A daily and 500m coupled evapotranspiration and gross primary production product across China during 2000-2020

Shaoyang He[1,2], Yongqiang Zhang[1*], Ning Ma[1], Jing Tian[1], Dongdong Kong[3], Changming Liu[1]

[1]Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China
[2]University of Chinese Academy of Sciences, Beijing 100040, China
[3]Department of Atmospheric Science, School of Environmental Studies, China University of Geosciences, Wuhan 430074, China

Correspondence to: Yongqiang Zhang (zhangyq@igsnrr.ac.cn)

Abstract. Accurate high-resolution actual evapotranspiration (ET) and gross primary production (GPP) information is essential for understanding the large-scale water and carbon dynamics. However, substantial uncertainties exist in the current ET and GPP datasets in China because of insufficient local ground measurements used for model constrain. This study utilizes a water-carbon coupled model, Penman-Monteith-Leuning Version 2 (PML-V2), to estimate 500m ET and GPP at a daily scale. The parameters of PML-V2(China) were well-calibrated against observations of 26 eddy covariance flux towers across nine plant functional types in China, indicated by Nash-Sutcliffe Efficiency (*NSE*) of 0.75 and Root Mean Square Error (*RMSE*) of 0.69 mm d^{-1} for daily ET respectively, and *NSE* of 0.82 and *RMSE* of 1.71 g C m^{-2} d^{-1} for daily GPP. The model estimates get a small *bias* of 6.28% and a high *NSE* of 0.82 against water-balance annual ET estimates across 10 major river basins in China. Further evaluations suggest that the newly developed product is better than other typical products (MOD16A2, SEBAL, GLEAM, MOD17A2H, VPM, and EC-LUE) in estimating both ET and GPP. Moreover, PML-V2(China) accurately monitors the intra-annual variations in ET and GPP in the croplands with a dual-cropping system. The new data showed that during 2001-2018, the annual GPP and water use efficiency experienced a significant (*p* < 0.001) increase (8.99 g C m^{-2} yr^{-2} and 0.02 g C mm^{-1} H$_2$O yr^{-1}, respectively), but annual ET showed a non-significant (*p* > 0.05) increase (0.43 mm yr^{-2}). This indicates that vegetation in China exhibits a huge potential for carbon sequestration with little cost in water resources. The PML-V2(China) product provides a great opportunity for academic communities and various agencies for scientific studies and applications, freely available at http://dx.doi.org/10.11888/Terre.tpdc.272389 (Zhang and He, 2022).

1 Introduction

Terrestrial evapotranspiration (ET) and photosynthesis (or gross primary productivity, GPP) are indispensable processes in hydrological and carbon cycles at global and regional scales. Forming the second largest water flux after precipitation in the terrestrial hydrological cycle, ET is the sum of plant transpiration (E$_c$), evaporation from the soil (E$_s$), and canopy evaporation

30   from precipitation interception (E_i). Photosynthesis rate, tightly coupled to E_c by leaf stomata, is a key indicator of plant growth and provides food and fibre for human society.

    In recent decades, numerous studies have been carried out to map ET at regional, continental, and global scales. Particular credit goes to remote sensing (RS)-based models that provide diagnostic ET estimates with a relatively high spatiotemporal continuity and reasonable biophysical significance. For example, Miralles et al. (2011b) and Martens et al. (2017) developed

35   a global and daily ET product employing the Global Land Evaporation Amsterdam Model 3.0a (hereafter GLEAM) based on the Priestley-Taylor (P-T) equation. Although GLEAM does a good job in temporal resolution, it has a coarse spatial resolution of 0.25°. Another widely used semi-empirical formula is Penman-Monteith (PM) equation. Mu et al. (2011) generated the MOD16A2 ET with 500m and 8-day resolutions based on PM equation as one global product of Moderate Resolution Imaging Spectroradiometer (MODIS). Leuning et al. (2008) developed the Penman-Monteith-Leuning (PML) model that described the

40   physical characteristics of canopy-soil water loss by improving the surface conductance (G_s) formulations and Zhang et al. (2010) estimated PML-based ET at 0.05° and 8-day resolution across the Australian continent. Cheng et al. (2021) produced a 1-km and daily ET dataset across China by the Surface Energy Balance Algorithm for Land (SEBAL), but only evaluated the SEBAL product at eight EC sites of three land cover types.

    Apart from methods and products for the ET estimated above, numerous approaches have been used to estimate GPP, such as

45   the enzyme kinetic (process-based) models (Houborg et al., 2013; Arain et al., 2006; Grant et al., 2005; Hanson et al., 2004; Medvigy et al., 2009), light use efficiency (LUE) principle (Liu et al., 2003; Turner et al., 2003; Yuan et al., 2007; Running et al., 2015; Zhang et al., 2017; Zheng et al., 2020), statistical methods (Potter et al., 1993; Hilker et al., 2008; Zhang et al., 2020b) and machine learning methods (Wolanin et al., 2019; Joiner and Yoshida, 2020; Huang et al., 2021). Among them, the LUE principle is well known because of its simple structure, strong portability, and relatively high temporal cover of inputs. Running

50   et al. (2015) recently updated the global GPP product of MODIS (hereafter MOD17A2H) at 500m and 8-day resolutions using the LUE principle. Zhang et al. (2017) and Zhang et al. (2021) also mapped global GPP for 2000 - 2019 dubbed Vegetation Photosynthesis Model GPP V20 (hereafter VPM) with the same spatiotemporal resolution as MOD17A2H using an improved LUE model adding the energy absorbed by chlorophyll. Zheng et al. (2020) generated a global GPP dataset at 0.05° and 8-day intervals by a revised LUE model (hereafter EC-LUE) integrating the atmospheric CO_2 concentration.

55   Although significant efforts have been put into estimating ET and GPP, there are barely any coupled products available in China, which meet the requirement of high temporal ($\leq$ 1 day) and high spatial ($\leq$ 500m) resolutions simultaneously that is necessary to detect variations of the eco-hydrological cycle in diverse and large areas for a long term precisely (Table 1). For instance, products with low temporal resolutions are erratic to detect subtle seasonal changes in areas seriously affected by human activities and in arid regions, such as irrigated farmland with a dry climate (Bodner et al., 2015) and an evergreen

60   broad-leaf Mediterranean forest during severe summer drought (Liu et al., 2015). On the other hand, products with high temporal resolutions like GLEAM can monitor the diurnal variability of ET, but their low spatial resolutions limit the effectiveness in fine-scale environment applications (Gevaert and García-Haro, 2015).

Secondly, the phenomenon of ignoring the water-carbon coupling process frequently appearing in the existing products has brought systematic errors. The photosynthesis and transpiration are coupled by the plant stomatal control on both water and

65 carbon exchange between the land ecosystem and the atmosphere (Xiao et al., 2013; Zhang et al., 2019). As indicated in Table 1, MODIS ET and MODIS GPP products are independent of each other, and cannot ensure similar biophysical characteristics of vegetation in the same place. Furthermore, using ET and GPP from different products can lead to large uncertainty in analysing the interaction between ET and GPP such as water use efficiency of ecosystem (WUE is the ratio of GPP to ET). In that case, it is necessary to build a coupled ET and GPP model considering the water-carbon coupling process. Zhang et al.

70 (2019) developed the second generation of PML (i.e., PML-V2 model) that estimates G_s using a water-carbon coupled model and mapped global ET and GPP at 500m and 8-day resolutions in 2002-2017.

More importantly, previous studies utilized sparse ground observations in China (as shown in Table 1) that covered few terrestrial ecosystems for models' calibration and validation, resulting in improper input parameters, and making it difficult to obtain more reliable estimates of ET and GPP in diverse landcover types (Heinsch et al., 2006, Anon, 2013). Although the

75 eddy-covariance (EC) flux sites have provided consecutive measurements of water and carbon fluxes since the early 1990s (Xiao et al., 2013; Wofsy et al., 1993; Baldocchi et al., 2001), the EC observed data in China remain much sparser than those in North America and Europe, and most of them are not publicly available, impeding a national scale constraint of RS-based models for improving ET and GPP estimates (Villarreal and Vargas, 2021; Chu et al., 2017). For instance, GLEAM which only employed 8 EC sites over China overestimates ET at a large scale, especially for evergreen needleleaf forest, evergreen

80 broadleaf forest and mixed forest (Li et al., 2018). Therefore, the uncertainty in estimating ET and GPP is large, and it requests sufficient EC flux sites to calibrate and validate ET and GPP models for better local and regional applications.

In addition, understanding the spatial and temporal pattern of ET and GPP is particularly important for China, the largest contributor to the absolute growth of greenhouse gases that directly induces global warming over the past decade (Minx et al., 2021). On the other hand, China owns huge carbon sequestration potential of terrestrial ecosystems for slowing accumulation

85 of atmospheric carbon dioxide and mitigating climate change. Additionally, given China's water shortage, it is crucial to clearly understand water budgets and transportation (Ma et al., 2020). Therefore, it becomes vital to estimate GPP and ET accurately across China under changes in climate and land cover types.

Facing the above challenges, this study utilizes a water-carbon coupled and remote sensing-based model, PML-V2 that is constrained against the most comprehensive observations in China (i.e. , 26 EC observations across nine plant functional types

90 (PFTs)), to generate daily and 500m ET and GPP gridded products from 2000 to 2020. We then test whether the newly developed product for China outperforms PML-V2(global) and other mainstream products (i.e., MOD16A2, SEBAL, GLEAM, MOD17A2H, VPM, and EC-LUE), and investigate annual change and spatial pattern in ET and its components: plant transpiration (E_c), evaporation from the soil (E_s), and canopy evaporation from precipitation interception (E_i) , GPP and WUE in 2001-2018 across China. It is noted that that the years of 2019-2020 have not been selected for trend analysis. This is because

95 that a different forcing dataset was used to drive PML-V2 (details are provided in section 2.2.2).

The novelties of this study mainly include:

(i) Observation data from 26 EC flux stations in nine PFTs across China are employed for constraining the PML-V2 calibration for estimating ET and GPP;

(ii) The country-specific meteorological forcing, i.e., the China Meteorological Forcing Dataset (CMFD), is used to drive the PML-V2 in China, which is more accurate than those forcings extracted from global forcing products;

(iii) The PML-V2(China) product is generated with a daily resolution, compared to the previous global product with a temporal resolution of 8-day; and

(v) Improving intra-annual ET and GPP dynamics for various ecosystems, particularly for the cropland ecosystem, which provides more accurate estimates and monitoring of agricultural water consumption, compared to other mainstream products.

2 Materials and methods

2.1 Description of the PML-V2 model

PML-V2 model is a water–carbon coupled diagnostic biophysical model. Compared to the old version that does not calculate GPP and the effect of CO_2 on evapotranspiration, PML-V2 couples a photosynthesis model (Farquhar et al., 1980) and an improved canopy stomatal conductance model (Yu et al., 2004) with the Penman–Monteith (P-M) equation to estimate GPP and E_c collectively (Gan et al., 2018). Zhang et al. (2019) further improved PML-V2 by incorporating the vapor pressure deficit constraint to GPP that is then used to constrain canopy conductance and E_c. The detailed descriptions of PML-V2 are provided in the supplement.

2.2 Model input data

2.2.1 Remote sensing data

MODIS collections with 500m spatial resolutions from February 26, 2000 to December 31, 2020 (hereafter 2000-2020) are used, which includes the LAI from MOD15A2H.006 (Myneni et al., 2015), the albedo from MCD43A3.006 product (Schaaf and Wang, 2015), and the surface emissivity from MOD11A2.006 (Wan et al., 2015). The smoothed LAI inputs for PML-V2(China) utilized the weighted Whittaker smoother with dynamic lambdas instead of a constant lambda to eliminate missing or unreliable pixels due to noise contaminations by snow, shadow, cloud, etc, compared with the PML-V2(Global) (Kong et al., 2019; Zhang et al., 2019). The improved LAI can better express peaks and seasonal changes. The albedo and surface emissivity inputs were gap-filled by the linear interpolation of the nearest good quality points (Zhang et al., 2019). If there were not enough good-quality points close to the point with missing value, it was filled by the historical averaged values for the same grid. Besides, the PML-V2 model needs land cover types to accurately estimate ET and GPP in different terrestrial ecosystems. Here we used the International Geosphere-Biosphere Program (IGBP) layer of MCD12Q1.006 land cover product (Sulla-Menashe et al., 2019) during 2000-2020 since IGBP classification is annually continuous and has acceptable accuracy in China when compared with other land cover products (Feng and Bai, 2019).

2.2.2 Meteorological data

The meteorological inputs of PML-V2 include specific humidity, air pressure, air temperature, wind speed, precipitation, downward longwave radiation, downward shortwave radiation, and land surface temperature. In this study, the main meteorological data came from the CMFD from 2000 to 2018 (He et al., 2020), which has spatial and temporal resolutions of 0.1° and 3-hr, respectively. Generated through the fusion of five RS or reanalysis datasets and 753 China Meteorological Administration stations, the CMFD dataset shows the best accuracy among most available meteorological datasets, and is widely employed in hydrological and land surface modelling of China (Zhang et al., 2019; Wang et al., 2020). The Global Land Data Assimilation System Version 2.1 (GLDAS-2.1) forcing data with 0.25° and 3-hr resolutions (Beaudoing and Rodell, 2016) were used for 2019-2020, as the CMFD was not available during such a period. Since the CMFD and the GLDAS-2.1 are different meteorological forcing datasets for driving the PML-V2 model, it is necessary to correct bias of 2019-2020. Here, a widely used methodology: delta change (i.e., DC, also called change factor), was selected for bias-correction (Anandhi et al., 2011; Teutschbein and Seibert, 2012; Rasmussen et al., 2012; Hempel et al., 2013; Beck et al., 2018; Haro-Monteagudo et al., 2020). The underlying idea of the DC method is to use simulated future anomalies (i.e., GLDAS-2.1 in this study) for a perturbation of observed data (i.e., CMFD) rather than to use the simulations of future conditions directly. For each grid cell, we bias-corrected the daily meteorological data during 2019-2020 by monthly scaling factors. Specifically, A multiplicative approach applied for precipitation (Eqs.1 to 2).

$$P_{(i,j)} = \Delta P_{(j)} \cdot P^*_{(i,j)}; i = 1,2,...,31; j = 1,2,...,12 \tag{1}$$

where $P_{(i,j)}$ is the precipitation corrected by the relative change factor for day i and month j; and $P^*_{(i,j)}$ is the multi-year daily mean observed precipitation at the ith day, jth month in the historical or reference period (2000-2018 in this study). $\Delta P_{(j)}$ is the change factor that can be formulated as:

$$\Delta P_{(j)} = \overline{P_{fut}} / \overline{P_{his}} ; j = 1,2,...,12 \tag{2}$$

where $\overline{P_{his}}$ is the daily mean precipitation of the jth month in the historic simulation (i.e., 2000-2018) and $\overline{P_{fut}}$ is similar but the average of the future simulation (i.e., 2019 or 2020). The multiplicative approach is also used for the variables: radiation, air pressure, and wind speed, whereas an additive approach was used to adjust temperature and specific humidity (Eqs 3 to 4).

$$X_{(i,j)} = \Delta X_{(j)} + X^*_{(i,j)}; i = 1,2,...,31; j = 1,2,...,12 \tag{3}$$

in which $\Delta X_{(j)}$ is calculated as:

$$\Delta X_{(j)} = \overline{X_{fut}} - \overline{X_{his}} ; j = 1,2,...,12 \tag{4}$$

The land surface temperature during 2000-2020 was from the Land component of the fifth generation of European ReAnalysis (i.e., ERA5-Land; Muñoz-Sabater et al., 2021) with spatial and temporal resolutions of 0.1° and 1-hr, respectively. Note that the above meteorological data were first aggregated into the daily scale, followed by resampling into 500m by the bilinear interpolation method (Zhang et al., 2019). The atmospheric CO_2 concentration data came from the National Oceanic and

Atmospheric Administration. The pro-processing of remote sensing data and meteorological data for model inputs is summarised in Figure 1.

160 **2.3 ET and GPP from eddy-covariance observations**

We collated EC flux towers and automatic weather stations (AWSs) data from 26 sites across China (Fig. 2 and Table 2) and generated the high-quality ET and GPP observed for calibration and validation of PML-V2. These data came from the following sources: FLUXNET2015 (Pastorello et al., 2020), the National Tibetan Plateau Data Center (Ma et al., 2020), the Heihe Integrated Observatory Network (Liu et al., 2011; Liu et al., 2018), and the Chinese Terrestrial Ecosystem Flux Research

165 Network (ChinaFLUX) (Yu et al., 2006). These 26 sites encompass nine major PFTs in China, including two in evergreen needleleaf forests, one in evergreen broadleaf forests, one in mixed forests, one in open shrublands, one in savannas, eight in grasslands, three in wetlands, seven in croplands, and two in barren sparse vegetation. The observation variables, including air temperature, relative humidity, incoming shortwave radiation, latent heat flux, sensible heat flux, and net ecosystem exchange, were collected from the interval of 0.5-hour or 1-hour to prepare for gap filling and flux partitioning (Reichstein et al., 2005;

170 Wutzler et al., 2018). Considering certain gaps exist in the original half-hourly or hourly latent heat (LE) and net ecosystem exchange (NEE) fluxes data, we employed the marginal distribution sampling method (Reichstein et al., 2005) to fill these gaps using the station-observed air temperature, relative humidity, and solar radiation data. Subsequently, we partitioned NEE into gross carbon uptake (GPP) and respiration of ecosystem according to the night-time-based method of Reichstein et al. (2005). Because any gap-filling of EC data may introduce extra uncertainties, we only used the days during which the

175 percentage of the original observed and good-quality gap-filled data was no less than 60% in the present study. Note that the data of the sites from the ChinaFLUX (i.e., CN-CBF, CF-HBG_S01, CF-HBG_W01, CF-NMG, CF-QYF, and CF-YCA) and the FLUXNET2015 (i.e., CN-Cng, CN-Du2, and CN-HaM) have already been gap-filled by original data providers. Therefore, they were used directly in this study. Note that while energy imbalance does exist in many EC sites, correcting such a problem may also introduce more uncertainties (Foken, 2008). Therefore, we used the observed LE directly in the present study.

180 **2.4 Basin-scale water balance-based evapotranspiration data**

The water balance method is generally regarded as a simple and accurate approach for calculating land evapotranspiration at the basin-scale (Liu et al., 2016). Here we used the water-balance-based evapotranspiration (ET_{wb}, mm) of 10 major basins across China to evaluate PML-V2 ET estimates at a basin scale, that is,

$$ET_{wb} = Prcp - Q - TWSC, \tag{5}$$

185 where $Prcp$, Q, and $TWSC$ (all with a unit of mm) are basin-wide precipitation, runoff, and change of terrestrial water storage at an annual scale, respectively. Among them, $Prcp$ and Q are the annual values of ten major river basins in China from 2003 to 2013, including the Hai, Huai, Liao, Northwest, Pearl, Songhua, Southeast, Southwest, Yangtze, and Yellow (Fig. 2), from the National Water Resources Bulletin, which is extensively used in water resources calculation (Miao et al., 2022) and assessment (Yang et al., 2004; Xie et al., 2018). $TWSC$ was quantified using three Gravity Recovery and Climate Experiment

190   (GRACE) products (Landerer and Swenson, 2012; Landerer, F. 2021) including the NASA Jet Propulsion Laboratory, the GeoForschungsZentrum Potsdam and the Center for Space Research, which are available since April 2002 at a monthly scale. For reducing the uncertainties, this study used the mean values of these three products. We regarded the differences in the terrestrial water storage anomaly between two consecutive Decembers as the annual *TWSC*. Note that the ET_{wb} was only calculated from 2003 to 2013 in the present study since the December values from GRACE was not available in 2014.

195 **2.5 Model calibration and model validation**

The 11 parameters of the PML-V2 model for each PFT were calibrated and cross-validated against 26 EC sites by a global optimisation method - genetic algorithm (GA). The GA generates a randomly initialized population, and then evaluates the fitness of solutions according to its objective function. As generations iterate, the population includes more appropriate solutions, and eventually it will converge (Holland, 1992; Konak et al., 2006). Specifically, we applied the GA algorithm with

200   a population size of 1000 and number generations of 50. All EC-observed ET and GPP data within a PFT are used to minimize the following objective function (F_{opt}):

$$F_{opt} = 2 - NSE_{ET} - NSE_{GPP} = \frac{\sum_{i=1}^{N}(ET_{est}-ET_{obs})^2}{\sum_{i=1}^{N}\left(ET_{obs}-\overline{ET_{obs}}\right)^2} + \frac{\sum_{i=1}^{N}(GPP_{est}-GPP_{obs})^2}{\sum_{i=1}^{N}\left(GPP_{obs}-\overline{GPP_{obs}}\right)^2} \qquad (6)$$

where NSE_{ET} and NSE_{GPP} are the Nash-Sutcliffe Efficiency of the daily ET and the daily GPP, respectively. The subscripts *est* and *obs* stand for the estimated and the observed, respectively. In this way, each of the nine PFTs gained a unique set with

205   11 calibrated parameter values, illustrated in Table S1.

The 'leave-one-out' cross-validation method was utilized to evaluate robustness of the PML-V2 model (Zhang et al., 2019). For each PFT, the data from one "ungauged" observation was excluded from the optimization while the data from all other observations at the same PFT were used for model calibration to obtain the simulated at the "ungauged" position. All nine PFTs were actualized in this way. Note that the PFT including EBF, MF, OSH, and SAV only has one ground site (Table 2).

210   Therefore, it is appropriate to divide the data in each of the four sites into two sub-groups for cross-validation. The CF-CBF and the CF-HBG_S01 covering from 2003 to 2010, were divided into two sub-groups, each of which had 4 years: 2003-2006 and 2007-2010. While both the BNXJL and YJGRHG only covered one year and were divided into two sub-groups by a two-day time step, separately. After that, the daily estimates in the cross-validation mode were against the daily observation from the 26 stations to explore the model transferability from known observations to any location.

215   **2.6 Model performance metrics**

We assessed the performance of calibration and cross-validation of PML-V2 (and other seven mainstream ET and GPP products) against the observed sites or water-balance basins utilizing the following four metrics:

$$NSE_X = 1 - \frac{\sum_{i=1}^{N}(X_{est}-X_{obs})^2}{\sum_{i=1}^{N}\left(X_{obs}-\overline{X_{obs}}\right)^2}, \qquad (7)$$

$$R_X = \frac{\sum_{i=1}^{N}(X_{est}-\overline{X_{est}})(X_{obs}-\overline{X_{obs}})}{\sqrt{\sum_{i=1}^{N}(X_{est}-\overline{X_{est}})^2 \times \sum_{i=1}^{N}(X_{obs}-\overline{X_{obs}})^2}}, \tag{8}$$

$$RMSE_X = \sqrt{\frac{\sum_{i=1}^{N}(X_{est}-X_{obs})^2}{N}}, \tag{9}$$

$$Bias_X = \frac{\sum_{i=1}^{N}(X_{est}-X_{obs})}{N \times \overline{X_{obs}}}, \tag{10}$$

where NSE, R, $RMSE$, and $Bias$ are the Nash-Sutcliffe Efficiency, the correlation coefficient, the Root Mean Square Error, and the ratio of the difference between the estimated and the observed to the observed average. The subscript X represents ET or GPP; the subscripts est and obs stand for the estimated and the observed, respectively.

3 Results

3.1 Model calibration and model validation

The simulated ET and GPP from the calibrated PML-V2(China) were first evaluated against EC measurements of 26 flux sites at a daily scale (Fig. 3). Overall, PML-V2(China) shows an excellent performance in estimating daily ET and daily GPP, as evidenced by the NSE (0.75 and 0.82, respectively), R (0.88 and 0.90, respectively), $RMSE$ (0.69 mm d^{-1} and 1.71 g C m^{-2} d^{-1}, respectively), and $Bias$ (-5.81% and -2.3%, respectively). For the mean values of each site, the simulated daily ET and daily GPP show higher NSE (≥ 0.87) and R (≥ 0.93) values (Fig. 3).

PML-V2(China) is only slightly degraded from calibration to cross-validation, indicated by slightly declined performance in ET and GPP (Fig. 3). For daily ET, the NSE and R values decrease by 0.06 and 0.04, respectively, from the calibration mode to the cross-validation mode. Correspondingly, the $RMSE$ and $Bias$ of ET in the cross-validation mode increase by 0.08 mm d^{-1} and 3.5%, respectively. For daily GPP, the NSE and R values in the cross-validation mode reduce by 0.11 and 0.06, respectively; the $RMSE$ and $Bias$ increase by 0.45 g C m^{-2} d^{-1} and 1.79%, respectively. A similarly slight degradation is applied to their site means. These results demonstrate that the PML-V2(China) is robust for estimating daily ET and daily GPP across large regions, and suitable for generating good quality daily ET and daily GPP data for China.

Figure 4 further summarises PML-V2(China) performance at 26 flux sites across nine PFTs. The estimates of ET and GPP from the model calibration show high consistency with the EC-observed values in all terrestrial biomes. For daily ET (Fig. 4a), the NSE values vary from the range of 0.36 ~ 0.82, the $RMSE$ 0.39 ~ 0.88 mm d^{-1}, and $Bias$ -10.09% ~ -0.21%. For daily GPP (Fig. 4b), the ranges of statistical metrics become 0.41 ~ 0.91 for NSE, 0.3 ~ 3.19 g C m^{-2} d^{-1} for $RMSE$, and -10.52% ~ 3.26% for $Bias$. In terms of cross-validation, nine PFTs all showed slight declines in the statistical metrics when compared to those in the calibration mode. For daily ET, the declines in NSE values are less than 0.14 in most PFTs except BSV and ENF, whose NSE decreased by 0.36 and 0.33, respectively. As expected, $RMSE$ values all increased to some extent in all PFTs (ranging from 0.002 to 0.305 mm d^{-1}) when compared with those in calibration mode. The $Bias$ values in the cross-validation mode were almost identical to those in the calibration mode for most PFTs except WET and ENF of which the absolute value of

Bias increased by 10.59% and 17.42%, respectively (Fig. 4a). From calibration to cross-validation, the degradation of BSV, ENF, and WET is more serious than that for the remaining PFTs, which is mainly caused by the small samples (2, 2, and 3, respectively) for ET estimates. Regarding daily GPP, the *NSE* values all degraded by less than 0.04 for most PFTs except BSV, GRA and WET where there exists 0.21 ~ 0.32 *NSE* degradation. In the meantime, the declines in *R* values are all within 0.19. Regarding *RMSE*, the increases are particularly marginal for most PFTs except WET with an increase of 1.58 g C m^{-2} d^{-1} (Fig. 4b). The above PFT tests suggest that the present PML-V2(China), with parameter values being calibrated against 26 EC flux stations, does perform satisfactorily in estimating both ET and GPP across different PFTs in China.

To investigate the model performance at each EC site, this study also compares the variations in daily ET and GPP between PML-V2(China) in calibration mode and the EC observations (Fig. 5, Fig. 6, and Table S2). Overall, the estimates from PML-V2(China) show similar amplitude and phase to the EC observations, indicating that it performs well in capturing the seasonal phase of ET and GPP in most flux sites. From Figure 5, the PML-V2(China) ET reveals a single peak in each annual cycle for most flux sites, except those with double-cropping systems such as CF-YCA, DXZ, and GTZ cropland sites. In particular, the ET values are consistent with observed data at the desert site (e.g., HZZHMZ and QZ-NAMORS) with *NSE* ranging from 0.41 to 0.48, indicating that model performs well in the sparse vegetation area. The measurements fluctuate higher than PML-V2(China) ET at peak areas for most sites. For the site mean ET, the difference values between PML-V2(China) and EC observations range from -0.46 to 0.21, with the minimum at QZ-NAMORS and the maximum at CN-HaM.

For daily GPP, the model also performs well in depicting the seasonal variation. However, in certain stations (e.g., DMCJZ) the peak GPP values within a year appear to be underestimated. In terms of the cropland flux sites, the GPP also shows double peaks within a year because of the double-cropping system (e.g., winter wheat and maize rotation), which is similar to ET. This is especially apparent for the GTZ, DXZ, and CF-YCA flux sites located in the North China Plain. For the site mean GPP, the discrepancies between the model and EC observations mainly range from -0.66 to 0.96 g C m^{-2} d^{-1} for most flux sites except CF-YCA, MYZ, and DXZ, where the differences exceed 1.3 g C m^{-2} d^{-1}.

3.2 Comparing with other products

3.2.1 Comparisons at a plot-scale using EC observed data

To explore whether ET and GPP of PML-V2(China) simulations are more accurate than the previous products, we also evaluated ET and GPP accuracy from its global version, MOD16A2, MOD17A2H, and other five widely available ET or GPP products against EC observations from the 26 sites at a daily or 8-day resolution. In this study, PML-V2(China) uses its cross-validated simulations to compare with other products instead of the calibration results to avoid introducing a priori knowledge. Additionally, to compare at a consistent time resolution, PML-V2(China) estimates in cross-validation mode need to be up-scaled to an 8-day average or remain a daily scale, depending on the temporal resolution of the comparison products. Specifically, PML-V2(China) is compared to GLEAM and SEBAL at the daily scale, compared to PML-V2(Global), MOD16A2, MOD17A2H, EC-LUE, and VPM at the 8-day scale.

Table 3 provides a direct comparison of model performance among varieties of ET or GPP products against observations overall from 26 ground stations. It is evident that PML-V2(China) excels other state-of-the-art ET or GPP products, presented by *NSE* being 0.12 ~ 7.76 higher for ET and 0.07 ~ 0.79 higher for GPP, *R* being 0.07 ~ 0.68 higher for ET and 0.04 ~ 0.51 higher for GPP, and *RMSE* being 0.15 ~ 3.62 mm d^{-1} lower for ET and 0.24 ~ 1.98 g C m^{-2} d^{-1} lower for GPP. Specifically, at a daily scale, PML-V2(China) ET exhibits the highest *NSE* value of 0.66, followed by GLEAM (0.44) and SEBAL (-7.10). PML-V2(China) daily ET achieves the highest *R* (0.84), followed by GLEAM (0.69) and SEBAL (0.16); correspondingly it obtains the smallest *RMSE* (0.33 mm d^{-1}), followed by GLEAM (1.04 mm d^{-1}) and SEBAL (3.94 mm d^{-1}). SEBAL is the worst performer, although its *Bias* of the closest to 0 (Table 3) because it is far away from the observed values significantly yielding a *Bias* value over 50% or less than -50% among five PFTs (Fig. 7). At the 8-day scale, PML-V2(China) outperforms PML-V2(Global) and MOD16A2 for estimating ET, with the highest value of *NSE* (0.74), *R* (0.87), and the lowest *RMSE* (0.66 mm d^{-1}). Moreover, PML-V2(China) has also the best performance in estimating 8-daily GPP, followed by PML-V2(Global), MOD17A2H, VPM, and EC-LUE, indicated by three statistics: *NSE*, *R*, and *RMSE*. In summary, PML-V2(China) performs well when compared with other mainstream ET or GPP products in China.

Figure 7 displays the performance comparison of four ET products with PML-V2(China) under nine PFTs. The simulated ET by PML-V2(China) has greater *NSE* and *R*, and less *RMSE* values than the other four products in most PFTs, especially in EBF, SAV and WET. All models have poor performance with *NSE* being lower than 0 except for PML-V2(China) in EBF and SAV. But PML-V2(China) is not the best in both ENF and BSV. In BSV, most models perform poorly, rendered by *NSE* being lower than 0 except for GLEAM (0.50) and PML-V2(China) (0.07 for the daily scale; 0.11 for the 8-day scale). And only SEBAL achieves worse results than PML-V2(China) in ENF. As shown in Fig. 8, PML-V2(China) performs significantly better than other advanced products in simulating the GPP of CRO, MF, ENF, EBF, SAV, and BSV, producing higher *NSE*, *R*, and lower *RMSE* and *Bias*. While PML-V2(China) ranked second in GRA (following MOD17A2H), OSH (following PML-V2(Global)), and WET (following PML-V2(Global)). Synthetically, PML-V2(China) successfully captures the sites' seasonality in most PFTs compared to the high-resolution ET/GPP datasets currently available.

3.2.2 Comparisons at the basin-scale using ET$_{wb}$

In addition to testing the model at a plot-scale, Figure 9 (a-e) presents the ET validations from PML-V2(China) and four ET products based on RS against the annual ET$_{wb}$ in the 10 major river basins of China during 2003-2013. It illustrates that PML-V2(China) shows the best performance among them, as indicated by the highest *NSE* (0.82) and the lowest *RMSE* (69.59 mm yr^{-1}) and *Bias* (6.28%) values. This is closely followed by the GLEAM and PML-V2(global) with *NSE* values of 0.36 and 0.26, respectively. However, MOD16A2 and SEBAL tend to overestimate ET in the majority of basins with much smaller *NSE* values of -0.21 and 0.02, respectively, which are consistent with the performance evaluations using the EC observation shown in Fig.7. Above basin-wide evaluations, together with the plot-scale validations against EC observation data, demonstrate PML-V2(China) overall performing best among the tested products selected in this study.

Fig. 9 (f) illustrates an inter-basin comparison of the 11-year mean ET of 2003–2013 within the five ET products. PML-V2(China) performs well in most basins (*Bias*es within ±15%) except for the Northwest and Southwest Basins, where ET is overestimated by 24.96% and 61.57%, respectively. Even so, PML-V2(China) performs best in the Southeast River Basin with *Bias* of -3.80%, which is still better than the five ET products selected herewith *Bias* ranging from 15.93% to 62.42%. Although PML-V2(China) overestimates ET in the Northwest to a large extent, it performs best relative to its global version (84.88%) and another ET product MOD16A2 (152.70%).

3.3 Spatial patterns and annual variations of ET, E_c, E_i, E_s, GPP, and WUE

Fig. 10 illustrates the spatial distribution of the multi-year (2001-2018) mean annual ET and three components (i.e., E_c, E_i, and E_s) from PML-V2(China) across China. In general, the ET shows an increasing gradient from the northwest to the southeast (Fig. 10a). High annual ET (> 900 mm yr^{-1}) is mainly located in the water bodies, Hainan Island, and western Taiwan, while most parts of the Northwest River Basin exist the low annual ET (< 100 mm yr^{-1}), especially in the western Inner Mongolia and Gansu, and southern Xinjiang. Annual ET experiences a statistically insignificant increasing trend during 2001-2018 with a tendency of 0.43 mm yr^{-2} ($p > 0.05$). On the whole, the mean annual ET over China is 392.12 ± 10.67 mm yr^{-1} (mean $\pm$ standard deviation) over the last 18 years. For three components, E_c and E_i products display a similar spatial distribution with annual ET, while high E_s values (> 400 mm yr^{-1}) are mainly scattered in higher soil moisture content areas including the Tibet Plateau, Pearl River Delta, and Yangtze River Delta (Fig. 10b-d). High E_c (> 600 mm yr^{-1}) and E_i (> 80 mm yr^{-1}) values overall occur in the tropical and subtropical forests (e.g., Southwest River and Pearl River Basins), but low E_c (< 50 mm yr^{-1}) and E_i (< 5 mm yr^{-1}) values in the Northwest River Basin except for the Tianshan, Altai, and Qilian mountains. Especially, low E_i values also appear in the cropland areas, such as the Northeast Plain, North China Plain, and Sichuan Basin (Fig. 10c). For annual variation over China, E_c and E_i increase significantly during 2001-2018 with a rate of 0.91 and 0.16 mm yr^{-2} ($p < 0.001$), respectively. However, annual E_s shows a declining trend with an insignificant rate of -0.69 mm yr^{-2} ($p < 0.05$).

The mean annual GPP shows similar spatial patterns compared to the mean annual ET, as indicated by Fig. 10a and Fig. 11a. High annual GPP (> 2000 g C m^{-2} yr^{-1}) mainly occurs in the tropical and subtropical forests and North China Plain where there exists the double-cropping system, but low annual GPP (< 100 g C m^{-2} yr^{-1}) in the arid zones such as the Northwest River Basin. On average, the multi-year GPP over China is 721.62 ± 51.83 g C m^{-2} yr^{-1}, and interannual change displays a steady rise trend with a rate of 8.99 g C m^{-2} yr^{-2} ($p < 0.001$) since 2001. Using the coupled estimation of PML-V2(China) model, we study WUE (GPP divided by ET) during 2001-2018 across China (Fig. 11b). This result indicates the high annual WUE (>3 g C mm^{-1} H$_2$O) occurring in the forests and cropland, particularly in Northeast China and North China Plain. The annual variation of WUE is similar to that of GPP, with a significant increasing trend (Slope $= 0.02$ g C mm^{-1} H$_2$O yr^{-1}, $p < 0.001$).

4 Discussion

4.1 Magnitude and trend in annual ET and GPP over China

For annual ET over China, the multi-year (2001-2018) mean annual ET from PML-V2(China) is 392.12 ± 10.67 mm yr^{-1} (Fig. 10a). This result is overall consistent with the country-wide averaged annual ET estimated by the machine learning method (Yin et al., 2021: 397.65 mm yr^{-1} for 2000-2018) and land surface models (Ma et al., 2019a: 395.34 mm yr^{-1} for 2001-2012), and slightly higher than MOD16A2 ET about 359.61 ± 59.52 mm yr^{-1} for 2001-2018 (Cheng et al., 2021). But they are all less than the annual ET about 482.27 ± 192.31 mm yr^{-1} from SEBAL for 2001-2018 (Cheng et al., 2021). Furthermore, previous studies by Ren et al. (2015) and Wang et al. (2012a) show that the long-term mean precipitation and runoff in China are about 720 mm yr^{-1} and 280 mm yr^{-1}, respectively. Hence, it is believed that an annual ET less than 440 mm could be reasonable in China (Ma et al., 2019a). The annual mean GPP over China from our results is 721.62 ± 51.83 g C m^{-2} yr^{-1} during 2001-2018 (Fig. 11a), which is lower than that of Jia et al. (2020) (771 g C m^{-2} yr^{-1}), and higher than those of Yao et al. (2018) and Ma et al. (2019b) (690 g C m^{-2} yr^{-1} and 710 g C m^{-2} yr^{-1}, respectively). These differences may associate with the distinctions in the time window and data sources (Jia et al., 2020).

The annual ET displays a statistically insignificant increasing trend from 2001 to 2018, which is consistent with the calculated ET using the Budyko equation (Feng et al., 2018; Su et al., 2022). In terms of annual GPP, we found that there is a significant ($p < 0.001$) increasing trend with a rate of 8.99 g C m^{-2} yr^{-2} during 2001-2018, in line with some other studies (Ma et al., 2019b; Yao et al., 2018 ; Ma et al., 2018). The most likely reason for the remarkable rise in GPP is the effect of ecological restoration projects in China (Tong et al., 2018). In fact, a large number of ecological restoration projects have been conducted since the 1990s, such as the Grain for Green project (Cao et al., 2009). These findings also confirmed that a significant increase in vegetation growth occurred in China over the past years, which agreed well with Ma et al. (2018).

4.2 Advantages of this new dataset

The multi-scale testing using EC observations and water balance showed that accuracy in ET and GPP by the present PML-V2(China) is better than the global product of PML-V2 and other mainstream ET or GPP models (Table 3). The reasons may be twofold. The first is that the water-carbon coupled process is particularly important for estimating ET and GPP since the water and carbon process are highly coupled by the stomatal aperture at the leaf level. This result was also supported by Xiao et al. (2013) and Zhang et al. (2019) in their recent studies. Second, this study employed 26 EC observations to calibrate the PML-V2 in China, which shows better accuracy than the previous global-scale ET and GPP estimates that were obtained using few EC observations to constrain the parameters. This indicates that more local observations will facilitate the improvement of ET and GPP estimates at regional and national scales. In fact, although the EC sites of MOD16A2 (72 EC sites), GLEAM (91 EC sites), and PML-V2(Global) (95 EC sites) are more than in this study, there are only 0, 8, and 8 sites in China, respectively (Table 1). In particular, the SEBAL model only used eight EC sites for three PFTs (i.e., forests, cropland, and grassland).

In addition to the advantages of the overall accuracy in ET and GPP in the present study, the PML-V2(China) showed its strong ability to reveal the characteristics of the water consumption from the croplands that have the double-cropping system.

375 The GTZ in Hebei, DXZ in Beijing, and CF-YCA in Shandong are the only three observed sites with winter wheat - summer maize rotation cropping systems. We compared the intra-annual variation of the simulated ET and GPP between PML-V2(China) and other products against the EC-observed values at the three cropland sites (Fig. 12). In theory, when the winter wheat is harvested, the ET or GPP should decline to their valley values in June, which often occur between the two peaks (i.e., the reproductive growth stage of winter wheat and summer maize, respectively) within a given year. With this in mind, it can

380 be seen that PML-V2(China) has improved its ability when compared to its global version, as has been indicated by its better performance in capturing the time when the lowest ET and GPP values emerged. This is mainly because an improved weighted Whittaker smoother was carried out to get better quality of LAI, as described in Section 2.2.1. While the GLEAM is also able to detect the time when the valley values appeared, it underestimates ET evidently during the wheat growing season. In terms of the SEBAL and MOD16A2, both have much poorer performances in detecting such intra-annual variations in ET. Regarding

385 other GPP models, only MOD17A2H can catch the time when the valley values appear. However, it substantially underestimates GPP in winter both wheat and summer maize growing seasons. Moreover, this study also estimated the ability of the simulated ET to identify the crop phenology at the regional scale (Fig. S1). We extracted the cropland with peaks and identified the dates of peaks appearing within a year at each pixel by a faster peak detection algorithm (Liu et al., 2020). Taking the typical double-cropping system as an example, we quantified the cropping intensity in croplands (Fig. S1a1) and identified

390 the dates of the first peak and the second peak appearing in one year (Fig. S1a2, a3). To verify the reliability of the results, we mapped the double-cropping cropland areas of winter wheat and summer maize rotations (Fig. S1b1), the heading dates distribution of winter wheat and summer maize (Fig. S1b2, b3) in 2015 based on the crop phenological dataset (Luo et al., 2020). The croplands with a double-cropping show similar spatial patterns, as indicated by Fig. S1a1 and b1. In particular, we also compared the first ET peak date (Fig. S1a2) with the heading date of winter wheat (Fig. S1b2) in 2015. The first ET peak

395 date (i.e., day of the year (DOY)) is mainly between DOY 120 and 150, occurring after the heading date of winter wheat about DOY 100 to 130. During the entire growth period of winter wheat, the ET intensity was the highest from the heading date to the filling date, although the critical periods of water demand for winter wheat are the "jointing date-heading date-filling date" (He et al., 2022). This is consistent with our result that the ET peak appears slightly later than the crop heading date. Similarly, the heading date of summer maize also occurs early than the second ET peak date (Fig. S1a3, b3), which further proves the

400 ability of the simulated ET in evaluating crop phenology.

4.3 Implications of PML-V2(China)

Based on the substantial advantages discussed above, PML-V2(China) has great implications and application prospects. For instance, daily outputs from PML-V2(China) can be better used by the agricultural and water sectors for operational applications. Timely access to daily data at the regional or national scale helps the Ministry of Agriculture and Water Resources

405 to develop better policies. Indeed, there is a remarkable relationship between soil water content and ET (Graf et al., 2014;

Brust et al., 2021), so getting daily ET information accurately is of great significance for soil water depletion assessment, irrigation system design, and water resources management in agricultural areas, such as in the North China Plain. On the other hand, this dataset has better simulations of carbon consequences and water use efficiency, which is important for carbon-neutron policy (Yang et al., 2022). Specifically, for 2001-2018, the annual GPP and water use efficiency experienced a significant increase (8.99 g C m^{-2} yr^{-2} and 0.02 g C mm^{-1} H_2O yr^{-1}, respectively), but annual ET showed a non-significant increase (0.43 mm yr^{-2}). This indicates that vegetation in China exhibits a huge potential for carbon sequestration with little cost in water resources, which plays an important role in the global carbon cycle.

4.4 Uncertainties

4.4.1 Eddy-covariance method and water-balance formula

Although PML-V2(China) showed relatively good performance when compared to the EC sites and water-balance-based evapotranspiration (ET_{wb}), there exist several uncertainties related to the observed data (i.e., flux sites and ET_{wb}). First, the EC technique, considered as a standard way to measure surface fluxes (Aubinet et al., 1999; Liu et al., 2011; Baldocchi, 2014), meets some issues as well including corrections when processing the turbulence data and from energy non-closure problem. The corrections, such as the spike detection, lag correction of H_2O and CO_2 based on the vertical wind, coordinating rotation, corrections for density fluctuation, and frequency response correction, had been pre-processed before the investigators shared data. Nevertheless, there is evidence that diverse data processing designs may lead to errors of 10% ~ 15% (Mauder et al., 2007). Besides, systematic bias in device, the loss from the contribution of low-frequency eddies to energy transmission and the ability to capture larger eddies and the secondary circulations, could cause the energy-imbalance problem (Liu et al., 2011). Hence, there are usually two schemes to deal with the energy non-closure issue for EC users, that is, to perform energy closure correction (Cheng et al., 2021), or to maintain the original four-component data including latent heat (LE), sensible heat (H), soil heat flux (G) and net radiation (R_n) (Zhang et al., 2019; Ma and Zhang, 2022). We chose the second method in this study, considering that (i) forcing energy closure will introduce new errors artificially, and (ii) most Chinese EC observation towers lack G and R_n. The observed ET calculated from latent heat flux of the site without energy closure correction will be slightly less than the real value, resulting in a smaller ET simulated by the model determined by calibration using the sites. This phenomenon is not fully reflected in comparing with the basin ET based on the water-balance calculation that only in the Pearl and Southeast River Basin PML-V2(China) underestimates multi-year mean ET (Fig. 9f), because the ET_{wb} used for assessing the simulation performance of models also needs to be explored its accuracy.

Second, the inconsistency of the grid cell and EC footprint could also result in uncertainty when compared rudely to the measurements. Generally, the EC towers have a footprint of 100 - 1000 m^2, which is usually decided by tower height and heterogeneity of the underlying surface (Liu et al., 2016; Xu et al., 2017). For example, the footprint of the forest sites is larger than the grassland and wetland sites (Chen et al., 2012). In this study, PML-V2(China) model is first calibrated at 500m grid cell avoiding the inconsistency issue to some extent. However, there still exists a mismatch between the grid cell center and

EC sites. In this regard, higher spatial resolution products or additional ground observations at relevant scales would be beneficial for the cross-validation of the modelling grid cell (Ma et al., 2019a). Note also that the limited flux sites for some PFTs may introduce extra uncertainties for model parameters since only one site was available for the OSH, SAV, EBF, and MF in this study. In fact, we artificially construct multiple site samples in the above PFTs sites, utilizing the characteristics of the long time series of these sites. Nevertheless, the terrestrial biome of the sole flux site maybe not typical in other climate zones for the same PFT (Cheng et al., 2021). Therefore, more flux sites for the same PFT are necessary to calibrate the model. Third, ET derived from water-balance over China was invested as a reference at the basin-scale, although the results may be affected by some sources of uncertainty. For instance, the applicability of water-balance relies on its formula composition. In this paper, the water storage change (in Equation 5) from GRACE was included for purpose of reducing the uncertainties in estimating annual ET_{wb}. Those studies not using TWSC in ET_{wb} may not explain the decreasing groundwater such as in the Hai River Basin due to the human water extraction, which is not conducive to the credibility of the verification results (Cheng et al., 2021). In addition, the precision of its input data also affects the reliability of ET_{wb} (Mao and Wang, 2017). Precipitation, as the main source for ET, impacts ET_{wb} to varying extent. Nevertheless, precipitation data are derived from observations of field rain gauge network and its usability relies on the intensive and high-quality ground observations, which makes the Prcp estimates from statistics of stations worse in the less populated remote regions or areas having highly various topography, particularly in the west China (Immerzeel et al., 2015; Tang et al., 2016; Zhong et al., 2019). Zhong et al. (2019) evaluated three precipitation products in China and found that there was a slight overestimation in the west of China and an obvious underestimation in the west-Tibet Plateau. Accordingly, it could be stated that ET_{wb} was overestimated in Pearl and the Southeast and underestimated in the Southwest River basin using Eq. 7.

4.4.2 Input data

While the daily ET and GPP of the PML-V2(China) product (in the calibration mode) simulated well against 26 flux sites overall (Fig. 3) and in most PFTs (Fig. 4), PML-V2(China) in calibration is degraded compared to its cross-validation (Fig. 4), such as in GRA. In this study, we got one parameter set for the GRA type by employing eight sites, including the sparse grassland (QZ-BJ) and the dense grassland (QZ-NAMORS). Although it may be appropriate to use diverse parameter values for estimating ET and GPP by further dividing the grassland type into finer land types, this comes at the expense of ignoring the possible interannual changes in land types because few LUCC maps with fine classifications and annual resolutions simultaneously (Ma and Zhang, 2022).

PML-V2(China) mainly used the remote sensing and meteorological data (e.g., MODIS, CMFD, GLDAS-2.1, and ERA5) as the inputs (see section 2.2). However, there are still some uncertainties in these data (Zhang et al., 2019; Cheng et al., 2021). For example, we used the land cover datasets (MCD12Q1.006) as the PFTs data across China. However, there exist misclassification issues for MCD12Q1 because of spectral confusion (e.g., savannas and grasslands) and coarse resolution (e.g., the mixed pixel of cropland and natural vegetation) (Zhang et al., 2019; Liang et al., 2015; Adzhar et al., 2022). Moreover, LAI is a critical variable describing vegetation growth, and its temporal changes affect stomatal conductance and further affect

the transpiration rate from the vegetation canopy. Using LAI, PML-V2(China) model simulated E_c and E_i components at the canopy scale. To avoid noise issues caused by clouds, shadows, snow, and so on, MODIS LAI in this study has been smoothed by the weighted Whittaker smoother which can deal with underestimation and inefficiency issues (Kong et al., 2019). However, there are still underestimates in the sparse vegetation areas. This may explain why the ET and GPP estimates are poor in BSV (Fig. 7 and Fig. 8).

Additionally, downscaling uncertainties could be also introduced by the bilinear interpolation method which has been applied to minimize the footprint impact of coarse meteorological inputs, such as CMFD (Fig. 1). This approach depends only on nearby grid cells to downscale, which could neglect the other relative supports. For instance, precipitation is not only related to the surrounding precipitation but also location and terrain (e.g., elevation and aspect) (Yue et al., 2020). Chao et al. (2018) found that gridded precipitation products in the high-altitude regions are far below what is inversely inferred by glacier mass balances. Consequently, the geographically weighted regression method coupled with a weighting function could work well to interpolate meteorological data (Chao et al., 2018).

5 Conclusions

This study developed a daily, 500m ET and GPP data product (PML-V2(China)) using the locally calibrated water-carbon coupled model, PML-V2. The model has been well-calibrated against observations at 26 flux sites across nine plant functional types, and it performs satisfactorily in the cross-validation mode. More importantly, the plot- and basin-scale evaluations suggest that the newly developed product outperforms not only the global version of PML-V2 but also other mainstream RS-based ET and/or GPP products. With such a new product, we investigated spatial patterns and trends in ET and its components (E_c, E_i, E_s), GPP, and WUE from 2001 to 2018 across China. In short, the present PML-V2(China) product has the following advantages: (i) the water output is tightly constrained by carbon flux; (ii) it has high spatial and temporal resolutions simultaneously; (iii) it obtains the improved accuracy in ET and GPP across different plant functional types because of the optimal parameter sets for China by exploiting 26 EC sites; and (iv) showing the better ability to reveal the ET and GPP for the croplands with the double-cropping system. In summary, we provide a novel daily and 500m resolution ET and GPP product across China, which can be used by research communities and various water and ecological departments for operational applications.

Data availability statement

The product named PML-V2(China) with daily and 500m resolutions from February 26, 2000 to December 31, 2020 is freely available at the National Tibetan Plateau Data Center (http://dx.doi.org/10.11888/Terre.tpdc.272389, Zhang and He, 2022).

Code availability statement

500 The core source code of PML-V2(China) is available through the public GitHub repository (https://github.com/SylviaHeee/PML-V2-China).

Author contributions

SH and YZ designed the research. SH collected the EC observations and other input data. DK provided the improved LAI data. SH wrote the paper. YZ, NM, JT, DK and CL revised the paper.

505 **Competing interests**

The authors declared that they have no conflict of interest.

Acknowledgements

We gratefully acknowledge all people participating in EC measurements and uploading the observed data to available platforms (i.e., the National Tibetan Plateau Data Center, the Heihe Integrated Observatory Network, theChinaFLUX, and 510 FLUXNET2015). We are grateful to the Google Earth Engine team for the data distribution and online computing. We thank NOAA for providing CO_2 concentration at ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_mm_gl.txt. We appreciate the China Meteorological Forcing Dataset shared at https://doi.org/10.6084/m9.figshare.c.4557599. Thanks to the Ministry of Water Resources of the People's Republic of China for providing the basin-wide precipitation and runoff data from the National Water Resources Bulletin at http://szy.mwr.gov.cn/gbsj/index.html.

515 **Financial support**

This study was supported by the National Key R&D Program of China (2022YFC3002804), the Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0208), the National Natural Science Foundation of China (41971032), the CAS International Partnership Program (183311KYSB20200015), and "Science for a Better Development of Inner Mongolia" Program of the Bureau of Science and Technology of the Inner Mongolia Province (KJXM-EEDS-2020005).

520 **References**

Adzhar, R., Kelley, D. I., Dong, N., George, C., Torello Raventos, M., Veenendaal, E., Feldpausch, T. R., Phillips, O. L., Lewis, S. L., Sonké, B., Taedoumg, H., Schwantes Marimon, B., Domingues, T., Arroyo, L., Djagbletey, G., Saiz, G., and

Gerard, F.: MODIS Vegetation Continuous Fields tree cover needs calibrating in tropical savannas, Biogeosciences, 19, 1377–1394, https://doi.org/10.5194/bg-19-1377-2022, 2022.

525 Anandhi, A., Frei, A., Pierson, D. C., Schneiderman, E. M., Zion, M. S., Lounsbury, D., and Matonse, A. H.: Examination of change factor methodologies for climate change impact assessment, Water Resources Research, 47, W03501, https://doi.org/10.1029/2010WR009104, 2011.

Anon: Evaluation of MODIS gross primary productivity for Africa using eddy covariance data, Remote Sensing of Environment, 131, 275–286, https://doi.org/10.1016/j.rse.2012.12.023, 2013.

530 Arain, M. A., Yuan, F., and Black, T. A.: Soil–plant nitrogen cycling modulated carbon exchanges in a western temperate conifer forest in Canada, Agricultural and Forest Meteorology, 140, 171-192, https://doi.org/10.1016/j.agrformet.2006.03.021, 2006.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, Ü., Moncrieff, J., Foken, T., Kowalski, A. S., Martin, P. H., Berbigier, P., and Bernhofer, C.: Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology, Advances in 535 ecological research, 113-175, https://doi.org/10.1016/S0065-2504(08)60018-5, 1999.

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., U, K. T. P., 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, Bulletin of 540 the American Meteorological Society, 82, 2415–2434, https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001.

Baldocchi, D.: Measuring fluxes of trace gases and energy between ecosystems and the atmosphere – the state and future of the eddy covariance method, Global Change Biology, 20, 3600-3609, https://doi.org/10.1111/gcb.12649, 2014.

Beaudoing, H. and Rodell, M. J. G., Maryland: GLDAS Noah Land Surface Model L4 monthly 0.25 x 0.25 degree V2. 1, 545 NASA/GSFC/HSL: Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed [2017-03-23] at https://doi.org/10.5067/SXAVCZFAQLNO, 2016.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen-Geiger climate classification maps at 1-km resolution, Sci Data, 5, 180214, https://doi.org/10.1038/sdata.2018.214, 2018.

Bodner, G., Nakhforoosh, A., and Kaul, H.-P.: Management of crop water under drought: a review, Agron. Sustain. Dev., 35, 550 401-442, https://doi.org/10.1007/s13593-015-0283-4, 2015.

Brust, C., Kimball, J. S., Maneta, M. P., Jencso, K., He, M., and Reichle, R. H.: Using SMAP Level-4 soil moisture to constrain MOD16 evapotranspiration over the contiguous USA, Remote Sensing of Environment, 255, 112277, https://doi.org/10.1016/j.rse.2020.112277, 2021.

Cao, S., Chen, L., and Yu, X.: Impact of China's Grain for Green Project on the landscape of vulnerable arid and semi-arid 555 agricultural regions: a case study in northern Shaanxi Province, Journal of Applied Ecology, 46, 536–543, https://doi.org/10.1111/j.1365-2664.2008.01605.x, 2009.

Chao, L., Zhang, K., Li, Z., Zhu, Y., Wang, J., and Yu, Z.: Geographically weighted regression based methods for merging satellite and gauge precipitation, Journal of Hydrology, 558, 275–289, https://doi.org/10.1016/j.jhydrol.2018.01.042, 2018.

Chen, B., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Black, T. A., Arain, M. A., Barr, A. G., Bourque, C. P.-A.,
560 Flanagan, L. B., Lafleur, P. M., McCaughey, J. H., and Wofsy, S. C.: Characterizing spatial representativeness of flux tower eddy-covariance measurements across the Canadian Carbon Program Network using remote sensing and footprint analysis, Remote Sensing of Environment, 124, 742–755, https://doi.org/10.1016/j.rse.2012.06.007, 2012.

Chen, S., Chen, J., Lin, G., Zhang, W., Miao, H., Wei, L., Huang, J., and Han, X.: Energy balance and partition in Inner Mongolia steppe ecosystems with different land use types, Agricultural and Forest Meteorology, 149, 1800-1809,
565 https://doi.org/10.1016/j.agrformet.2009.06.009, 2009.

Cheng, M., Jiao, X., Li, B., Yu, X., Shao, M., and Jin, X.: Long time series of daily evapotranspiration in China based on the SEBAL model and multisource images and validation, Earth System Science Data, 13, 3995-4017, https://doi.org/10.5194/essd-13-3995-2021, 2021.

Chu, H., Baldocchi, D. D., John, R., Wolf, S., and Reichstein, M.: Fluxes all of the time? A primer on the temporal
570 representativeness of FLUXNET, Journal of Geophysical Research: Biogeosciences, 122, 289–307, https://doi.org/10.1002/2016JG003576, 2017.

Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W.: Regional evaporation estimates from flux tower and MODIS satellite data, Remote Sensing of Environment, 106, 285-304, https://doi.org/10.1016/j.rse.2006.07.007, 2007.

Dong, G., Guo, J., Chen, J., Sun, G., Gao, S., Hu, L., and Wang, Y.: Effects of spring drought on carbon sequestration,
575 evapotranspiration and water use efficiency in the songnen meadow steppe in northeast China, Ecohydrol., 4, 211–224, https://doi.org/10.1002/eco.200, 2011.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, 1980.

Feng, M. and Bai, Y: A global land cover map produced through integrating multi-source datasets, Big Earth Data, 3, 191-219,
580 https://doi.org/10.1080/20964471.2019.1663627, 2019.

Feng, T., Su, T., Ji, F., Zhi, R., and Han, Z.: Temporal Characteristics of Actual Evapotranspiration Over China Under Global Warming, Journal of Geophysical Research: Atmospheres, 123, 5845-5858, https://doi.org/10.1029/2017JD028227, 2018.

Foken, T.: The energy balance closure problem: an overview, Ecol Appl, 18, 1351-1367, https://doi.org/10.1890/06-0922.1, 2008.

585 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/10.1002/eco.1974, 2018.

Gevaert, C. M. and García-Haro, F. J.: A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion, Remote Sensing of Environment, 156, 34-44, https://doi.org/10.1016/j.rse.2014.09.012, 2015.

590 Graf, A., Bogena, H. R., Drüe, C., Hardelauf, H., Pütz, T., Heinemann, G., and Vereecken, H.: Spatiotemporal relations between water budget components and soil water content in a forested tributary catchment, Water Resources Research, 50, 4837-4857, https://doi.org/10.1002/2013WR014516, 2014.

Grant, R. F., Arain, A., Arora, V., Barr, A., Black, T. A., Chen, J., Wang, S., Yuan, F., and Zhang, Y.: Intercomparison of techniques to model high temperature effects on CO2 and energy exchange in temperate and boreal coniferous forests, 595 Ecological Modelling, 188, 217–252, https://doi.org/10.1016/j.ecolmodel.2005.01.060, 2005.

Hanson, P. J., Amthor, J. S., Wullschleger, S. D., Wilson, K. B., Grant, R. F., Hartley, A., Hui, D., Hunt, E. R., Jr, Johnson, D. W., Kimball, J. S., King, A. W., Luo, Y., McNulty, S. G., Sun, G., Thornton, P. E., Wang, S., Williams, M., Baldocchi, D. D., and Cushman, R. M.: Oak Forest Carbon and Water Simulations: Model Intercomparisons and Evaluations Against Independent Data, Ecological Monographs, 74, 443–489, https://doi.org/10.1890/03-4049, 2004.

600 Hao, Y., Wang, Y., and Yu, G.: A dataset of carbon and water fluxes over Xilinhot temperate steppe in Inner Mongolia (2003 – 2010) [dataset], Science Data Bank, https://doi.org/10.11922/sciencedb.996, 2020.

Haro-Monteagudo, D., Palazón, L., and Beguería, S.: Long-term sustainability of large water resource systems under climate change: A cascade modeling approach, Journal of Hydrology, 582, 124546, https://doi.org/10.1016/j.jhydrol.2020.124546, 2020.

605 He, H., Wu, Z., Li, D., Zhang, T., Pan, F., Yuan, H., Jiang, S., Shi, Z., Yang, S., and Wang, F.: Characteristics of Winter Wheat Evapotranspiration in Eastern China and Comparative Evaluation of Applicability of Different Reference Evapotranspiration Models, J Soil Sci Plant Nutr, 22, 2078–2091, https://doi.org/10.1007/s42729-022-00795-y, 2022.

He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., and Li, X.: The first high-resolution meteorological forcing dataset for land process studies over China, Sci Data, 7, 25, https://doi.org/10.1038/s41597-020-0369-y, 2020.

610 Heinsch, F. A., Maosheng Zhao, Running, S. W., Kimball, J. S., Nemani, R. R., Davis, K. J., Bolstad, P. V., Cook, B. D., Desai, A. R., Ricciuto, D. M., Law, B. E., Oechel, W. C., Hyojung Kwon, Hongyan Luo, Wofsy, S. C., Dunn, A. L., Munger, J. W., Baldocchi, D. D., Liukang Xu, Hollinger, D. Y., Richardson, A. D., Stoy, P. C., Siqueira, M. B. S., Monson, R. K., Burns, S. P., and Flanagan, L. B.: Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations, IEEE Trans. Geosci. Remote Sensing, 44, 1908–1925, 615 https://doi.org/10.1109/TGRS.2005.853936, 2006.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J., and Piontek, F.: A trend-preserving bias correction – the ISI-MIP approach, Earth Syst. Dynam., 4, 219–236, https://doi.org/10.5194/esd-4-219-2013, 2013.

Hilker, T., Coops, N. C., Wulder, M. A., Black, T. A., and Guy, R. D.: The use of remote sensing in light use efficiency based models of gross primary production: A review of current status and future requirements, Science of The Total Environment, 620 404, 411–423, https://doi.org/10.1016/j.scitotenv.2007.11.007, 2008.

Holland, J. H.: Genetic Algorithms, Scientific American, 267, 66–73, 1992.

Houborg, R., Cescatti, A., Migliavacca, M., and Kustas, W. P.: Satellite retrievals of leaf chlorophyll and photosynthetic capacity for improved modeling of GPP, Agricultural and Forest Meteorology, 177, 10–23, https://doi.org/10.1016/j.agrformet.2013.04.006, 2013.

625 Huang, Y., Nicholson, D., Huang, B., and Cassar, N.: Global Estimates of Marine Gross Primary Production Based on Machine Learning Upscaling of Field Observations, Global Biogeochemical Cycles, 35, e2020GB006718, https://doi.org/10.1029/2020GB006718, 2021.

Hui, Y., Dehua, Q., Yiping, Z., Qinghai, S., Xuehai, F., Liqing, S., Yuntong, L., Wenjun, Z., Liguo, Z., Xiaobao, D., Yan, L., Yun, D., and Donghai, Y.: An observation dataset of carbon and water fluxes in Xishuangbanna rubber plantations from 2010

630 to 2014 [dataset], Science Data Bank, https://doi.org/10.11922/sciencedb.j00001.00123, 2021.

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 CO_2 fluxes in Asia

635 using a standardized database of eddy covariance measurements, remote sensing data, and support vector regression, Journal of Geophysical Research: Biogeosciences, 122, 767-795, https://doi.org/10.1002/2016jg003640, 2017.

Immerzeel, W. W., Wanders, N., Lutz, A. F., Shea, J. M., and Bierkens, M. F. P.: Reconciling high-altitude precipitation in the upper Indus basin with glacier mass balances and runoff, Hydrol. Earth Syst. Sci., 19, 4673-4687, https://doi.org/10.5194/hess-19-4673-2015, 2015.

640 Jia, B., Luo, X., Cai, X., Jain, A., Huntzinger, D. N., Xie, Z., Zeng, N., Mao, J., Shi, X., Ito, A., Wei, Y., Tian, H., Poulter, B., Hayes, D., and Schaefer, K.: Impacts of land use change and elevated CO_2 on the interannual variations and seasonal cycles of gross primary productivity in China, Earth Syst. Dynam., 11, 235-249, https://doi.org/10.5194/esd-11-235-2020, 2020.

Joiner, J. and Yoshida, Y.: Satellite-based reflectances capture large fraction of variability in global gross primary production (GPP) at weekly time scales, Agricultural and Forest Meteorology, 291, 108092,

645 https://doi.org/10.1016/j.agrformet.2020.108092, 2020.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, Journal of

650 Geophysical Research, 116, https://doi.org/10.1029/2010jg001566, 2011.

Kato, T., Tang, Y., Gu, S., Hirota, M., Du, M., Li, Y., and Zhao, X.: Temperature and biomass influences on interannual changes in CO_2 exchange in an alpine meadow on the Qinghai-Tibetan Plateau, Global Change Biology, 12, 1285–1298, https://doi.org/10.1111/j.1365-2486.2006.01153.x, 2006.

Konak, A., Coit, D. W., and Smith, A. E.: Multi-objective optimization using genetic algorithms: A tutorial, Reliability

655 Engineering & System Safety, 91, 992–1007, https://doi.org/10.1016/j.ress.2005.11.018, 2006.

Kong, D., Zhang, Y., Gu, X., and Wang, D.: A robust method for reconstructing global MODIS EVI time series on the Google Earth Engine, ISPRS Journal of Photogrammetry and Remote Sensing, 155, 13–24, https://doi.org/10.1016/j.isprsjprs.2019.06.014, 2019.

Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resources Research, 660   48, https://doi.org/10.1029/2011wr011453, 2012.

Landerer, F.: TELLUS_GRAC_L3_CSR_RL06_LND_v04. Ver. RL06 v04 [dataset]. PO.DAAC, CA, USA. Accessed [2022-03-04] at https://doi.org/10.5067/TELND-3AC64, 2021.

Leuning, R., Zhang, Y. Q., Rajaud, A., Cleugh, H., and Tu, K.: A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman-Monteith equation, Water Resources Research, 44, 665   https://doi.org/10.1029/2007WR006562, 2008.

Li, S., Wang, G., Sun, S., Chen, H., Bai, P., Zhou, S., Huang, Y., Wang, J., and Deng, P.: Assessment of Multi-Source Evapotranspiration Products over China Using Eddy Covariance Observations, Remote Sensing, 10, https://doi.org/10.3390/rs10111692, 2018.

Li, X., Li, X., Li, Z., Ma, M., Wang, J., Xiao, Q., Liu, Q., Che, T., Chen, E., Yan, G., Hu, Z., Zhang, L., Chu, R., Su, P., Liu, 670   Q., Liu, S., Wang, J., Niu, Z., Chen, Y., Jin, R., Wang, W., Ran, Y., Xin, X., and Ren, H.: Watershed Allied Telemetry Experimental Research, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/10.1029/2008JD011590, 2009.

Liang, D., Zuo, Y., Huang, L., Zhao, J., Teng, L., and Yang, F.: Evaluation of the Consistency of MODIS Land Cover Product (MCD12Q1) Based on Chinese 30 m GlobeLand30 Datasets: A Case Study in Anhui Province, China, ISPRS International Journal of Geo-Information, 4, 2519–2541, https://doi.org/10.3390/ijgi4042519, 2015.

675   Liu, J., Rambal, S., and Mouillot, F.: Soil Drought Anomalies in MODIS GPP of a Mediterranean Broadleaved Evergreen Forest, Remote Sensing, 7, 1154–1180, https://doi.org/10.3390/rs70101154, 2015.

Liu, L., Xiao, X., Qin, Y., Wang, J., Xu, X., Hu, Y., and Qiao, Z.: Mapping cropping intensity in China using time series Landsat and Sentinel-2 images and Google Earth Engine, Remote Sensing of Environment, 239, 111624, https://doi.org/10.1016/j.rse.2019.111624, 2020.

680   Liu, S., Bliss, N., Sundquist, E., and Huntington, T. G.: Modeling carbon dynamics in vegetation and soil under the impact of soil erosion and deposition, Global Biogeochemical Cycles, 17, https://doi.org/10.1029/2002GB002010, 2003.

Liu, S., Li, X., Xu, Z., Che, T., Xiao, Q., Ma, M., Liu, Q., Jin, R., Guo, J., Wang, L., Wang, W., Qi, Y., Li, H., Xu, T., Ran, Y., Hu, X., Shi, S., Zhu, Z., Tan, J., Zhang, Y., and Ren, Z.: The Heihe Integrated Observatory Network: A Basin-Scale Land Surface Processes Observatory in China, Vadose Zone Journal, 17, 180072, https://doi.org/10.2136/vzj2018.04.0072, 2018.

685   Liu, S., Xu, Z., Song, L., Zhao, Q., Ge, Y., Xu, T., Ma, Y., Zhu, Z., Jia, Z., Zhang, F.: Upscaling evapotranspiration measurements from multi-site to the satellite pixel scale over heterogeneous land surfaces, Agricultural and Forest Meteorology, 230, 97-113, https://doi.org/10.1016/j.agrformet.2016.04.008, 2016.

Liu, S., Xu, Z., Zhu, Z., Jia, Z., and Zhu, M.: Measurements of evapotranspiration from eddy-covariance systems and large aperture scintillometers in the Hai River Basin, China, Journal of Hydrology, 487, 24-38, https://doi.org/10.1016/j.jhydrol.2013.02.025, 2013.

Liu, W., Wang, L., Zhou, J., Li, Y., Sun, F., Fu, G., Li, X., and Sang, Y.-F.: A worldwide evaluation of basin-scale evapotranspiration estimates against the water balance method, Journal of Hydrology, 538, 82-95, https://doi.org/10.1016/j.jhydrol.2016.04.006, 2016.

Luo, Y., Zhang, Z., Chen, Y., Li, Z., and Tao, F.: ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000–2015 based on leaf area index (LAI) products, Earth System Science Data, 12, 197–214, https://doi.org/10.5194/essd-12-197-2020, 2020.

Ma, J., Xiao, X., Miao, R., Li, Y., Chen, B., Zhang, Y., and Zhao, B.: Trends and controls of terrestrial gross primary productivity of China during 2000–2016, Environ. Res. Lett., 14, 084032, https://doi.org/10.1088/1748-9326/ab31e4, 2019.

Ma, J., Xiao, X., Zhang, Y., Doughty, R., Chen, B., and Zhao, B.: Spatial-temporal consistency between gross primary productivity and solar-induced chlorophyll fluorescence of vegetation in China during 2007–2014, Science of The Total Environment, 639, 1241–1253, https://doi.org/10.1016/j.scitotenv.2018.05.245, 2018.

Ma, N. and Zhang, Y.: Increasing Tibetan Plateau terrestrial evapotranspiration primarily driven by precipitation, Agricultural and Forest Meteorology, 317, https://doi.org/10.1016/j.agrformet.2022.108887, 2022.

Ma, N., Szilagyi, J., Zhang, Y., and Liu, W.: Complementary-Relationship-Based Modeling of Terrestrial Evapotranspiration Across China During 1982–2012: Validations and Spatiotemporal Analyses, Journal of Geophysical Research: Atmospheres, 124, 4326-4351, https://doi.org/10.1029/2018jd029850, 2019a.

Ma, T., Sun, S., Fu, G., Hall, J. W., Ni, Y., He, L., Yi, J., Zhao, N., Du, Y., Pei, T., Cheng, W., Song, C., Fang, C., and Zhou, C.: Pollution exacerbates China's water scarcity and its regional inequality, Nature Communications, 11, 650, https://doi.org/10.1038/s41467-020-14532-5, 2020.

Ma, Y., Hu, Z., Xie, Z., Ma, W., Wang, B., Chen, X., Li, M., Zhong, L., Sun, F., Gu, L., Han, C., Zhang, L., Liu, X., Ding, Z., Sun, G., Wang, S., Wang, Y., and Wang, Z.: A long-term (2005–2016) dataset of hourly integrated land–atmosphere interaction observations on the Tibetan Plateau, Earth Syst. Sci. Data, 12, 2937-2957, https://doi.org/10.5194/essd-12-2937-2020, 2020.

Mao, Y. and Wang, K.: Comparison of evapotranspiration estimates based on the surface water balance, modified Penman-Monteith model, and reanalysis data sets for continental China, Journal of Geophysical Research: Atmospheres, 122, 3228-3244, https://doi.org/10.1002/2016JD026065, 2017.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development, 10, 1903-1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Mauder, M., Desjardins, R. L., and MacPherson, I.: Scale analysis of airborne flux measurements over heterogeneous terrain in a boreal ecosystem, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006JD008133, 2007.

Medvigy, D., Wofsy, S., Munger, J., Hollinger, D., and Moorcroft, P.: Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2, Journal of Geophysical Research: Biogeosciences, 114, https://doi.org/10.1029/2008JG000812, 2009.

Miao, C., Gou, J., Fu, B., Tang, Q., Duan, Q., Chen, Z., Lei, H., Chen, J., Guo, J., Borthwick, A. G. L., Ding, W., Duan, X., Li, Y., Kong, D., Guo, X., and Wu, J.: High-quality reconstruction of China's natural streamflow, Science Bulletin, 67, 547-556, https://doi.org/10.1016/j.scib.2021.09.022, 2022.

Minx, J. C., Lamb, W. F., Andrew, R. M., Canadell, J. G., Crippa, M., Döbbeling, N., Forster, P. M., Guizzardi, D., Olivier, J., Peters, G. P., Pongratz, J., Reisinger, A., Rigby, M., Saunois, M., Smith, S. J., Solazzo, E., and Tian, H.: A comprehensive and synthetic dataset for global, regional, and national greenhouse gas emissions by sector 1970–2018 with an extension to 2019, Earth System Science Data, 13, 5213-5252, https://doi.org/10.5194/essd-13-5213-2021, 2021.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrology and Earth System Sciences, 15, 453-469, https://doi.org/10.5194/hess-15-453-2011, 2011b.

Miralles, D., De Jeu, R., Gash, J., Holmes, T., Dolman, A. J.: Magnitude and variability of land evaporation and its components at the global scale, Hydrology and Earth System Sciences, 15, 967-981, https://doi.org/10.5194/hess-15-453-2011, 2011a.

Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W.: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sensing of Environment, 111, 519-536, https://doi.org/10.1016/j.rse.2007.04.015, 2007.

Mu, Q., Zhao, M., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sensing of Environment, 115, 1781-1800, https://doi.org/10.1016/j.rse.2011.02.019, 2011.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349-4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Myneni, R., Knyazikhin, Y., Park, T.: MOD15A2H MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006 [dataset]. NASA EOSDIS Land Processes DAAC. Accessed [2022-04-18] at https://doi.org/10.5067/MODIS/MOD15A2H.006, 2015.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A.: Terrestrial ecosystem production: a process model based on global satellite and surface data, Global Biogeochemical Cycles, 7, 811-841, https://doi.org/10.1029/93GB02725, 1993.

Rasmussen, J., Sonnenborg, T. O., Stisen, S., Seaby, L. P., Christensen, B. S. B., and Hinsby, K.: Climate change effects on irrigation demands and minimum stream discharge: impact of bias-correction method, Hydrology and Earth System Sciences, 16, 4675–4691, https://doi.org/10.5194/hess-16-4675-2012, 2012.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, 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, Global Change Biology, 11, 1424-1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005.

Ren, G., Zhan, Y., Ren, Y., Chen, Y., Wang, T., Liu, Y., and Sun, X.: Spatial and temporal patterns of precipitation variability over mainland China: I: climatology, Advances in Water Science, 26, 299-310, https://doi.org/10.14042/j.cnki.32.1309.2015.03.001, 2015.

Running, S., Mu, Q., Zhao, M.: MOD17A2H MODIS/Terra Gross Primary Productivity 8-Day L4 Global 500m SIN Grid V006 [dataset]. NASA EOSDIS Land Processes DAAC. Accessed at https://doi.org/10.5067/MODIS/MOD17A2H.006, 2015.

Schaaf, C., Wang, Z.: MCD43A1 MODIS/Terra + Aqua BRDF/Albedo Model Parameters Daily L3 Global - 500 m V006 [dataset]. NASA EOSDIS Land Processes DAAC. Accessed at https://doi.org/10.5067/MODIS/MCD43A1.006, 2015.

Su, T., Feng, T., Huang, B., Han, Z., Qian, Z., Feng, G., Hou, W., and Dong, W.: Long-term mean changes in actual evapotranspiration over China under climate warming and the attribution analysis within the Budyko framework, International Journal of Climatology, 42, 1136-1147, https://doi.org/10.1002/joc.7293, 2022.

Sulla-Menashe, D., Gray, J. M., Abercrombie, S. P., and Friedl, M. A.: Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product, Remote Sensing of Environment, 222, 183-194, https://doi.org/10.1016/j.rse.2018.12.013, 2019.

Tang, G., Ma, Y., Long, D., Zhong, L., and Hong, Y.: Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales, Journal of Hydrology, 533, 152-167, https://doi.org/10.1016/j.jhydrol.2015.12.008, 2016.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, Journal of Hydrology, 456–457, 12–29, https://doi.org/10.1016/j.jhydrol.2012.05.052, 2012.

Tong, X., Brandt, M., Yue, Y., Horion, S., Wang, K., Keersmaecker, W. D., Tian, F., Schurgers, G., Xiao, X., Luo, Y., Chen, C., Myneni, R., Shi, Z., Chen, H., and Fensholt, R.: Increased vegetation growth and carbon stock in China karst via ecological engineering, Nat Sustain, 1, 44–50, https://doi.org/10.1038/s41893-017-0004-x, 2018.

Turner, D. P., Urbanski, S., Bremer, D., Wofsy, S. C., Meyers, T., Gower, S. T., and Gregory, M.: A cross-biome comparison of daily light use efficiency for gross primary production, Global Change Biology, 9, 383–395, https://doi.org/10.1046/j.1365-2486.2003.00573.x, 2003.

Villarreal, S. and Vargas, R.: Representativeness of FLUXNET Sites Across Latin America, Journal of Geophysical Research: Biogeosciences, 126, https://doi.org/10.1029/2020JG006090, 2021.

Wan, Z., Hook, S., Hulley, G.: MYD11A2 MODIS/Aqua Land Surface Temperature/ Emissivity 8-Day L3 Global 1 km SIN Grid V006 [dataset]. NASA EOSDIS Land Processes DAAC. Accessed at https://doi.org/10.5067/MODIS/MYD11A2.006, 2015.

Wang, F., Chen, J. M., Gonsamo, A., Zhou, B., Cao, F., and Yi, Q.: A two-leaf rectangular hyperbolic model for estimating GPP across vegetation types and climate conditions, Journal of Geophysical Research: Biogeosciences, 119, 1385-1398, https://doi.org/10.1002/2013JG002596, 2014.

Wang, G., Zhang, J., Jin, J., Pagano, T., Calow, R., Bao, Z., Liu, C., Liu, Y., Yan, X.: Assessing water resources in China using PRECIS projections and a VIC model, Hydrol. Earth Syst. Sci., 16, 231-240, https://doi.org/10.5194/hess-16-231-2012, 2012a.

Wang, S., Li, J., Zhang, B., Lee, Z., Spyrakos, E., Feng, L., Liu, C., Zhao, H., Wu, Y., and Zhu, L.: Changes of water clarity in large lakes and reservoirs across China observed from long-term MODIS, Remote Sensing of Environment, 247, 111949, https://doi.org/10.1016/j.rse.2020.111949, 2020.

Wang, X., Ma, M., Li, X., Song, Y., Tan, J., Huang, G., Zhang, Z., Zhao, T., Feng, J., Ma, Z., Wei, W., and Bai, Y.: Validation of MODIS-GPP product at 10 flux sites in northern China, International Journal of Remote Sensing, 34, 587-599, https://doi.org/10.1080/01431161.2012.715774, 2012b.

Wen, X., Yu, G., Sun, X., Li, Q., Liu, Y., Zhang, L., Ren, C., Fu, Y., and Li, Z.: Soil moisture effect on the temperature dependence of ecosystem respiration in a subtropical Pinus plantation of southeastern China, Agricultural and Forest Meteorology, 137, 166–175, https://doi.org/10.1016/j.agrformet.2006.02.005, 2006a.

Wofsy, S. C., Goulden, M. L., Munger, J. W., Fan, S. M., Bakwin, P. S., Daube, B. C., Bassow, S. L., and Bazzaz, F. A.: Net Exchange of CO2 in a Mid-Latitude Forest, Science, 260, 1314-1317, https://doi.org/10.1126/science.260.5112.1314, 1993.

Wolanin, A., Camps-Valls, G., Gómez-Chova, L., Mateo-García, G., van der Tol, C., Zhang, Y., and Guanter, L.: Estimating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations, Remote Sensing of Environment, 225, 441-457, https://doi.org/10.1016/j.rse.2019.03.002, 2019.

Wutzler, T., Lucas-Moffat, A., Migliavacca, M., Knauer, J., Sickel, K., Šigut, L., Menzer, O., and Reichstein, M.: Basic and extensible post-processing of eddy covariance flux data with REddyProc, Biogeosciences,15, 5015-5030, https://doi.org/10.5194/bg-15-5015-2018, 2018.

Xiao, J., Sun, G., Chen, J., Chen, H., Chen, S., Dong, G., Gao, S., Guo, H., Guo, J., Han, S., Kato, T., Li, Y., Lin, G., Lu, W., Ma, M., McNulty, S., Shao, C., Wang, X., Xie, X., Zhang, X., Zhang, Z., Zhao, B., Zhou, G., and Zhou, J.: Carbon fluxes, evapotranspiration, and water use efficiency of terrestrial ecosystems in China, Agricultural and Forest Meteorology, 182-183, 76-90, https://doi.org/10.1016/j.agrformet.2013.08.007, 2013.

Xie, P., Wu, Z., Sang, Y.-F., Gu, H., Zhao, Y., and Singh, V. P.: Evaluation of the significance of abrupt changes in precipitation and runoff process in China, Journal of Hydrology, 560, 451-460, https://doi.org/10.1016/j.jhydrol.2018.02.036, 2018.

Xu, F., Wang, W., Wang, J., Xu, Z., Qi, Y., Wu, Y.: Area-averaged evapotranspiration over a heterogeneous land surface: aggregation of multi-point EC flux measurements with a high-resolution land-cover map and footprint analysis, Hydrol. Earth Syst. Sci., 21, 4037-4051, https://doi.org/10.5194/hess-21-4037-2017, 2017.

Yang, D., Li, C., Hu, H., Lei, Z., Yang, S., Kusuda, T., Koike, T., and Musiake, K.: Analysis of water resources variability in the Yellow River of China during the last half century using historical data, Water Resources Research, 40, https://doi.org/10.1029/2003wr002763, 2004.

825 Yang, F., Sarathchandra, C., Liu, J., Huang, H., Gou, J.-Y., Li, Y., Mao, X., Wen, H., Zhao, J., Yang, M., Homya, S., and Prueksakorn, K.: How fern and fern allies respond to heterogeneous habitat — a case in Yuanjiang dry-hot valley, Communicative & Integrative Biology, 14, 248–260, https://doi.org/10.1080/19420889.2021.2007591, 2021.

Yang, Y., Shi, Y., Sun, W., Chang, J., Zhu, J., Chen, L., Wang, X., Guo, Y., Zhang, H., Yu, L., Zhao, S., Xu, K., Zhu, J., Shen, H., Wang, Y., Peng, Y., Zhao, X., Wang, X., Hu, H., Chen, S., Huang, M., Wen, X., Wang, S., Zhu, B., Niu, S., Tang, Z., Liu,
830 L., and Fang, J.: Terrestrial carbon sinks in China and around the world and their contribution to carbon neutrality, Sci. China Life Sci., 65, 861–895, https://doi.org/10.1007/s11427-021-2045-5, 2022.

Yao, Y., Wang, X., Li, Y., Wang, T., Shen, M., Du, M., He, H., Li, Y., Luo, W., Ma, M., Ma, Y., Tang, Y., Wang, H., Zhang, X., Zhang, Y., Zhao, L., Zhou, G., and Piao, S.: Spatiotemporal pattern of gross primary productivity and its covariation with climate in China over the last thirty years, Global Change Biology, 24, 184–196, https://doi.org/10.1111/gcb.13830, 2018.

835 Yin, L., Tao, F., Chen, Y., Liu, F., and Hu, J.: Improving terrestrial evapotranspiration estimation across China during 2000–2018 with machine learning methods, Journal of Hydrology, 600, https://doi.org/10.1016/j.jhydrol.2021.126538, 2021.

Yu, G.-R., Wen, X.-F., Sun, X.-M., Tanner, B. D., Lee, X., and Chen, J.-Y.: Overview of ChinaFLUX and evaluation of its eddy covariance measurement, Agricultural and Forest Meteorology, 137, 125-137, https://doi.org/10.1016/j.agrformet.2006.02.011, 2006.

840 Yu, H., Qi, D., Zhang, Y., Sha, L., Liu, Y., Zhou, W., Deng, Y., and Song, Q.: An observation dataset of carbon and water fluxes in Xishuangbanna rubber plantations from 2010 to 2014 [dataset], Science Data Bank, https://cstr.cn/31253.11.sciencedb.j00001.00123, 2021.

Yu, Q., Zhang, Y., Liu, Y., and Shi, P.: Simulation of the Stomatal Conductance of Winter Wheat in Response to Light, Temperature and CO2 Changes, Annals of Botany, 93, 435–441, https://doi.org/10.1093/aob/mch023, 2004.

845 Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L. L., Baldocchi, D., Bernhofer, C., Gholz, H., Goldstein, A. H., Goulden, M. L.: Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes, Agricultural and Forest Meteorology, 143, 189-207, https://doi.org/10.1016/j.agrformet.2006.12.001, 2007.

Yue, T., Zhao, N., Liu, Y., Wang, Y., Zhang, B., Du, Z., Fan, Z., Shi, W., Chen, C., and Zhao, M.: A fundamental theorem for eco-environmental surface modelling and its applications, Science China Earth Sciences, 63, 1092-1112,
850 https://doi.org/10.1007/s11430-019-9594-3, 2020.

Zhang, F., Li, H., Zhao, L., Zhang, L., Chen, Z., Zhu, J., Xu, S., Yang, Y., Zhao, X., and Yu, G.: An observation dataset of carbon, water and heat fluxes of alpine wetland in Haibei (2004–2009) [dataset], Science Data Bank, https://doi.org/10.11922/sciencedb.1010, 2020a.

Zhang, J.-H., Han, S.-J., Yu, G.-R.: Seasonal variation in carbon dioxide exchange over a 200-year-old Chinese broad-leaved Korean pine mixed forest, Agricultural and Forest Meteorology, 137, 150-165, https://doi.org/10.1016/j.agrformet.2006.02.004, 2006a.

Zhang, Y. and He, S.: PML-V2(China): evapotranspiration and gross primary production (2000.02.26-2020.12.31) [dataset], National Tibetan Plateau Data Center, http://dx.doi.org/10.11888/Terre.tpdc.272389, 2022.

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 Sensing of Environment, 222, 165–182, https://doi.org/10.1016/j.rse.2018.12.031, 2019.

Zhang, Y., Xiao, X., Wu, X., Zhou, S., Zhang, G., Qin, Y., and Dong, J.: A global moderate resolution dataset of gross primary production of vegetation for 2000-2016, Sci Data, 4, 170165, https://doi.org/10.1038/sdata.2017.165, 2017.

Zhang, Y., Xiao, X., Wu, X., Zhou, S., Zhang, G., Qin, Y., and Dong, J.: Global gross primary production from vegetation photosynthesis model for 2000-2016 [dataset], PANGAEA, https://doi.org/10.1594/PANGAEA.879560, 2017.

Zhang, Y., Xiao, X., Wu, X., Zhou, S., Zhang, G., Qin, Y., Dong, J., and Doughty, R.: Global gross primary production from vegetation photosynthesis model for 2017-2019 [dataset], PANGAEA, https://doi.org/10.1594/PANGAEA.928381, 2021.

Zhang, Z., Arnault, J., Wagner, S., Laux, P., and Kunstmann, H.: Impact of lateral terrestrial water flow on land-atmosphere interactions in the Heihe River Basin in China: Fully coupled modeling and precipitation recycling analysis, Journal of Geophysical Research: Atmospheres, 124, 8401-8423, https://doi.org/10.1029/2018JD030174, 2019.

Zhang, Z., Zhang, Y., Zhang, Y., Gobron, N., Frankenberg, C., Wang, S., and Li, Z.: The potential of satellite FPAR product for GPP estimation: An indirect evaluation using solar-induced chlorophyll fluorescence, Remote Sensing of Environment, 240, 111686, https://doi.org/10.1016/j.rse.2020.111686, 2020b.

Zhao, F., Li, F., Zhan, C., Zhang, L., and Chen, Z.: A carbon and water fluxes dataset over farmland ecosystem of winter wheat and summer corn in Yucheng (2003–2010) [dataset], Science Data Bank, https://doi.org/10.11922/sciencedb.j00001.20002, 2021.

Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., Chen, J. M., Ju, W., Zhang, L., and Yuan, W.: Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017, Earth System Science Data, 12, 2725-2746, https://doi.org/10.5194/essd-12-2725-2020, 2020.

Zhong, R., Chen, X., Lai, C., Wang, Z., Lian, Y., Yu, H., and Wu, X.: Drought monitoring utility of satellite-based precipitation products across mainland China, Journal of Hydrology, 568, 343-359, https://doi.org/10.1016/j.jhydrol.2018.10.072, 2019.

Table 1: Summary of typical ET and GPP products with high temporal and/or high spatial resolutions.

Variable	Dataset abbreviation	Spatial resolution	Temporal resolution	Temporal coverage	Principle or model	EC evaluation	Reference
ET	MOD16A2	500m	8-day	2001-present	PM	72 EC sites in AmeriFlux (no using sites in China)	Mu et al. (2011)
ET	SEBAL	1000m	daily	2001-2018	one-source model of surface energy balance residual	8 EC sites in China	Cheng et al. (2021)
ET	GLEAM	0.25°	daily	1980-2020	P-T	91 global EC sites (including 8 sites in China)	Miralles et al. (2011a) and Martens et al. (2017)
GPP	MOD17A2H	500m	8-day	2000-present	LUE	Not performed.	Running et al. (2015)
GPP	VPM	500m	8-day	2000-2019	LUE	113 global EC sites (including 8 sites in China)	Zhang et al. (2017), Zhang et al. (2017); Zhang et al. (2021)
GPP	EC-LUE	0.05°	8-day	1982-2018	LUE	95 global EC sites (including 7 sites in China)	Zheng et al. (2020)
ET, GPP	PML-V2 (Global)	500m	8-day	2000-2020	PML-V2	95 global EC sites (including 8 sites in China)	Zhang et al. (2019)
ET, GPP	PML-V2 (China)	500m	daily	2000-2020	PML-V2	26 EC sites in China	This study

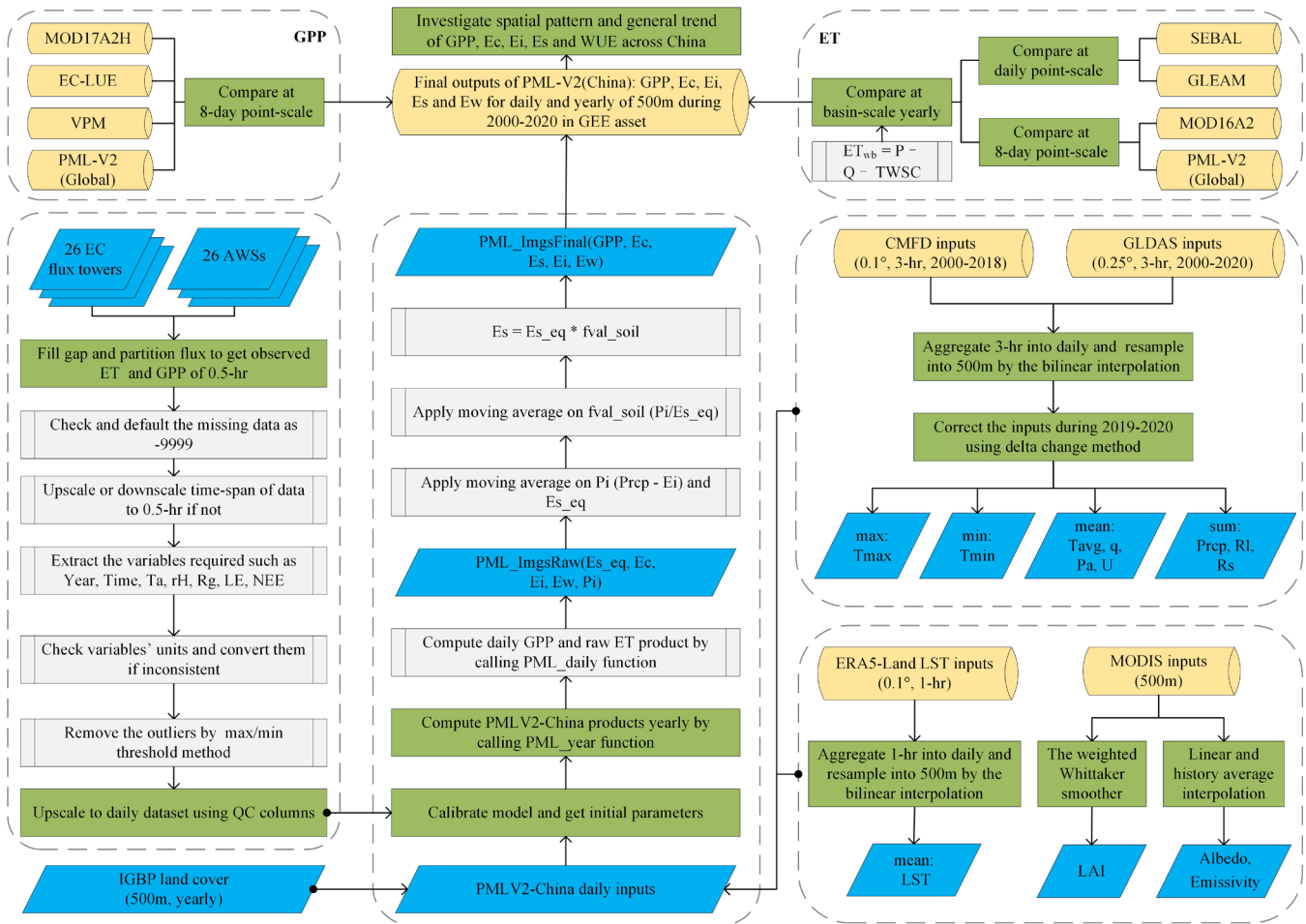



Figure 1: Flowchart of EC flux and AWSs data pre-processing and PML-V2 model processing which is used to convert RS images and meteorological forcing images into GPP, E_c, E_i, E_s, E_w, and ET. For pre-processing part: NEE (Net Ecosystem Exchange, µmol m^{-2} s^{-1}), LE (Latent heat, W m^{-2}), Rg (incoming radiation, W m^{-2}), rH (relative humidity, %), Ta (air temperature, °C), and QC (Quality Control). For the PML-V2 model part: Tmax (daily maximum temperature, °C), Tmin (daily minimum temperature, °C), Tavg (daily mean temperature, °C), Pa (atmosphere pressure, kPa), U (wind speed at 10-m height, m s^{-1}), q (specific humidity, kg kg^{-1}), Prcp (precipitation, mm d^{-1}), Rl (inward longwave solar radiation, W m^{-2}), Rs (inward shortwave solar radiation, W m^{-2}), Pi (the difference of Prcp and Ei, mm d^{-1}), Es_eq (equilibrium evaporation, mm d^{-1}), ET_w (evaporation from water body, snow and ice, mm d^{-1}) and GEE (© Google Earth Engine).

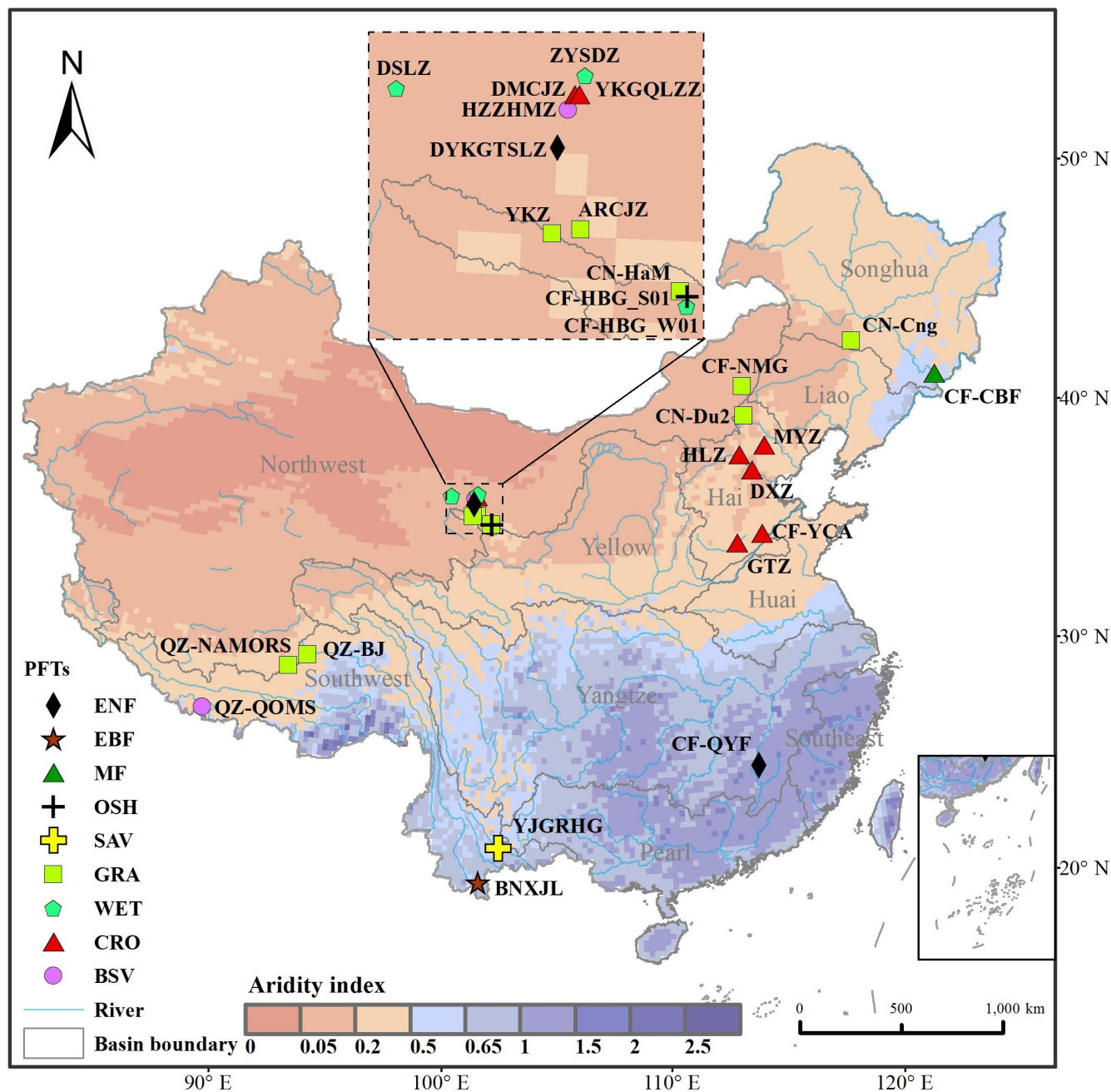



Figure 2: Geographical locations of 26 EC flux towers for nine major IGBP PFTs, the main rivers, and the ten major river basins in China. Overlain are 20-year mean annual aridity index (AI) values during 2001-2020 using GLDAS-2.1, that is, the ratio of annual precipitation to Penman potential evapotranspiration. PFTs shown in legend are ENF (Evergreen Needleleaf Forests), EBF (Evergreen Broadleaf Forests), MF (Mixed Forests), OSH (Open Shrublands), SAV (Savannas), GRA (Grasslands), WET (Permanent Wetlands), CRO (Croplands), and BSV (Barren Sparse Vegetation).

Table 2: Details of 26 EC flux towers employed in this study. Note that AP indicates mean annual precipitation and AT refers to mean annual temperature in its observed period.

Site code	Site name	IGBP	Latitude (°E)	Longitude (°N)	AP (mm yr-1)	AT (°C)	Time coverage	References
ARCJZ	Arou	GRA	38.0473	100.4643	521	-2.7	2013-2017	Liu et al. (2018)
BNXJL	Xishuangbanna rubber	EBF	21.9000	101.2667	1765	22.1	2013	Yu et al. (2021)
CF-CBF	Chinaflux Changbai forest	MF	42.4025	128.0958	608	4.3	2003-2010	Zhang et al. (2006a)
CF-HBG_S01	Chinaflux Haibei grassland	OSH	37.6653	101.3311	610	-5.9	2003-2010	Hui et al. (2021)
CF-HBG_W01	Chinaflux Haibei wetland	WET	37.6086	101.3269	616	-3.9	2004-2006	Zhang et al. (2020a)
CF-NMG	Chinaflux Neimengu grassland	GRA	43.3233	116.4036	387	1.2	2004	Hao et al. (2020)
CF-QYF	Chinaflux Qianyanzhou forest	ENF	26.7414	115.0581	1490	19.3	2004-2006	Wen et al. (2006)
CF-YCA	Chinaflux Yucheng	CRO	36.8290	116.5702	602	14.8	2006-2007	Zhao et al. (2021)
CN-Cng	Changling	GRA	44.5934	123.5092	364	6.5	2007-2010	Dong et al. (2011)
CN-Du2	Duolun_grassland (D01)	GRA	42.0467	116.2836	388	3.0	2006-2008	Chen et al. (2009)
CN-HaM	Haibei Alpine Tibet site	GRA	37.6975	101.2733	534	-4.0	2002-2004	Kato et al. (2006)
DMCJZ	Daman	CRO	38.8555	100.3722	163	9.2	2017	Liu et al. (2018)
DSLZ	Dashalong	WET	38.8399	98.9406	346	-8.3	2015-2018	Liu et al. (2018)
DXZ	Daxing	CRO	39.6213	116.4271	547	12.7	2010	Liu et al. (2013)
DYKGTSLZ	Dayekouguantan forest	ENF	38.5337	100.2502	228	0.2	2010-2011	Li et al. (2009)
GTZ	Guantao	CRO	36.5150	115.1274	433	14.0	2008	Liu et al. (2013)
HLZ	Huailai	CRO	40.3491	115.7880	377	10.2	2014	Liu et al. (2013)

HZZHMZ	Huazhaizi Desert Steppe	BSV	38.7659	100.3201	167	8.7	2017	Liu et al. (2018)
MYZ	Miyun	CRO	40.6308	117.3233	584	9.0	2008	Liu et al. (2013)
QZ-BJ	Tibetan Plateau BJ	GRA	31.3688	91.8988	460	0.2	2011-2013	Ma et al. (2020)
QZ-NAMORS	Tibetan Plateau NAMORS	GRA	30.7730	90.9632	405	-0.3	2008-2009	Ma et al. (2020)
QZ-QOMS	Tibetan Plateau QOMS	BSV	28.3607	86.9491	199	1.2	2015	Ma et al. (2020)
YJGRHG	Yuanjiang dry-hot valley	SAV	101.2667	21.9000	876	20.2	2014	Yang et al., (2021)
YKGQLZZ	Yingke	CRO	38.8569	100.4103	85	8.3	2011	Liu et al. (2018)
YKZ	Yakou	GRA	38.0142	100.2421	484	-1.2	2016-2018	Liu et al. (2018)
ZYSDZ	Zhangye wetland	WET	38.9751	100.4464	146	8.8	2013-2018	Liu et al. (2018)

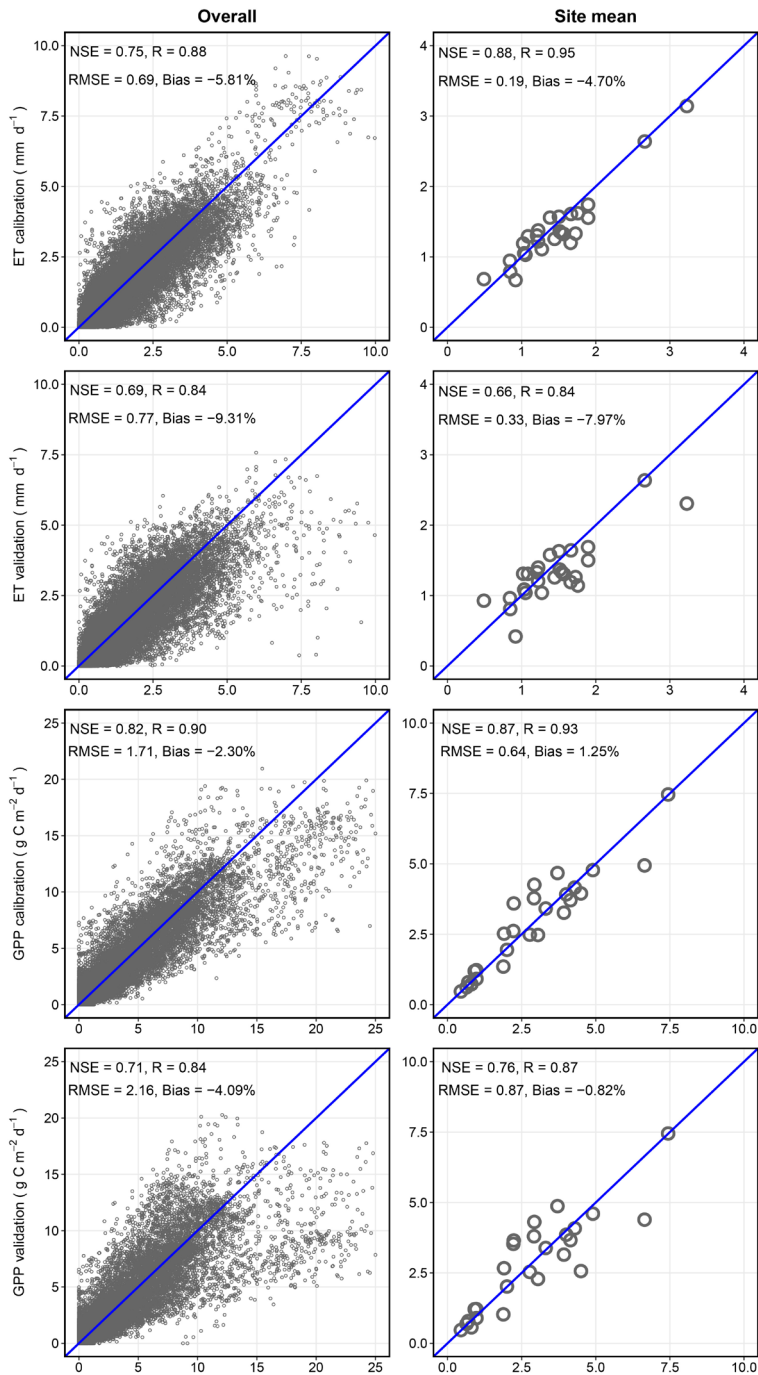



Figure 3: Scatter plots between the observed ET and GPP against PML-V2(China) simulations in calibration and cross-validation modes: daily comparisons in the left panels and site mean comparison in the right panels.

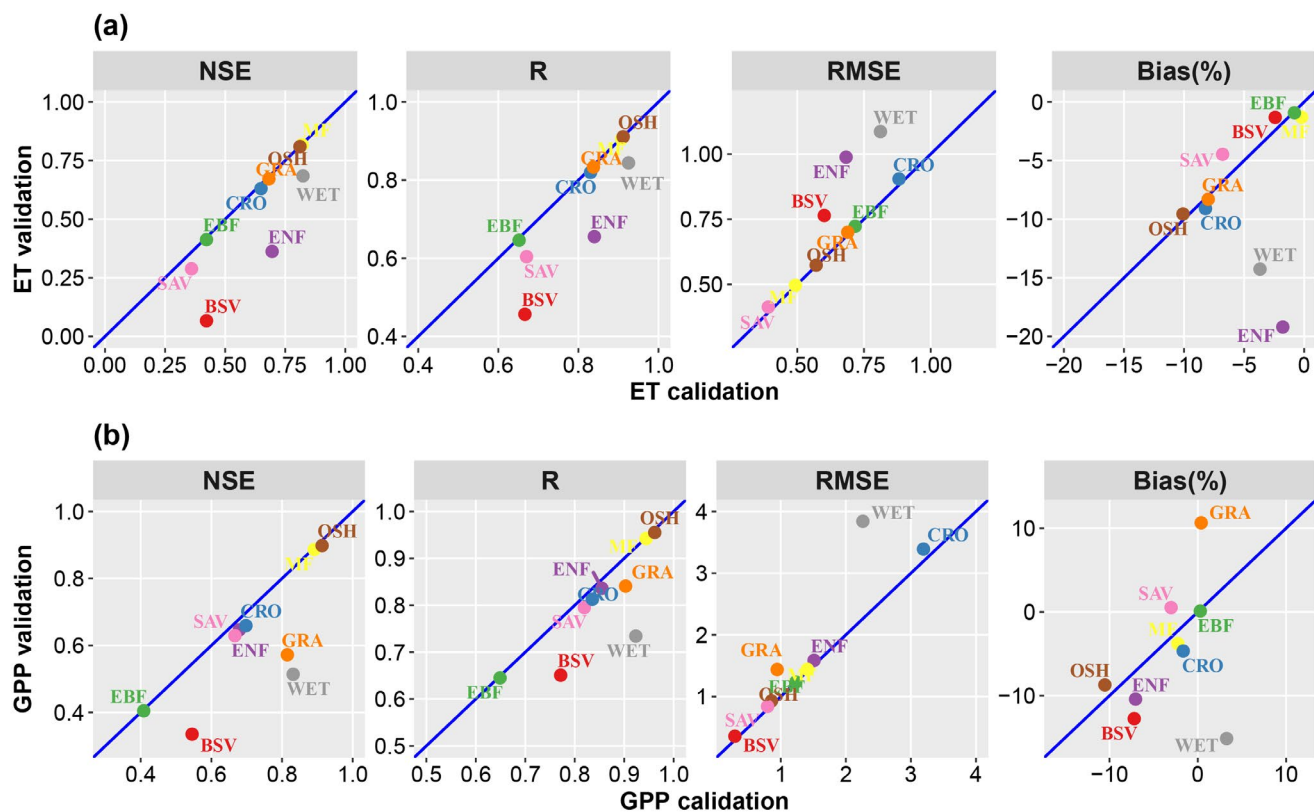



Figure 4: Comparison of ET and GPP between PML-V2(China) model calibration and validation across ten PFTs.

910

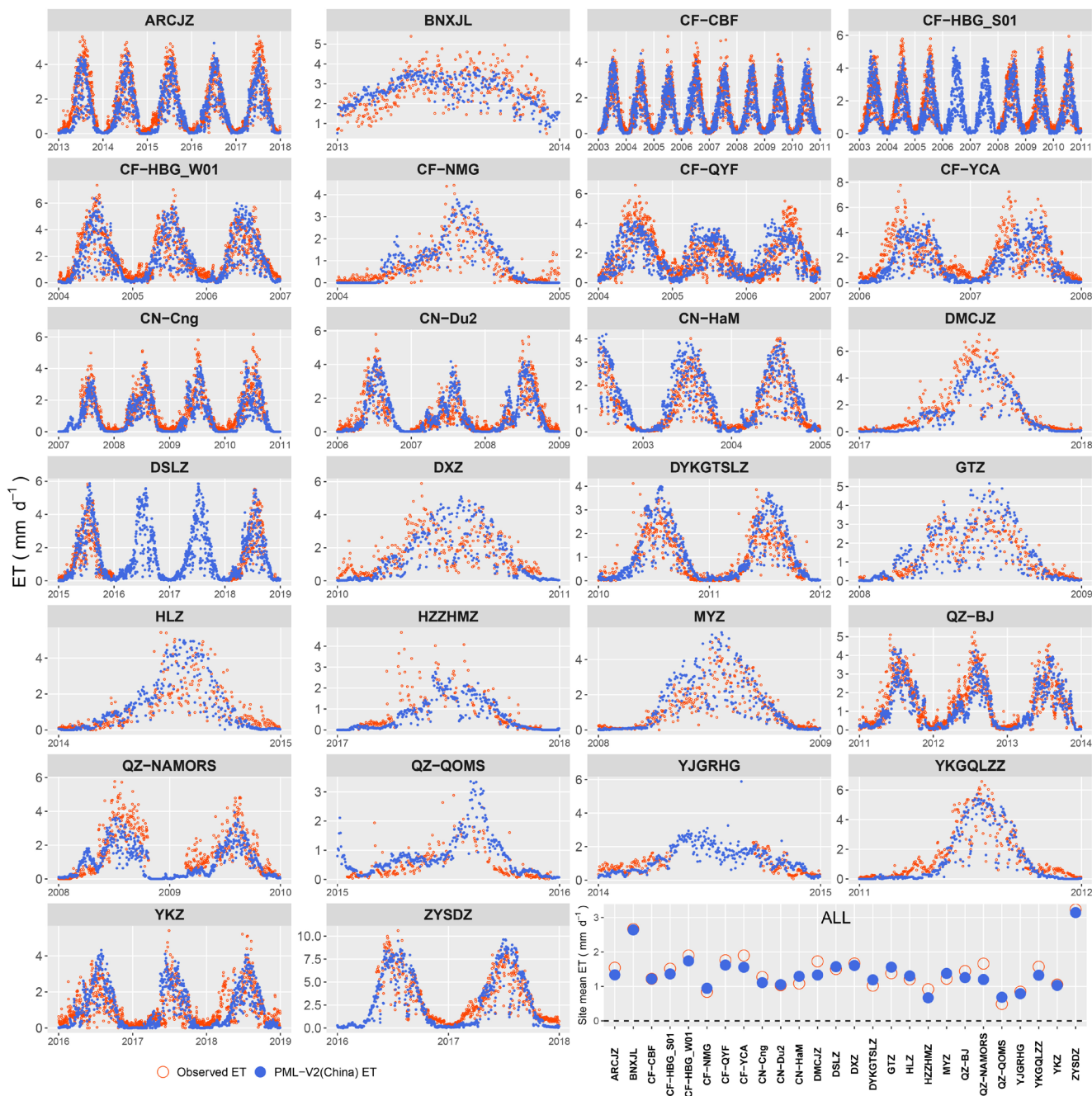



Figure 5: The daily ET simulated by PML-V2(China) in calibration mode and the observed daily ET variation in time series from 26 EC sites (see Figure 2) across China. 'ALL' represents the site mean value for each EC site.

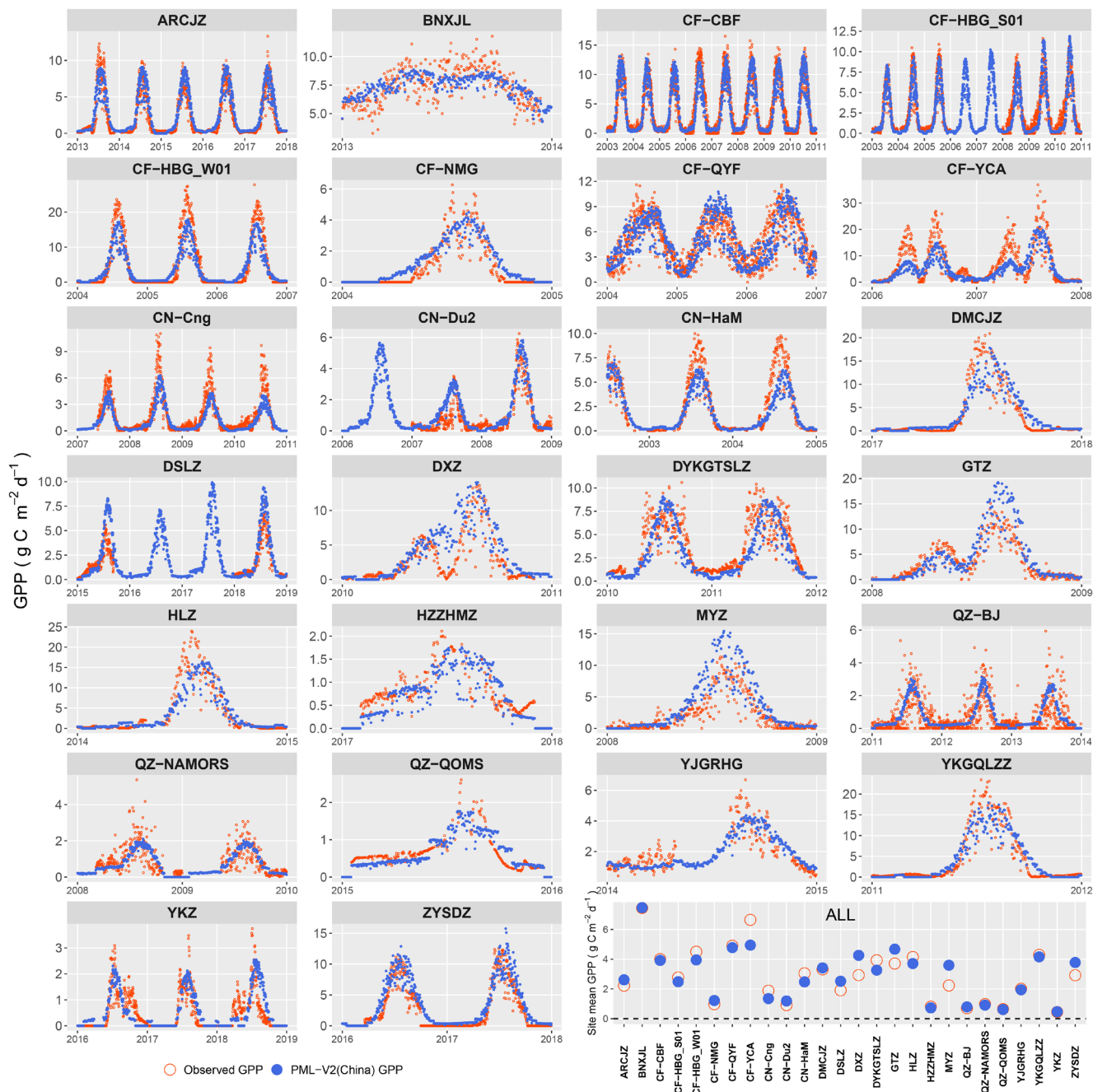



Figure 6: The daily GPP simulated by PML-V2(China) in calibration mode and the observed daily GPP variation in time series from 26 EC (see Figure 2) sites over China. 'ALL' represents the site mean value for each EC site.

Table 3: Statistical indicators of PML-V2(China) and other models for simulating ET and GPP at 26 EC flux towers. *NSE* and *R*
values are unitless. The unit of *RMSE* for ET is mm d^{-1} while it is g C m^{-2} d^{-1} for GPP. The unit of *Bias* is %.

Scale	Variable	Models	*NSE*	*R*	*RMSE*	*Bias*
daily	ET	PML-V2(China)	0.66	0.84	0.33	-7.97
		GLEAM	0.44	0.69	1.04	-14.45
		SEBAL	-7.10	0.16	3.95	5.31
8-day	ET	PML-V2(China)	0.74	0.87	0.66	-11.54
		PML-V2(Global)	0.62	0.80	0.81	-5.05
		MOD16A2	0.37	0.63	1.07	-10.90
daily	GPP	PML-V2(China)	0.76	0.87	0.87	-0.82
8-day	GPP	PML-V2(China)	0.75	0.87	1.93	-6.51
		PML-V2(Global)	0.68	0.82	2.17	-1.74
		MOD17A2H	0.49	0.78	2.74	-38.79
		EC-LUE	-0.04	0.35	3.91	-41.91
		VPM	0.21	0.60	3.41	-8.21

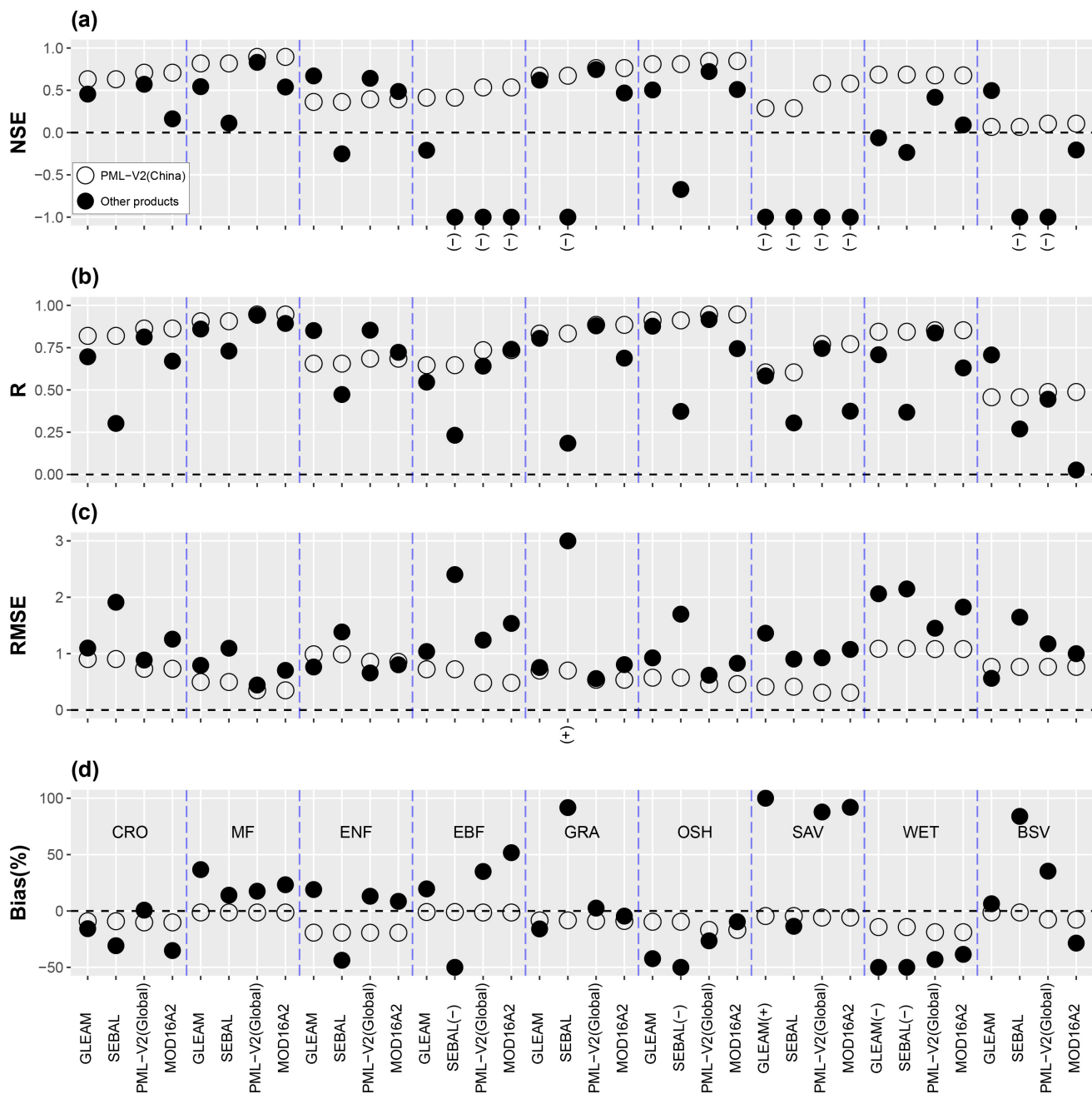



Figure 7: Statistical indicators of PML-V2(China) and other models for estimating ET at each PFTs. Open and solid dots represent PML-V2(China) estimated ET in cross-validation mode and other models. For the daily temporal resolution of GLEAM and SEBAL, PML-V2(China) is also daily scale; while for 8-day resolution of PML-V2(Global) and MOD16A2, the referred PML-V2(China) is upscaled to 8-day. Note that '(+)' indicates the model's simulation statistics dot is more than the upper bound while '(-)' indicates the model's simulation statistics dot is less than the low bound.

930 **Figure 8: Statistical indicators of PML-V2(China) and other models for estimating GPP at each PFTs. Open and solid dots represent PML-V2(China) estimated GPP in cross-validation mode and other models. PML-V2(China) is upscaled to 8-day to compare with the 8-day resolution of PML-V2(Global), MOD17A2H, EC-LUE, and VPM. Note that '(+)' indicates the model's simulation statistics dot is more than the upper bound while '(-)' indicates the model's simulation statistics dot is less than the low bound.**

935

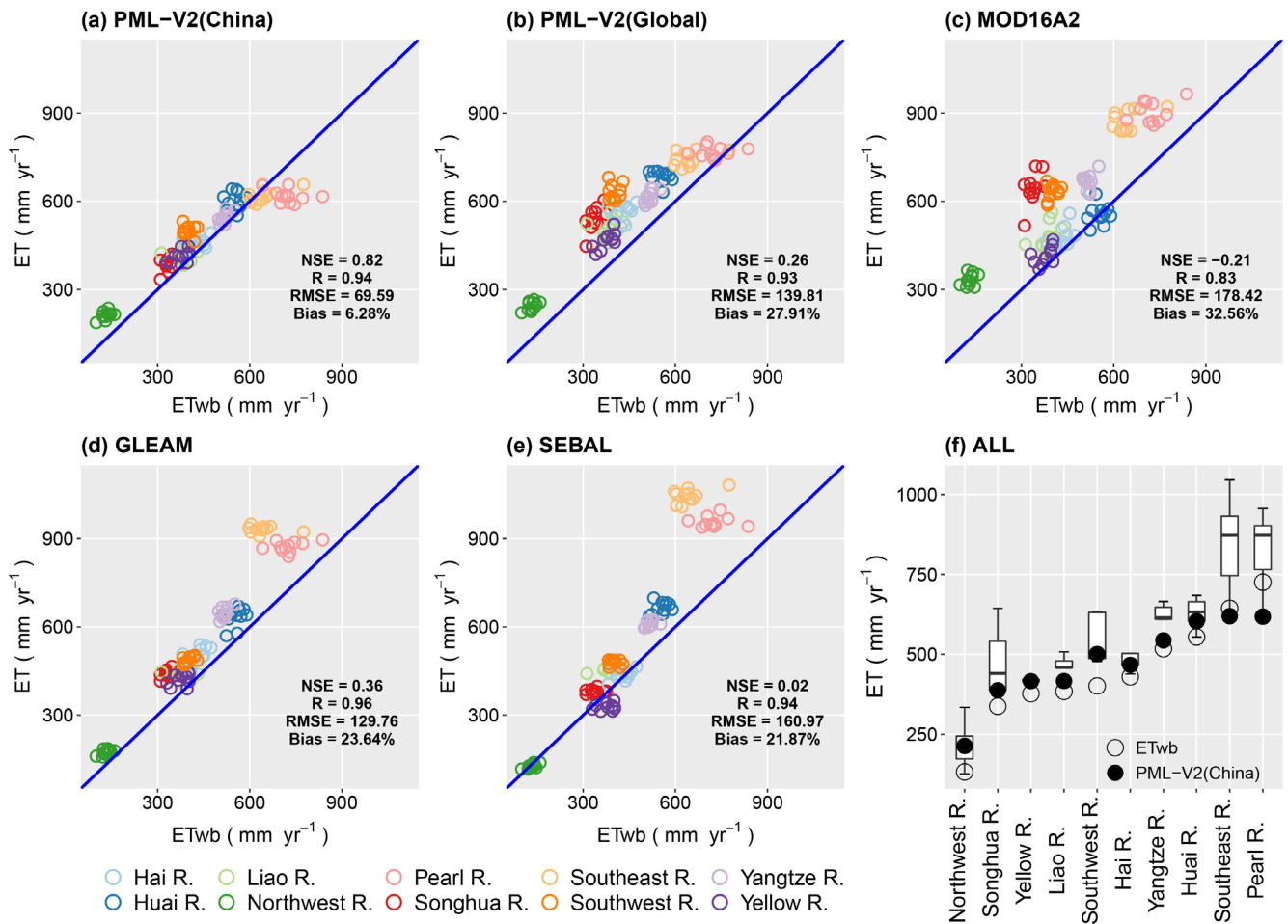



Figure 9: Annual evapotranspiration (ET) of (a) PML-V2(China), (b) PML-V2(Global), (c) MOD16A2, (d) GLEAM, and (e) SEBAL plotted against the water-balanced derived ET (ET_wb) values for ten major river basins over China during 2003–2013. The boxplot in (f) shows a multi-year mean of the five ET products above per river basin.

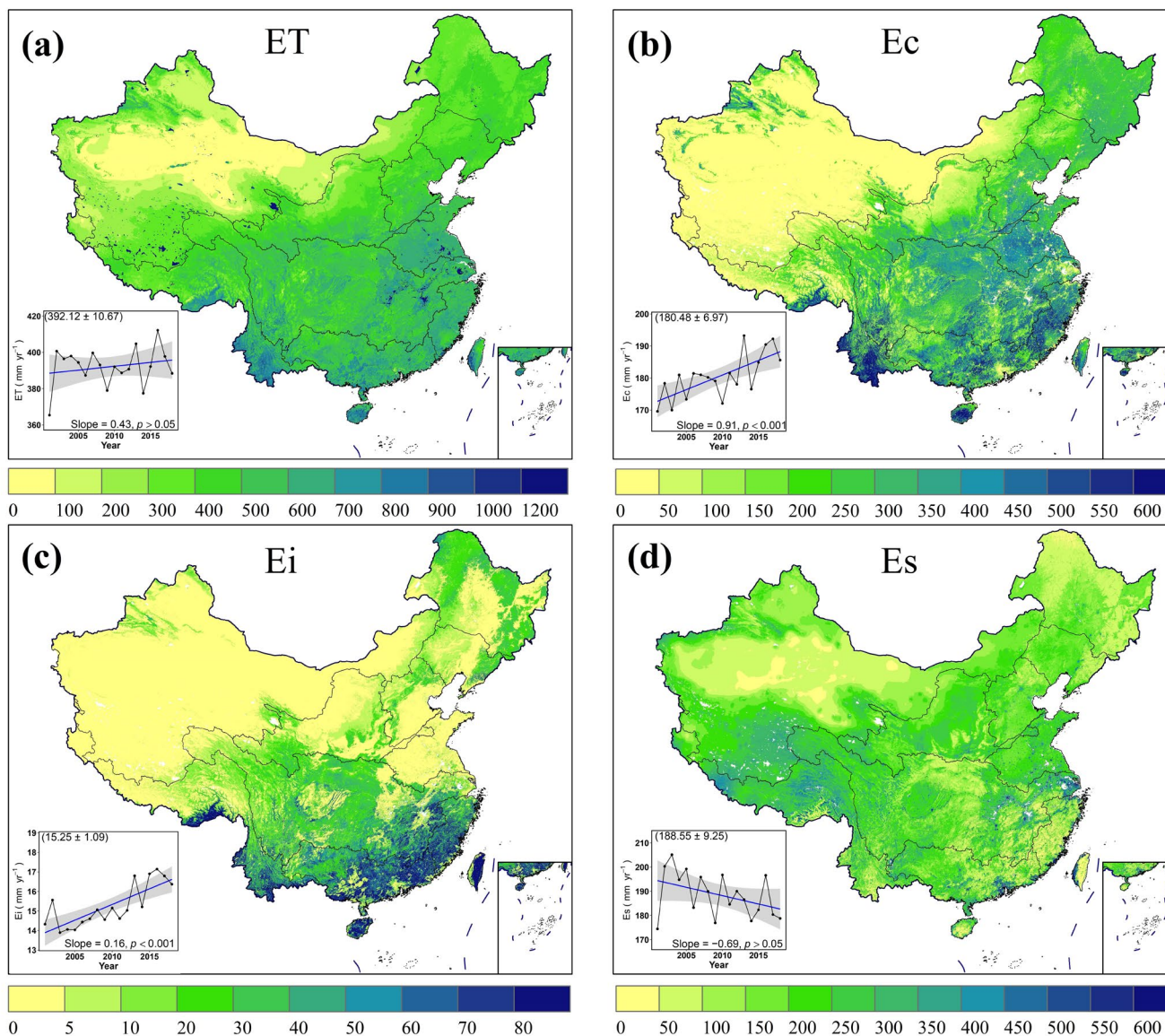
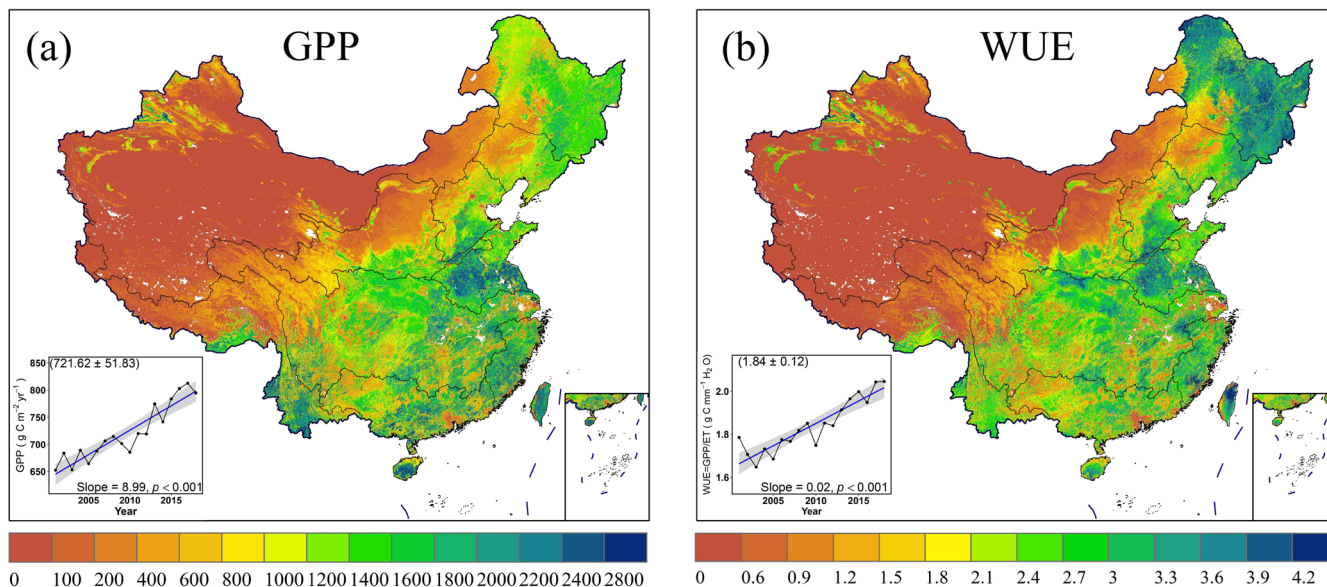



Figure 10: Spatial pattern of mean annual ET, E_c, E_i, E_s and their annual variation during 2001–2018. In all insets, the shaded areas represent the 95% confidence interval based on the linear regression modelling. The number in the parentheses of each inset is mean ± standard deviation of the annual simulated variables during the 18 years.

945 **Figure 11: Spatial pattern of mean annual GPP, WUE and their annual variation during 2001–2018. In all insets, the shaded areas represent the 95% confidence interval based on the linear regression modelling. The number in the parentheses of each inset is mean ± standard deviation of the annual simulated variables during the 18 years.**

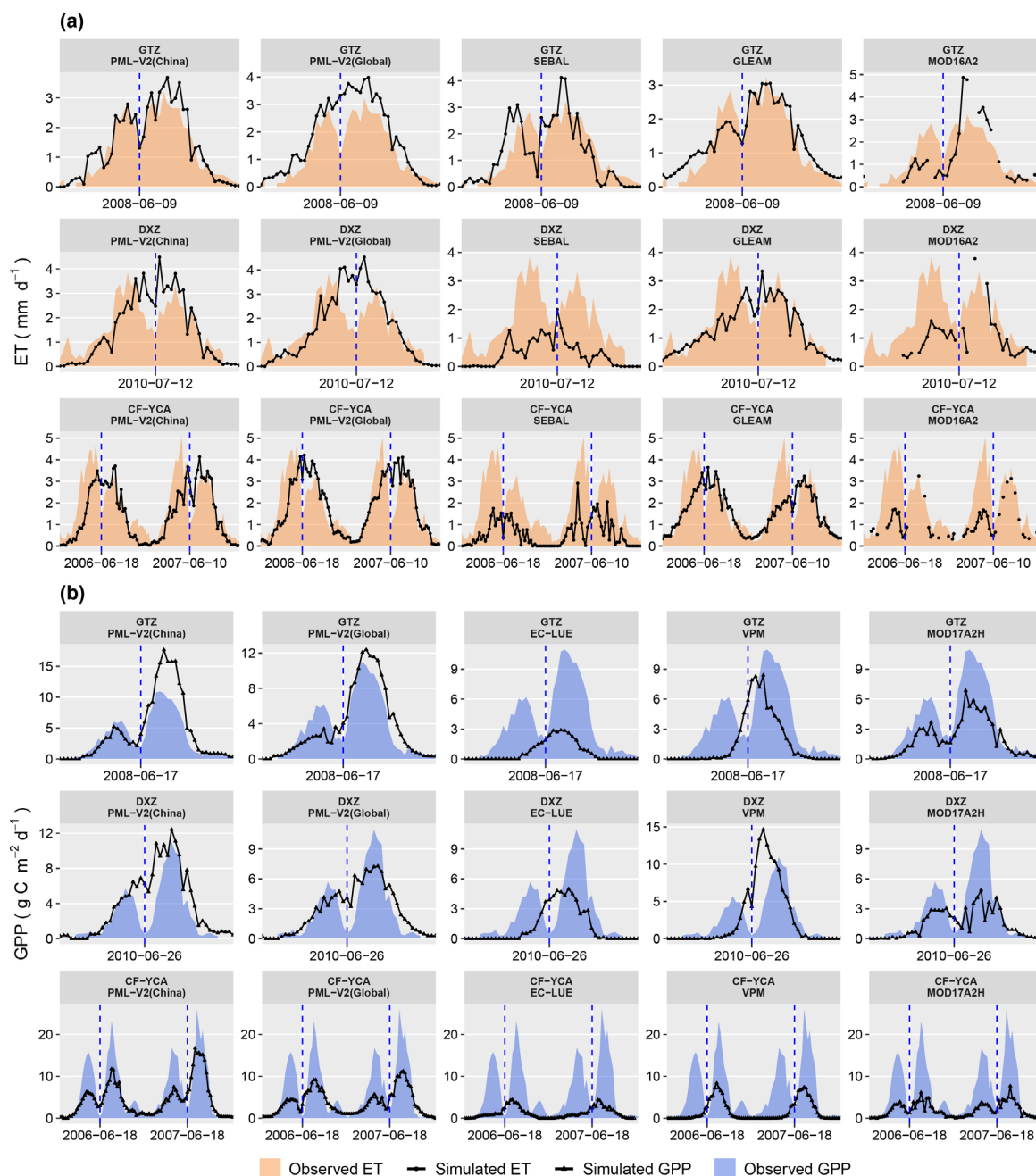



Figure 12: The intra-annual variation of (a) ET at three crop-rotation stations between the observed and the simulated by PML-V2(China) in validation mode, PML-V2(Global), SEBAL, GLEAM, and MOD16A2, respectively; (b) GPP at three crop-rotation stations between the observed and the simulated by PML-V2(China) in validation mode, PML-V2(Global), EC-LUE, VPM, and MOD17A2H, respectively. The blue dotted lines above pass through the lowest values between the two peaks of the observed ET or GPP per year. Note that all variables are averages every 8 days although units are per day.