Supporting Information for

EGO: a global 0.05° hourly GPP dataset for monitoring diurnal photosynthesis dynamics

Xi Liu[1], Xing Li[1,*], Dalei Hao[2], Jingfeng Xiao[3], Yanan Zhou[1], Cenliang Zhao[1], Zikang Diao[1], Fuqiang Qu[1], Shangrong Lin[1], Xiangzhuo Liu[4], Zhaoying Zhang[5], Xinjie Liu[6], Helin Zhang[7]

1 School of Geography and Planning, Sun Yat-sen University, Guangzhou 510006, China

2 Atmospheric, Climate, & Earth Sciences Division, Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99354, USA

3 Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH 03824, USA

4 INRAE, Bordeaux Sciences Agro, UMR 1391 ISPA, Villenave-d'Ornon, France

5 International Institute for Earth System Sciences, Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing University, Nanjing, Jiangsu 210023, China

6 Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, 100094 Beijing, China

7 Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul 08826, Republic of Korea

* Correspondence: Xing Li (lixing58@mail.sysu.edu.cn);

Contents of this file

Text S1 to S6

Figures S1 to S20

Tables S1 to S3

Text S1. Conversion of timestamps to local solar time

To analyze diurnal photosynthetic dynamics with physical consistency across global sites, it is essential to align observations to the sun's position rather than local time. We converted all original timestamps to Local Solar Time (LST) (Zerubavel et al.,1982). In LST, 12:00 (solar noon) corresponds to the time when the sun reaches its highest point in the sky at the specific location. This conversion removes phase shifts caused by the site's longitude relative to its time zone central meridian. The calculation of LST involves two main corrections: the Equation of Time (EoT) and the Longitude Correction (LC). The general formula is expressed as:

$$LST = t_{local} + \frac{LC+EoT}{60} \qquad \text{(Eq.S1)}$$

where t_{local} is the local time. The EoT accounts for the irregularity of the sun's apparent motion due to the earth's elliptical orbit and axial tilt. It is calculated as:

$$EOT = 9.87 * \sin(2\Phi) - 7.53 * \cos(\Phi) - 1.5 * \sin(\Phi) \qquad \text{(Eq.S2)}$$

$$\Phi = 2\pi * (DoY - 81)/365 \qquad \text{(Eq.S3)}$$

where DoY is the day of the year (1–365 or 366), and Φ is the day angle in radians. λ_{site} is the longitude. For datasets provided in UTC (e.g., ERA-5 Land), the reference meridian λ_{site} is 0. The conversion simplifies to:

$$LST = t_{UTC} + \frac{4\times\lambda_{site}+EoT}{60} \qquad \text{(Eq.S4)}$$

For datasets provided in local time (e.g., FLUXNET2015 and AmeriFlux), we first identified the time zone offset (TZ_{offset}) based on the site's metadata. The reference meridian for the local time zone is determined as $\lambda_{ref} = 15° * TZ_{offset}$.

$$LST = t_{local} + \frac{4\times(\lambda_{site}-15\times TZ_{offset})+EoT}{60} \qquad \text{(Eq.S5)}$$

Alternatively, t_{local} can first be converted to UTC ($t_{local} = t_{local} - TZ_{offset}$), and then the formula Eq.S4 is applied. By applying these corrections, all diurnal cycles were centered around solar noon, enabling a rigorous cross-site comparison of diurnal metrics (e.g., peak time and centroid).

Text S2. Calculation of VPD

To compute vapor pressure deficit (VPD) from ERA5-Land hourly air temperature (Ta) and dew point temperature (Td), we first ensure both variables share the same units. ERA5-Land typically provides temperature in kelvin. We converted K to °C by $T_C = T_K - 273.15$. Compute saturation vapor pressure e_s at the air temperature using the Magnus–Tetens approximation (Yuan et al., 2019):

$$e_{s(T)} = 0.6108 \times \exp\left(\frac{17.27 \times T}{T+237.3}\right) \quad \text{(Eq.S6)}$$

and compute actual vapor pressure e_a from the dew point temperature:

$$e_{a(T_d)} = 0.6108 \times \exp\left(\frac{17.27 \times T_d}{T_d+237.3}\right) \quad \text{(Eq.S7)}$$

Both $e_{s(T)}$ and $e_{a(Td)}$ are in kilopascals (kPa). The VPD is the difference:

$$VPD = e_{s(T)} - e_{a(T_d)} \quad \text{(Eq.S8)}$$

Apply the calculations for every hourly time step and grid cell.

Text S3. Calculation of EVI and LSWI

Use the BRDF-corrected surface reflectance from MODIS MCD43C4 to compute the two vegetation indices as follows.

First, convert the stored reflectance values using the product's scale factor. Use the bands corresponding to red (R), near-infrared (NIR), blue (B) and a shortwave-infrared (SWIR) band. Then compute Enhanced Vegetation Index (EVI) (Zeng et al., 2022):

$$EVI = 2.5\frac{\rho_{NIR}-\rho_R}{\rho_{NIR}+6\rho_R-7.5\rho_B+1} \qquad \text{(Eq.S9)}$$

EVI reduces soil and atmospheric influences and is less prone to saturation in dense canopies than NDVI (Huete et al., 2002); it therefore better captured canopy greenness and photosynthetic capacity at high biomass.

Land Surface Water Index (LSWI):

$$LSWI = \frac{\rho_{NIR}-\rho_{SWIR}}{\rho_{NIR}+\rho_{SWIR}} \qquad \text{(Eq.S10)}$$

LSWI is sensitive to leaf and surface water content (canopy liquid water and soil moisture influence), making it useful for detecting moisture stress and short-term changes in plant water status (Zhang et al., 2025).

Text S4. Interpretation of sub-grid strategy

As seen from FigS1., scenario (a) serves as the baseline, with a composition of F: 60%, G: 10%, and C: 30%. Scenarios (b) and (c) represent two distinct land-use change trajectories originating from (a): in (b), cropland expands significantly (F: 50%, G: 5%, C: 45%), while in (c), grassland replaces substantial portions of other types (F: 50%, G: 45%, C: 5%). Despite these drastic sub-grid variations, the dominant vegetation type in both (b) and (c) remains "Forest" (50%). Consequently, a traditional upscaling approach relying solely on the dominant class would fail to detect any land-use change between (a), (b), and (c), erroneously treating them as stable forest ecosystems. A "land-use change" is only triggered in the dominant-class scheme in scenario (d) (F: 40%, G: 15%, C: 45%), where the forest fraction further declines and cropland finally exceeds forest to become the dominant type. In contrast, our fractional coverage method explicitly captures the quantitative shifts in vegetation composition across all scenarios. By resolving these sub-grid dynamics—even when the dominant class remains unchanged—our approach enables the model to incorporate finer-scale land surface information and more accurately reflect the impact of land-use changes.

Text S5. Calculation of R^2, RMSE and Bias

To evaluate model performance, three statistical metrics were employed: the coefficient of determination (R^2), the root mean square error (RMSE), and the mean bias error (Bias).

The coefficient of determination (R^2) quantifies how well the predicted values capture the variability of observed data (Li and Xiao, 2019) and is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad \text{(Eq.S11)}$$

Where y_i and $\hat{y}_i$ denote the observed and predicted values, respectively, and $\bar{y}_i$ is the mean of observed values. Higher R^2 values (closer to 1) indicate a stronger correspondence between model outputs and observations, reflecting good predictive capability.

The root mean square error (RMSE) measures the overall magnitude of prediction errors (Li et al., 2025) and is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad \text{(Eq.S12)}$$

It provides an estimate of the average deviation between predictions and observations. A lower RMSE indicates higher model accuracy and consistency.

The mean bias error (Bias) evaluates the systematic offset between predictions and observations (Li et al., 2025) and is computed as:

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i}) \qquad \text{(Eq.S13)}$$

A positive Bias suggests model overestimation, whereas a negative Bias indicates underestimation.

Together, these three metrics provide a comprehensive evaluation framework: R^2 reflects the model's explanatory power, RMSE quantifies overall error magnitude, and Bias reveals systematic tendencies. This combination ensures a balanced assessment of both accuracy and reliability in model performance.

Text S6. Evaluation of the sufficiency of diurnal metrics

To evaluate whether the proposed three metrics are sufficient to characterize diurnal GPP curves, we design four comparative experiments (Case 1–4). For each experiment, we generated several diurnal GPP curves using a Gaussian-based function with random perturbations and retained the 100 curves whose computed diurnal metrics matched those of the reference curve. The mean aggregated curve from each case was then analyzed to assess whether the corresponding metric combination could reconstruct the reference diurnal GPP curve. The experimental settings were designed as follows: controlling only mean GPP (Case 1), controlling both mean GPP and ΔGPP (Case 2), controlling both ΔGPP and MDI (Case 3), and controlling mean GPP, ΔGPP, and MDI simultaneously (Case 4). Comparisons of the reconstructed curves with the reference curve reveals the specific contribution of each metric.

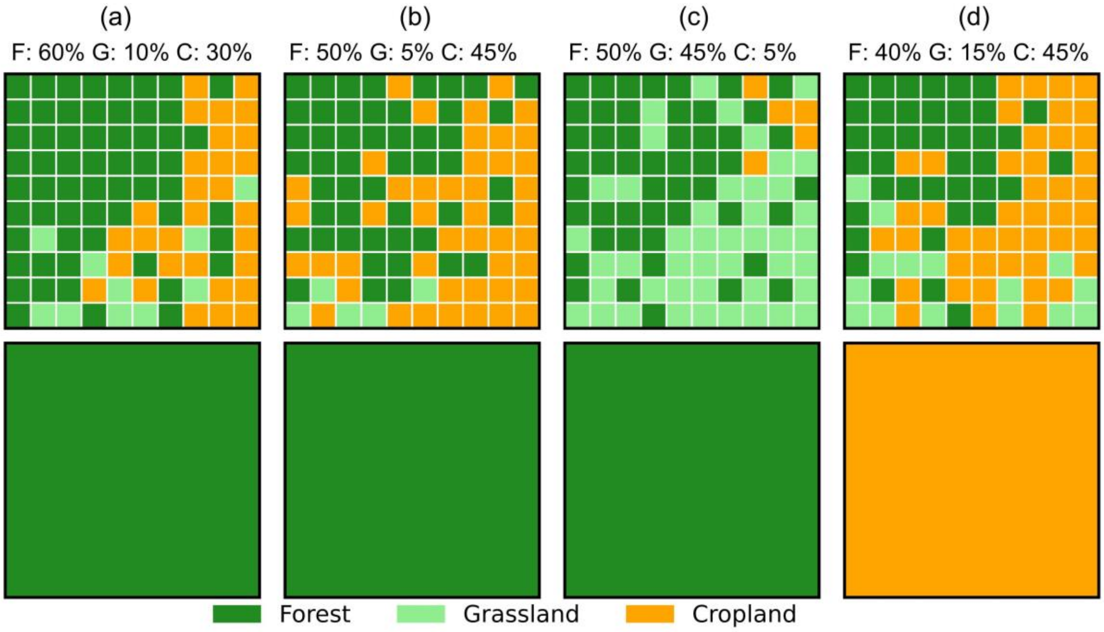

Figure S1. Four hypothetical scenarios (a)–(d) of vegetation composition within a single 0.05° grid cell, comprising varying fractions of 500 m pixels classified as Forest (F), Grassland (G), and Cropland (C).

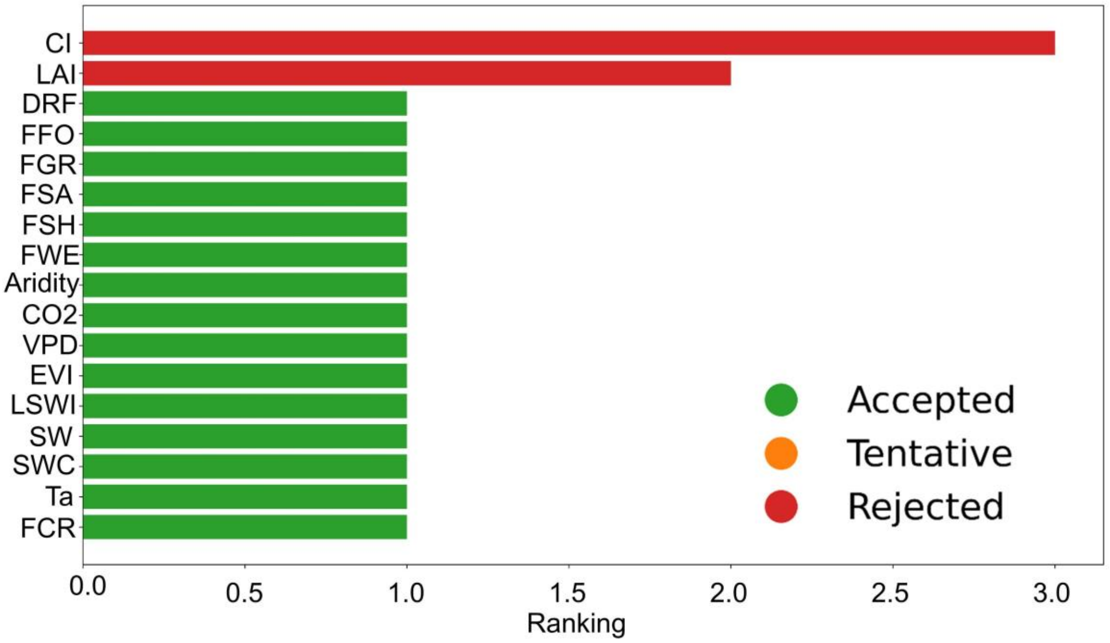

Figure S2. Boruta feature selection. The CI (clumping index) and LAI (leaf area index) were discarded due to their lower importance than shadow feature (Red legend).

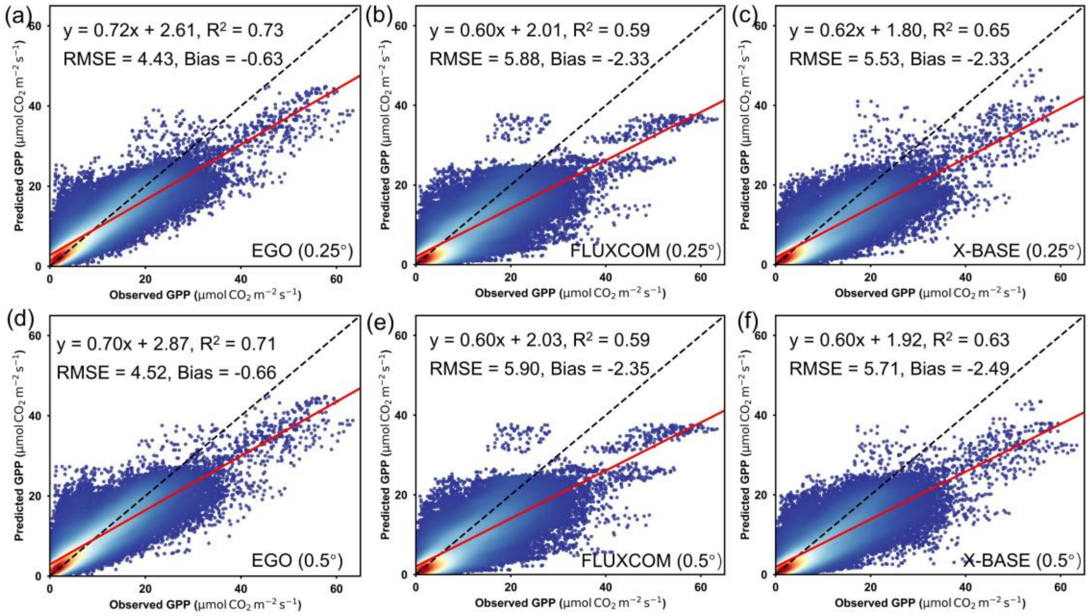

Figure S3. Performance comparison of EGO, FLUXCOM, and X-BASE GPP products against site observations at harmonized spatial resolutions. To ensure a fair comparison, all products were spatially resampled to match specific target resolutions using the bilinear interpolation method. Panels (a)–(c) display the accuracy evaluation when all products are unified to a 0.25° spatial resolution, which is consistent with the native resolution of X-BASE. Panels (d)–(f) present the corresponding performance when all datasets are unified to a 0.5° spatial resolution, consistent with the native resolution of FLUXCOM. The red solid lines represent the linear regression fits, and the black dashed lines denote the 1:1 reference line.

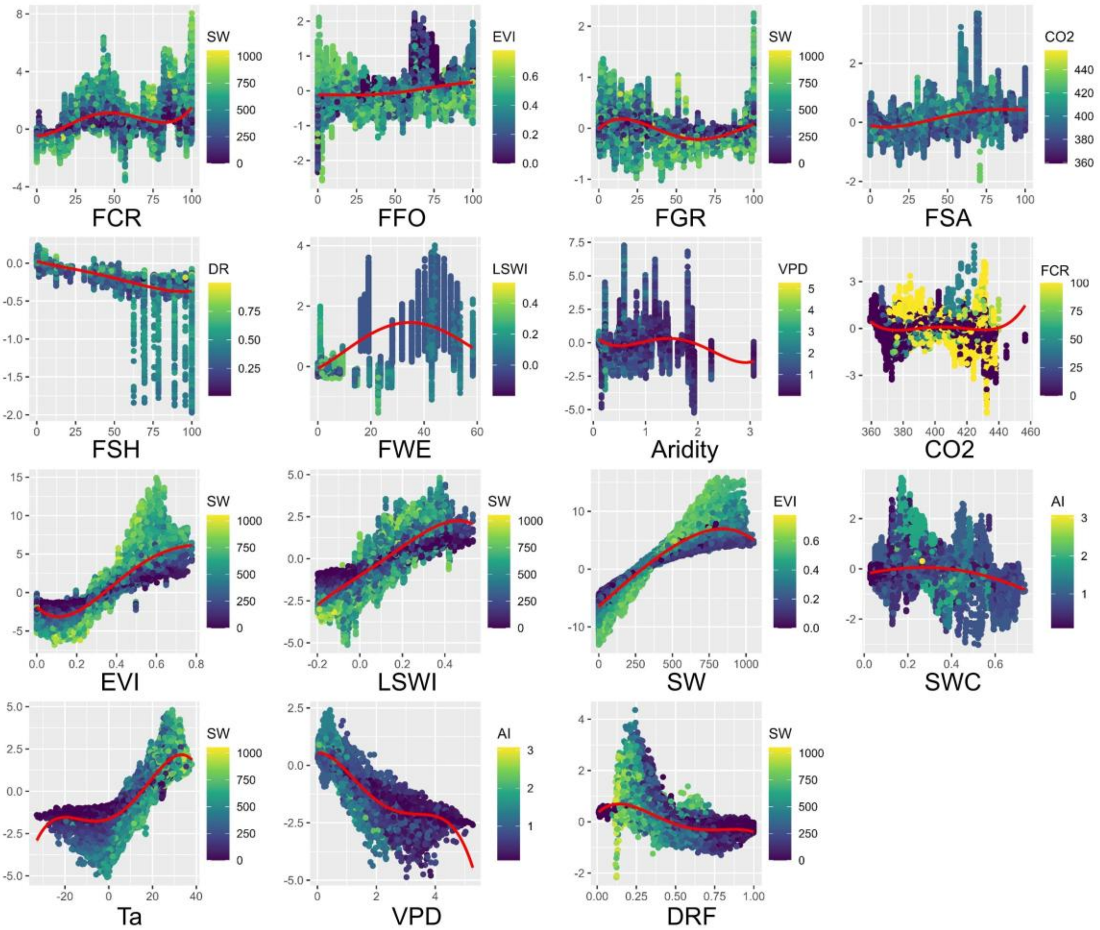

Figure S4. Interaction between each driver and their corresponding SHAP values.

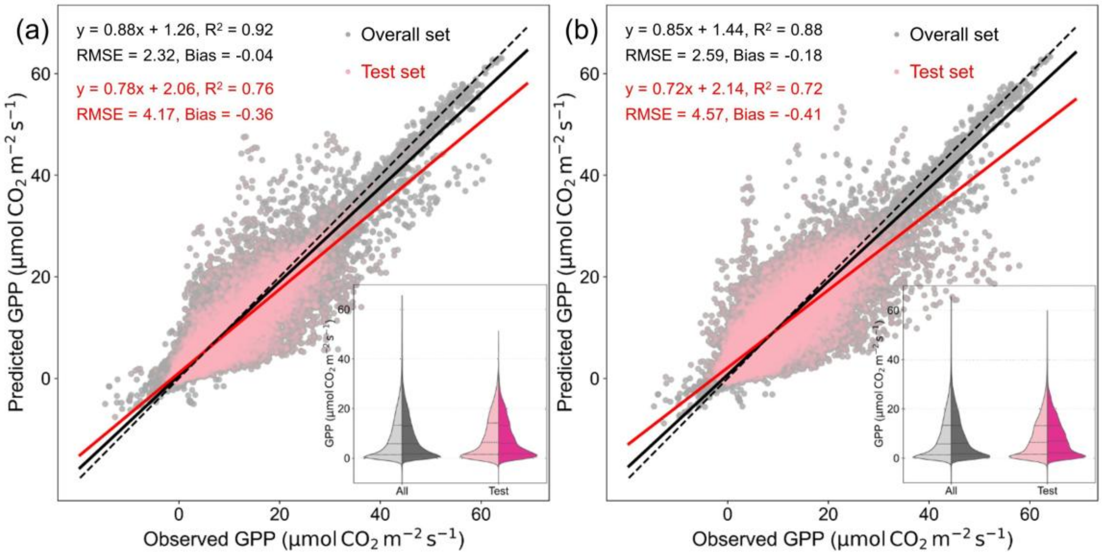

Figure S5. Performance of CKML-GPP of training models (a) with and (b) without aridity index.

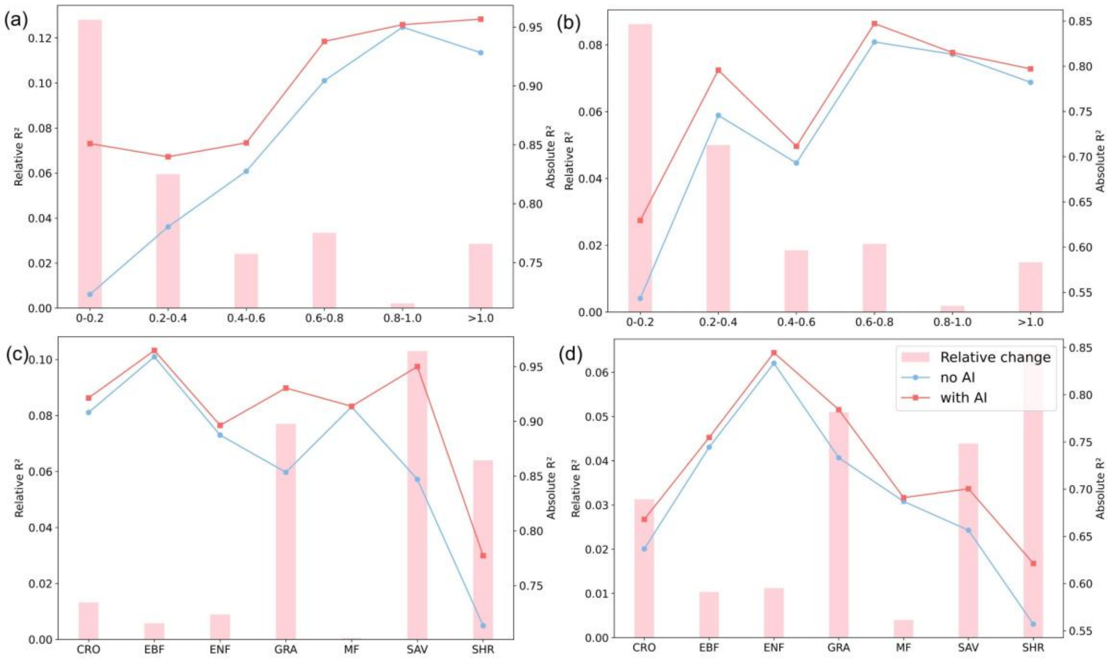

Figure S6. Relative improvement of R² across (a,b) aridity gradients and (c,d) vegetation types when incorporated aridity index into predictor. Red and blue line represented the evaluation metrics with and without aridity index. Pink bar was relative change. (a, c) for training set, and (b, d) for test set.

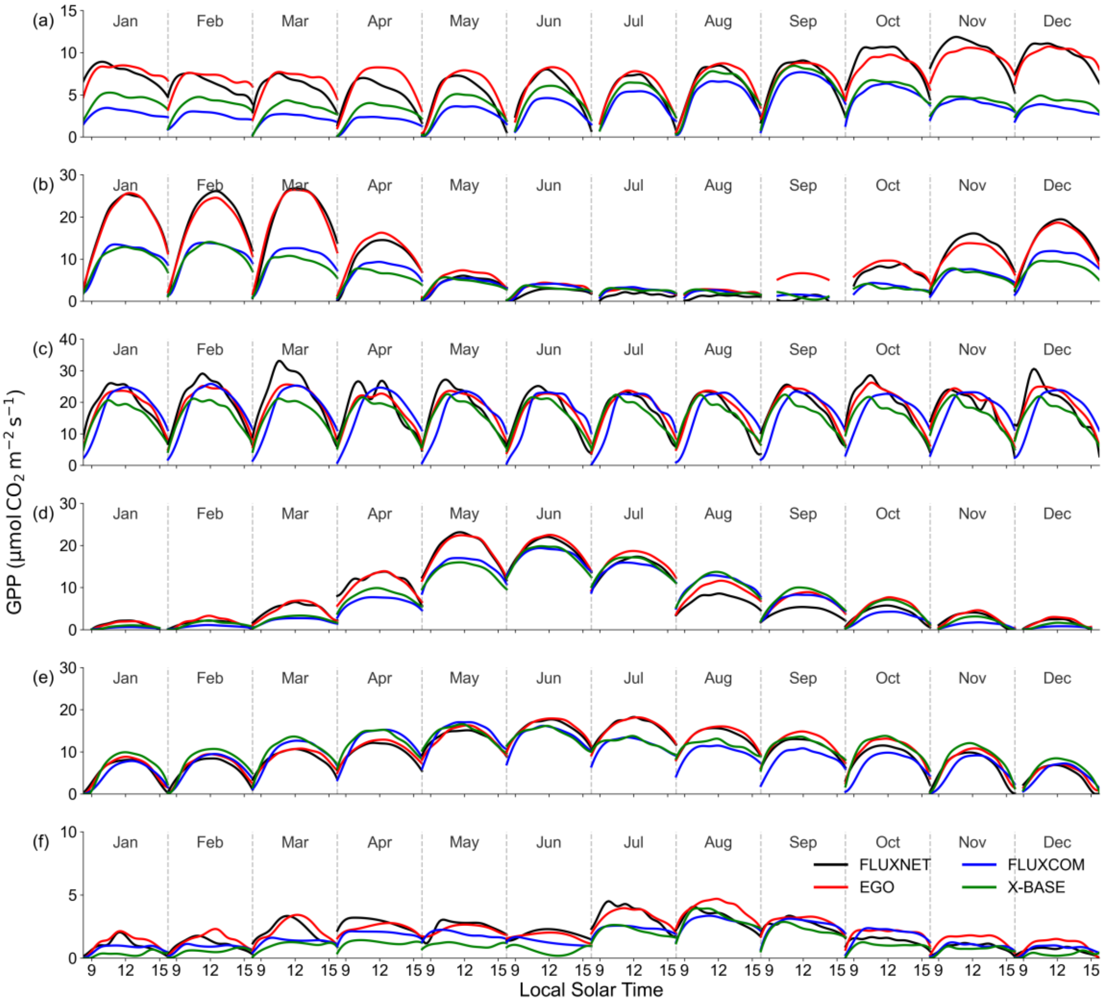

Figure S7. Comparison of diurnal GPP dynamics between site observations and products across diverse biomes for each month of the year. Panels (a)–(f) illustrate the multi-year average diurnal GPP variations across different months at six representative flux sites: (a) savannas (AU-Gin), (b) grassland (AU-DaP), (c) broadleaf forest (BR-Sa3), (d) cropland (DE-Kli), (e) needleleaf forest (US-Blo), and (f) shrubland (US-SRC). The black solid lines represent the in-situ FLUXNET observations. The colored lines denote the estimations from different global hourly GPP products: EGO (red), FLUXCOM (blue), and X-BASE (green).

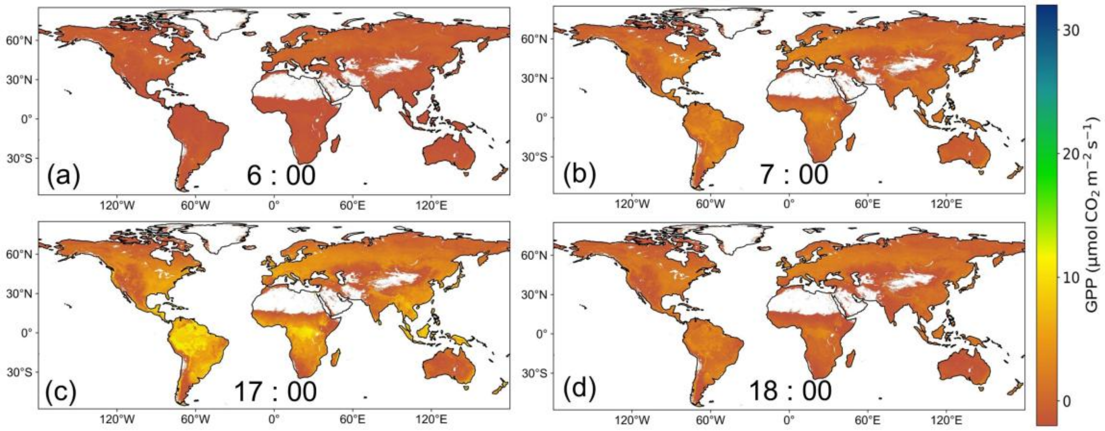

Figure S8. Global diurnal pattern of EGO GPP averaged over the growing seasons from 2000 to 2022, a-i represents the spatial distribution of GPP in 06:00, 07:00, 17:00, and 18:00.

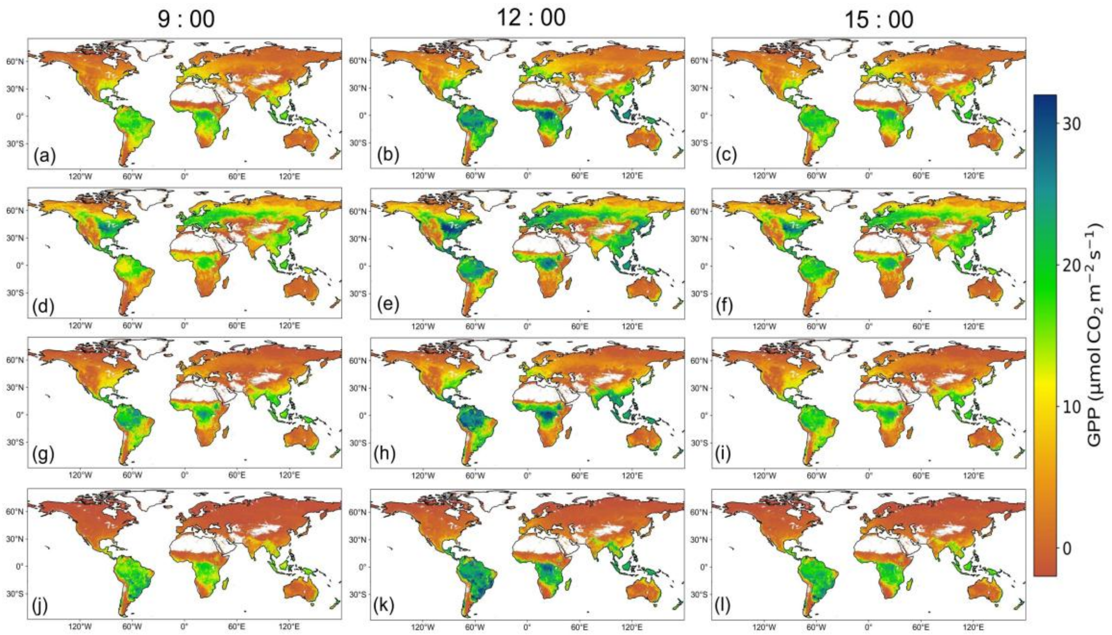

Figure S9. Global diurnal pattern of EGO GPP from 2000 to 2022 in different seasons. (a-c) for March to May, (d-f) for June to August, (g-i) for September to November, and (j-l) for December to February.

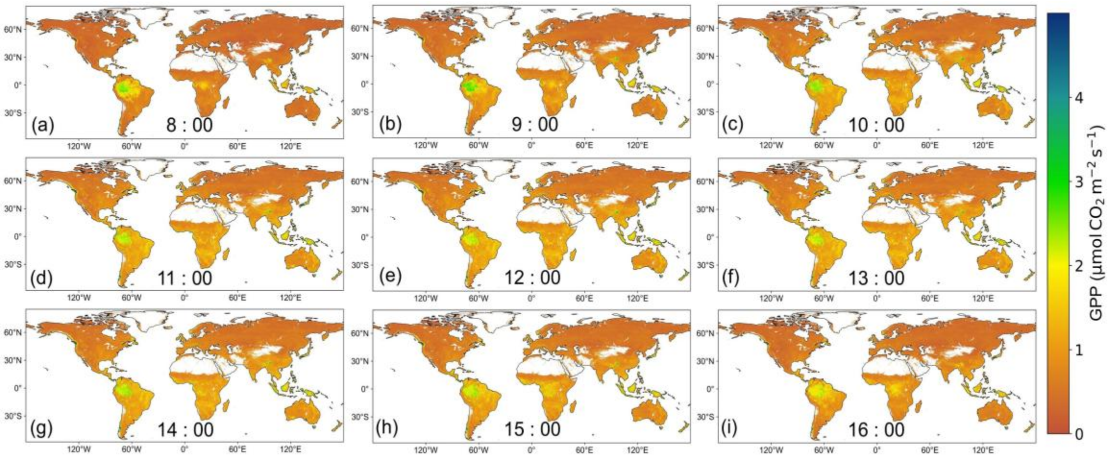

Figure S10. Global diurnal pattern of standard deviation of hourly GPP of EGO at 1-hour intervals of local solar time.

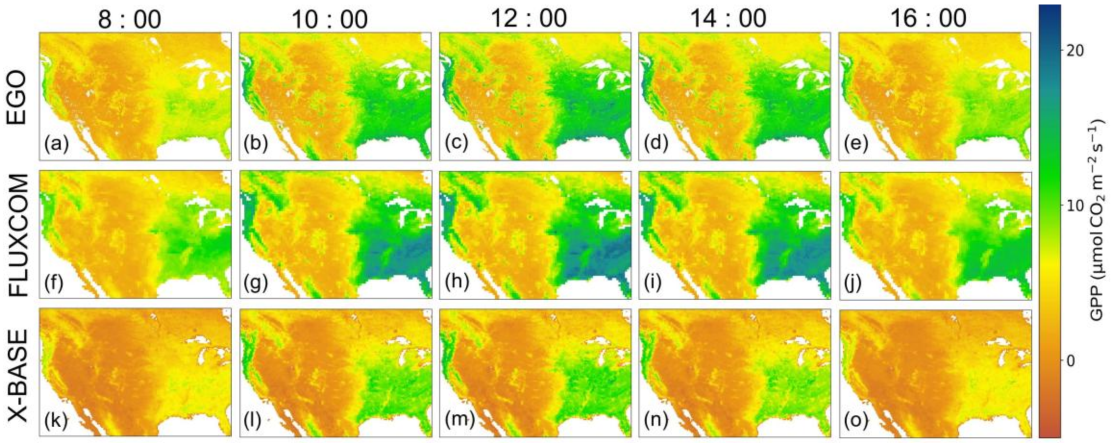

Figure S11. Detailed comparison of diurnal GPP among products in the U.S. Columns 1-5 show average GPP maps from 08:00 to 16:00 at 2 h intervals averaged over the growing seasons from 2000 to 2022, and rows 1-4 correspond to EGO, Fluxcom and X-BASE, respectively.

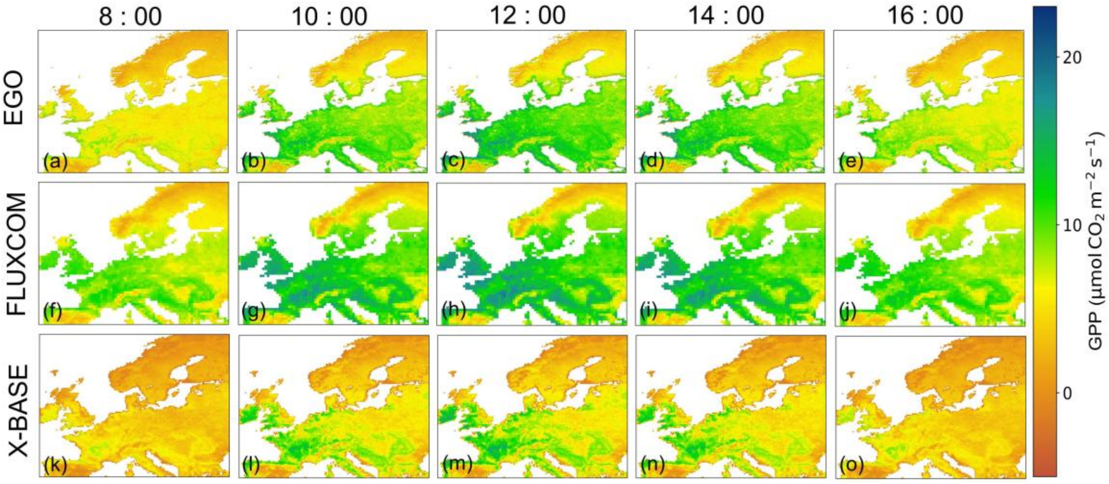

Figure S12. Detailed comparison of diurnal GPP among products in the European. Columns 1-5 show average GPP maps from 08:00 to 16:00 at 2 h intervals averaged over the growing seasons from 2000 to 2022, and rows 1-4 correspond to EGO, Fluxcom and X-BASE, respectively.

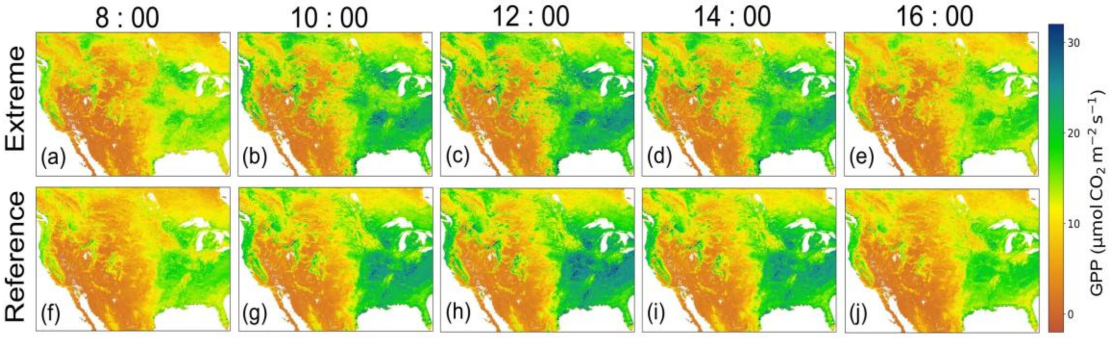

Figure S13. Spatial distribution of GPP in the U.S. during June 2021 (a–e), averaged reference year June (f–j).

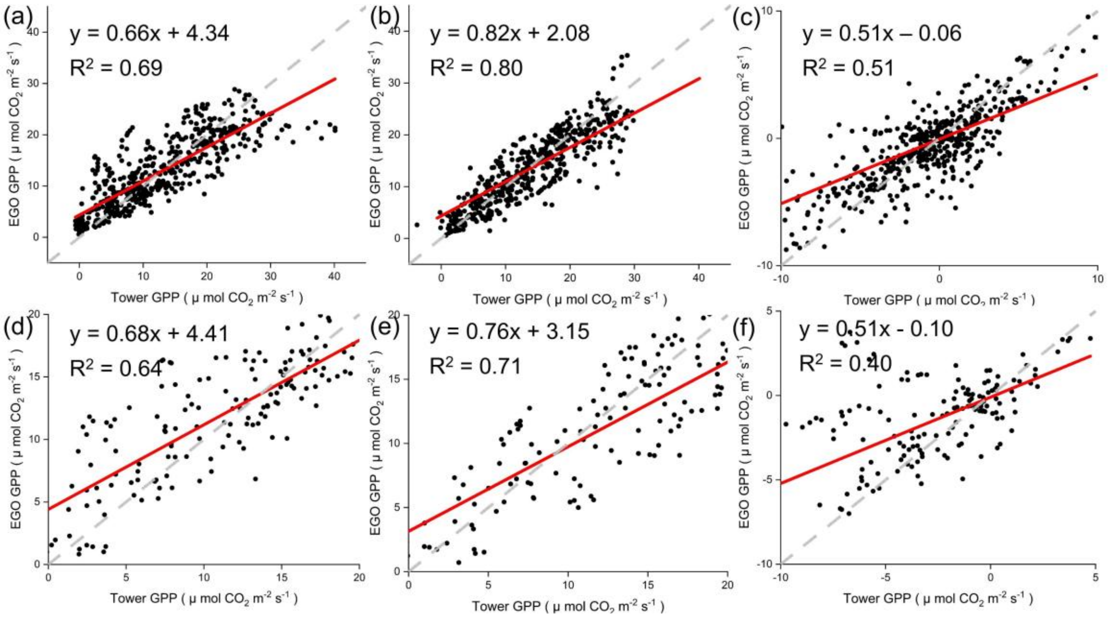

Figure S14. The fitting accuracy of EGO GPP with the corresponding 39 sites GPP in U.S. in (a) June 2021, (b) June of reference year, and (c) anomaly GPP. Similarly, 12 sites GPP in Europe in (d) August 2003, (b) August of reference year, and (f) anomaly GPP.

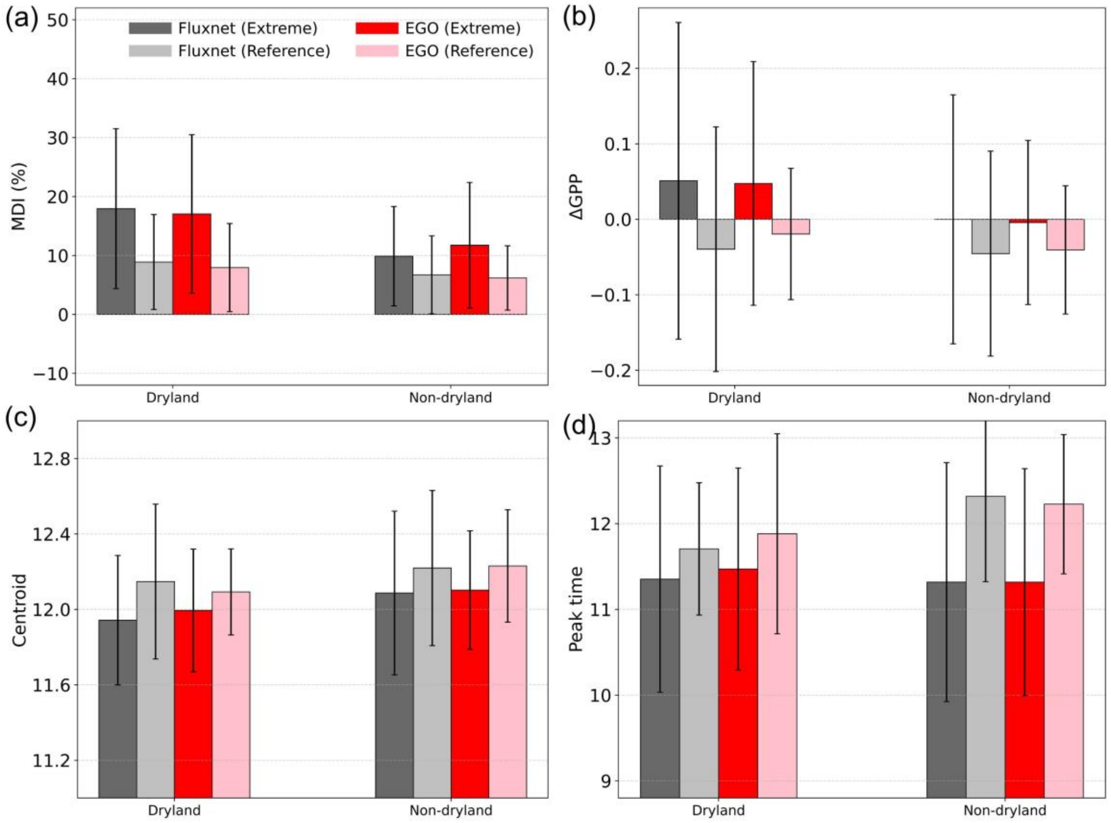

Figure S15. Boxplot of diurnal metrics of site GPP and EGO GPP in dryland and non-dryland in U.S. (a) for MDI, (b) for ΔGPP, (c) for Centroid and (d) for Peak time.

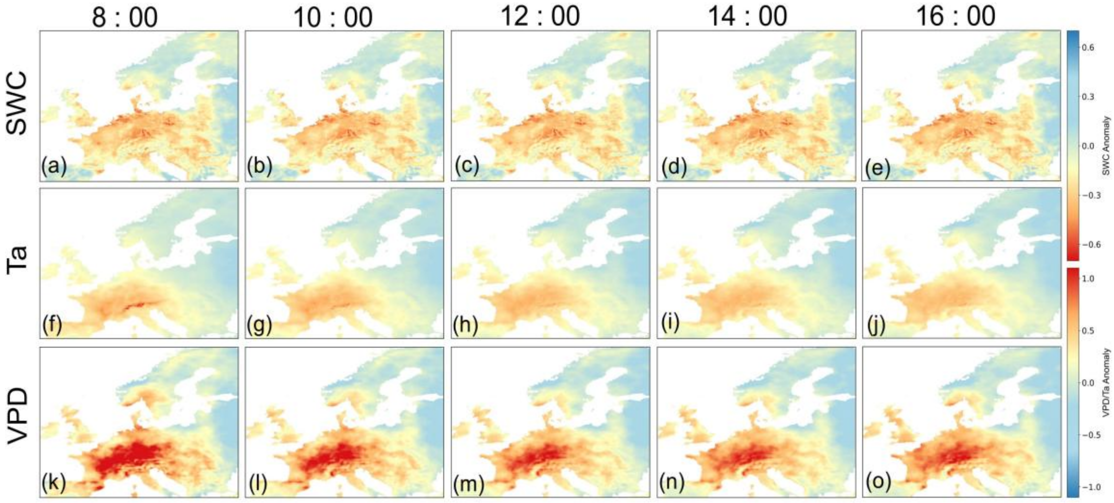

Figure S16. Spatial distribution of anomaly of SWC (a-e), Ta (f-j) and VPD (k-o) in the European during August 2003.

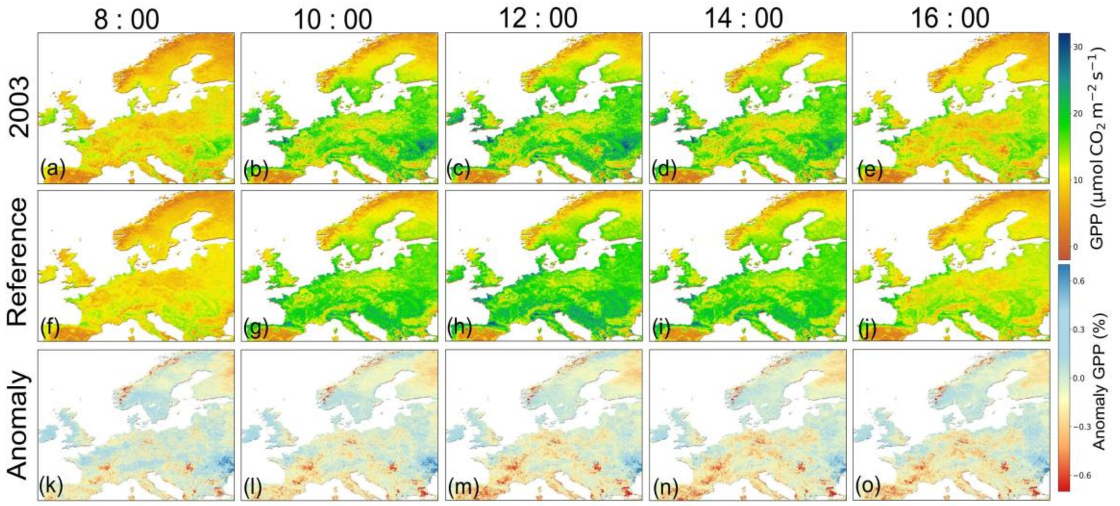

Figure S17. Spatial distribution of GPP in the European during August 2003 (a–e), averaged normal year August (f–j), and their anomalies (k–o). Anomaly = (GPP_{2003} – GPP_{normal}) / GPP_{normal}.

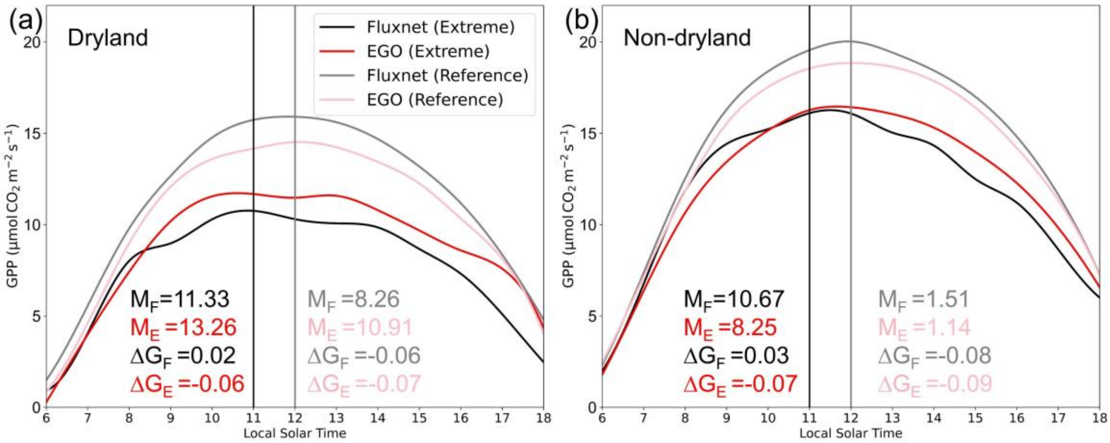

Figure S18. Mean diurnal GPP cycles aggregated for (a) dryland and (b) non-dryland sites in the European. Black and red lines denote average GPP cycles for August 2003, while gray and pink lines represent averaged years for August (Reference). The vertical line represents the peak time. M and ΔG indicate MDI and ΔGPP; F and E denote flux-tower observations and EGO GPP, respectively. Dark and light colors correspond to 2003 and reference-year metrics.

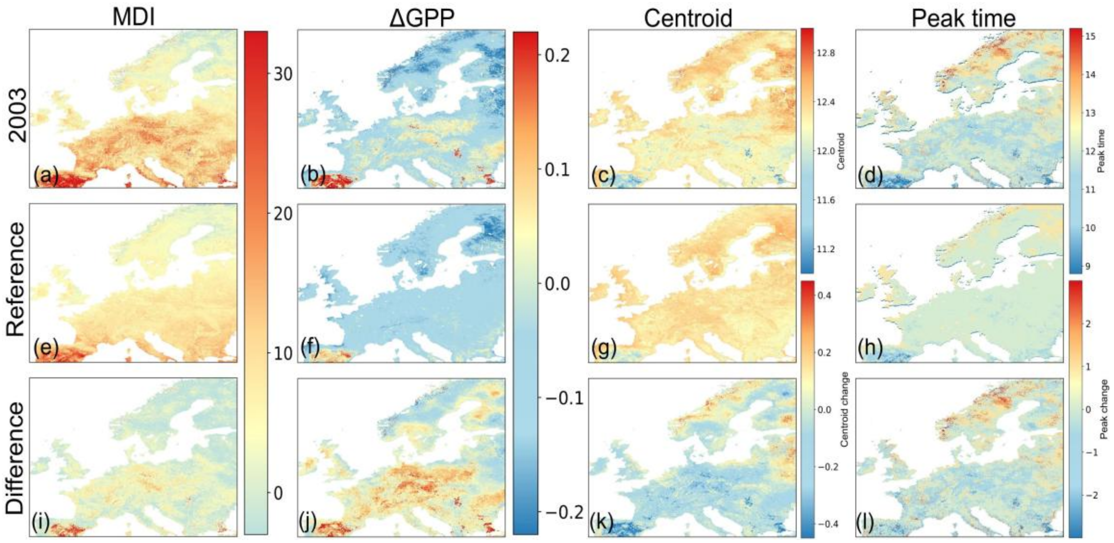

Figure S19. Spatial pattern of the diurnal metrics in the European during August 2003 (a-d) and averaged normal year August (e-h), with (i-l) showing their differences (extreme–normal).

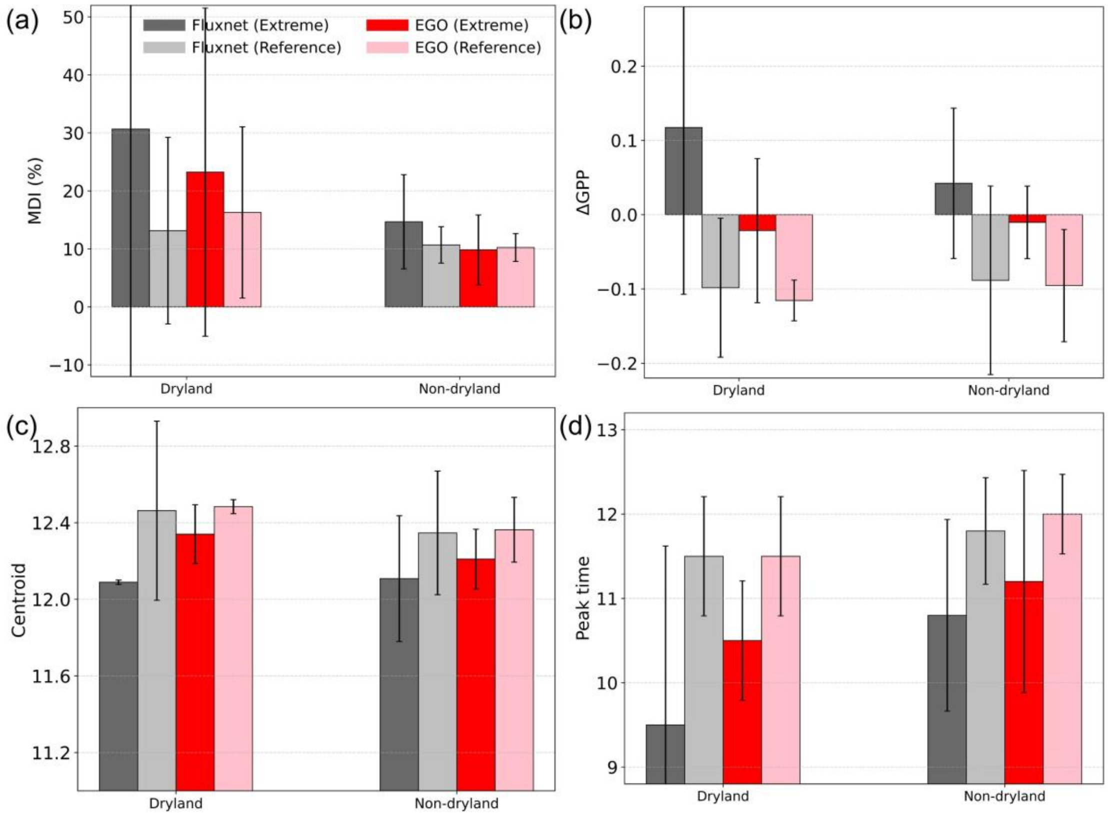

Figure S20. Boxplot of diurnal metrics of site GPP and EGO GPP in dryland and non-dryland in European. (a) for MDI, (b) for ΔGPP, (c) for Centroid and (d) for Peak time.

TableS1. The information of flux sites used in this study. The reclassification of IGBP vegetation types followed these rules: evergreen needleleaf forest (ENF), evergreen broadleaf forest (EBF), deciduous needleleaf forest (DNF), deciduous broadleaf forest (DBF), and mixed forest (MF) were grouped into Forest; closed shrubland and open shrubland were merged into Shrubland; woody savannas and savannas were combined into Savannas. In addition, Grassland, Cropland, and Wetland were retained, while non-vegetated pixels were excluded.

Station	LAT	LONG	IGBP	Reclass	AI
AR-Vir	-28.2395	-56.1886	ENF	Forest	0.991926
AU-DaP	-14.0633	131.3181	GRA	Grassland	0.500888
AU-Emr	-23.8587	148.4746	GRA	Grassland	0.278671
AU-Gin	-31.3764	115.7138	SAV	Savannas	0.319696
AU-GWW	-30.1913	120.6541	SAV	Savannas	0.106255
AU-How	-12.4943	131.1523	SAV	Savannas	0.753779
AU-Rig	-36.6499	145.5759	GRA	Grassland	0.366454
AU-Rob	-17.1175	145.6301	EBF	Forest	1.829711
AU-Stp	-17.1507	133.3502	GRA	Grassland	0.232459
AU-Tum	-35.6566	148.1517	EBF	Forest	1.208532
AU-Wac	-37.4259	145.1878	EBF	Forest	1.120573
AU-Wom	-37.4222	144.0944	EBF	Forest	0.850607
AU-Ync	-34.9893	146.2907	GRA	Grassland	0.230271
BE-Bra	51.3076	4.5198	MF	Forest	0.947956
BE-Lon	50.5516	4.7462	CRO	Cropland	0.96105
BE-Vie	50.3049	5.9981	MF	Forest	1.386077
BR-Npw	-16.498	-56.412	SAV	Savannas	0.8674
BR-Sa1	-2.8567	-54.9589	EBF	Forest	1.432958
BR-Sa3	-3.018	-54.9714	EBF	Forest	1.418789
CA-ARB	52.695	-83.9452	WET	Wetland	0.9311
CA-Ca1	49.8673	-125.334	ENF	Forest	1.8893
CA-Ca2	49.8705	-125.291	ENF	Forest	1.8893
CA-Cbo	44.3184	-79.9341	DBF	Forest	0.9675
CA-ER1	43.6405	-80.4123	CRO	Cropland	1.0457
CA-Gro	48.2167	-82.1556	MF	Forest	0.950414
CA-LP1	55.1119	-122.841	ENF	Forest	0.9346
CA-Man	55.8796	-98.4808	ENF	Forest	0.712348
CA-NS1	55.8792	-98.4839	ENF	Forest	0.712252
CA-NS2	55.9058	-98.5247	ENF	Forest	0.715733
CA-NS3	55.9117	-98.3822	ENF	Forest	0.721134
CA-NS4	55.9144	-98.3806	ENF	Forest	0.721903
CA-NS5	55.8631	-98.485	ENF	Forest	0.714291
CA-Oas	53.6289	-106.198	DBF	Forest	0.554919
CA-Obs	53.9872	-105.118	ENF	Forest	0.566818
CA-SF1	54.485	-105.818	ENF	Forest	0.61319
CA-TP3	42.7068	-80.3483	ENF	Forest	0.991783
CA-TPD	42.6353	-80.5577	DBF	Forest	0.992744

CG-Tch	-4.2892	11.6564	SAV	Savannas	1.078185
CH-Cha	47.2102	8.4104	GRA	Grassland	1.328297
CH-Dav	46.8153	9.8559	ENF	Forest	1.258241
CH-Lae	47.4783	8.3644	MF	Forest	1.431653
CH-Oe2	47.2864	7.7337	CRO	Cropland	1.268436
CL-SDF	-41.883	-73.676	EBF	Forest	3.0706
CN-Cha	42.4025	128.0958	MF	Forest	0.763632
CN-Cng	44.5934	123.5092	GRA	Grassland	0.314945
CN-Dan	30.4978	91.0664	GRA	Grassland	0.30837
CN-Din	23.1733	112.5361	EBF	Forest	1.450652
CN-Du2	42.0467	116.2836	GRA	Grassland	0.32201
CN-Du3	42.0551	116.2809	GRA	Grassland	0.316103
CN-HaM	37.37	101.18	GRA	Grassland	0.587822
CN-Sw2	41.7902	111.8971	GRA	Grassland	0.15036
CZ-BK1	49.5021	18.5369	ENF	Forest	1.43765
DE-Geb	51.0997	10.9146	CRO	Cropland	0.584699
DE-Hai	51.0792	10.4522	DBF	Forest	1.094869
DE-Kli	50.8931	13.5224	CRO	Cropland	1.002298
DE-Lkb	49.0996	13.3047	ENF	Forest	1.944474
DE-Lnf	51.3282	10.3678	DBF	Forest	1.18699
DE-Obe	50.7867	13.7213	ENF	Forest	1.059658
DE-RuR	50.6219	6.3041	GRA	Grassland	1.377442
DE-RuS	50.8659	6.4471	CRO	Cropland	0.814898
DE-Seh	50.8706	6.4497	CRO	Cropland	0.813893
DE-Tha	50.9626	13.5651	ENF	Forest	0.885046
DK-Fou	56.4842	9.5872	CRO	Cropland	1.013091
ES-Amo	36.8336	-2.2523	SHR	Shrubland	0.128633
ES-LJu	36.9266	-2.7521	SHR	Shrubland	0.368305
FI-Hyy	61.8474	24.2948	ENF	Forest	1.060804
FI-Let	60.6418	23.9595	ENF	Forest	0.973933
FR-Fon	48.4764	2.7801	DBF	Forest	0.684879
FR-Gri	48.8442	1.9519	CRO	Cropland	0.674678
FR-Pue	43.7413	3.5957	EBF	Forest	0.48814
GF-Guy	5.2788	-52.9249	EBF	Forest	1.805171
GH-Ank	5.2685	-2.6942	EBF	Forest	1.279149
GL-ZaH	74.4733	-20.5503	GRA	Grassland	0.695722
IT-BCi	40.5237	14.9574	CRO	Cropland	0.802622
IT-CA2	42.3772	12.026	CRO	Cropland	0.245579
IT-Col	41.8494	13.5881	DBF	Forest	0.753786
IT-Cp2	41.7043	12.3573	EBF	Forest	0.596353
IT-Cpz	41.7052	12.3761	EBF	Forest	0.600799
IT-La2	45.9542	11.2853	ENF	Forest	0.544502
IT-Lav	45.9562	11.2813	ENF	Forest	0.536086
IT-Ren	46.5869	11.4337	ENF	Forest	0.772143

IT-SR2	43.732	10.2909	ENF	Forest	0.753013
IT-SRo	43.7279	10.2844	ENF	Forest	0.749533
IT-Tor	45.8444	7.5781	GRA	Grassland	1.503514
JP-MBF	44.3869	142.3186	DBF	Forest	1.635463
JP-SMF	35.2617	137.0788	MF	Forest	1.648311
MY-PSO	2.973	102.3062	EBF	Forest	1.322133
NL-Loo	52.1666	5.7436	ENF	Forest	1.085121
RU-Cok	70.8291	147.4943	SHR	Shrubland	0.58537
RU-Fyo	56.4615	32.9221	ENF	Forest	0.948624
RU-Ha1	54.7252	90.0022	GRA	Grassland	0.553207
US-Akn	33.3825	-81.5653	MF	Forest	0.7577
US-AR1	36.4267	-99.42	GRA	Grassland	0.297143
US-AR2	36.6358	-99.5975	GRA	Grassland	0.302365
US-ARM	36.6058	-97.4888	CRO	Cropland	0.491714
US-Bar	44.0646	-71.2881	DBF	Forest	1.2812
US-Bi1	38.0992	-121.499	CRO	Cropland	0.222
US-Bi2	38.1091	-121.535	CRO	Cropland	0.2291
US-Blo	38.8953	-120.633	ENF	Forest	0.896422
US-BZS	64.6963	-148.324	ENF	Forest	0.4647
US-CF1	46.7815	-117.082	CRO	Cropland	0.4916
US-CF2	46.784	-117.091	CRO	Cropland	0.4916
US-CF3	46.7551	-117.126	CRO	Cropland	0.4736
US-CF4	46.7518	-117.129	CRO	Cropland	0.4736
US-Cop	38.09	-109.39	GRA	Grassland	0.189553
US-CRT	41.6285	-83.3471	CRO	Cropland	0.742022
US-Cst	33.0442	-91.9204	ENF	Forest	0.9489
US-DFC	43.3448	-89.7117	CRO	Cropland	0.7681
US-EML	63.8784	-149.254	SHR	Shrubland	0.7239
US-Fwf	35.4454	-111.772	GRA	Grassland	0.2881
US-Ha1	42.5378	-72.1715	DBF	Forest	1.099351
US-Ho1	45.2041	-68.7402	ENF	Forest	1.1078
US-Ho2	45.2091	-68.747	ENF	Forest	1.1229
US-IB2	41.8406	-88.241	GRA	Grassland	0.812575
US-ICh	68.6068	-149.296	SHR	Shrubland	0.5515
US-ICt	68.6063	-149.304	SHR	Shrubland	0.5566
US-Jo1	32.582	-106.635	SHR	Shrubland	0.1249
US-Jo2	32.5849	-106.603	SHR	Shrubland	0.1249
US-KFS	39.0561	-95.1907	GRA	Grassland	0.6859
US-KLS	38.7745	-97.5684	GRA	Grassland	0.4507
US-Kon	39.0824	-96.5603	GRA	Grassland	0.5742
US-Lin	36.3566	-119.092	CRO	Cropland	0.1554
US-Me2	44.4526	-121.559	ENF	Forest	0.717831
US-Me3	44.3154	-121.608	ENF	Forest	0.412902
US-Me5	44.4372	-121.567	ENF	Forest	0.633482

US-Me6	44.3233	-121.608	ENF	Forest	0.413754
US-MMS	39.3232	-86.4131	DBF	Forest	0.929459
US-Mo1	39.2298	-92.1167	CRO	Cropland	0.748
US-Mo3	39.2322	-92.1493	CRO	Cropland	0.748
US-MOz	38.7441	-92.2	DBF	Forest	0.7403
US-Myb	38.0499	-121.765	WET	Wetland	0.201679
US-NC1	35.8118	-76.7119	ENF	Forest	0.8775
US-NC3	35.799	-76.656	ENF	Forest	0.8902
US-Ne1	41.1651	-96.4766	CRO	Cropland	0.580269
US-Ne2	41.1649	-96.4701	CRO	Cropland	0.581517
US-Ne3	41.1797	-96.4397	CRO	Cropland	0.579032
US-Oho	41.5545	-83.8438	DBF	Forest	0.761
US-OWC	41.3795	-82.5125	WET	Wetland	0.7759
US-PFa	45.9459	-90.2723	MF	Forest	0.905195
US-Ro1	44.7143	-93.0898	CRO	Cropland	0.7004
US-Ro2	44.7288	-93.0888	CRO	Cropland	0.7004
US-Ro5	44.691	-93.0576	CRO	Cropland	0.711
US-Ro6	44.6946	-93.0578	CRO	Cropland	0.711
US-Seg	34.3623	-106.702	GRA	Grassland	0.1095
US-Ses	34.3349	-106.744	SHR	Shrubland	0.1131
US-Sne	38.0369	-121.755	GRA	Grassland	0.2113
US-SP1	29.7381	-82.2188	ENF	Forest	0.8091
US-SRC	31.9083	-110.84	SHR	Shrubland	0.155234
US-SRG	31.7894	-110.828	GRA	Grassland	0.22216
US-Syv	46.242	-89.3477	MF	Forest	0.955047
US-Ton	38.4309	-120.966	SAV	Savannas	0.342592
US-Tw2	38.0969	-121.637	CRO	Cropland	0.21349
US-Tw3	38.1152	-121.647	CRO	Cropland	0.217776
US-Twt	38.1087	-121.653	CRO	Cropland	0.217052
US-UMB	45.5598	-84.7138	DBF	Forest	0.818674
US-UMd	45.5625	-84.6975	DBF	Forest	0.81652
US-Var	38.4133	-120.951	GRA	Grassland	0.335051
US-WCr	45.8059	-90.0799	DBF	Forest	0.921557
US-Whs	31.7438	-110.052	SHR	Shrubland	0.149825
US-Wi0	46.6188	-91.0814	ENF	Forest	0.845314
US-Wi1	46.7305	-91.2329	DBF	Forest	0.842827
US-Wi2	46.6869	-91.1528	ENF	Forest	0.868687
US-Wi3	46.6347	-91.0987	DBF	Forest	0.86199
US-Wi4	46.7393	-91.1663	ENF	Forest	0.847938
US-Wi5	46.6531	-91.0858	ENF	Forest	0.853917
US-Wi8	46.7223	-91.2524	DBF	Forest	0.850212
US-Wi9	46.7385	-91.0746	ENF	Forest	0.858191
US-Wkg	31.7365	-109.942	GRA	Grassland	0.169065
US-xAB	45.7624	-122.33	ENF	Forest	2.2572

US-xAE	35.4106	-99.0588	GRA	Grassland	0.3495
US-xBR	44.0639	-71.2873	DBF	Forest	1.2812
US-xCL	33.4012	-97.57	GRA	Grassland	0.4702
US-xCP	40.8155	-104.746	GRA	Grassland	0.2149
US-xDC	47.1617	-99.1066	GRA	Grassland	0.3745
US-xDL	32.5417	-87.8039	MF	Forest	0.9132
US-xGR	35.689	-83.5019	DBF	Forest	1.0297
US-xHA	42.5369	-72.1727	DBF	Forest	1.0855
US-xHE	63.8757	-149.213	SHR	Shrubland	0.709
US-xJR	32.5907	-106.843	SHR	Shrubland	0.1169
US-xKA	39.1104	-96.6129	GRA	Grassland	0.5704
US-xKZ	39.1008	-96.5631	GRA	Grassland	0.5742
US-xML	37.3783	-80.5248	DBF	Forest	0.9891
US-xNG	46.7697	-100.915	GRA	Grassland	0.3338
US-xSE	38.8901	-76.56	DBF	Forest	0.8226
US-xSR	31.9107	-110.836	SHR	Shrubland	0.1422
US-xST	45.5089	-89.5864	DBF	Forest	0.8892
US-xTA	32.9505	-87.3933	ENF	Forest	0.9781
US-xTR	45.4937	-89.5857	DBF	Forest	0.8823
US-xUN	46.2339	-89.5373	MF	Forest	0.9666
US-xWD	47.1282	-99.2414	GRA	Grassland	0.3706

TableS2. Global datasets used in the CKML-GPP model for the production of EGO GPP. Here, we displayed the original resolution of all the data source. In the study, we unified the temporal resolution to hourly and the spatial resolution to 0.05°. Additionally, we reaffirm the abbreviations each feature: incoming shortwave radiation (SW), enhanced vegetation index (EVI), air temperature (Ta), land surface water index (LSWI), vapor pressure deficit (VPD), aridity index (Aridity), diffuse radiation fraction (DRF), atmospheric CO_2 concentration (CO2) and soil water content (SWC). FCR, FFO, FSA, FGR, FEW, and FSH denoting as the fraction of Cropland, Forest, Savannas, Grassland, Wetland and Shrubland in every 0.05° grid.

Variable	**Year**	**Temporal**	**Spatial**	**Resource**
Aridity	\	\	0.05°	Global-AI_PET_v3
Ta	2000-2022	Hourly	0.1°	ERA5-Land
VPD	2000-2022	Hourly	0.1°	ERA5-Land
SWC	2000-2022	Hourly	0.1°	ERA5-Land
SW	2000-2022	Hourly	0.1°	ERA5-Land
CO2	2000-2022	3-hourly	3°, 2°	CarbonTracker
EVI	2000-2022	Daily	0.05°	MCD43C4
LSWI	2000-2022	Daily	0.05°	MCD43C4
DRF	2000-2022	Hourly	0.5°,0.625°	MERRA2
FFO	2000-2022	Yearly	0.05°	MCD12Q1
FSH	2000-2022	Yearly	0.05°	MCD12Q1
FSA	2000-2022	Yearly	0.05°	MCD12Q1
FGR	2000-2022	Yearly	0.05°	MCD12Q1
FCR	2000-2022	Yearly	0.05°	MCD12Q1
FWE	2000-2022	Yearly	0.05°	MCD12Q1

TableS3. Multicollinearity test used the variance inflation factor (VIF). F means VIF < 10, indicating satisfy the collinearity criterion and could be used in subsequent modeling.

ID	Feature	Collinearity
1	CO2	F
2	FCR	F
3	SWC	F
4	FFO	F
5	FWE	F
6	Aridity	F
7	FSA	F
8	DRF	F
9	SW	F
10	FGR	F
11	FSH	F
12	LSWI	F
13	EVI	F
14	VPD	F
15	Ta	F

References

Zerubavel, Eviatar. “The Standardization of Time: A Sociohistorical Perspective.” American Journal of Sociology, vol. 88, no. 1, 1982, pp. 1–23. JSTOR, http://www.jstor.org/stable/2779401.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G.: Overview of the radiometric and biophysical performance of the MODIS vegetation indices, Remote Sens. Environ., 83, 195-213, https://doi.org/DOI 10.1016/S0034-4257(02)00096-2, 2002.

Li, W., Yao, Z., Qu, Y., Yang, H., Song, Y., Song, L., Wu, L., and Cui, Y.: A benchmark dataset for global evapotranspiration estimation based on FLUXNET2015 from 2000 to 2022, Earth Syst. Sci. Data, 17, 3835-3855, https://doi.org/10.5194/essd-17-3835-2025, 2025.

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, Remote Sens., 11, 21, https://doi.org/10.3390/rs11212563, 2019.

Yuan, W., Zheng, Y., Piao, S., Ciais, P., Lombardozzi, D., Wang, Y., Ryu, Y., Chen, G., Dong, W., Hu, Z., Jain, A. K., Jiang, C., Kato, E., Li, S., Lienert, S., Liu, S., Nabel, J. E. M. S., Qin, Z., Quine, T., Sitch, S., Smith, W. K., Wang, F., Wu, C., Xiao, Z., and Yang, S.: Increased atmospheric vapor pressure deficit reduces global vegetation growth, Sci. Adv., 5, 12, https://doi.org/10.1126/sciadv.aax1396, 2019.

Zeng, Y., Hao, D., Huete, A., Dechant, B., Berry, J., Chen, J. M., Joiner, J., Frankenberg, C., Bond-Lamberty, B., Ryu, Y., Xiao, J., Asrar, G. R., and Chen, M.: Optical vegetation indices for monitoring terrestrial ecosystems globally, Nat. Rev. Earth Environ., 3, 477-493, https://doi.org/10.1038/s43017-022-00298-5, 2022.

Zhang, Q., Wei, S., Chu, X., Zhang, X., Gong, Z., Wang, X., Song, W., Song, Y., Xie, B., and Han, G.: Climate and vegetation jointly determine the interannual variation of net ecosystem CO2 fluxes over 12 years in a restored coastal wetland, Agric. For. Meteorol., 373, 10, https://doi.org/10.1016/j.agrformet.2025.110760, 2025.