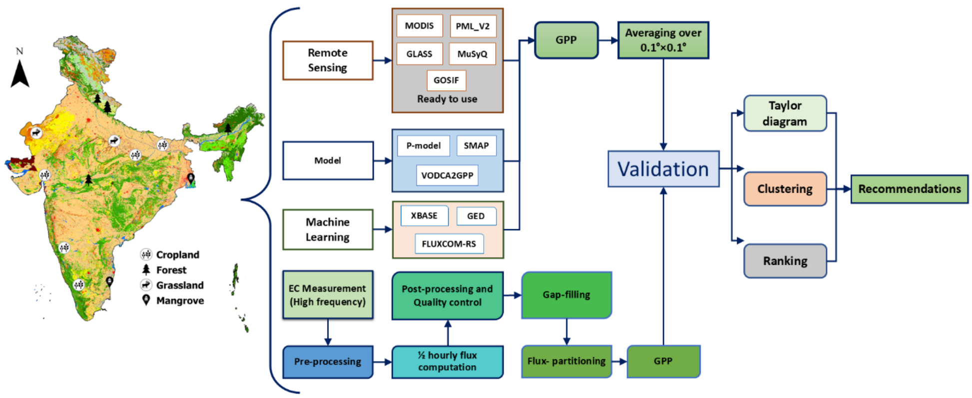
Pramit Kumar Deb Burman[1,2,*], Amol Kale[1], Yogesh K. Tiwari[1], Suraj Reddy Rodda[3], Sandipan Mukherjee[4], Pulakesh Das[5], Arunima Jayachandran[1,2], Dipankar Sarma[1], Palingamoorthy Gnanamoorthy[6], Nirmali Gogoi[7], Somnath Baidya Roy[8], Ganapati S. Bhat[9], Kazuhito Ichii[10,11]

[1]Centre for Climate Change Research, Indian Institute of Tropical Meteorology, Ministry of Earth Sciences, Pune, India
[2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad - 201002, India
[3]National Remote Sensing Centre, Indian Space Research Organisation, Hyderabad, India
[4]Ladakh Regional Centre, GB Pant National Institute of Himalayan Environment, Leh, India
[5]Department of Agriculture, Madhya Pradesh State Electronics Development Corporation, Noida, India
[6]Yunnan Key Laboratory of Forest Ecosystem Stability and Global Change, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, China
[7]Department of Environmental Sciences, Tezpur University, Tezpur, India
[8]Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, New Delhi, India
[9]Centre for Atmospheric and Oceanic Sciences, Indian Institute of Science, Bangalore, India
[10]Center for Environmental Remote Sensing (CEReS), Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba 263-8522, Japan
[11]Graduate School of Science and Engineering, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

Correspondence to: Pramit Kumar Deb Burman (pramit.cat@tropmet.res.in)

Abstract. Gross primary productivity (GPP) represents the total carbon fixed by plants via photosynthesis, which is a fundamental component of the terrestrial carbon cycle. Yet its accurate estimation remains highly uncertain in regions with strong climate variability and ecosystem heterogeneity. In particular, monsoon-driven ecosystems, such as those across India, represent one of the most challenging and under-evaluated environments for global GPP products due to pronounced hydroclimatic variability, diverse vegetation types, and intensive land management practices, which introduce significant uncertainties in regional and global carbon budgets. Here, we conduct the first systematic and independent benchmarking of 11 widely used global GPP datasets using eddy covariance (EC) measurements spanning 14 sites representing the major land cover types in India, namely croplands, forests, grasslands, and mangroves. These datasets are generated using diverse methods, including remote sensing, process-based and light-use efficiency models, and machine learning (ML) algorithms. By integrating multi-metric statistical evaluation, Taylor diagram analysis, and hierarchical clustering, we assess the capability of these datasets and rank them to capture site-level variability in GPP. Our results reveal substantial ecosystem-specific discrepancies among these products. ML-based datasets (e.g., FLUXCOM-RS and XBASE) consistently perform well in croplands and forests, whereas process-based and microwave-derived products exhibit variable or degraded performance, particularly in grasslands. Notably, all datasets show reduced skill in mangrove ecosystems, highlighting fundamental limitations in representing the complex hydrological and biophysical processes. These findings indicate that global GPP products, although widely used, exhibit systematic biases under monsoon regimes and management-intensive landscapes. Our study demonstrates that monsoon-driven ecosystems provide a critical testbed for evaluating global carbon cycle products and reveals key

limitations in current GPP estimation approaches. The results provide guidance for selecting appropriate datasets for regional applications and underscore the need for improved representation of hydrological variability and ecosystem complexity in next-generation GPP products and informed decision-making for carbon management and climate mitigation.



Graphical abstract.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

1. Introduction

Gross Primary Productivity (GPP), is a central component of the terrestrial carbon cycle, representing the total carbon fixed by vegetation through photosynthesis (Chapin et al., 2006). Accurate estimation of GPP is essential for understanding ecosystem functioning, quantifying carbon budgets (Friedlingstein et al., 2023), and supporting climate mitigation strategies, where the signatory nations to the Paris Climate Accord are required to report their national carbon inventories in regular intervals, as the intended nationally determined contribution (INDC) (Hönle et al., 2019; MoEFCC, 2015, 2020; Ravindranath et al., 2017), wherein GPP plays a central role. An accurate GPP measurement is also needed for ecosystem health monitoring (Sarkar et al., 2022), quantifying the carbon cycle (Malhi et al., 2021), estimating crop productivity, ensuring food security, and assessing economic growth (Beer et al., 2010; Pan et al., 2011; Running et al., 2004).

Multiple methods are available for estimating GPP. The most commonly used method is the eddy covariance (EC), which measures the ecosystem-atmosphere exchanges of mass, momentum, and energy in-situ at sub-hourly time intervals for extended periods (Aubinet et al., 2012; Baldocchi et al., 2001). Global EC networks like AmeriFlux, AsiaFlux, ICOS, and OzFlux monitor ecosystem-atmosphere carbon fluxes in North and South America, Asia, Europe, and Australia and New Zealand, respectively (Beringer et al., 2016; Novick et al., 2018; Rebmann et al., 2018; Yamamoto et al., 2005). The EC data provide precise CO_2 flux measurements, needed to validate GPP models and remotely sensed GPP products.

The most common bottom-up approach used for GPP estimation is the Radiation Use Efficiency (RUE) method proposed by Monteith, (1972), which estimates GPP based on radiation absorbed by plants. This method relies on the Light Use Efficiency (LUE), which represents the plant's efficiency in converting the absorbed photosynthetically active radiation (PAR) to biomass through photosynthesis. Process-based models simulate photosynthesis based on the plant physiological traits and photosynthetic pathways (Ball et al., 1987; Farquhar et al., 1980). GPP is also computed from satellite-measured reflectances, which bear the mark of plants' photosynthetic activities, leaf temperature, chlorophyll, and water contents (Bannari et al., 1995; Gilabert et al., 2015; Huang et al., 2019). Solar-induced fluorescence (SIF) is a direct proxy of GPP, as it represents the energy emitted by chlorophyll during photosynthesis (Frankenberg et al., 2011). Recent advancements in Machine Learning (ML), and Deep Learning (DL) techniques have significantly enhanced the ability to upscale GPP estimates at global and regional scales (Agarwal et al., 2024; Gaber et al., 2024; Liao et al., 2023; Reichstein et al., 2019; Taye, 2023; Nelson et al., 2024). These techniques combine ground-based and remotely sensed data with model simulations, and show proficiency in detecting complex patterns often missed by traditional methods (Mohamed and Klaus, 2024).

Over the past decades, numerous global GPP datasets have been developed using the aforementioned approaches, enabling continuous monitoring of ecosystem productivity at large spatial scales. These products are derived using diverse methodological frameworks, each characterized by distinct assumptions and sensitivities to environmental conditions. Despite these advances, substantial uncertainties remain in global GPP estimates (Wild et al., 2022), which arise from simplified assumptions in light-use efficiency models, parameterization limitations in process-based models, and biases in machine learning approaches related to training data representativeness (Bloomfield et al., 2023). Additionally, most of the validation efforts have relied on EC flux tower networks

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

concentrated in temperate and boreal regions, leaving other climatic regimes insufficiently evaluated (Turner et al., 2006; Coops et al., 2007a). Consequently, the performances of global GPP products remain highly biased towards a few specific environmental conditions.

Monsoon-driven ecosystems represent one of the most challenging yet under-evaluated environments for GPP estimation. Strong seasonal variability in precipitation, frequent cloud cover, and rapid vegetation responses introduce nonlinear dynamics that are difficult to capture with existing frameworks. In addition, intensive land management practices, such as multiple cropping cycles in agricultural systems, and complex ecosystems such as mangroves, further complicate the relationship between remotely sensed signals and actual carbon uptake. India, characterized by diverse land cover types and a dominant monsoon climate, provides a unique and critical testbed for evaluating global GPP datasets under these conditions.

The current study includes fourteen EC flux sites, located in different land use land cover (LULC) classes spanning diverse ecosystems across India, to validate, objectively assess, and rank multiple global GPP products. This evaluation against EC measurements in India serves as an independent, unique, and challenging test of these gridded datasets, as tower-based measurements were not available during their development. In particular, the present study aims to (1) quantify the performance of global GPP products under monsoon-driven conditions, (2) identify ecosystem-dependent strengths and limitations of different modeling approaches, and (3) provide guidance for the selection and improvement of GPP datasets in regions characterized by strong climatic variability and complex land surface processes.

2. Materials and methods

All the abbreviations and acronyms used in this work are listed in Table 1.

2.1 Study area

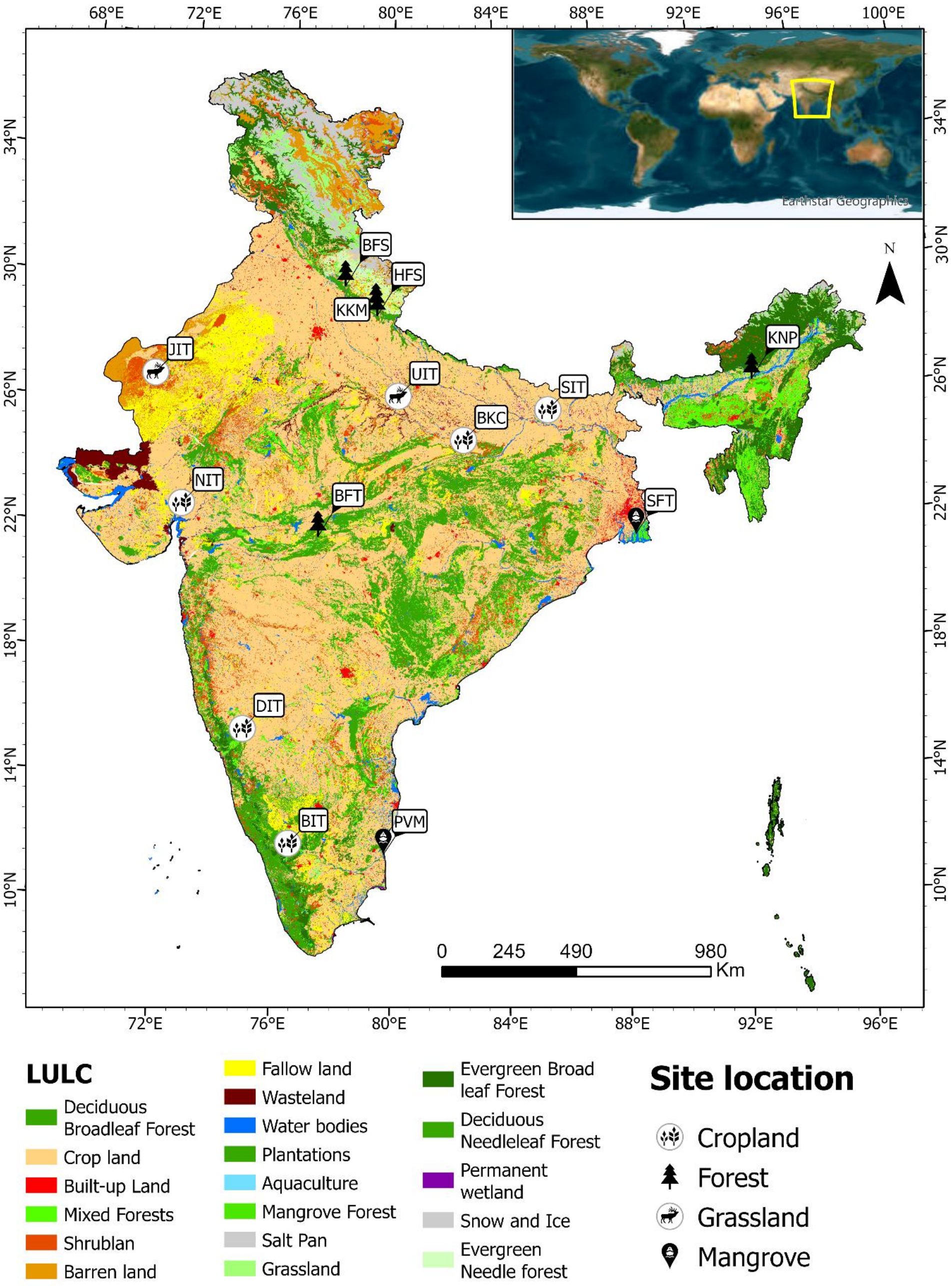
The locations of the flux tower sites are shown in Figure 1, and site-wise characteristics and features are given in Table 2. These sites represent a variety of LULCs prevalent across India. Further, those sites that fall under the same broad category [for example, forest (BFT, BFS, and KNP)] may have different climates, receive different amounts of annual precipitation, and therefore, have different plant types, and their GPPs are expected to have different magnitudes and variabilities (check Deb Burman et al., (2025) for a detailed review of these).

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Table 1. List of the abbreviations and acronyms used in the current study work.

Acronyms	Definition	Unit
Adj. R^2	Adjusted coefficient of determination	--
CO_2 fluxes	Carbon dioxide flux	µmol mol
GPP	Gross primary productivity	gC m^{-2} d^{-1}
LAI	Leaf area index	m^2 m^{-2}
LUE	Light Use Efficiency	gC M J^{-1}
MAP	Mean Annual Precipitation	mm
MAT	Mean Annual Temperature	°C
NDVI	Normalized Difference Vegetation Index	--
NEE	Net ecosystem exchange	µmol m^{-2} s^{-1}
PAR	Photosynthetically Active Radiation	µmol m^{-2} s^{-1}
r	Pearson's correlation coefficient	--
R^2	Coefficient of determination	--
rH	Relative humidity	%
T_{air}	Air temperature	°C
VPD	Vapor pressure deficit	hPa
Abbreviation	**Definition**	
APAR	Absorbed photosynthetically active radiation	
AVHRR	Advanced Very High-Resolution Radiometer	
CI	Clearness index	
DL	Deep Learning	
DSR	Downward solar radiation	
EC	Eddy covariance	
ERA5	5th generation European reanalysis data	
ETM+	Enhanced Thematic Mapper Plus	
fAPAR	Fraction of absorbed photosynthetically active radiation	
GED	Global Environmental Database	
GEE	Google Earth Engine	
GLASS	Global LAnd Surface Satellite	
GLDAS	Global Land Data Assimilation System	
GOSIF	Global OCO-2 based Solar-induced Chlorophyll Fluorescence	
GPP_{EC}	GPP from EC measurement	
IGBP	International Geosphere-Biosphere Programme	
IRS	India Remote Sensing satellites	
LISS-III	Linear Imaging Self-Scanning Sensor III	
MAE	Mean Absolute Error	
ML	Machine Learning	
MODIS	Moderate Resolution Imaging Spectroradiometer	
MuSyQ	Multisource data Synergized Quantitative	
NDVI	Normalized Difference Vegetation Index	
OCO-2	Orbiting Carbon Observatory-2	
PFT	Plant Functional Type	
PML	Penman-Monteith-Leuning Evapotranspiration	
QC	Quality control	
RF	Random Forest	
RMSE	Root Mean Squared Error	
SD	Standard Deviation	
SIF	solar-induced chlorophyll fluorescence	
VIIRS	Visible Infrared Imaging Radiometer Suite	
VODCA2GPP	Vegetation Optical Depth Climate Archive	
WUE	Water Use Efficiency	

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026



120

Figure 1. The eddy covariance flux tower sites used in this work are shown on the map of India with LULC classification following Roy et al. (2015).

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Table 2. Characteristics of eddy covariance sites from different ecosystems in India. AMSL: above mean sea level; MAT: mean annual temperature; MAP: mean annual precipitation.

Serial No.	Name of the site	Site code	Location Details			Land cover	MAT (°C)	MAP (mm)	Flux footprint (m)	Surface energy budget closure (%)	Period of measurement availability	Reference
			Latitude (°N)	Longitude (°E)	Elevation (m AMSL)							
1.	Berambadi, Karnataka	BIT	11.76	76.59	876	Cropland	24.4	818	55	81	Jan 2016 – Dec 2018	(Morrison et al., 2019)
2.	Barkachha, Uttar Pradesh	BKC	25.06	82.59	192		24.1	948	245	41	June 2014 – March 2016	(Deb Burman et al., 2020)
3.	Dharwad, Karnataka	DIT	15.5	74.99	709		24.7	964	45	93	Jan 2016 – Dec 2017	(Morrison et al., 2019)
4.	Nawagam, Gujarat	NIT	22.8	72.57	43		27.0	916	130	71	June 2016 – Nov 2018	(Bhat et al., 2020)
5.	Samastipur, Bihar	SIT	26	85.67	69		24.3	1033	145	66	June 2016 – Nov 2018	(Bhat et al., 2020)
6.	Barkot, Uttarakhand	BFS	30.11	78.21	448	Forest	21.9	1115	--	74	Jan 2016 – Dec 2018	(Watham et al., 2020)
7.	Betul, Madhya Pradesh	BFT	21.86	77.42	519		24.7	1208	750	72	Jan 2016 – Dec 2018	(Rodda et al., 2021)
8.	Haldwani, Uttarakhand	HFS	29.15	79.42	323		23.2	943	--	57	Jan 2016 – Dec 2018	(Watham et al., 2020)
9.	Kosi Watershed, Uttarakhand	KKM	29.38	79.37	1256		13.7	1136	170	77	March 2014 – Dec 2016	(Deb Burman et al., 2021)
10.	Kaziranga National Park, Assam	KNP	26.58	93.1	95		23.8	1725	425	83	Jan 2016 – Dec 2018	(Deb Burman et al., 2021)

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

11.	Jaisalmer, Rajasthan	JIT	26.99	71.34	216	**Grassland**	26.8	187	165	82	Jan 2016 – Oct 2018	(Bhat et al., 2020)
12.	Kanpur, Uttar Pradesh	UIT	26.51	80.22	148		24.3	733	53	81	Jan 2016 – Dec 2017	(Morrison et al., 2019)
13.	Pichavaram Mangrove, Tamil Nadu	PVM	11.43	79.79	14	**Mangrove**	28.5	1267	150	79	Oct 2017 – Sept 2018	(Gnanamoorthy et al., 2020)
14.	Sundarbans, West Bengal	SFT	21.82	88.62	21		26.6	1783	160	70	Jan 2014 – March 2016	(Rodda et al., 2016)

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

2.2 Data

2.2.1 Land Use Land Cover (LULC)

The LULC data from Roy et al. (2016), available at https://daac.ornl.gov/VEGETATION/guides/Decadal_LULC_India.html by NASA's Oak Ridge National Laboratory Distributed Active Archive Centre (ORNL DAAC), has been used in this work. The most recent data available in this dataset includes the LULC map of India for 2005 at 100 m resolution. This product is derived from Enhanced Thematic Mapper Plus (ETM+) sensors of Landsat 5 at 30 m spatial resolution and Resourcesat Linear Imaging Self-Scanning Sensor III (LISS-III) of India Remote Sensing satellites (IRS) data at 23.5 m spatial resolution. The data were classified according to the International Geosphere-Biosphere Program (IGBP) classification scheme based on visual image interpretation.

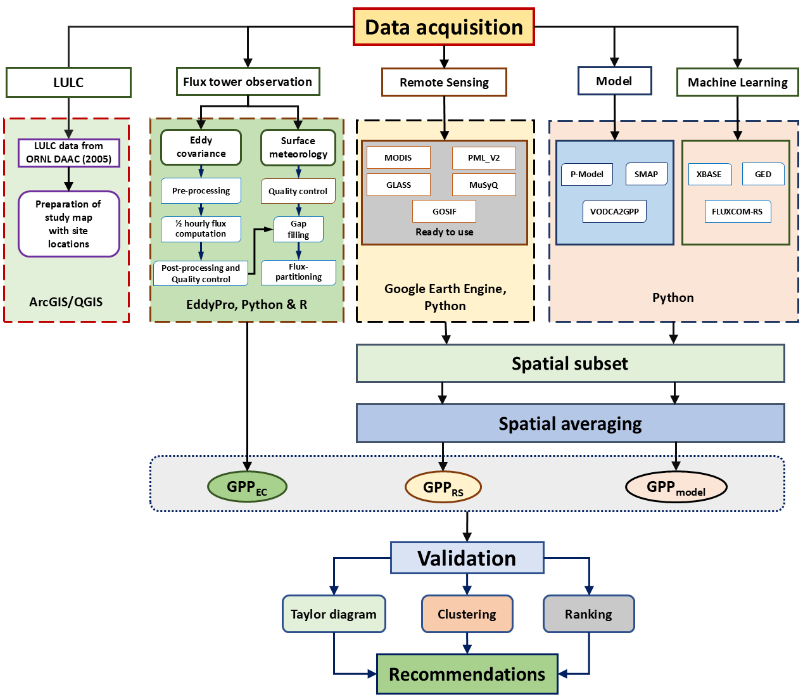


Figure 2. Schematic representation of the workflow implemented in the present study.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

2.2.2 In-situ GPP

Tower-based EC measurements at the following locations are used in this work: Barambadi (BIT), Barkachha (BKC), Dharwad (DIT), Nawagam (NIT), Samastipur (SIT), Barkot (BFS), Betul (BFT), Haldwani (HFS), Kosi Watershed (KKM), Kaziranga National Park (KNP), Jaisalmer (JIT), Kanpur (UIT), Pichavaram (PVM), and Sundarbans (SFT). They span different LULC types, as listed in Table 2. EC data processing involves multiple steps, including pre- and post-processing quality control, gap-filling, and flux-partitioning (Figure 2). Instrumentation details and flux estimation procedure are described for the individual sites in Deb Burman et al., (2025). Ancillary meteorological, soil, and plant physiological measurements were made at multiple heights and depths on some of the flux towers and are described in the corresponding references listed in Table 2.

2.2.3 Remote sensing measurements

We use 11 GPP datasets from various remote sensing databases, summarized in Table 3. These data are generated at different spatio-temporal scales.

2.2.3.1 GPP data products

2.2.3.1.1 MODIS GPP

The Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard the Terra and Aqua satellite constellation by NASA provides the longest record of biophysical products, including GPP. MOD17A2H version 6 GPP product (Running and Zhao, 2015), is taken from the Google Earth Engine (GEE) platform (https://earthengine.google.com/platform/). This data has 8-day temporal and 500 m spatial resolutions. The data product includes GPP, Net Photosynthesis (PSN), and Quality Control (QC) layers.

2.2.3.1.2. PML GPP

The Penman-Monteith-Leuning Evapotranspiration (Version V2) product was developed by Zhang et al., (2019), which combined five ready-to-use bands, including GPP at 500 m spatial and 8-day temporal resolutions using a coupled diagnostic biophysical model. This model was built on the GEE platform, using leaf area index (LAI), albedo, and emissivity from MODIS and Global Land Data Assimilation System (GLDAS) meteorological forcing data as model input variables.

2.2.3.1.3. GOSIF GPP

The Global OCO-2 based Solar-Induced Chlorophyll Fluorescence (GOSIF) is derived from the SIF observations by the Orbiting Carbon Observatory-2 (OCO-2). This database has 0.05° spatial and 8-day temporal resolutions and has been accessed from the Global Ecology Data Repository by the Global Ecology Group, University of New Hampshire (Li and Xiao, 2019). All the metadata of the aforementioned databases are given in **Error! Reference source not found.**3.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Table 3. Characteristics of different datasets used for current study.

Name of Dataset	Temporal Resolution	Spatial Resolution	Earth Engine Snippet /Version	Data availability	Algorithm used	Selected Bands	Source / Link to the dataset	Reference
Satellite								
Moderate Resolution Imaging Spectroradiometer	8-day	500 m	(MOD17A2H)	2021-01-01 to present	LUE	Gpp, Psn_QC, PsnNet	GEE link: https://developers.google.com/earth-engine/datasets/catalog/MODIS_061_MOD17A2H Dataset provider: https://lpdaac.usgs.gov/products/mod17a2hv061/	(Running, S., Mu, Q., Zhao, 2021)
Penman-Monteith-Leuning Evapotranspiration V2	8-day	500 m	PML_V2	2000-02-26 to 2020-12-26	PML RS	GPP	GEE link: https://developers.google.com/earth-engine/datasets/catalog/CAS_IGSNRR_PML_V2_v017 Dataset provider: https://github.com/gee-hydro/gee_PML	(Gan et al., 2018; Zhang et al., 2016, 2019)
GLASS	8-day	0.05 degree	--	1982 to 2018			https://glass.hku.hk/	(Yuan et al., 2010)
MUSYQ	8-day	0.05 degree	--	1981 to 2018			https://zenodo.org/records/3996814	(Wang et al., 2021)
Model								
A global dataset of solar-induced chlorophyll fluorescence (GOSIF)	8-day	0.05 degree	v2	2001 to 2023	Cubist regression tree model	GPP	http://data.globalecology.unh.edu/data/GOSIF-GPP_v2/	(Li and Xiao, 2019)
Soil Moisture Active Passive Data (SMAP)	1-day	9 km	Version 4	31 March 2015 to 8 November 2022	L-band microwave observations, land cover and vegetatio	GPP	Data accessed: https://search.earthdata.nasa.gov/search?q=SPL4CMDL%20V007	(Kimball, J. S., 2021)

					n inputs from MODIS, VIIRS, and GEOS-5 land model			
P-model Site-level data	1-day	0.05 degree	Version v1.0	1982-01-06 to 2016-12-27	LUE	GPP	Data: https://zenodo.org/records/1423484 Article: https://gmd.copernicus.org/articles/13/1545/2020/	(Stocker, 2018; Stocker et al., 2020)
Vegetation Optical Depth Climate Archive (VODCA 2GPP)	8-day	0.25 ° degree	V2	1988 to 2020	Utilize microwave remote sensing estimates of vegetation optical depth (VOD) to estimate GPP.	GPP	https://researchdata.tuwien.ac.at/records/1k7aj-bdz35	(Wild et al., 2022b)
Machine learning								
FLUXCOM-RS	8-day	0.0833°	Version 4	2001 to 2019	Used machine learning to merge energy flux measurements from FLUXNET eddy covariance towers with remote sensing and meteorological data to estimate global gridded net radiation, latent and sensible heat and their	GPP	http://www.fluxcom.org	(Jung et al., 2019)

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

					uncertainties			
Center for Global Environmental Research (CGER), National Institute for Environmental Studies (GED)	10-day	0.1 degree	ver.2020.2	1999 to 2019	Upscaling of site observations using Random Forest	GPP	https://db.cger.GED.go.jp/DL/10.17595/20200227.001.html.en	(A Data-driven Upscale Product of Global Gross Primary Production, Net Ecosystem Exchange and Ecosystem Respiration, ver.2020.2)
XBASE	1-day	0.25 degree	--	2001 to 2020	FLUXCOM-X, an advanced deep learning model with enhanced flexibility and technical capabilities	GPP	Data: https://meta.icos-cp.eu/collections/zfwf1Ak2I7OlziGDTX8Xl6_T Article: https://bg.copernicus.org/articles/21/5079/2024/ Integrated Carbon Observation System (ICOS)	(Nelson et al., 2024)

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

2.2.3.1.4. GLASS GPP

An improved version of the LUE model with four input variables: NDVI, PAR, Tair, and the Bowen ratio (the ratio of sensible and latent heat flux), was developed as part of the Global LAnd Surface Satellite (GLASS) project (Yuan et al., 2010). A global GPP product at 0.05° spatial and 8-daily temporal resolution using the input data from AVHRR was produced on DoY 363 in 2019 and disseminated as a Level 3 product, which has been used in this work.

2.2.3.1.5. MuSyQ GPP

Using the FPAR and LAI records from GLASS as inputs, a global GPP product, MuSyQ, was developed at 0.05° spatial and 8-daily temporal resolutions using a LUE approach. However, an improved parameterization was achieved by using the clearness index (CI), ERA-Interim meteorological data, and other environmental factors (Wang et al., 2021).

2.2.4 GPP based on Model simulations

2.2.4.1 Vegetation models

2.2.4.1.1 P-model

The P-model (version 1.0) computes daily GPP for the C3 vegetation in any climate following the combined Farquhar-–von Caemmerer–Berry photosynthesis model (Wang et al., 2017). It predicts a multi-day average LUE, accounts for the low-temperature effects on the intrinsic quantum yield, and has an empirical soil moisture stress factor. P-model incorporates an optimality principle for balancing carbon assimilation and transpiration, powered by meteorological and site-specific fAPAR data, and its output is compared against the GPP estimates obtained from the global Fluxnet network (Stocker et al., 2020). The global and site-scale sensitivity studies are conducted on hydrometeorological and biophysical variables during the parameter calibration. No distinction is made among the croplands, with two fAPAR datasets, having different estimating techniques taken from MODIS and GLASS, used in the model simulation.

2.2.4.2 Coupled models

2.2.4.2.1. SMAP

The Soil Moisture Active Passive (SMAP) is a global GPP database generated by integrating multiple satellite-based observations, such as SMAP L-band microwave observations, land cover, and vegetation inputs from the MODIS, Visible Infrared Imaging Radiometer Suite (VIIRS), and the Goddard Earth Observing System Model, Version 5 (GEOS-5) land model assimilation system. The global plant functional type (PFT) classification data were taken from MODIS (MCD12Q1 version 6) and fAPAR (MOD15A2) at 500 m resolution, SMAP Level-4 soil moisture data (SPL4SMGP) at 9 km resolution, pre-processed global, daily averaged meteorology data from the GEOS-5 Forward Processing (FP) system at 0.25° degree resolution, as mentioned in the use guide of the database. This data is accessed from NASA's Earthdata portal (https://www.earthdata.nasa.gov/).

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

2.2.4.2.2 VODCA2GPP

The Vegetation Optical Depth Climate Archive (VODCA2), developed by Wild et al. (2022), represents the first-ever global database of microwave remote-sensing vegetation optical depth (VOD)- derived GPP at 0.25°×0.25° spatial resolution. It is produced following the novel sink-driven GPP estimation approach by Teubner et al., (2019, 2021), and consequently trained and evaluated with in situ observations at the FLUXNET stations. This GPP data (VODCA2GPP) is accompanied by an uncertainty metric showing regions with less reliable estimates. This dataset is accessed from the TU Data Repository of TU Wien.

2.2.5 Machine Learning-based GPP products

2.2.5.1 FLUXCOM

The FLUXCOM dataset combines Fluxnet EC measurements with satellite data using machine learning algorithms to estimate global GPP (Jung et al., 2019). It provides 147 products at half-hourly and monthly intervals, including high-resolution (0.083°) MODIS remote sensing data and 0.5° meteorological data. The global carbon and water cycle evaluations are improved by the gap-filled flux estimates in the dataset, and validations against independent data demonstrate strong performance across ecosystems. In this work, the FLUXCOM-RS version of the FLUXCOM data has been used.

2.2.5.2 GED

This product comprises a 10-day average of global GPP at a spatial resolution of 0.1°×0.1° developed by upscaling the FLUXNET 2015 database to a global scale using the Random Forest (RF) algorithm, covering latitudes from 60°S to 80°N (A Data-driven Upscale Product of Global Gross Primary Production, Net Ecosystem Exchange and Ecosystem Respiration, ver.2020.2). Daily GPP values from this database were extracted from the GPP_NT_VUT_REF variable. The predictor variables for GPP estimation include LAI, fAPAR, air temperature (T_{air}), relative humidity (rH), and incoming solar radiation (DSR). LAI and fAPAR were sourced from the Copernicus Global Land Service, while T_{air}, rH, and DSR were from ERA5. This approach offers new insights for studying global GPP (Pastorello et al., 2020). The GPP data for the current study were acquired from the Global Environmental Database of the National Institute for Environmental Studies, Japan (https://dx.doi.org/10.17595/20200227.001).

2.2.5.3 X-BASE

X-BASE provides the first globally upscaled GPP product using the gradient-boosted regression trees algorithm at 0.25° spatial and daily temporal resolutions, using the FLUXCOM-X methodology, which is more flexible and adaptive than the previous generations of FLUXCOM; the inputs include EC measurements from the Fluxnet, meteorological variables from the ERA5 reanalysis data, vegetation indices, land surface temperature, and land use records from MODIS (Nelson et al., 2024).

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

2.2.6 Spatial subsetting

Different GPP products have different spatial resolutions, and each was spatially subset within a common 0.1°×0.1° area surrounding each flux tower. Subsequently, they were spatially averaged to create their time series records at each site. Some of the ready-to-use GPP products required additional quality control. For MODIS, the Psn_QC was set to 0 to retain only the highest-quality pixels. A common unit of gC m^{-2} is used for all the GPP products commensurate with the EC GPP record. Figure 2 shows the flowchart of data processing and analysis used in this study.

2.2.7. Statistical analysis

2.2.7.1 Linear regression

EC data are available daily, whereas GPP products listed above have different temporal resolutions as shown in Table 4. The daily EC data are aggregated to match the temporal resolution of the corresponding dataset before linearly regressing at each site to obtain slope, intercept, coefficient of determination (R^2), adjusted coefficient of determination (Adj. R^2), and p-value.

2.2.7.2 Taylor diagram

The Taylor diagram, first proposed by Taylor (2001), is a graphical method that can comprehensively assess the performance of multiple datasets objectively, including various statistical metrics, including correlation, standard deviation, and variance (Gleckler et al., 2008; Taylor et al., 2012; Knutti and Sedláček, 2013). It represents different model values as points on a polar plot where the radial distance from the origin indicates the ratio of their standard deviation to that of the reference dataset, with points close to 1 indicating better agreement in variability. The azimuthal angle from the reference axis refers to the correlation coefficient: smaller angles signify higher correlation. Additionally, the distance from the reference axis quantifies the root mean square error (RMSE) between datasets, providing a measure of overall agreement (Namadi et al., 2022). Data are normalized with their standard deviation (SD), to facilitate a consistent assessment (Taylor, 2001). The standardized records of GPP estimated from EC measurements (GPP_{EC}) records are used as references.

2.2.7.3 Hierarchical clustering

Hierarchical clustering, proposed by Johnson, (1967), is a clustering algorithm used to group similar objects/data based on their characteristics and resemblance to each other. The merging algorithm in hierarchical clustering assesses the similarity between data points within each category by computing distances between them and all other data points. Smaller distances between data points indicate more significant similarity (Siegel and Wagner, 2022). This technique constructs a dendrogram by either iteratively merging smaller clusters into larger ones (agglomerative approach) or recursively splitting larger clusters into smaller ones (divisive approach), providing a detailed visualization of the nested relationships among the data points (Ding and He, 2002; Murtagh, 2015). Clustering starts by computing the Euclidean distance (Danielsson, 1980) between samples X and Y. The two closest classes are merged into a new class, followed by calculating distances between this new class and all other classes. This merging process continues iteratively until all samples are consolidated into one class. In the

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

current study, the hierarchical clustering was initiated using the Ward cluster method (Ward, 1963), and the correlation distance type (Sokal, 1958, 1970) has been deployed to generate the cluster based on the similarity matrix.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Table 4. Descriptive statistical summary for all EC tower sites.

Major LULC Class	**Cropland**					**Forest**					**Grassland**		**Mangroves**	
Entity	**BIT**	**BKC**	**DIT**	**NIT**	**SIT**	**BFS**	**BFT**	**HFS**	**KKM**	**KNP**	**JIT**	**UIT**	**PVM**	**SFT**
Total no. of samples (*N*)	511	388	493	511	505	511	510	510	511	508	504	510	365	511
Minimum	0.14	0.02	0.003	0.31	0.07	1.29	0.27	1.48	0.97	0.16	0.06	0.65	0.10	1.79
Maximum	10.94	7.32	18.30	18.97	29.19	15.33	30.33	14.97	13.42	16.74	11.30	26.74	6.17	7.68
Mean	2.90	1.78	3.62	8.97	7.63	8.68	6.91	7.02	5.74	7.07	2.54	7.29	3.79	4.44
Standard Deviation	1.84	1.76	4.47	3.99	4.52	2.82	6.93	3.22	1.84	3.85	2.02	4.89	0.93	1.11
Median	2.42	1.10	1.77	8.76	7.11	7.74	3.55	6.13	5.58	6.92	1.88	5.88	3.81	4.18
SE of mean	0.08	0.09	0.20	0.18	0.20	0.12	0.31	0.14	0.08	0.17	0.09	0.22	0.05	0.05
Variance	3.38	3.08	19.95	15.89	20.48	7.95	48.01	10.38	3.40	14.84	4.07	23.95	0.87	1.23
Coefficient of Variation	0.63	0.98	1.23	0.44	0.59	0.32	1.00	0.46	0.32	0.54	0.79	0.67	0.25	0.25
Mean absolute Deviation	1.36	1.37	3.20	3.35	3.62	2.39	5.47	2.77	1.45	3.22	1.40	3.87	0.78	0.88
Lower 95% CI of Mean	2.74	1.61	3.22	8.62	7.24	8.44	6.31	6.74	5.58	6.73	2.37	6.87	3.70	4.34
Upper 95% CI of Mean	3.06	1.96	4.01	9.32	8.03	8.93	7.52	7.30	5.90	7.41	2.72	7.72	3.89	4.54

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

2.2.7.4 Statistical metrics for ranking the database

The following statistical metrics were used in this work to assess the performance of the gridded datasets: Mean Absolute Error (MAE) or the mean absolute difference between predicted and observed (EC) GPP values (Müller and Guido, 2016), Root Mean Squared Error (RMSE), (Jiang and Ryu, 2016; Joiner et al., 2018; Kumar et al., 2020; Turner et al., 2003), Pearson's correlation (r), which measures the robustness of the linear relationship between two datasets (Pearson, 1895; Bo et al., 2022; Gao et al., 2021), and the coefficient of determination (R^2), measuring the proportion of variance in the dependent variable, which is predictable from the independent variables in the regression model (Pearson, 1901).

2.2.7.5 Objective ranking of databases

A systematic approach was used to calculate a composite score based on the metrics described above to evaluate and rank the performance of different GPP datasets. First, all these metrics were normalized to ensure inter-comparability. We applied general normalization for r (Pearson_r_norm) and R^2 (R^2_norm), which denotes higher values are better, and reverse normalization for MAE (MAE_norm) and RMSE (RMSE_norm), which suggests that lower values are better.

$$\text{General normalization for r and } R^2 = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

$$\text{Reverse normalization MAE and RMSE} = \frac{X_{min} - X}{X_{max} - X_{min}} \quad (3)$$

Where X is the original, X_{max} is the maximum, and X_{min} is the minimum value of the metric.

After normalizing the individual metrics, we calculated a combined score for each dataset by averaging the normalized values of MAE, RMSE, Pearson, and R^2 as follows:

$$\text{Combined Score} = \frac{\text{MAE\_norm} + \text{RMSE\_norm} + \text{Pearson\_r\_norm} + R^2\text{\_norm}}{4}$$

Various researchers used R^2 to rank different datasets (Coops et al., 2007; Wu et al., 2010; Yuan et al., 2014; Ichii et al., 2017; Yu et al., 2018; Sinha et al., 2021). We ranked the datasets based on their combined score, where a higher score reflects better overall performance. This method has an edge over using any single metric in ensuring the performance assessment encompasses multiple aspects.

3 Results and Discussion

3.1 Site features

We have considered fourteen flux tower sites, of which five belong to cropland and forest, two belong to grassland, and two belong to mangroves. These sites are shown in Figure 1 and detailed in Table 2, including the location, local climate, soil, land use types, and EC instrumentation. For brevity, more site-specific details are provided in the supplementary section S1.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

3.2 Observed GPP variability from eddy covariance measurements

The descriptive statistics of the GPP_{EC} at all sites are mentioned in Table 4 and presented in Figure 3. The key variation features at these sites are described in detail in Deb Burman et al. (2025).

3.2.1 Cropland

Among the cropland sites (Fig. 3), the mean GPP_{EC} measurement at BIT was 2.90 ± 1.84 gC m^{-2} d^{-1}, ranging from 0.14 gC m^{-2} d^{-1} to 10.94 gC m^{-2} d^{-1}. The mean GPP_{EC} at BKC was 1.78 ± 1.76 gC m^{-2} d^{-1}, ranging from 0.02 to 7.32 gC m^{-2} d^{-1}. The mean GPP_{EC} at DIT was 3.62 ± 4.47 gC m^{-2} d^{-1}, which varied from 0.003 to 18.30 gC m^{-2} d^{-1}. At NIT, the mean GPP_{EC} was 8.97 ± 3.99 gC m^{-2} d^{-1}. The mean GPP_{EC} at SIT was 7.63 ± 4.52 gC m^{-2} d$^{-1,}$ varied from 0.07 to 29.19 gC m^{-2} d^{-1} (Fig. 3). Such ranges of variability are atypical for cropland ecosystems (Ukkola et al., 2021; Liu et al., 2023).

3.2.2 Forest

Among the forested sites (Fig. 3), GPP_{EC} ranged from 1.29 to 15.33 gC m^{-2} d^{-1} with a mean of 8.68 ± 2.82 gC m^{-2} d^{-1} at BFS.

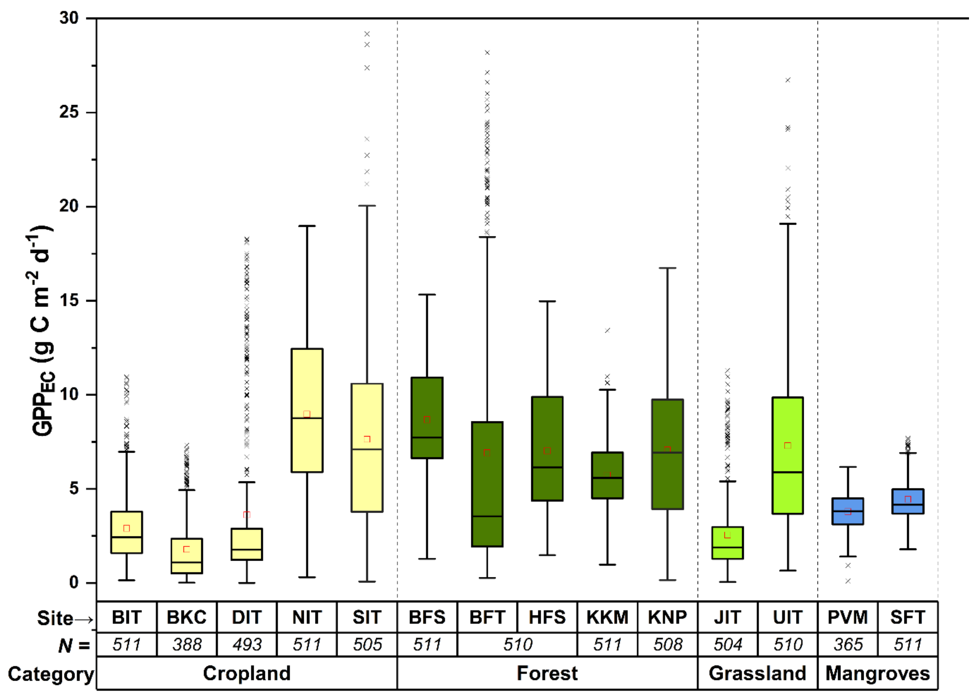


Figure 3. Observed variability of GPP_{EC} at the study sites.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

At BFT, the mean GPP_{EC} was 6.91 ± 6.93 gC m^{-2} d^{-1}, ranging from 0.27 to 30.33 gC m^{-2} d^{-1}; wherein, the GPP was the lowest during the summer and rose quickly after the rainy season due to leaf appearance. The highest GPP was noted in August, which started to decline in September due to a decrease in soil moisture availability for photosynthesis, followed by subsequent temperature and solar radiation declines during the winter season (Rodda et al., 2021). BFT is located in Central India, where Uchale et al., (2023) also noticed the large-scale fluctuations in GPP. This area is positioned within the core monsoon zone, and hence, the biosphere native to this region undergoes strong drying and wetting episodes every year. This, in turn, results in high interannual fluctuation in carbon uptake in this region.

HFS is a mixed plantation with a mean GPP_{EC} of 7.02 ± 3.22 gC m^{-2} d^{-1}, which ranged from 1.48 to 14.97 gC m^{-2} d^{-1}. Watham et al. (2014) highlighted the role of radiation and VPD in regulating the carbon uptake at HFS. The seasonal dynamics were primarily influenced by phenology and environmental factors such as soil moisture and temperature. Sett et al., (2023) estimated water use efficiency (WUE) using GPP and ET for four forest sub-types in HFS: teak, mixed plantation, poplar, and pulpwood.

KKM exhibited a mean GPP_{EC} measurement of 5.74 ±1.84 gC m^{-2} d^{-1}, ranging from 0.97 to 13.42 gC m^{-2} d^{-1}. Deb Burman et al., (2021) studied the ecosystem-atmosphere carbon and water fluxes in the KKM and examined the factors for the variability of GPP values, like surface conductance, along with its intra-annual variability and seasonal cycle of leaf-area, which are the main factors controlling the seasonal patterns of surface-atmosphere coupling and the ecosystem. The KNP GPP_{EC} had a mean value of 7.07 ± 3.85 gC m^{-2} d^{-1} (Fig. 3); the carbon uptake at this site is controlled by differential radiation and water availability at different times of the year (Sarma et al., 2021).

3.2.3 Grassland

Among these sites (Fig. 3), JIT recorded a mean GPP_{EC} of 2.54 ± 2.02 gC m^{-2} d^{-1}. The rainfall variability drives the worst drought conditions and vegetation degradation, especially in arid and semi-arid climates like JIT, where slight departure from the mean may become critical (Santra and Chakraborty, 2012). Here, GPP_{EC} values ranged from 0.06 to 11.30 gC m^{-2} d^{-1}. Lower GPP values may occur during dry periods (Rajan et al., 2013), or after the grazing events, while higher values typically correlate with peak vegetation growth during the monsoon season (Bai and Li, 2022). This variability underscores the rapidly dynamic nature of carbon cycling in arid grassland ecosystems.

The mean GPP_{EC} at UIT was 7.29 ± 4.89 gC m^{-2} d^{-1} (Fig. 3). The moderate variability in carbon fluxes at UIT may be due to outflow from a nearby canal increasing soil moisture (Bhat et al., 2020), which results in maximum carbon assimilation occurring between August and September. The GPP values ranged widely from 0.65 to 26.74 gC m^{-2} d^{-1}. Lower EC values might occur during dry periods or after disturbances, whereas higher values typically coincide with active vegetation growth and favorable weather conditions.

3.2.4 Mangrove

The mean GPP_{EC} at PVM varied from 0.10 to 6.17 gC m^{-2} d^{-1} with a mean of 3.79 ± 0.93 gC m^{-2} d^{-1} (Fig. 3). At SFT, the mean GPP_{EC} is 4.44 ± 1.11 gC m^{-2} d^{-1}. Seasonal environmental changes in temperature, VPD, rainfall

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

and photosynthetic active radiation regulated this variability (Ghosh et al., 2017; Rodda et al., 2022), along with tidal dynamics (Neogi et al., 2016), and ecological factors such as canopy structure and hydrology (Wu et al., 2019) that are characteristic of mangrove ecosystems in the Sundarbans. Overall, the GPP_{EC} values at SFT ranged from 1.79 to 7.68 gC m^{-2} d^{-1} (Fig. 3).

3.3 Performance evaluation of the global GPP datasets

The linear regressions of the remotely sensed, modelled, and ML-developed GPP products with the in-situ measured GPP values at each site are shown in Supplementary Figure S1, with the statistical parameters listed in Table S1. The performance of each dataset is described below, arranged alphabetically.

3.3.1. FLUXCOM-RS

The FLUXCOM-RS database performs well, statistically significant at the 95% level, at most of the sites (Table S1). Its correlation (R^2) values at these sites range from 0.001 to 0.79 (Table S1). However, the mangrove site at SFT is an exception, where the p-value of the FLUXCOM-RS database is 0.70, indicating the lack of any significant relationship with the GPP_{EC}. This is supported by the R^2 value at this site, which is nearly zero, suggesting the model does not explain any of the variability in the data. The FLUXCOM-RS GPP has been widely used in global and regional climate and carbon cycle studies (Jung et al., 2020). However, significant uncertainties persist in this database over mangroves.

3.3.2. GED

The GED dataset for cropland and forest sites shows statistically significant relationships (p-value <0.05) (Table S1), with strong positive slopes between GED and GPP_{EC}, affirming this dataset's ability to reliably estimate GPP in grassland environments. In contrast to this, the GED dataset does not perform satisfactorily at the mangrove sites (p-value > 0.05), highlighting the requirement for alternative approaches or supplementary data sources to improve the GPP estimation of mangroves. According to Zeng et al., (2020), the GED database relies on the Random Forest (RF) method and has trouble handling undersampled areas; this could potentially explain its poor performance over mangroves, which are mostly located in tropical and subtropical belts and have been historically under-explored compared to the other vegetation classes.

3.3.3. GOSIF

The GOSIF provides proxies for monitoring land photosynthesis, but its sparse coverage remains a bottleneck for mapping finer-resolution GPP globally (Li and Xiao, 2019). The GOSIF database performed significantly well at BIT, BKC, and DIT, with R^2 values ranging from 0.34 to 0.61 (Table S1). Though the relationship is statistically significant at the site NIT, the R^2 value is relatively low (0.20), suggesting the GOSIF database captures some variations in GPP values, with some other factors being unaccounted for in its computation. At SIT, the performance is not statistically significant (p-value = 0.09), with a low R^2 value (0.04), suggesting the weak performance of this database in the croplands.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

3.3.4. MODIS

The MODIS performed well at all cropland sites, exhibiting statistically significant relations between GPP_{EC} and MODIS-derived GPP ($p < 0.05$) (Table S1). However, low slope values suggest MODIS's limited sensitivity to GPP trends, with low R^2 values reflecting that MODIS GPP accounts for only a small portion of the variability in cropland ecosystems. In forests, the dataset generally shows significant correlations with GPP ($p < 0.05$), except for KNP, where a p-value of 0.98 indicates a weak correlation (Table S1).

Over grasslands and mangroves, MODIS performed with mixed results; UIT and PVM show significant relationships ($p < 0.05$), while JIT and SFT do not ($p = 0.52$ and 0.34, respectively) (Table S1). These highlight location-specific limitations in MODIS's productivity calculation (Table S1). Our findings are supported by Lele et al., (2021), wherein the spatial resolution of MODIS may not be sufficiently high to capture the heterogeneous spectral responses of mangroves. MODIS also struggles to capture vegetation dynamics post-monsoon, particularly in forest sites like KNP (Deb Burman et al., 2017). Wang et al., (2017) noted that though the latest MODIS GPP product improved significantly in its resolution, there are several other sources of error related to the MODIS algorithm, which impairs the GPP estimation.

3.3.5. PML_V2

The PML_V2 GPP exhibits high adjusted R^2 values (0.57 and 0.61) with significant p-values (<0.05) at BIT and DIT croplands (Table S1). At NIT and SIT, the R^2 values are lower (0.12–0.13), reflecting modest explanatory power. The PML_V2 GPP record has an insignificant p-value (0.28) and a low R^2 (0.03) at BKC, indicating poor performance. Among the forests, it performs well at HFS and BFT (adjusted R^2: 0.56), but worse at KKM and KNP (lower R^2 values of 0.11 and 0.13, respectively), although significant ($p<0.05$). Its correlation with EC GPP is weak but significant in grasslands (JIT, UIT) (adjusted R^2: 0.07; p-value <0.05). Over mangroves, PML has mixed performance – moderate at PVM (R^2: 0.23, p-value < 0.05), while poor at SFT (R^2: 0.00, p-value: 0.84) (Table S1). However, Zhang et al., (2019) noted this data to perform well across 95 global flux sites, aligning closely with observed GPP_{EC}. Hence, our findings underscore the need for further calibration of PML_V2, especially for grassland and mangrove ecosystems, with a location-specific approach.

3.3.6. P-model

Regression analysis of P-model data shows varied performance across sites. At KKM (forest), a significant relationship ($p < 0.05$) is observed with a low adjusted R^2 of 0.04, indicating limited performance with observed variability (Table S1). Conversely, BFS and SFT show non-significant p-values (0.39 and 0.13, respectively) and near-zero adjusted R^2 values, reflecting the model's poor performance in capturing GPP variability at these sites, which limited effectiveness.

3.3.7. SMAP

The regression analysis of GPP_{EC} vs. SMAP GPP shows mixed results. It performs well at the croplands with robust and significant fits (Table S1), except at SIT. In forest sites, it is mostly significant with high to moderate adjusted R^2 values (Table S1). SMAP GPP performs significantly at both the grasslands; however, at

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

JIT, it demonstrates a weaker correlation with the EC GPP (Table S1). It performs fairly well at PVM but very poorly at the other mangrove at SFT with a non-significant p-value (0.96).

3.3.8. VODCA2GPP

VODCA2GPP performed moderately but significantly well at the cropland sites, except SIT, with a non-significant p-value of 0.58 and a low adjusted R^2 (-0.01). In BFS and BFT forest sites, it showed strong performance (adjusted R^2 values being 0.42 and 0.70, respectively). At the same time, it recorded a low and insignificant performance (adjusted R^2: -0.011, p-value = 0.88) at JIT, making its performance unacceptable over the grasslands (Table S1).

3.3.9. GLASS

The overall performance of GLASS was good, except SIT and SFT, with non-significant p-values of 0.37 and 0.49, respectively (Table S1). Over DIT, BFT, and HFS, it showed high adjusted R^2 values of 0.629, 0.737, and 0.571, respectively (Table S1).

3.3.10. MuSyQ

MuSyQ showed non-significant correlations at SIT (p-value = 0.3), JIT (p-value = 0.79), and PVM (p-value = 0.16), while its performance at other sites has been acceptable (Table S1).

3.3.11. XBASE

XBASE showed significant correlations at all the sites, and the adjusted R^2 values have also been high; however, it showed low adjusted R^2 values at BKC, NIT, SIT, UIT, PVM, and SFT (Table S1).

3.4 Model Performance based on Pearson's correlation

In the correlation heatmap plot (Figure S2), the colour intensity (red-to-blue) indicates the strength of the correlation coefficient. This helps in understanding the strength and direction of relationships. Each cell in the heatmap represents the correlation between two variables. The values in the cells are annotated to show the exact correlation coefficients, which range from -1 to 1. A value of 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, and 0 indicates no correlation.

3.5 Model performance based on the Taylor diagram

At the BIT cropland site, FLUXCOM-RS GPP correlated significantly with GPP_{EC} (r = 0.86, RMSE = 0.61 gC m^{-2} d^{-1} and SD = 1.22 gC m^{-2} d^{-1}); GOSIF GPP had the lowest RMSE (0.57 gC m^{-2} d^{-1}) and a strong correlation (r = 0.82) (Fig. 4a). The RMSE and SD of VODCA2GPP with GPP_{EC} at this site were 1.28 gC m^{-2} d^{-1} and 1.92 gC m^{-2} d^{-1}, respectively. GOSIF performed best at BKC, with the lowest RMSE (0.78 gC m^{-2} d^{-1}) and a strong correlation (r = 0.68), followed by VIIRS (RMSE = 0.89 gC m^{-2} d^{-1} and r = 0.64) (Fig. 4b). The maximum RMSE at BKC (2.36 gC m^{-2} d^{-1}) was registered by VODCA2GPP, which even had a negative correlation (r = -0.37), highlighting its poor performance. Here, MODIS GPP had the lowest SD (0.01 gC m^{-2} d^{-1}) but a marginal RMSE (1.0 gC m^{-2} d^{-1}) and correlation (r = 0.53).

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

At DIT, FLUXCOM-RS performed well, with an RMSE of 0.47 gC m^{-2} d^{-1} and a significant correlation (r = 0.88) with GPP_{EC}. The lowest SD at this site was recorded by MODIS (±0.01 gC m^{-2} d^{-1}) with moderate correlation (r = 0.84); r was highest for GED (0.76) (Fig. 4c). SMAP GPP did not perform well at this site, with both the highest SD and RMSE (2.72 and 2.02 gC m^{-2} d^{-1}, respectively). Similarly, at NIT, FLUXCOM-RS had a significant r of 0.70 with an RMSE of 1.1 gC m^{-2} d^{-1} (Fig. 4d).

A moderate response regarding the correlation coefficients has been noticed at SIT. The highest response (not significant) was r = 0.54, with the FLUXCOM-RS having RMSE = 0.99 gC m^{-2} d^{-1}. The lowest and highest RMSE at SIT were recorded for NOAA (0.92 gC m^{-2} d^{-1}) and VODCA2GPP (2.74 gC m^{-2} d^{-1}), respectively (Fig. 4e). FLUXCOM-RS consistently shows strong correlations and low RMSE at BIT and DIT, while GOSIF exhibits good performance at BKC and NIT, respectively. However, VODCA2GPP and SMAP display high RMSE and poor correlation at several sites, indicating that not all databases are equally reliable for validation in these regions.

Among the forests, GED performed best at BFS, with the highest r of 0.94, despite having a moderate RMSE (0.97 gC m^{-2} d^{-1}) and a higher SD (±1.88 gC m^{-2} d^{-1}) (Fig. 5a). FLUXCOM-RS exhibited a high r (0.89), a moderate SD (1.19 gC m^{-2} d^{-1}), and the lowest RMSE (0.53 gC m^{-2} d^{-1}) here. At BFT, the GOSIF demonstrates the best performance with a high r (0.94), a relatively moderate SD (1.13 gC m^{-2} d^{-1}), and a low RMSE (0.37 gC m^{-2} d^{-1}) (Fig. 5b). VODCA2GPP showed a strong r (0.90) with GPP_{EC}. However, it had a large SD and RMSE (3.0 gC m^{-2} d^{-1} and 2.09 gC m^{-2} d^{-1}, respectively) (Fig. 5b).

FLUXCOM-RS performs best at HFS with a high r (0.94), moderate SD (1.89 gC m^{-2} d^{-1}), and RMSE (0.99 gC m^{-2} d^{-1}). indicating substantial agreement with the EC data (Fig. 5c). In contrast, VODCA2GPP had the poorest performance, with a weak r (0.14), coupled with both the highest SD and RMSE (2.35 gC m^{-2} d^{-1} and 2.38 gC m^{-2} d^{-1}, respectively). At KKM, GED demonstrates the best performance among the datasets evaluated, with a substantial r (0.76), a relatively low SD (0.92 gC m^{-2} d^{-1}), and RMSE (0.67 gC m^{-2} d^{-1}) (Fig. 5d). At KNP, GOSIF was the top performer, having a significant r (0.84) with the lowest RMSE and moderate SD (0.56 gC m^{-2} d^{-1} and 0.64 gC m^{-2} d^{-1}, respectively) (Fig. 5d). GOSIF was seconded by SMAP at KNP with a high r (0.80), moderate RMSE and SD (0.71 gC m^{-2} d^{-1} and ±1.20 gC m^{-2} d^{-1}, respectively). The VODCA2GPP had a negative correlation r (-0.65) with GPP_{EC}, which was associated with the highest SD and RMSE (1.74 gC m^{-2} d^{-1} and 2.45 gC m^{-2} d^{-1}, respectively).

At the grassland site at JIT, GED performed best with a high correlation (r = 0.69), low SD, and RMSE (0.47 gC m^{-2} d^{-1} and 0.76 gC m^{-2} d^{-1}) (Fig. 6a). VODCA2GPP had a poor performance here with a low r but high SD and RMSE (1.13 gC m^{-2} d^{-1} and 1.49 gC m^{-2} d^{-1}). Whereas, at the other grassland UIT, GOSIF perfromed best with r ~ 0.75, SD ~ 0.84 gC m^{-2} d^{-1}, and RMSE of ~ 0.7 gC m^{-2} d^{-1} (Fig. 6b). The enhanced performance of GOSIF over grasslands is supported by Pandey et al. (2024).

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

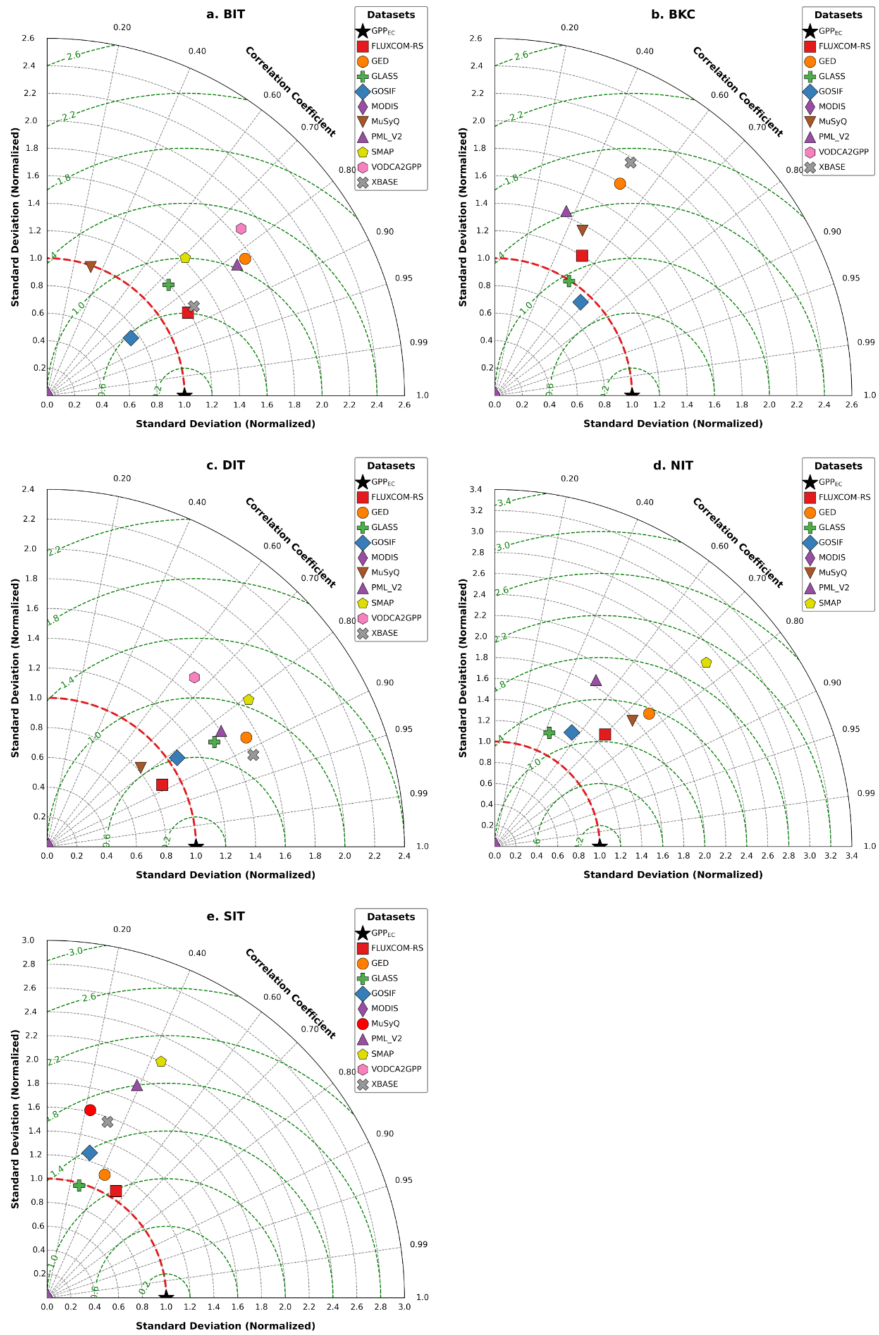


Figure 4. Taylor diagrams of GPP from gridded datasets at the cropland sites used in this work, namely (a) BIT, (b) BKC, (c) DIT, (d) NIT and (e) SIT. GPP_{EC} data is used as the reference.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

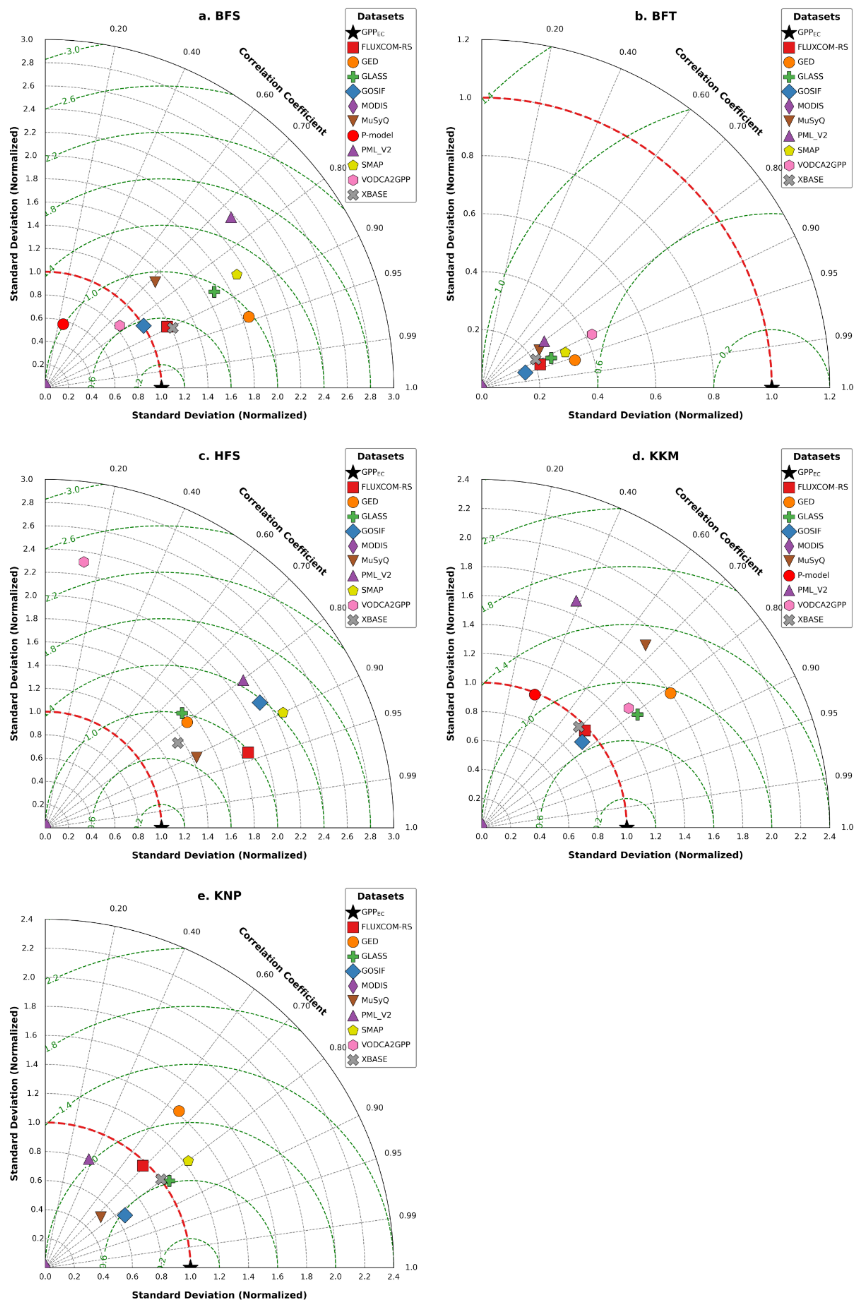


Figure 5. Taylor diagrams of GPP from gridded datasets at the forest sites used in this work, namely (a) BFS, (b) BFT, (c) HFS, (d) KKM, and (e) KNP. GPP$_{EC}$ data is used as the reference.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

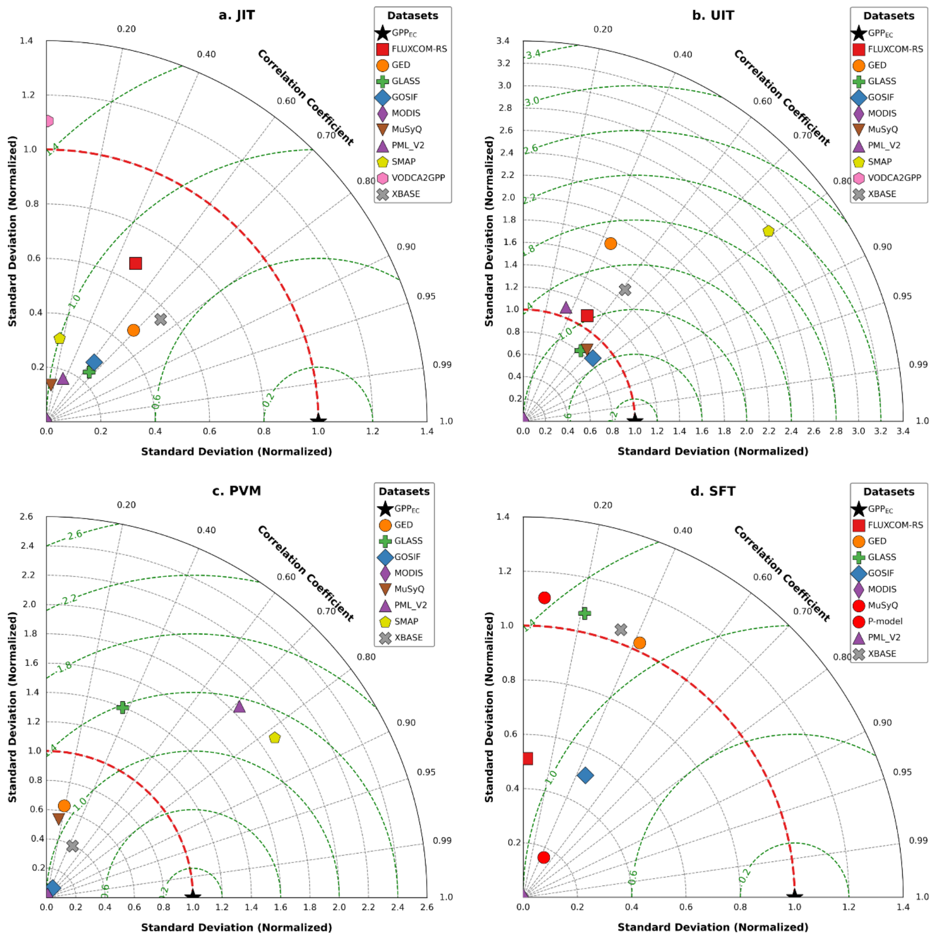


Figure 6. Taylor diagrams of GPP from gridded datasets at the grassland and mangrove sites used in this work, namely (a) JIT, (b) UIT, (c) PVM, and (d) SFT. GPP_{EC} data is used as the reference.

At the PVM mangrove, SMAP performed well with a high correlation (r = 0.82), a relatively higher SD, and RMSE; NOAA performed poorly with a negative correlation (r = -0.18) and moderate RMSE (1.00 gC m^{-2} d^{-1}) (Fig. 6a). At SFT, GOSIF performed well based on the highest r (0.45), relatively low SD (0.51 gC m^{-2} d^{-1}), and RMSE (0.89 gC m^{-2} d^{-1}) (Fig. 6b).

3.6 Clustering of the global GPP datasets

The circular cluster dendrogram for croplands (Fig. 7) shows that at BIT, the EC data are clustered with GLASS, FLUXCOM-RS, GED, GOSIF, XBASE, and PML_V2, indicating a similar pattern in GPP estimation. In contrast, MODIS, VODCA2GPP, and SMAP are relegated to a separate cluster, suggesting a moderate deviation (Fig. 7a). MuSyQ forms the third cluster, indicating a very distinct GPP estimation approach (Fig. 7a). At BKC,

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

none of the gridded GPP data clusters with the EC data, indicating that none of the datasets closely replicate the GPP measurement here (Fig. 7b). The second cluster contains GLASS, PML_V2, MuSyQ, XBASE, FLUXCOM-RS, GED, GOSIF, and MODIS, suggesting that these datasets have similar GPP estimation patterns among themselves, which, however, diverges from the site measurements. VODCA2GPP is placed in the third cluster here, showing a moderate association with EC (Fig. 7b).

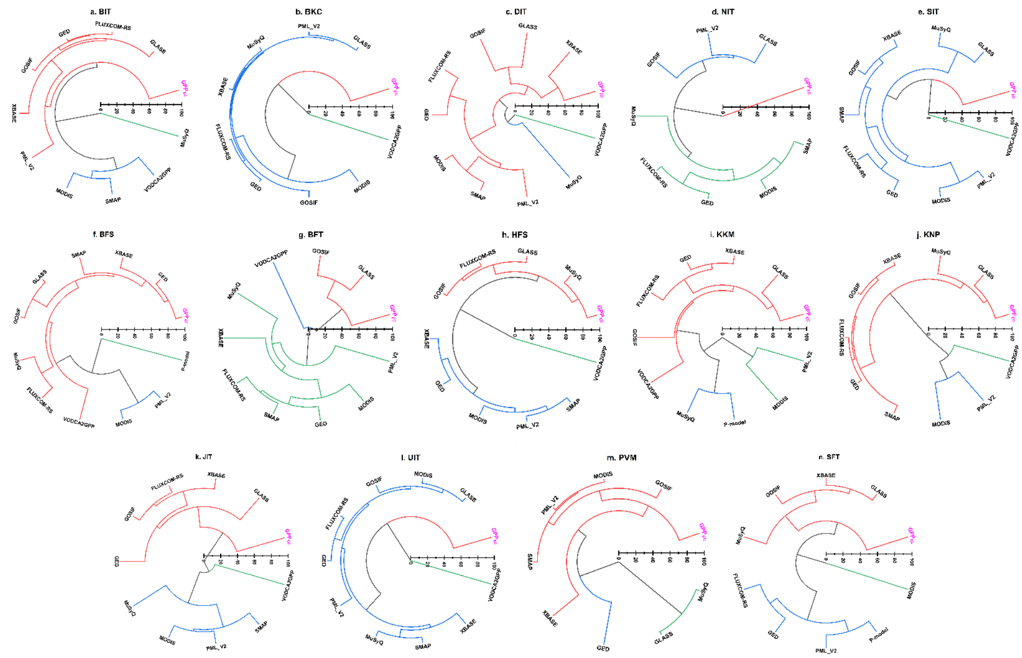
At DIT, the first cluster includes XBASE, GLASS, GED, MODIS, SMAP, and PML_V2, with EC, whereas MuSyQ, and VODCA2GPP form the second and third clusters, respectively, indicating lowered association with EC (Fig. 7c). At NIT, none of the gridded GPP datais in the same cluster as EC, like BKC (Fig. 7d). At SIT, three clusters are formed, with the first one including only EC, while GLASS, MuSyQ, XBASE, GOSIF, XMAP, FLUXCOM-RS, GED, MODIS, and PML_V2 form the second cluster; VODCA2GPP is the sole member of the third cluster here (Fig. 7e).

GED, XBASE, SMAP, GLASS, GOSIF, MuSyQ, FLUXCOM-RS, VODCA2GPP, MODIS, and PML_V2 form cluster 1 at BFS, with EC being most closely aligned with GED (Fig. 7f). This association stems from a similar GPP estimation method implemented in both approaches, particularly the nighttime flux-partitioning (Reichstein et al., 2005). P_Model forms the other cluster at this site (Fig. 7f). At BFT, GLASS and GOSIF form the first cluster with EC, whereas VODCA2GPP forms the second, with rest all being relegated to the third cluster (Fig. 7g). At HFS, MuSyQ, GLASS, FLUXCOM-RS, and GOSIF form the first cluster with EC; XBASE, GED, MODIS, PML_V2, and SMAP form the second cluster and VODCA2GPP is relegated to the final cluster (Fig. 7h). At KKM, the primary cluster includes GLASS, XBASE, GED, FLUXCOM-RS, GOSIF, and VODCA2GPP, with MuSyQ and P_model forming the next cluster; MODIS and PML_V2 form the last cluster (Fig. 7i). At KNP, GLASS, MuSyQ, XBASE, GOSIF, FLUXCOM-RS, GED, and SMAP form the first cluster with EC; MuSyQ, XBASE, GED, and SMAP are members of the same cluster; MODIS and PML_V2A form the second cluster, whereas VODCA2GPP is relegated to the most distant cluster (Fig. 7j).

Among the grassland sites, GLASS, XBASE, FLUXCOM-RS, GOSIF, and GED form the primary cluster at JIT (Fig. 7k). The second cluster here includes MUSYQ, MODIS, PML_V2, and SMAP, while the third cluster contains only VODCA2GPP (Fig. 7k). On contrary to this, at another grassland site, UIT, none of the datasets is in the first cluster with EC; GLASS, XBASE, FLUXCOM-RS, and GED belong to the next cluster; cluster 3 here includes only VODCA2GPP (Fig. 7l).

Cluster 1 at the PVM mangrove included GOSIF, MODIS, PML_V2, SMAP, and XBASE; GED forms the second cluster here whereas, GLASS and MUSYQ form the last cluster (Fig. 7m). At SFT, GLASS, XBASE, GOSIF, and MUSYQ form the first cluster containing EC; FLUXCOM-RS, GED, PML_V2, and P_Model form the next cluster; MODIS forms the third cluster here (Fig. 7n). EC data stand for the ground truth of measurement and, by construction, falls in the first cluster at any site. If EC data are unavailable at any site, datasets in the first cluster can be used as the best approximation, followed by those in the second cluster, and so on.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026



550

Figure 7. Hierarchical cluster analysis of the gridded datasets at all the sites used in this work, namely (a) BIT, (b) BKC, (c) DIT, (d) NIT, (e) SIT, (f) BFS, (g) BFT, (h) HFS, (i) KKM, (j) KNP, (k) JIT, (l) UIT, (m) PVM, and (n) SFT. The distance from the origin of the plot to the end of the circle represents the similarity difference between databases) Each coloured line in the plot represents the number of clusters formed.

3.7 Overall ranking of the GPP databases

All the aforementioned global GPP databases are objectively ranked, as summarized in Figure 8. According to this ranking, Fluxcom consistently performed well across most sites, ranking first at BIT, DIT, and NIT, and maintaining strong performance at other sites, ranking from second to seventh. XBASE also ranked first at three forest sites (BFS, HFS, and KKM), whereas it ranked from second to eighth at other sites (Fig. 8).

The GED dataset also demonstrated robust performance, ranking first at JIT, and second at NIT, BFS, and BIT (Fig. 8). Its performance was moderately good at other sites, with ranks ranging from second to sixth. GOSIF showed variability in its performance but excelled at BKC and BKC, where it secured the top rank (Fig. 8). At other sites, its ranks ranged from second to eighth, except at KNP, where it was tenth. GLASS ranked between second and eighth at all the sites, indicating a moderately good performance (Fig. 8).

The MODIS ranked first in SIT; however, it showed a mixed performance with higher ranks (indicating lower performance) at other sites (Fig. 8). The PML_V2 ranked second at PVM, but did not show remarkable performance at the remaining sites (Fig. 8). Similarly, MuSyQ ranked first at KNP and second at UIT, without noteworthy performance at most of the sites (Fig. 8). SMAP achieved the top rank at PVM; however, its performance at other sites was lower (Fig. 8). VODCA2GPP did not perform well (Fig. 8).

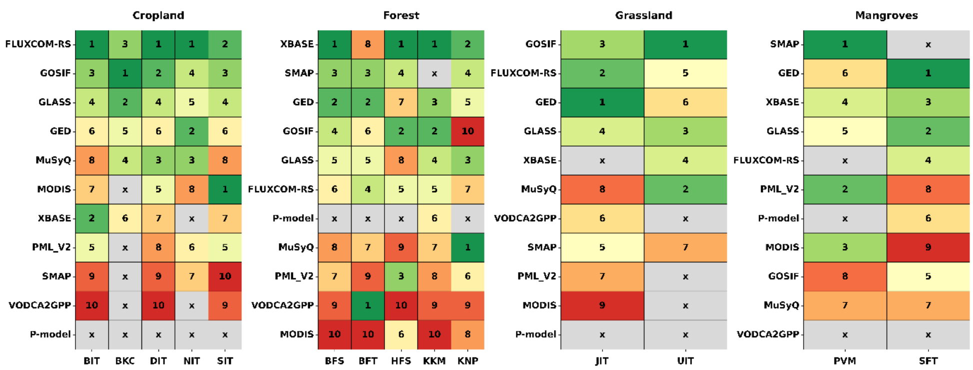


Figure 8. Site-wise ranking of the gridded GPP databases, based on MAE, RMSE, Pearson's correlation, and R^2. The number in each cell represents the rank of the dataset at each site, and the colour shows the rank position (lowered from green to red).

4. Conclusion

The present study objectively ranked the performance of global GPP product measurements for the first time in the Indian subcontinent, following a comprehensive performance assessment against their ground measurements at several flux towers representing four of India's major land cover types. The EC GPP measurements across these sites revealed substantial variability in carbon fluxes influenced by crop growth stages, forest composition, climatic conditions, and agricultural practices (Deb Burman et al., 2025), underscoring the importance of recognizing and supervising carbon dynamics in diverse ecosystems for sustainable agriculture,

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

effective carbon sequestration, and climate change mitigation strategies. The performance evaluation of global GPP datasets following our multi-tiered approach reveals significant variability in the accuracy of different models across various ecosystems. Models like FLUXCOM-RS and XBASE demonstrate strong correlations with EC-derived GPP in cropland and forest sites, but struggle in mangrove and some grassland regions. The results highlight the need for further refinement and the potential use of machine learning to enhance GPP estimation, especially in biologically complex environments like mangroves.

The Taylor diagram analysis reveals varying performance among GPP databases across different sites. FLUXCOM-RS and GOSIF demonstrate consistently strong performance with high correlations and low RMSE at multiple sites, making them reliable choices for GPP modelling. Conversely, VODCA2GPP generally showed weaker performance, with high RMSE and low correlation, indicating less reliability for accurate GPP estimation in the evaluated regions. The cluster analysis reveals that various datasets exhibit different patterns in GPP estimation across sites. Notably, GOSIF strongly aligns with EC data at mangrove and forest sites, suggesting their effectiveness as alternatives when ground-based measurements are unavailable. In contrast, MODIS displays unique patterns, indicating the need for careful selection based on site-specific data characteristics.

Overall, FLUXCOM-RS is the best dataset for cropland sites due to its consistently high performance across multiple evaluation metrics, including low error rates and strong correlation with observed data, whereas XBASE retains this distinction for the forests. GOSIF does the overall best job for the grasslands.

5. Data availability

All the gridded data used in this work are available open-access and have been detailed in the text. The EC GPP data at KNP are taken from Deb Burman et al., (2024) which are freely available at https://dx.doi.org/10.5281/zenodo.10102999. The EC GPP data at PVM are taken from Gnanamoorthy et al., (2020) and are freely available at https://data.mendeley.com/datase ts/vwbn2gcgv3/1. The EC GPP data at DIT, BIT and UIT are available open access at https://doi.org/10.5285/78c64025-1f8d-431c-bdeb-e69a5877d2ed and https://doi.org/10.5285/c5e72461-c61f-4800-8bbf-95c85f74c416. The EC GPP data at BIT and SIT are available on request from Suraj Reddy Rodda (surajreddy_r@nrsc.gov.in). Ther EC GPP data at KKM are available on request from Sandipan Mukherjee ((andipan@gbpihed.nic.in) upon approval of the competent authority at GB Pant National Institute of Himalayan Environment. The EC GPP data at BFS and HFS are taken from Watham et al., (2020). The EC GPP data at BKC, NIT, SIT, and JIT are available on request from Pramit Kumar Deb Burman (pramit.cat@tropmet.res.in).

6. Author contributions

Conceptualization: PKDB. Data curation: PKDB, SRR, SM, DS, NG, GSB. Formal analysis: PKDB, AK. Funding acquisition: PKDB, YKT. Investigation: PKDB, AK, AJ, PD. Methodology: PKDB, AK. Project administration: PKDB, YKT, NG. Software: AK, PKDB, PD. Supervision: PKDB. Validation: AJ. Visualization: AK, PKDB. Writing – original draft: PKDB, AK. Writing – review and editing: SBR, AJ, GSB, KI.

7. Competing interests

The authors declare no competing interest.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

8. Acknowledgements

We are grateful to the Director of IITM, Pune, for his all-possible support and encouragement for the MetFlux-India project. IITM is an autonomous research institute fully funded by MoES, the Government of India. INCOMPASS was a collaborative Indo-UK research initiative focused on the monsoon, jointly funded by MoES in India under the Monsoon Mission and the Natural Environment Research Council (NERC) in the UK. We acknowledge the Directors of GBPNIHE, Almora, and CSIR-4PI (formerly CSIR-CMMACS), Bangalore, for setting up the Kosi Watershed flux tower and providing access to its data for research purposes. We also express our sincere gratitude to the Directors of ISRO Space Applications Centre (SAC), Ahmedabad, and the Central Arid Zone Research Institute, Jodhpur, as well as the Vice-Chancellors of Dr. Rajendra Prasad Central Agricultural University (RPCAU), Pusa, and Anand Agricultural University for their support. We also extend our appreciation to the forest departments of Assam and Tamil Nadu for granting permission and providing logistical support for conducting the research. The Sundarbans and Betul flux towers were funded through the ISRO Geosphere-Biosphere Programme. We thank Martin Jung and Jacob Nelson at the Max Planck Institute of Biogeochemistry, Jena, Germany for their assistance with the FLUXCOM-RS and XBASE data.

9. Financial support

No funding was received for this study.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

References

Agarwal, G., Burman, P. K. D., Kosamkar, P. K., and Kulkarni, V. Y.: Predicting Gross Primary Productivity of the Forest Ecosystems using Machine Learning Techniques: A Review of Existing Approaches, IOP Conf. Ser. Earth Environ. Sci., 1285, 012014, https://doi.org/10.1088/1755-1315/1285/1/012014, 2024.

Aubinet, M., Vesala, T., and Papale, D.: Eddy covariance: a practical guide to measurement and data analysis, Springer Science & Business Media, Dordrecht Netherlands, 2012.

Bai, Y. and Li, S.: Growth peak of vegetation and its response to drought on the Mongolian Plateau, Ecol. Indic., 141, 109150, https://doi.org/10.1016/j.ecolind.2022.109150, 2022.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, U. K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, Bull. Am. Meteorol. Soc., 82, 2415–2434, https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001.

Ball, J. T., Woodrow, I. E., and Berry, J. A.: A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis under Different Environmental Conditions, Prog. Photosynth. Res., 221–224, https://doi.org/citeulike-article-id:8423355, 1987.

Bandopadhyay, S., Pal, L., and Das, R. D.: Predicting gross primary productivity and PsnNet over a mixed ecosystem under tropical seasonal variability: a comparative study between different machine learning models and correlation-based statistical approaches, J. Appl. Remote Sens., 15, 14523, https://doi.org/10.1117/1.jrs.15.014523, 2021.

Bannari, A., Morin, D., Bonn, F., and Huete, A. R.: A review of vegetation indices, Remote Sens. Rev., 13, 95–120, https://doi.org/10.1080/02757259509532298, 1995.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate, Science (80-.)., 329, 834–838, https://doi.org/10.1126/science.1184984, 2010.

Beringer, J., Hutley, L. B., McHugh, I., Arndt, S. K., Campbell, D., Cleugh, H. A., Cleverly, J., De Dios, V. R., Eamus, D., Evans, B., Ewenz, C., Grace, P., Griebel, A., Haverd, V., Hinko-Najera, N., Huete, A., Isaac, P., Kanniah, K., Leuning, R., Liddell, M. J., MacFarlane, C., Meyer, W., Moore, C., Pendall, E., Phillips, A., Phillips, R. L., Prober, S. M., Restrepo-Coupe, N., Rutledge, S., Schroder, I., Silberstein, R., Southall, P., Sun Yee, M., Tapper, N. J., Van Gorsel, E., Vote, C., Walker, J., and Wardlaw, T.: An introduction to the Australian and New Zealand flux tower network - OzFlux, Biogeosciences, 13, https://doi.org/10.5194/bg-13-5895-2016, 2016.

Bhat, G. S., Morrison, R., Taylor, C. M., Bhattacharya, B. K., Paleri, S., Desai, D., Evans, J. G., Pattnaik, S., Sekhar, M., Nigam, R., Sattar, A., Angadi, S. S., Kacha, D., Patidar, A., Tripathi, S. N., Krishnan, K. V. M., and Sisodiya, A.: Spatial and temporal variability in energy and water vapour fluxes observed at seven sites on the Indian subcontinent during 2017, Q. J. R. Meteorol. Soc., 146, 2853–2866, https://doi.org/10.1002/qj.3688, 2020.

Bloomfield, K. J., Stocker, B. D., Keenan, T. F., and Prentice, I. C.: Environmental controls on the light use efficiency of terrestrial gross primary production, Glob. Chang. Biol., 29, 1037–1053, https://doi.org/https://doi.org/10.1111/gcb.16511, 2023.

Bo, Y., Li, X., Liu, K., Wang, S., Zhang, H., Gao, X., and Zhang, X.: Three Decades of Gross Primary Production (GPP) in China: Variations, Trends, Attributions, and Prediction Inferred from Multiple Datasets and Time Series Modeling, https://doi.org/10.3390/rs14112564, 2022.

Chapin, F. S., Woodwell, G. M., Randerson, J. T., Rastetter, E. B., Lovett, G. M., Baldocchi, D. D., Clark, D. A., Harmon, M. E., Schimel, D. S., Valentini, R., Wirth, C., Aber, J. D., Cole, J. J., Goulden, M. L., Harden, J. W., Heimann, M., Howarth, R. W., Matson, P. A., McGuire, A. D., Melillo, J. M., Mooney, H. A., Neff, J. C.,

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Houghton, R. A., Pace, M. L., Ryan, M. G., Running, S. W., Sala, O. E., Schlesinger, W. H., and Schulze, E. D.: Reconciling carbon-cycle concepts, terminology, and methods, https://doi.org/10.1007/s10021-005-0105-7, 2006.

Coops, N. C., Black, T. A., Jassal, R. P. S., and Trofymow, J. T., Morgenstern, K.: Comparison of MODIS, eddy covariance determined and physiologically modelled gross primary production (GPP) in a Douglas-fir forest stand, Remote Sens. Environ., 107, 385–401, https://doi.org/10.1016/j.rse.2006.09.010, 2007a.

Coops, N. C., Black, T. A., Jassal, R. (Paul) S., Trofymow, J. A. (Tony., and Morgenstern, K.: Comparison of MODIS, eddy covariance determined and physiologically modelled gross primary production (GPP) in a Douglas-fir forest stand, Remote Sens. Environ., 107, 385–401, https://doi.org/10.1016/j.rse.2006.09.010, 2007b.

Danielsson, P. E.: Euclidean distance mapping, Comput. Graph. Image Process., 14, 227–248, https://doi.org/10.1016/0146-664X(80)90054-4, 1980.

Deb Burman, P. K., Sarma, D., Williams, M., Karipot, A., and Chakraborty, S.: Estimating Gross Primary Productivity of a tropical forest ecosystem over north-east India using LAI and meteorological variables, J. Earth Syst. Sci., 126, 1–16, https://doi.org/10.1007/s12040-017-0874-3, 2017.

Deb Burman, P. K., Shurpali, N. J., Chowdhuri, S., Karipot, A., Chakraborty, S., Lind, S. E., Martikainen, P. J., Chellappan, S., Arola, A., Tiwari, Y. K., Murugavel, P., Gurnule, D., Todekar, K., and Prabha, T. V: Eddy covariance measurements of CO2 exchange from agro-ecosystems located in subtropical (India) and boreal (Finland) climatic conditions, J. Earth Syst. Sci., 129, 43, https://doi.org/10.1007/s12040-019-1305-4, 2020.

Deb Burman, P. K., Launiainen, S., Mukherjee, S., Chakraborty, S., Gogoi, N., Murkute, C., Lohani, P., Sarma, D., and Kumar, K.: Ecosystem-atmosphere carbon and water exchanges of subtropical evergreen and deciduous forests in India, For. Ecol. Manage., 495, 119371, https://doi.org/10.1016/j.foreco.2021.119371, 2021.

Deb Burman, P. K., Bhat, G., Tiwari, Y. K., Morrison, R., Reddy Rodda, S., Mukherjee, S., Dadhwal, V., Turner, A. G., Das, P., Agarwal, G., Sarma, D., Mutyala, P., Gogoi, N., Gnanamoorthy, P., Paleri, S., and Desai, D.: Ecosystem-atmosphere exchanges of carbon dioxide, water vapour and energy in India: A synthesis of insights from eddy covariance measurements, Agric. For. Meteorol., 372, 110730, https://doi.org/10.1016/j.agrformet.2025.110730, 2025.

Ding, C. and He, X.: Cluster Merging and Splitting in Hierarchical Clustering Algorithms, 139–146 pp., https://doi.org/10.1109/ICDM.2002.1183896, 2002.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, 1980.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J. E., Toon, G. C., Butz, A., Jung, M., Kuze, A., and Yokota, T.: New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity, Geophys. Res. Lett., 38, 1–6, https://doi.org/10.1029/2011GL048738, 2011.

Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., and others: Global Carbon Budget 2023, Earth Syst. Sci. Data, 15, 5301–5369, https://doi.org/10.5194/essd-15-5301-2023, 2023.

Gaber, M., Kang, Y., Schurgers, G., and Keenan, T.: Using automated machine learning for the upscaling of gross primary productivity, Biogeosciences, 21, 2447–2472, https://doi.org/10.5194/bg-21-2447-2024, 2024.

Gan, R., Zhang, Y., Shi, H., Yang, Y., Eamus, D., Cheng, L., Chiew, F. H. S., and Yu, Q.: Use of satellite leaf area index estimating evapotranspiration and gross assimilation for Australian ecosystems, Ecohydrology, 11, e1974, https://doi.org/https://doi.org/10.1002/eco.1974, 2018.

Gao, H., Liu, S., Lu, W., Smith, A. R., Valbuena, R., Yan, W., Wang, Z., Xiao, L., Peng, X., Li, Q., Feng, Y., McDonald, M., Pagella, T., Liao, J., Wu, Z., and Zhang, G.: Global Analysis of the Relationship between Reconstructed Solar-Induced Chlorophyll Fluorescence (SIF) and Gross Primary Production (GPP), Remote Sens., 13, 2824, https://doi.org/10.3390/rs13142824, 2021.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Ghosh, M. K., Kumar, L., and Roy, C.: Climate Variability and Mangrove Cover Dynamics at Species Level in the Sundarbans, Bangladesh, https://doi.org/10.3390/su9050805, 2017.

Gilabert, M. A., Moreno, A., Maselli, F., Martínez, B., Chiesi, M., Sánchez-Ruiz, S., García-Haro, F. J., Pérez-Hoyos, A., Campos-Taberner, M., Pérez-Priego, O., Serrano-Ortiz, P., and Carrara, A.: Daily GPP estimates in Mediterranean ecosystems by combining remote sensing and meteorological data, ISPRS J. Photogramm. Remote Sens., 102, 184–197, https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.01.017, 2015.

Gnanamoorthy, P., Selvam, V., Deb Burman, P. K., Chakraborty, S., Karipot, A., Nagarajan, R., Ramasubramanian, R., Song, Q., Zhang, Y., and Grace, J.: Seasonal variations of net ecosystem (CO2) exchange in the Indian tropical mangrove forest of Pichavaram, Estuar. Coast. Shelf Sci., 243, 106828, https://doi.org/10.1016/j.ecss.2020.106828, 2020.

Hönle, S. E., Heidecke, C., and Osterburg, B.: Climate change mitigation strategies for agriculture: an analysis of nationally determined contributions, biennial reports and biennial update reports, Clim. Policy, 19, 688–702, https://doi.org/10.1080/14693062.2018.1559793, 2019.

Huang, X., Xiao, J., and Ma, M.: Evaluating the Performance of Satellite-Derived Vegetation Indices for Estimating Gross Primary Productivity Using FLUXNET Observations across the Globe, https://doi.org/10.3390/rs11151823, 2019.

Ichii, K., Ueyama, M., Kondo, M., Saigusa, N., Kim, J., Alberto, M. C., Ardö, J., Euskirchen, E. S., Kang, M., Hirano, T., Joiner, J., Kobayashi, H., Marchesini, L. B., Merbold, L., Miyata, A., Saitoh, T. M., Takagi, K., Varlagin, A., Bret-Harte, M. S., Kitamura, K., Kosugi, Y., Kotani, A., Kumar, K., Li, S. G., Machimura, T., Matsuura, Y., Mizoguchi, Y., Ohta, T., Mukherjee, S., Yanagi, Y., Yasuda, Y., Zhang, Y., and Zhao, F.: 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. Biogeosciences, 122, 767–795, https://doi.org/10.1002/2016JG003640, 2017.

Jiang, C. and Ryu, Y.: Multi-scale evaluation of global gross primary productivity and evapotranspiration products derived from Breathing Earth System Simulator (BESS), Remote Sens. Environ., 186, 528–547, https://doi.org/https://doi.org/10.1016/j.rse.2016.08.030, 2016.

Johnson, S. C.: Hierarchical clustering schemes, Psychometrika, 32, 241–254, https://doi.org/10.1007/BF02289588, 1967.

Joiner, J., Yoshida, Y., Zhang, Y., Duveiller, G., Jung, M., Lyapustin, A., Wang, Y., and Tucker, C.: Estimation of Terrestrial Global Gross Primary Production (GPP) with Satellite Data-Driven Models and Eddy Covariance Flux Data, Remote Sens., 10, 1346, https://doi.org/10.3390/rs10091346, 2018.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M.: The FLUXCOM ensemble of global land-atmosphere energy fluxes, Sci. Data, 6, 1–14, https://doi.org/10.1038/s41597-019-0076-8, 2019.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., and Carvalhais, N.: Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach, Biogeosciences, 17, 1343–1365, 2020.

Kimball, J. S., et al.: SMAP L4 Global Daily 9 km EASE-Grid Carbon Net Ecosystem Exchange, NASA National Snow and Ice Data Center Distributed Active Archive Center, https://doi.org/https://doi.org/10.5067/L6C9EY1O8VIC, 2021.

Knutti, R. and Sedláček, J.: Robustness and uncertainties in the new CMIP5 climate model projections, Nat. Clim. Chang., 3, 369–373, https://doi.org/10.1038/nclimate1716, 2013.

Kumar, S. V, Holmes, T. R., Bindlish, R., de Jeu, R., and Peters-Lidard, C.: Assimilation of vegetation optical depth retrievals from passive microwave radiometry, Hydrol. Earth Syst. Sci., 24, 3431–3450, https://doi.org/10.5194/hess-24-3431-2020, 2020.

Lele, N., Kripa, M. K., Panda, M., Das, S. K., Nivas, A. H., Divakaran, N., Naik-Gaonkar, S., Sawant, A., Pattnaik, A. K., Samal, R. N., and others: Seasonal variation in photosynthetic rates and satellite-based GPP estimation over mangrove forest, Environ. Monit. Assess., 193, 1–20, https://doi.org/10.1007/s10661-021-

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

08846-0, 2021.

Li, X. and Xiao, J.: Mapping Photosynthesis Solely from Solar-Induced Chlorophyll Fluorescence: A Global, Fine-Resolution Dataset of Gross Primary Production Derived from OCO-2, https://doi.org/10.3390/rs11212563, 2019.

Liao, Z., Zhou, B., Zhu, J., Jia, H., and Fei, X.: A critical review of methods, principles and progress for estimating the gross primary productivity of terrestrial ecosystems, Front. Environ. Sci., 11, https://doi.org/10.3389/fenvs.2023.1093095, 2023.

Malhi, Y., Girardin, C., Metcalfe, D. B., Doughty, C. E., Aragão, L. E. O. C., Rifai, S. W., Oliveras, I., Shenkin, A., Aguirre-Gutiérrez, J., Dahlsjö, C. A. L., Riutta, T., Berenguer, E., Moore, S., Huasco, W. H., Salinas, N., da Costa, A. C. L., Bentley, L. P., Adu-Bredu, S., Marthews, T. R., Meir, P., and Phillips, O. L.: The Global Ecosystems Monitoring network: Monitoring ecosystem productivity and carbon cycling across the tropics, Biol. Conserv., 253, 108889, https://doi.org/https://doi.org/10.1016/j.biocon.2020.108889, 2021.

MoEF: Second National Communication to the United Nations Framework Convention on Climate Change: Philippines, Unfcc, New Delhi, 2000–2004 pp., 2014.

MoEFCC: India’ s Intended Nationally Determined Contributions – Towards Climate Justice Background, Unfccc/Indc, New Delhi, 1–28 pp., 2015.

MoEFCC: Fourth Biennial Update Report of India To The United Nations Framework Convention on Climate Change, Ministry of Environment, Forest and Climate Change, Government of India New …, New Delhi, 240 pp., 2020.

Mohamed, S. and Klaus, H.: AI and ML for Data-Driven Insights: Machine learning algorithms can analyze vast amounts of medical data to identify patterns and trends, 2024.

Monteith, J. L.: Solar Radiation and Productivity in Tropical Ecosystems, J. Appl. Ecol., 9, 747–766, https://doi.org/10.2307/2401901, 1972.

Morrison, R., Angadi, S. S., Cooper, H. M., Evans, J. G., Rees, G., Sekhar, M., Taylor, C., Tripathi, S. N., and Turner, A.: Energy and carbon dioxide fluxes, meteorology and soil physics observed at INCOMPASS land surface stations in India, 2016 to 2017, https://doi.org/10.5285/78c64025-1f8d-431c-bdeb-e69a5877d2ed, 2019.

Murtagh, F.: Brief history of cluster analysis, 21–30 pp., https://doi.org/10.1201/b19706, 2015.

Namadi, P., He, M., and Sandhu, P.: Salinity-constituent conversion in South Sacramento-San Joaquin Delta of California via machine learning, Earth Sci. Informatics, 15, 1749–1764, https://doi.org/10.1007/s12145-022-00828-1, 2022.

Narayanan, V., Cao, Y., Panda, P., Challapalle, N., Du, X., Kim, Y., Krishnan, G., Lee, C., Li, Y., Sun, J., Venkatesha, Y., Wang, Z., and Zheng, Y.: Overview of Recent Advancements in Deep Learning and Artificial Intelligence, 23–79, https://doi.org/10.1002/9781119853923.ch2, 2023.

Nelson, J. A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., and others: X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, FLUXCOM-X, Biogeosciences, 21, 5079–5115, 2024.

Neogi, S., Dey, M., Kabir, S. M. L., Masum, S., Kopprio, G., Yamasaki, S., and Lara, R.: Sundarban mangroves: diversity, ecosystem services and climate change impacts, Asian J. Med. Biol. Res., 2, https://doi.org/10.3329/ajmbr.v2i4.30988, 2016.

Novick, K. A., Biederman, J. A., Desai, A. R., Litvak, M. E., Moore, D. J. P., Scott, R. L., and Torn, M. S.: The AmeriFlux network: A coalition of the willing, Agric. For. Meteorol., 249, 444–456, https://doi.org/10.1016/j.agrformet.2017.10.009, 2018.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., and Hayes, D.: A large and persistent carbon sink in the world’s forests, Science (80-.)., 333, 988–993, https://doi.org/10.1126/science.1201609, 2011.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., Arkebauer, T., Arndt, S. K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L. B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Boike, J., Bolstad, P. V., Bonal, D., Bonnefond, J.-M., Bowling, D. R., Bracho, R., Brodeur, J., Brümmer, C., Buchmann, N., Burban, B., Burns, S. P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T. R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B. D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P. S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K. J., Cinti, B. De, Grandcourt, A. de, Ligne, A. De, De Oliveira, R. C., Delpierre, N., Desai, A. R., Di Bella, C. M., Tommasi, P. di, Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrêne, E., Dunn, A., Dušek, J., Eamus, D., Eichelmann, U., ElKhidir, H. A. M., Eugster, W., Ewenz, C. M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., et al.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Sci. Data, 7, 225, https://doi.org/10.1038/s41597-020-0534-3, 2020.

Pearson, K.: VII. Note on regression and inheritance in the case of two parents, Proc. R. Soc. London, 58, 240–242, https://doi.org/10.1098/rspl.1895.0041, 1895.

Pearson, K.: LIII. On lines and planes of closest fit to systems of points in space , London, Edinburgh, Dublin Philos. Mag. J. Sci., 2, 559–572, https://doi.org/10.1080/14786440109462720, 1901.

Rajan, N., Maas, S. J., and Cui, S.: Extreme drought effects on carbon dynamics of a semiarid pasture, Agron. J., 105, 1749–1760, https://doi.org/10.2134/agronj2013.0112, 2013.

Ravindranath, N. H., Chaturvedi, R. K., and Kumar, P.: Paris Agreement; research, monitoring and reporting requirements for India, Curr. Sci., 916–922, https://doi.org/10.18520/cs/v112/i05/916-922, 2017.

Rebmann, C., Aubinet, Ma., Schmid, H., Arriga, N., Aurela, M., Burba, G., Clement, R., Ligne, A. De, Fratini, G., Gielen, B., Grace, J., Graf, A., Gross, P., Haapanala, S., Herbst, M., Hörtnagl, L., Ibrom, A., Joly, L., Kljun, N., Kolle, O., Kowalski, A., Lindroth, A., Loustau, D., Mammarella, I., Mauder, M., Merbold, L., Metzger, S., Mölder, M., Montagnani, L., Papale, D., Pavelka, M., Peichl, M., Roland, M., Penélope, S.-O., Siebicke, L., Steinbrecher, R., Tuovinen, J.-P., Vesala, T., Wohlfahrt, G., and Franz, D.: ICOS eddy covariance flux-station site setup: a review, Int. Agrophysics, 32, 471–494, https://doi.org/10.1515/intag-2017-0044, 2018.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J. M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm, Glob. Chang. Biol., 11, 1424–1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and others: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Rodda, S., Thumaty, K., Jha, C., and Dadhwal, V.: Seasonal Variations of Carbon Dioxide, Water Vapor and Energy Fluxes in Tropical Indian Mangroves, Forests, 7, 35, https://doi.org/10.3390/f7020035, 2016.

Rodda, S. R., Thumaty, K. C., Praveen, M. S. S., Jha, C. S., and Dadhwal, V. K.: Multi-year eddy covariance measurements of net ecosystem exchange in tropical dry deciduous forest of India, Agric. For. Meteorol., 301–302, 108351, https://doi.org/10.1016/j.agrformet.2021.108351, 2021.

Rodda, S. R., Thumaty, K. C., Fararoda, R., Jha, C. S., and Dadhwal, V. K.: Unique characteristics of ecosystem CO2 exchange in Sundarban mangrove forest and their relationship with environmental factors, Estuar. Coast. Shelf Sci., 267, 107764, https://doi.org/10.1016/j.ecss.2022.107764, 2022.

Roy, P. S., Meiyappan, P., Joshi, P. K., Kale, M. P., Srivastav, V. K., Srivasatava, S. K., Behera, M. D., Roy, A., Sharma, Y., and Ramachandran, R. M.: Decadal land use and land cover classifications across India, 1985, 1995, 2005. ORNL DAAC, Oak Ridge, Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/1336, 2016.

Running, S., Mu, Q., Zhao, M.: MODIS/Aqua Gross Primary Productivity 8-Day L4 Global 500m SIN Grid V061, NASA EOSDIS L. Process. DAAC, 660, 2021.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Running, S. W. and Zhao, M.: Daily GPP and Annual NPP (MOD17A2/A3) products NASA Earth Observing System MODIS Land Algorithm - User's guide V3, 28 pp., 2015.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., and Hashimoto, H.: A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production, https://doi.org/10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2, 2004.

Santra, P. and Chkraborty, A.: Analysis of Seasonal and Annual Change of Vegetation in the Indian Thar Desert Using Modis Data, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., XXXVIII-8/, 175–178, https://doi.org/10.5194/isprsarchives-xxxviii-8-w20-175-2011, 2012.

Sarkar, D. P., Shankar, B. U., and Parida, B. R.: Machine learning approach to predict terrestrial gross primary productivity using topographical and remote sensing data, Ecol. Inform., 70, 101697, https://doi.org/10.1016/j.ecoinf.2022.101697, 2022.

Sett, T., Nandy, S., Patel, N. R., Padalia, H., Srinet, R., and Watham, T.: Spatio-temporal dynamics of water use efficiency over forest ecosystems using time series satellite data and carbon flux measurements, For. Ecol. Manage., 548, 121385, https://doi.org/https://doi.org/10.1016/j.foreco.2023.121385, 2023.

Siegel, A. F. and Wagner, M. R.: Chapter 19 - Statistical (Machine) Learning: Using Complex Models With Large Data Sets, edited by: Siegel, A. F. and Wagner, M. R. B. T.-P. B. S. (Eighth E., Academic Press, 571–599, https://doi.org/https://doi.org/10.1016/B978-0-12-820025-4.00019-1, 2022.

Sinha, S. K., Padalia, H., Patel, N. R., and Chauhan, P.: Modelling sun-induced fluorescence for improved evaluation of forest carbon flux (GPP): Case study of tropical deciduous forest, India, Ecol. Modell., 449, 109552, https://doi.org/10.1016/j.ecolmodel.2021.109552, 2021.

Sokal, R. R.: A statistical method for evaluating systematic relationships, Univ Kans Sci Bull, 38, 1409–1438, 1958.

Sokal, R. R.: Biometry. The principles and practice of statistics in biological research., Syst. Zool., 19, 391–393, https://doi.org/10.2307/2412280, 1970.

Stocker, B.: GPP: Site-scale and global model outputs from P-model used for Stocker et al. (2019) Nature Geosci., https://doi.org/10.5281/zenodo.1423484, September 2018.

Stocker, B. D., Wang, H., Smith, N. G., Harrison, S. P., Keenan, T. F., Sandoval, D., Davis, T., and Prentice, I. C.: P-model v1.0: An optimality-based light use efficiency model for simulating ecosystem gross primary production, Geosci. Model Dev., 13, 1545–1581, https://doi.org/10.5194/gmd-13-1545-2020, 2020.

Taye, M. M.: Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions, https://doi.org/10.3390/computers12050091, 2023.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, https://doi.org/10.1029/2000JD900719, 2001.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bull. Am. Meteorol. Soc., 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Teubner, I. E., Forkel, M., Camps-Valls, G., Jung, M., Miralles, D. G., Tramontana, G., van der Schalie, R., Vreugdenhil, M., Mösinger, L., and Dorigo, W. A.: A carbon sink-driven approach to estimate gross primary production from microwave satellite observations, Remote Sens. Environ., 229, 100–113, https://doi.org/https://doi.org/10.1016/j.rse.2019.04.022, 2019.

Teubner, I. E., Forkel, M., Wild, B., Mösinger, L., and Dorigo, W.: Impact of temperature and water availability on microwave-derived gross primary production, Biogeosciences, 18, 3285–3308, https://doi.org/10.5194/bg-18-3285-2021, 2021.

Tramontana, G., Ichii, K., Camps-Valls, G., Tomelleri, E., and Papale, D.: Uncertainty analysis of gross primary production upscaling using Random Forests, remote sensing and eddy covariance data, Remote Sens. Environ., 168, 360–373, https://doi.org/10.1016/j.rse.2015.07.015, 2015.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Zhao, M., Running, S. W., Wofsy, S. C., Urbanski, S., Dunn, A. L., and Munger, J. W.: Scaling Gross Primary Production (GPP) over boreal and deciduous forest landscapes in support of MODIS GPP product validation, Remote Sens. Environ., 88, 256–270, https://doi.org/10.1016/j.rse.2003.06.005, 2003.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., Costa, M. H., Kirschbaum, A. A., Ham, J. M., Saleska, S. R., and Ahl, D. E.: Evaluation of MODIS NPP and GPP products across multiple biomes, Remote Sens. Environ., 102, 282–292, https://doi.org/10.1016/j.rse.2006.02.017, 2006.

Uchale, G., Deb Burman, P. K., Tiwari, Y. K., Datye, A., and Sarkar, A.: Investigating Terrestrial Carbon Uptake Over India Using Multimodel Simulations of Gross Primary Productivity and Satellite-Based Biophysical Product, J. Geophys. Res. Biogeosciences, 128, e2023JG007468, https://doi.org/10.1029/2023JG007468, 2023.

Wang, H., Prentice, I. C., Keenan, T. F., Davis, T. W., Wright, I. J., Cornwell, W. K., Evans, B. J., and Peng, C.: Towards a universal model for carbon dioxide uptake by plants, Nat. plants, 3, 734–741, https://doi.org/10.1038/s41477-017-0006-8, 2017a.

Wang, J., Sun, R., Zhang, H., Xiao, Z., Zhu, A., Wang, M., Yu, T., and Xiang, K.: New Global MuSyQ GPP/NPP Remote Sensing Products from 1981 to 2018, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 14, 5596–5612, https://doi.org/10.1109/JSTARS.2021.3076075, 2021.

Wang, L., Zhu, H., Lin, A., Zou, L., Qin, W., and Du, Q.: Evaluation of the Latest MODIS GPP Products across Multiple Biomes Using Global Eddy Covariance Flux Data, https://doi.org/10.3390/rs9050418, 2017b.

Ward, J. H.: Hierarchical Grouping to Optimize an Objective Function, J. Am. Stat. Assoc., 58, 236–244, https://doi.org/10.1080/01621459.1963.10500845, 1963.

Watham, T., Kushwaha, S. P. ., Patel, N. R., and Dadhwal, V. K.: Monitoring of carbon dioxide and water vapour exchange over a young mixed forest plantation using eddy covariance technique, Curr. Sci., 107, 858–866, 2014.

Watham, T., Srinet, R., Nandy, S., Padalia, H., Sinha, S. K., Patel, N. R., and Chauhan, P.: Environmental control on carbon exchange of natural and planted forests in Western Himalayan foothills of India, Biogeochemistry, 151, 291–311, https://doi.org/10.1007/s10533-020-00727-x, 2020.

Wild, B., Teubner, I., Moesinger, L., Zotta, R.-M., Forkel, M., van der Schalie, R., Sitch, S., and Dorigo, W.: VODCA2GPP – a new, global, long-term (1988–2020) gross primary production dataset from microwave remote sensing, Earth Syst. Sci. Data, 14, 1063–1085, https://doi.org/10.5194/essd-14-1063-2022, 2022a.

Wu, C., Munger, J. W., Niu, Z., and Kuang, D.: Comparison of multiple models for estimating gross primary production using MODIS and eddy covariance data in Harvard Forest, Remote Sens. Environ., 114, 2925–2939, https://doi.org/10.1016/j.rse.2010.07.012, 2010.

Wu, L., Kato, T., Sato, H., Hirano, T., and Yazaki, T.: Sensitivity analysis of the typhoon disturbance effect on forest dynamics and carbon balance in the future in a cool-temperate forest in northern Japan by using SEIB-DGVM, For. Ecol. Manage., 451, 117529, 2019.

Wu, Z., Boke-Olén, N., Fensholt, R., Ardo, J., Eklundh, L., and Lehsten, V.: Effect of climate dataset selection on simulations of terrestrial GPP: Highest uncertainty for tropical regions, PLoS One, 13, e0199383, https://doi.org/10.1371/journal.pone.0199383, 2018.

Yamamoto, S., Saigusa, N., Gamo, M., Fujinuma, Y., Inoue, G., and Hirano, T.: Findings through the AsiaFlux network and a view toward the future, J. Geogr. Sci., 15, 142–148, https://doi.org/10.1007/bf02872679, 2005.

Yu, T., Sun, R., Xiao, Z., Zhang, Q., Liu, G., Cui, T., and Wang, J.: Estimation of global vegetation productivity from Global LAnd Surface Satellite data, Remote Sens., 10, https://doi.org/10.3390/rs10020327, 2018.

Yuan, W., Liu, S., Yu, G., Bonnefond, J.-M., Chen, J., Davis, K., Desai, A. R., Goldstein, A. H., Gianelle, D., Rossi, F., Suyker, A. E., and Verma, S. B.: Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data, Remote Sens. Environ., 114, 1416–1431, https://doi.org/10.1016/j.rse.2010.01.022, 2010.

https://doi.org/10.5194/egusphere-2026-3597
Preprint. Discussion started: 14 July 2026

Yuan, W., Cai, W., Xia, J., Chen, J., Liu, S., Dong, W., Merbold, L., Law, B., Arain, A., Beringer, J., Bernhofer, C., Black, A., Blanken, P. D., Cescatti, A., Chen, Y., Francois, L., Gianelle, D., Janssens, I. A., Jung, M., Kato, T., Kiely, G., Liu, D., Marcolla, B., Montagnani, L., Raschi, A., Roupsard, O., Varlagin, A., and Wohlfahrt, G.: Global comparison of light use efficiency models for simulating terrestrial vegetation gross primary production based on the LaThuile database, Agric. For. Meteorol., 192–193, 108–120, https://doi.org/10.1016/j.agrformet.2014.03.007, 2014.

A Data-driven Upscale Product of Global Gross Primary Production, Net Ecosystem Exchange and Ecosystem Respiration, ver.2020.2:

Zeng, J., Matsunaga, T., Tan, Z.-H., Saigusa, N., Shirai, T., Tang, Y., Peng, S., and Fukuda, Y.: Global terrestrial carbon fluxes of 1999--2019 estimated by upscaling eddy covariance data with a random forest, Sci. data, 7, 313, https://doi.org/10.1038/s41597-020-00653-5, 2020.

Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H. S., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang, Y., Liu, Y. Y., Miralles, D. G., and Pan, M.: Multi-decadal trends in global terrestrial evapotranspiration and its components, Sci. Rep., 6, 19124, https://doi.org/10.1038/srep19124, 2016.

Zhang, Y., Kong, D., Gan, R., Chiew, F. H. S., McVicar, T. R., Zhang, Q., and Yang, Y.: Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017, Remote Sens. Environ., 222, 165–182, https://doi.org/10.1016/j.rse.2018.12.031, 2019.

Zhao, M. and Running, S. W.: Drought-Induced Reduction in Global Terrestrial Net Primary Production from 2000 Through 2009, Science (80-.)., 329, 940–943, https://doi.org/10.1126/science.1192666, 2010.

Zhu, M., Liu, S., Xia, Z., Wang, G., Hu, Y., and Liu, Z.: Crop growth stage GPP-driven spectral model for evaluation of cultivated land quality using GA-BPNN, https://doi.org/10.3390/agriculture10080318, 2020.