We thank both referees for their constructive comments to which we reply below followed by the track changes version of the manuscript.

Response to Referee 1’s Comments

General comments

This manuscript is very interesting and valuable in developing a global accurate ET dataset. Currently, there are many global or regional ET datasets, but their performances vary across different regions. This manuscript provides an insightful approach in processing these datasets ensemble. However, there are many procedures to be clarified to inform the readers.

» **Answer:** Thank you very much for your positive comments and suggestions which for sure significantly improved the manuscript.

Major comments

I don’t understand the meanings of “best first and second levels” and “levels one and two validation metrics”. These two phrases have appeared many times and are vitally important to understand the synthesis procedure. If I understand this correctly, the performance metrics in Tables 5-8 were used to select two or three best ET datasets and the new dataset is produced by averaging these two or three datasets. A figure of processing procedure may be helping.

» Answer: Thank you for pointing this out. The level one of validation metrics has the highest R and TS values and the lowest ME, RME, RMSE, and RRMSE while the level two of validation metrics has the highest R and TS values and the lowest ME, RME, RMSE, and RRMSE after level one. You are right, the performance metrics in **Tables 5-8** were used to select two or three best ET datasets and the new dataset is produced by averaging these two or three datasets. For that, **Fig. 1**, **that shown below**, was created to preview the synthesization method and is included in the revised manuscript under **Section 3.2**. Hence, we rewrote **Section 3.2** as follows:

“There are 6 validation metrics including R, TS, ME, RME, RMSE, and RRMSE. The validation values of 6 metrics are categorized into levels. The level one of validation metrics has the highest R and TS values and the lowest ME, RME, RMSE, and RRMSE while the level two of validation metrics has the highest R and TS values and the lowest ME, RME, RMSE, and RRMSE after level one. For that, R and TS sorted descending while ME, RME, RMSE, and RRMSE sorted ascending (**Fig.1a**) then the corresponding ET product of each validation metric saved in a new table to be used to fill in **Fig.1b**.

The current study proposes three steps to develop a synthesized global ET dataset. First, the ET datasets are compared based on 6 validated metrics to generate a matrix to indicate level one and two of the validation metrics of all ET products over all comparison levels (**Fig.1b**). For each level, there are 6 validation metrics in rows and 26 ET values of different time periods and underlying conditions in columns (comparison levels), including monthly average (01), annual average (02), monthly (January-December: 03-14), land cover types (15-19), climate classes (20-23), and elevation levels (24-26). Thus, the total number of cells is 156 for each level. Each cell in the matrix represents one of twelve ET products that belong to this level. Then, to select ET data for further synthesis, the number and percentage of ET product occurrence at matrix (**Fig.2b**) of level one and two were calculated (**Fig.1c**). ET products were ranked in descending order based on the occurrence percentage of levels one and two (the last column in **Fig.1c**). Finally, the first two or three highly ranked ET products were selected to incorporate into the ensemble ET. For that, the selected ET products were resampled to a comparable spatial resolution if needed, and the average was used as the synthesized ET value."

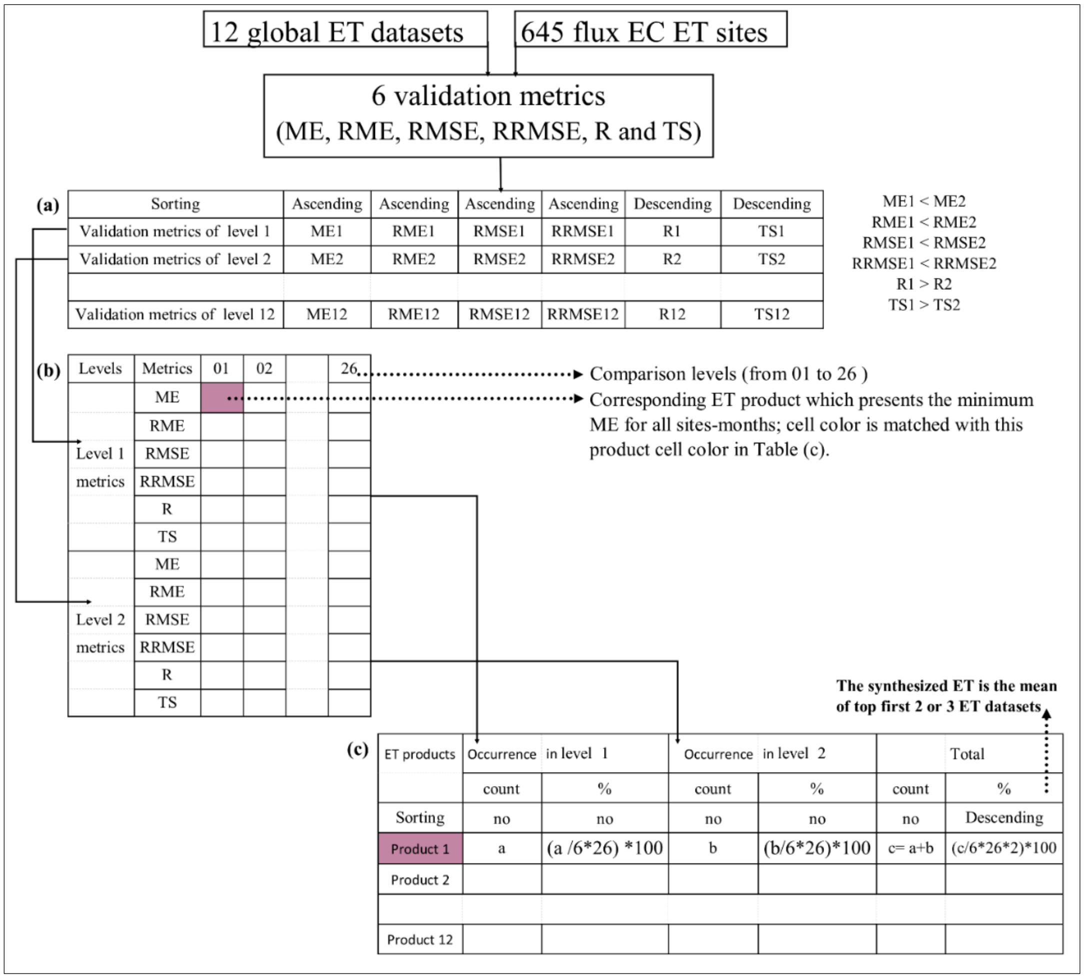

Fig.1. Flowchart of the synthesization method.

The second major problem is the validation data. By reading this manuscript, it could be confirmed that the observed EC ET data serve as validation data in evaluating and ranking the 12 ET datasets and also the validation data in evaluating the proposed Global Actual Evapotranspiration dataset. There could be an overfitting effect. It is like we use the same dataset as calibration data and validation data at the same time. Therefore, a cross-checking method should be applied. For example, 2/3 of the EC sites be used to evaluate the ET datasets and 1/3 of EC sites to validate.

» Answer: Thank you very much for your comment. You are right, it needs to split the in-situ data into 2 groups for calibration and validation. However, we do not calibrate ET products. We use in-situ data to see which one is performing better than others. Once ET products are selected, then we synthesize them into one and use in-situ data to validate to see if the synthesized data is better. Furthermore, From **Tables 1 and 3**, the flux EC ET sites, as well as the 12 ET products, are available in different periods. For evaluating each ET product the matched periods between EC sites and ET datasets were used (Xu et al. 2019; Li et al. 2018), that is why RME and RRMSE are included in the validation metrics. Further, the synthesized ET represented by the mean of PML and SSEBop about 60% (2003 to 2017) and the mean of NTSG and MOD16A2105 about 8% (2002-2002) indicating 68% of the synthesized ET are new data. For that, we used the matched period's method aiming to validate the new product under the same conditions of the experiment. We agreed with this method because we did not incorporate the flux EC ET data into the synthesized ET, it just serves as a ruler to prove to what extend each ET product works well in all comparison levels. Moreover, we used three regional ET datasets for comparison of consistent agreement over China, the United States, and the African continent to ensure the proposed product works well.

Another question that should be discussed is the scale problem. The EC sites normally work in a very limited area and can only present the ET condition of a small region. The related uncertainty should be discussed in the manuscript.

» Answer: Thank you for your very thoughtful comment. This is a common issue. The best way to validate the ET datasets is to use closure watershed water balances, however, these data set are quite a few. Flux EC ET data has its footprint, covering a larger area, but hard to match with a pixel. This issue still needs fundamental study. For that, we added **Lines 61-64**, as follows:

"Although flux EC ET is commonly flawed, particularly concerning energy balance closure at some sites (Foken, 2008; Helgason and Pomeroy, 2012), relatively short periods, and sparse spatial coverage, it is the most direct method for measuring the exchange between the surface and the atmosphere in different ecosystems (Foken et al., 2012; Baldocchi, 2014). Thus, site-pixel-level validation of certain ET products against flux EC ET as typically observed data has been performed by several studies in specific regions"

Minor comments

Line 11: What do you mean by “they produce different levels of uncertainties?”

» **Answer:** Thank you for that comment. We rewrote the sentence (**Lines 9-11**) to be clearer, as follows:

“Although it is difficult to estimate ET over a large scale and for a long time, there are several global ET datasets available with uncertainty associated with various assumptions regarding their algorithms, parameters, and inputs”.

In the abstract, the synthesization method should be indicated clearly.

» **Answer:** Thank you for your cogent advice. We agree and have indeed done that (**Lines 12-15**), as follows:

“Through a site-pixel evaluation of 12 global ET products over different time periods, land surface types, and conditions, the high performing products were selected for synthesis of the new dataset using a high-quality flux eddy covariance covering the entire globe.”

Line 74: check the time period.

» **Answer:** Thank you for pointing this out. We changed “**1998**-1995” to “**1989**-1995”.

Line 258: the title of subplot c.

» **Answer:** Thank you for pointing this out. We changed RMSE (mm): (**d**) to RMSE (mm): (**c**).

Line 371-392: Different datasets were selected due to data availability. That means for each period, for example before 2003 and 2003-2017, different datasets were used. My concern is that the ensemble means/variations may differ greatly. An adjustment in the period mean/variation should be considered.

» **Answer:** Thank you for your very thoughtful comment. Although we agree with you, this time series adjustment is very important and should be done in the future. Therefore, we have added Lines **549-551**, as follows:

“since different datasets were selected due to data availability, also future improvements may be focused on the adjustment of the ensemble means particularly for longterm pixel-based studies.”

Some tables and figure captions are similar. For example, Table 5 and Table 12. The major differences between them are the time period, which should be clearly indicated.

» **Answer:** We appreciate your advice. Tables and figures caption has been revised (**Figures 6 and 13; Tables 6-9, 12, 13**).

Response to Referee 2's Comments

The present article proposes a long-term synthesized ET product at a kilometer spatial resolution and monthly temporal resolution from 1982 to 2019. The authors made a trial application of GIS and remotely sensed data to reach the proposed aim of their study.

» Answer: Thank you very much for your positive comments and suggestions which for sure significantly improved the manuscript.

The presented article would be a good piece of work by supporting the conclusion with the obtained findings.

» Answer: Thank you for your very thoughtful comment. We have added the obtained findings to the Conclusions section as follows:

" The average annual ET from 1982–2019 is 567 mm $year^{-1}$. Although no product performed better in terms of all selected validation criteria in all classification levels, PML, GLDAS20, SSEBop, MOD16A2105, GLDAS21, SEBS, and NTSG are the sequence of their performances. The synthesized ET from PML, SSEBop, MOD16A2105 and NTSG agreed with the flux EC ET with R-values higher than 0.70, a maximum ME (RME) of 13.94 mm (17.13%) and a maximum RMSE (RRMSE) of 38.61 mm (47.45%) over 62% of all comparisons levels, as remotely sensed based ET product spanning from 1982 to 2019 with highest agreements, accuracies and lower biases over most of the land surface types and conditions. It performs well when compared with country-based and continental ET products over China, the United States and the African continent. However, the further synthesis of local ET products is encouraged if regional ET products are available.".

Synthesis of Global Actual Evapotranspiration from 1982 to 2019

Abdelrazek Elnashar[1,2,3], Linjiang Wang[1,2], Bingfang Wu[1,2*], Weiwei Zhu[1], Hongwei Zeng[1,2]

[1]State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100094, China
[2]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100049, China
[3]Department of Natural Resources, Faculty of African Postgraduate Studies, Cairo University, Giza, 12613, Egypt

Correspondence to: Bingfang Wu (wubf@aircas.ac.cn)

Abstract. As a linkage among water, energy, and carbon cycles, global actual evapotranspiration (ET) plays an essential role in agriculture, water resource management, and climate change. Although it is difficult to estimate ET over a large scale and for a long time, there are several global ET datasets available with ~~varied~~uncertainty associated with various assumptions regarding their algorithms, parameters, and inputs~~, and they produce different levels of uncertainties~~. In this study, we propose a long-term synthesized ET product at a kilometer spatial resolution and monthly temporal resolution from 1982 to 2019. Through a site-pixel ~~validation~~evaluation of ~~certain~~12 global ET products over different time periods, land surface types, and conditions, the high performing products were selected ~~through~~for synthesis of the new dataset using a high-quality flux eddy covariance covering the entire globe. According to the study results, Penman-Monteith Leuning (PML), operational Simplified Surface Energy Balance (SSEBop), Moderate Resolution Imaging Spectroradiometer (MODIS, MOD16A2105) and the Numerical Terradynamic Simulation Group (NTSG) ET products were chosen to create the synthesized ET set. The proposed product agreed well with flux EC ET over most of the all comparison levels, with a maximum ME (RME) of 13.94 mm (17.13%) and a maximum RMSE (RRMSE) of 38.61 mm (47.45%). Furthermore, the product performed better than local ET products over China, the United States, and the African continent and presented an ET estimation across all land cover classes. While no product can perform best in all cases, the proposed ET can be used without looking at other datasets and performing further assessments. Data are available on the Harvard Dataverse public repository through the following Digital Object Identifier (DOI): https://doi.org/10.7910/DVN/ZGOUED (Elnashar et al., 2020) as well as it is available as Google Earth Engine (GEE) application through this link: https://elnashar.users.earthengine.app/view/synthesizedet.

1. Introduction

Over most of the global land area, terrestrial evapotranspiration (ET) considers the second largest element of the hydrological cycle after precipitation (Waring and Running, 2007b;Bastiaanssen et al., 2014) and represents the linkage between water, energy, and carbon cycles (Gentine et al., 2019;Yang et al., 2016;Ferguson and Veizer, 2007) and ecosystem services (Almusaed, 2011;Yang et al., 2015;Revelli and Porporato, 2018).

Hence, the accurate estimation of global ET is essential for understanding the global hydrological cycle and water budgets (Oki and Kanae, 2006;Trenberth et al., 2007;Rodell et al., 2015), global drought (Sheffield et al., 2012;Naumann et al., 2018;Spinoni et al., 2019;Lu et al., 2019;Forootan et al., 2019), impacts of climate change (Waring and Running, 2007a;Zomer et al., 2008;Scheff and Frierson, 2014;Pan et al., 2015), climate change and global water resources (Arnell, 1999;Haddeland et al., 2014;Arnell and Lloyd-Hughes, 2014), global transboundary basin water scarcity (Degefu et al., 2018), water competition within a basin (Scott et al., 2001) and water stress/conflict within transboundary basins (Samaranayake et al., 2016;Munia et al., 2016;Bastiaanssen et al., 2014).

While precipitation and runoff, which are other paramount factors of the global water balance, can be directly measured by in situ weather stations and stream gauge networks as well as the availability of several datasets of remotely sensed precipitation (Funk et al., 2015;Ashouri et al., 2015;Huffman et al., 1997;Yamamoto and Shige, 2015), it is difficult to measure ET, especially at large spatial scales (Senay et al., 2012;Zhang et al., 2016).

Recently, several global ET datasets have become available for a variety of purposes, and they have been generated using remote sensing models, land surface models (LSM), and hydrological models (Trambauer et al., 2014;Li et al., 2018;Sörensson and Ruscica, 2018). There are many differences among these models concerning their algorithms, parameters, and inputs, and they produce different levels of uncertainty (Wang and Dickinson, 2012;Xu et al., 2019;Weerasinghe et al., 2020;Vinukollu et al., 2011a). The remote sensing model, which mainly focuses on thermal remote sensing and the energy balance equation, will be represented by MOD16A2 (Mu et al., 2011), PML (Zhang et al., 2019), SSEBop (Senay et al., 2013), SEBS (Chen et al., 2013), NTSG (Zhang et al., 2010), and GLEAM v3.3b (Miralles et al., 2011b). The land surface model uses quantitative methods to simulate the vertical exchanges of water and energy fluxes between the atmosphere and the land surface, as represented by GLDAS ET (Rodell et al., 2004), GLEAM v3.3a (Miralles et al., 2011b), and FLDAS (McNally et al., 2017). TerraClimate, which is a hydrological model, is based on a one-dimensional water balance approach (Abatzoglou et al., 2018). However, the availability of many datasets introduces challenges related to how users choose the appropriate dataset for their purposes (Wu et al., 2020).

Some studies have evaluated global ET products using an inferred estimate of ET obtained by subtracting discharge (Q) from precipitation (P), ET = P - Q, over global river basins (Zhang et al., 2010;Vinukollu et al., 2011a;Vinukollu et al., 2011b), continental river basins (Weerasinghe et al., 2020), transboundary river basins (Hofste, 2014), and national river basins (Zhong et al., 2020). Some, on the other hand, have used the ensemble ET product as observed data for evaluating certain ET products (Hofste, 2014;Trambauer et al., 2014;Andam-Akorful et al., 2015;Bhattarai et al., 2019).

~~Site-pixel-level validation of certain ET products against flux EC ET as typically observed data has been performed by several studies in specific regions (e.g., globally (Leuning et al., 2008;Zhang et al., 2010;Ershadi et al., 2014;Michel et al., 2016); Asia (Kim et al., 2012); South Africa (Majozi et al., 2017); Europe (Ghilain et al., 2011;Hu et al., 2015); North America (Jiménez et al., 2009;Mu et al., 2011); Europe and the United States (Miralles et al., 2011b); the United States (Vinukollu et al., 2011b;Velpuri et al., 2013;Xu et al., 2019); and China (Jia et al., 2012;Liu et al., 2013;Chen et al., 2014b;Tang et al., 2015;Yang et al., 2017;Li et al., 2018)).~~ Although flux EC ET is commonly flawed, particularly concerning energy balance closure at some sites (Foken, 2008;Helgason and Pomeroy, 2012), relatively short periods, and sparse spatial coverage, it is the most direct method for measuring the exchange between the surface and the atmosphere in different ecosystems (Foken et al., 2012;Baldocchi, 2014). Thus, site-pixel-level validation of certain ET products against flux EC ET as typically observed data has been performed by several studies in specific regions (e.g., globally (Leuning et al., 2008;Zhang et al., 2010;Ershadi et al., 2014;Michel et al., 2016);

Asia (Kim et al., 2012); South Africa (Majozi et al., 2017); Europe (Ghilain et al., 2011;Hu et al., 2015); North America (Jiménez et al., 2009;Mu et al., 2011); Europe and the United States (Miralles et al., 2011b); the United States (Vinukollu et al., 2011b;Velpuri et al., 2013;Xu et al., 2019); and China (Jia et al., 2012;Liu et al., 2013;Chen et al., 2014b;Tang et al., 2015;Yang et al., 2017;Li et al., 2018)).

Few previous studies have focused on merging certain ET products to create an ensemble ET product; for instance, (Vinukollu et al., 2011a;Mueller et al., 2013;Badgley et al., 2015). They used all ET products and created a merged product with a low spatial resolution. There are some global merged benchmarking evaporation products. Vinukollu et al. (2011a) generated an ensemble of six global ET datasets at a daily time scale and 0.5°×0.5° (≈55 km) spatial resolution for the period 1984–2007 using two surface radiation budget products and three models (i.e., surface energy balance, revised Penman-Monteith, and modified Priestley-Taylor). They reported that the ensemble simple mean value was reasonable; however, it was generally highly biased globally. Mueller et al. (2013) presented two monthly global ET products that differed in their input ET members and temporal coverage. The first dataset consisted of 40 datasets for the period ~~1998~~ 1989–1995, while the second dataset merged 14 datasets from 1989 to 2005. Their ET was derived from satellite and/or in situ observations (diagnostic) or calculated via LSM driven with observation-based forcing or output from atmospheric reanalysis. Hence, they provided four merged synthesis products, one including all datasets and three including datasets of each category (i.e., diagnostic, LSM, and reanalysis). They introduced the first benchmark products for global ET and found that its multi-annual variations showed realistic responses and were consistent with previous findings. Badgley et al. (2015) used a Priestly-Taylor Jet Propulsion Lab (PT-JPL) model with 19 different combinations of forcing data to produce global ET estimates from 1984 to 2006 at a 1°×1° (≈100 km) spatial resolution. The ensemble ET members changed according to the number of products available each year, which ranged between 4 and 12 members for 1999/2000 and 2001/2002, respectively. Their study focused on the uncertainty in global ET estimates resulting from each class of input forcing datasets.

However, from the aforementioned studies, we can report three findings: (1) no single ET product performed better than any other over different land surface types and conditions, (2) no one generated a single dataset for users, and (3) the created ensemble ET products relied on several individual ET products and were not based on the product with the best performance.

From our point of view, this work attempts to add to the growing scientific literature using a high-quality dataset from global flux towers for further validations and inter-comparison between different global ET products to understand their behavior within defined land cover types, elevation levels, and climatic classes. Moreover, we attempt to build an ensemble ET product that has a minimum level of uncertainty over as many conditions as possible. The study has two objectives: (1) to assess global ET products with in situ data derived from global flux towers across a variety of land surface types and conditions to gain a better understanding of the disparities among datasets and (2) to synthesize an ensemble global ET product with minimum uncertainties over more land surface types, climate systems, and monthly, annually and interannual time steps for a longer time.

2. Data

2.1. Evapotranspiration

Twelve global ET datasets were explored in the current study (Table 1 and Appendix A). Of them, 5 datasets used the Moderate Resolution Imaging Spectroradiometer (MODIS) as input, including two versions (V6 and V105) of Global Evapotranspiration Project (MOD16A2), Penman-Monteith Leuning ET (PML), the operational Simplified Surface Energy Balance ET (SSEBop) and the Surface Energy Balance System (SEBS). One dataset used the Advanced Very High-Resolution Radiometer (AVHRR) as input, including the Numerical Terradynamic Simulation Group (NTSG). The remainder mainly uses meteorological datasets as direct input, including field measurements such as TerraClimate and reanalysis datasets such as FLADS and GLADS. The algorithm used in 12 global ET datasets is mainly the Penman-Monteith model, except for FLADS and GLDAS, which use the LSM, and TerraClimate, which uses the soil water balance model. Priestley-Taylor is used to estimate evaporation from open water by NTSG while Penman evapotranspiration is used in PML for a water body, snow and ice evaporation. SSEBop, SEBS, NTSG, and GLEAM are individually managed, and other ET products, as well as elevation data, are available from GEE.

Table 1. Global ET products.

Product	Method	Satellite data	Meteorological data	Resolution Spatial	Resolution Temporal	Temporal coverage
MOD16A2 V6	P-M, SC	MODIS	GMAO	500 m	8 days	Jan 1, 2001 – Ongoing
MOD16A2 V105	P-M, SC	MODIS	GMAO	1 km	8 days	Jan 1, 2000 – Dec 31, 2014
PML	PML	MODIS	GLDAS V21	500 m	8 days	Jul 4, 2002 – Dec 27, 2017
SSEBop	P-M	MODIS	GDAS, PRISM	1 km	1 month	Jan 1, 2003 – Ongoing
SEBS	RS-SEB	MODIS, GLASS, GLAS	ERA-Interim	5 km	1 month	Jan 1, 2001 – Dec 31, 2010
NTSG	Modified P-M & P-T	AVHRR	NCEP/NCAR Reanalysis	8 km	1 month	Jan 1, 1982 – Dec 31, 2013
GLEAM 3.3b	P-T, SSF	Radiation & air temperature	Certain reanalysis data	0.25°	1 month	Jan 1, 2003 – Dec 31, 2018
GLEAM 3.3a	P-T, SSF	-	Certain reanalysis data	0.25°	1 month	Jan 1, 1980 – Dec 31, 2018
FLADS	LSM	MODIS-IGBP, UMD-AVHRR	MERRA-2, CHIRPS	0.10 °	1 month	Jan 1, 1982 – Dec 1, 2019
GLDAS V20	LSM	MCD12Q1, MOD44W, GTOPO30	NOAA/GDAS, GPCP, AGRMET	0.25°	3 hours	Jan 1, 1948 – Dec 31, 2010
GLDAS V21	LSM	MCD12Q1, MOD44W, GTOPO30	NOAA/GDAS, GPCP, AGRMET	0.25°	3 hours	Jan 1,2000 – Dec 23,2019
TerraClimate	SWB	Root zone storage capacity	WorldClim V1.4&2, CRU Ts4.0, JRA-55	0.25°	1 month	Jan 1, 1958 – Dec 1, 2018

Note: P-M: Penman-Monteith; PML: P-M Leuning; SC: Surface Conductance; P-T: Priestley-Taylor; RS-SEB: remotely sensed surface energy balance; LSM: land surface model; SWB: soil water balance; GMAO: Global Modelling and Assimilation Office for daily meteorological reanalysis data; GDAS: Global Data Assimilation System; PRISM: Parameter-elevation Regressions on Independent Slopes Model; GLASS: Global Land Surface Satellite; GLAS: Geoscience Laser Altimeter System; MERRA-2: Modern-Era Retrospective analysis for Research and Applications version 2; CHIRPS: Climate Hazards Group InfraRed Precipitation with Station data; RFE2: The African Rainfall Estimation version 2.0; NOAA: National Oceanic and Atmospheric Administration; GPCP: Global Precipitation Climatology Project; AGRMET: Agricultural Meteorological modeling system; CRU Ts4.0: Climate Research Unit time series data version 4.0; JRA-55: Japanese 55-year Reanalysis.

Three regional ET datasets were used for comparison of consistent agreement over China, the United States and the African continent (Table 2). Over China Mainland, The Complementary Relationship (CR) ET product was used (Ma et al., 2019); it is estimated monthly at a 0.1° (≈10 km) spatial resolution over 1982–2015 and can be retrieved from http://en.tpedatabase.cn/. For the United States, daily SSEBop was used (Savoca et al., 2013;Senay and Kagone, 2019). These data are produced at a 0.009°×0.009° (≈1 km) grid cell spatial resolution from 2000 to 2018 and can be downloaded from https://earlywarning.usgs.gov/ssebop/modis/daily/. Daily SSEBop aggregated to monthly time steps to be comparable with the synthesized ET temporal resolution. The Food and Agriculture Organization (FAO) Water Productivity through Open access of Remotely sensed derived ET product (FAO WaPOR version 2) was used for Africa (FAO, 2018, 2020). These data estimates are the sum of ET and interception, provided at a 0.002°×0.002° (≈250 m) spatial resolution with a monthly temporal resolution from 2009. WaPOR ET estimates are available through the following website: https://wapor.apps.fao.org/home/WAPOR_2/1/.

Table 2. Regional ET products.

Product	Method	Satellite data	Meteorological data	Resolution		Temporal coverage
				Spatial	Temporal	
CR	CR	MODIS	CMFD	10 km	1 month	Jan 1, 1982 – Dec 31, 2015
SSEBop	P-M	MODIS	NASA GDAS	1 km	1 day	Jan 1, 2000 – Dec 31, 2018
WaPOR	RS-SEB	MODIS	MERRA/GEOS-5, CHIRPS	250 m	1 month	Jan 1, 2009 – Ongoing

Note: CR: Complementary Relationship; P-M: Penman-Monteith; P-T: Priestley-Taylor; RS-SEB: remotely sensed surface energy balance; CMFD: China Meteorological Forcing Dataset; NASA GDAS: National Oceanic and Atmospheric Administration's (NOAA) Global Data Assimilation System; MERRA: Modern-Era Retrospective Analysis for Research and Applications; GEOS-5: Goddard Earth Observing System, Version 5; CHIRPS: Climate Hazards Group InfraRed Precipitation with Stations.

2.2. Flux EC data

Comprehensive flux EC ET data from 645 sites (Fig. 1 and Table 3), AmeriFlux; FluxNET; EuroFlux; AsiaFlux; and ChinaFlux, were collected and processed to examine the performance of different estimated ET products. The downloaded EC data are half-hourly text-type data, while the periods of flux EC ET ranged from 1 year (12 months) to 21 years (252 months) from 1994 to 2019. The gap-filling technique was applied to the downloaded in situ EC data (Reichstein et al., 2005). Different EC flux sites were spatially distributed on the heterogeneous underlying surface, corresponding to different land cover types according to the International Geosphere-Biosphere Programme (IGBP) classification system, which is recorded in each flux attribute data. The in-situ measured ET (mm day^{-1}) can be obtained by the half-hourly average latent heat flux (LE, W·m^{-2}s^{-1}) through Eq. (1), (Su, 2002):

$$\mathrm{ET} = \frac{\overline{\mathrm{LE}}}{\lambda} \times 3600 \times 24 \tag{1}$$

Where $\overline{\mathrm{LE}}$ (W·m^{-2}s^{-1}) is the daily average of the half-hourly average latent heat flux, and λ is the latent heat of evaporation. λ varies with air temperature in hydrologic or agricultural system modeling but only to a small extent (Walter et al., 2001), and the value acts directly on the accuracy of the estimated in situ measured ET. Considering that there are very limited impacts of the changes in air temperature on the estimated in-situ measured ET (Henderson-Sellers, 1984;Li et al., 2018), the constant value of 2.45 MJ kg^{-1} is fixed in the calculation above (Walter et al., 2001).

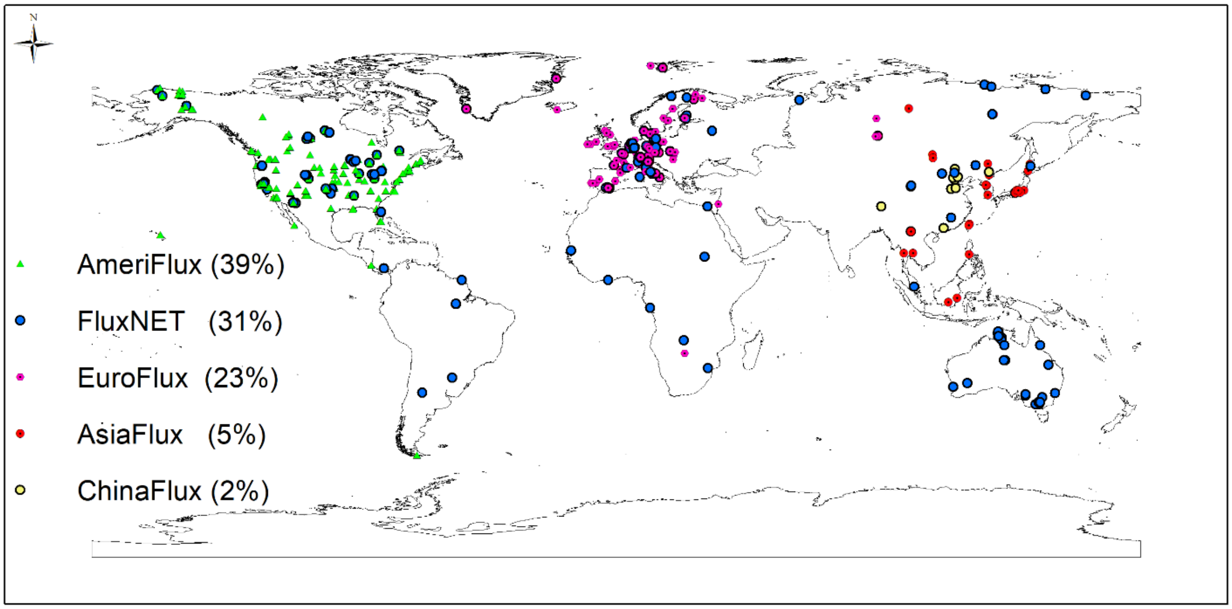

Figure 1. Spatial distribution of 645 in-situ flux EC sites across the world.

Table 3. Summary of 645 in-situ EC flux sites.

Flux	Sites number	Temporal coverage	Elevation range (m)	Underlying surface IGBP type	Website
AmeriFlux	249	1994–2019	-9 to 3199	ENF/EBF/DBF/MF/CSH/OSH/WSA/SAV/GRA/WET/CRO/SNO/BSV/WAT	ameriflux.lbl.gov
FluxNET	203	1994–2019	-10 to 4312	ENF/EBF/DNF/DBF/MF/CSH/OSH/WSA/SAV/GRA/WET/CRO	fluxnet.fluxdata
EourFlux	148	1996–2018	-4 to 2436	ENF/EBF/DBF/MF/CSH/OSH/WSA/SAV/GRA/WET/CRO/SNO	europe-fluxdata.eu
AsiaFlux	33	2000–2015	0 to 3308	ENF/EBF/DNF/DBF/MF/GRA/CRO/URB/WAT	asiaflux.net
ChinaFlux	12	2003–2017	26 to 4317	EBF/MF/GRA/CRO	chinaflux.org

Note: ENF: Evergreen Needleleaf Forests; EBF: Evergreen Broadleaf Forests; DBF: Deciduous Broadleaf Forests; MF: Mixed Forests; CSH: Closed Shrublands; OSH: Open Shrublands; WSA: Woody Savannas; SAV: Savannas; GRA: Grasslands; WET: Permanent Wetlands; CRO; Croplands; URB: Urban and Build-up Lands; SNO: Permanent Snow and Ice; BSV: Barren or Sparsely Vegetated Area; WAT: Water Bodies.

2.3. Aridity index

The mean global aridity index dataset was produced by (Zomer et al., 2008) using WorldClim global climate data. The aridity index was estimated as the mean annual precipitation divided by the mean annual potential evapotranspiration, and the latter was calculated by the Hargreaves equation. The spatial resolution was 0.0083°×0.0083° (≈1 km) grid cell (Trabucco and Zomer, 2018) and the data can be downloaded from the following website: https://cgiarcsi.community/data/global-aridity-and-pet-database/

2.4. Elevation data

The Shuttle Radar Topography Mission (SRTM) data were provided at a resolution of one arc-second and void-filled (Farr et al., 2007). For the geographic areas outside the SRTM coverage area, the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010), which have a resolution of 7.5 arc-seconds, were used (Danielson and Gesch, 2011).

3. Methods

3.1 Assessment

Because ET is highly variable in both space and time (Schaffrath and Bernhofer, 2013;Fisher et al., 2017), a comprehensive evaluation from different perspectives is required (Trambauer et al., 2014;McCabe et al., 2016;Li et al., 2018). For each flux tower location, the aridity index, elevation and estimated ET data were extracted. The aridity index was classified (Table 4) according to the United Nations Environment Programme definition (UNEP, 1997) into four classes (i.e., humid: 361 (56%), semiarid: 167 (26%), dry sub-humid: 82 (13%), and arid: 35 (5%)). Elevations were classified into three levels (i.e., <500 m: 452 (70%), 500 m–1500 m: 135 (21%), and >1500 m: 58 (9%)). Land cover included five types (i.e., forests: 349 (54%), grasslands: 128 (20%), croplands: 89 (14%), water bodies: 73 (11%), and others (barren land and permanent snow and ice): 6 (1%)). Accordingly, the following metrics were estimated using Eqs. (2-7):

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n} Y_i - X_i \tag{2}$$

$$\mathrm{RME} = \frac{\mathrm{ME}}{X} \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - X_i)^2}{n}} \tag{4}$$

$$\mathrm{RRMSE} = \frac{\mathrm{RMSE}}{X} \tag{5}$$

$$R = \frac{\sum_{i=1}^{n}[(Y_i - Y)(X_i - X)]}{\sqrt{\sum_{i=1}^{n}(Y_i - Y)^2}\sqrt{\sum_{i=1}^{n}(X_i - X)^2}} \tag{6}$$

$$\mathrm{TS} = \frac{4(1+R)}{\left(\mathrm{std} + \frac{1}{\mathrm{std}}\right)^2 (1 + R_0)} \tag{7}$$

Where ME is the mean error; RME is the relative mean error; RMSE is the root mean square error; RRMSE is the relative root mean square error; R is the correlation coefficient; TS is the Taylor score; n is the sample number; i is the i^{th} sample; X is the mean of the observed EC ET data; Y is the mean of different estimated ET data; std is the standard deviation of the estimated ET normalized by the standard deviation of the observed EC ET; and R_0 is the maximum theoretical R, with an R_0 value of 0.9976 (Taylor, 2001).

The magnitude of ME (the absolute value) is used as a bias indicator (Mu et al., 2011;Yang et al., 2017), while its sign indicates whether different ET products overestimate or underestimate the flux EC ET values. The accuracy of each ET product can be described by the RMSE (Miralles et al., 2011b;Hu et al., 2015). Moreover, the relative values of ME and RMSE are used for a fairer comparison between certain ET products among different regions and periods (Majozi et al., 2017). In addition, correlation coefficients (R values) are used to measure the strength of the relation between flux EC ET and different ET products (Ghilain et al., 2011;Hu et al., 2015), and with the aid of the Taylor score (TS), the overall performance of each product can be described well (Taylor, 2001;Mu et al., 2011). To rank each ET product, the lower ME, RME, RMSE, and RRMSE values and the higher R and TS values are desired; lower biases and higher accuracies.

Table 4. Climate classification according to the global aridity index values.

Aridity Index value	Climate class
<0.03	Hyper arid
0.03 – 0.20	Arid
0.20 – 0.50	semiarid
0.50 – 0.65	Dry sub-humid
>0.65	Humid

3.2 Synthesis method

There are 6 validation metrics including R, TS, ME, RME, RMSE, and RRMSE. The validation values of 6 metrics are categorized into levels. The level one of validation metrics has the highest R and TS values and the lowest ME, RME, RMSE, and RRMSE while the level two of validation metrics has the highest R and TS values and the lowest ME, RME, RMSE, and RRMSE after level one. For that, R and TS sorted descending while ME, RME, RMSE, and RRMSE sorted ascending (Fig. 2a) then the corresponding ET product of each validation metric saved in a new table to be used to fill in Fig. 2b.

The current study proposes three steps to develop a synthesized global ET dataset. First, the ET datasets are compared based on 6 validated metrics~~, in which~~ to generate a matrix ~~was developed~~ to indicate level one and two of the validation metrics of all ET products over all comparison levels~~, see Table 5. There~~ (Fig. 2b). For each level, there are ~~six~~6 validation ~~criteria~~metrics in rows ~~(i.e., ME (mm), RME (%), RMSE (mm), RRMSE (%), R, and TS)~~and 26 ET values of different periods and ~~26~~ underlying conditions in columns (comparison levels~~ in columns (i.e.,~~), including monthly average (01), annual average (02), monthly (January-December: 03~~-~~–14), land cover types (15~~-~~–19), climate classes (20~~-~~–23), and elevation levels (24~~-~~–26~~)). The~~). Thus, the total number of cells is 156~~.~~ for each level. Each cell in the matrix represents ~~a free competition between certain~~one of twelve ET products that belong to ~~occupy~~ this ~~cell based on each validation criterion.~~level. Then, ~~selecting~~to select ET data for further synthesis, ~~based on the magnitudes (absolute values) of each validation index of all ET products across all comparison classes (01-26), the best first~~the number and ~~second levels of ET products within each cell were selected; additionally, the count and percent of each~~ percentage of ET product ~~in all cells~~occurrence at matrix (Fig. 2b) of level one and two were calculated ~~to calculate the total count and percent from levels one and two, see Table 6. All~~(Fig. 2c). ET products ~~will be sorted~~were ranked in descending order based on the ~~total~~occurrence percentage of levels one and two~~.~~ (the last column in Fig. 2c). Finally, the first two or three highly ranked ET products were ~~incorporated~~selected to incorporate into the ensemble ET. For that, the selected ET products were resampled to a comparable spatial resolution if needed, and the average was used as the synthesized ET value.

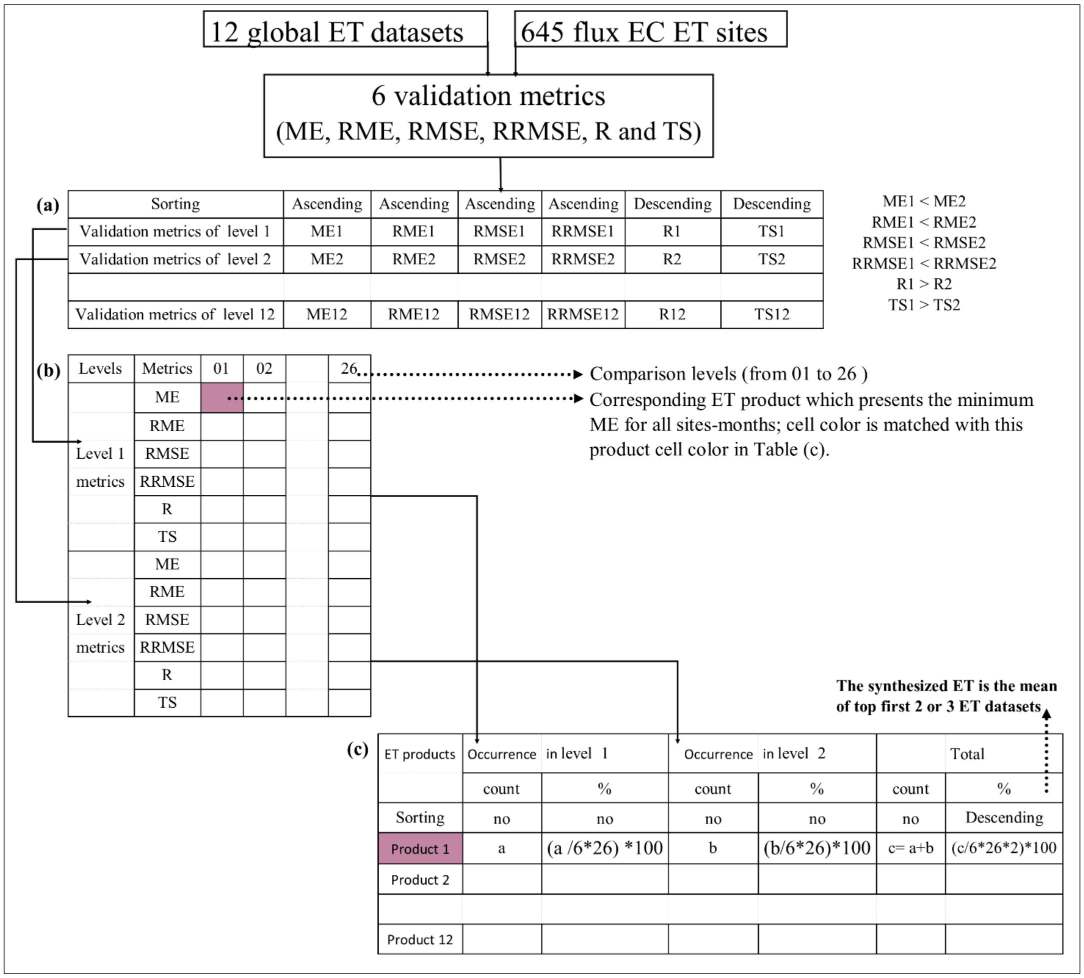

Figure 2. Flowchart of the synthesization method.

4. Results

4.1. Assessment of existing global ET datasets

Figure ~~2~~3 shows that seasonality exists and is captured well by all ET datasets, with some exceptions over barren land, permanent snow and ice, and arid areas (not shown). The maximum ET occurs during July and differs according to each ET dataset. Generally, MOD16A2 represents the minimum estimated ET across all conditions, while SSEBop represents the maximum ET across all conditions except over humid regions and at elevations between 500 m and 1500 m. From Figures (~~3, 5–11~~4, 6–12), the best-fitted linear regression line (blue-solid line) compared to the 1:1 line (red-dashed line), all ET datasets overestimate the flux EC ET in lower ET values and underestimate the flux EC ET in higher ET values with two exceptions. The first exception is over barren land and permanent snow and ice, where MOD16A2 underestimates and GLDAS21, GLEAM33a, and TerraClimate overestimate under both lower and higher ET values (not shown). Second, in dry sub-humid areas, SSEBop (Fig. 9c3) and GLDAS21 (Fig. 9e3) overestimate under both lower and higher ET values. Applying for the highest R (TS) and lowest error metrics role, MOD16A2 cannot present any role; additionally, only one contribution by the lowest RRMSE was found in February and the highest TS was found in March for TerraClimate and GLEAM33b, respectively.

4.1.2. Validation by all sites' monthly ET

Figure ~~3~~4 shows that only SEBS and MOD16A2 underestimate flux EC ET. PML is the dataset that best agrees with the observed ET, and it had the lowest RMSE (RRMSE). MOD16A2105 returned the smallest absolute ME, while SEBS yielded the smallest RME. Figure ~~4~~5 shows there are interannual differences between certain ET product performances. MOD16A2 shows negative MEs and RMEs for all months, with larger biases during March, April, and May, while FLDAS shows positive MEs and RMEs for all months, with larger biases during March, April May, June, and July. For other products, the ME and RME signs vary among months; for instance, the ME and RME values of GLDAS21 are negative (underestimated) during February, September, and November and positive (overestimated) in the remaining months, with larger biases during March, April, May, June, and July. The RMSE declines from January to February and then increases until July and declines again until November. The minimum RMSE values occur during February, November, and December, while the maximum values occur during June, July, and August.

For instance, the RMSE in July ranges from 36.28 mm to 52.41 mm for FLDAS and PML, respectively, while it ranges from 17.08 mm to 21.68 mm for PML and SEBS, respectively. RRMSE declines from January reaches its minimum in June and then increases again until December, except for SEBS in December. The highest values of RRMSE (>80%) occur in January, February, November, and December except for SEBS in December, while the lowest values (<60%) exist in June, July, and August. The R-value declines from January and reaches its minimum in May; it then increases starting in August. Except for MOD16A2, all products have an R-value greater than 0.60 during January, February, November, and December. SEBS has the lowest R-value during March, April, May, and June, while PML yields the highest R-value during all months except January and December. Except for MOD16A2 in February, which has a TS value above 0.60, as with the R-value, the TS declines from January, reaches its minimum in May, and then increases again starting in August. Figures ~~3 and~~ 4 and 5 show these products yield intra-annual ET variations but vary in their performance according to the selected validation metrics, which also vary among all months (from January to December).

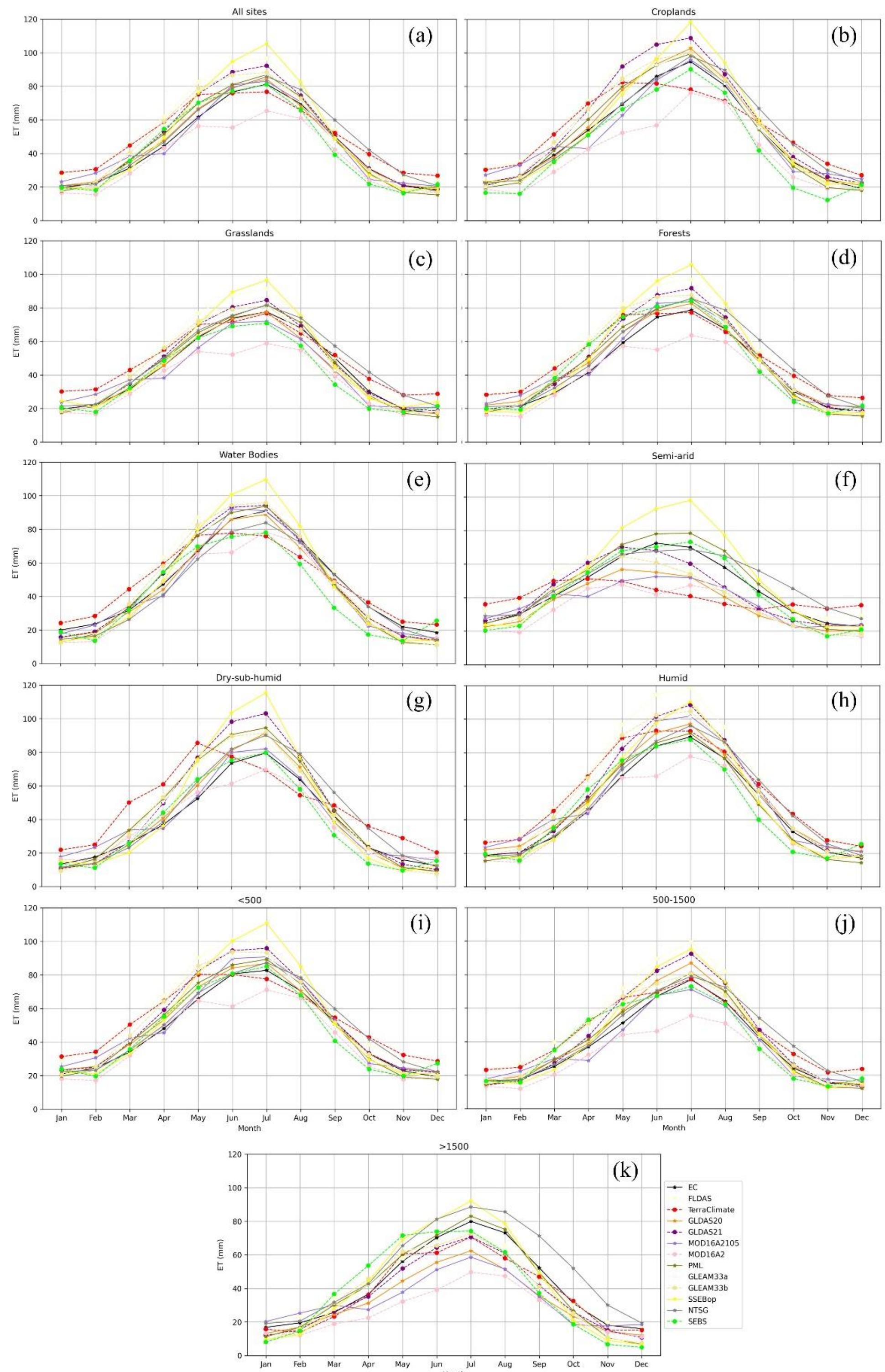

Figure ~~2~~3. Monthly average flux EC ET and 12 ET products over all flux sites (**a**), land cover types (croplands: (**b**); grasslands: (**c**); forests: (**d**); water bodies: (**e**)), climate classes (semiarid: (**f**); dry sub-humid: (**g**); humid: (**h**)), and elevation levels (<500 m: (**l**), 500 m-1500 m: (**j**), and >1500m: (**k**)).

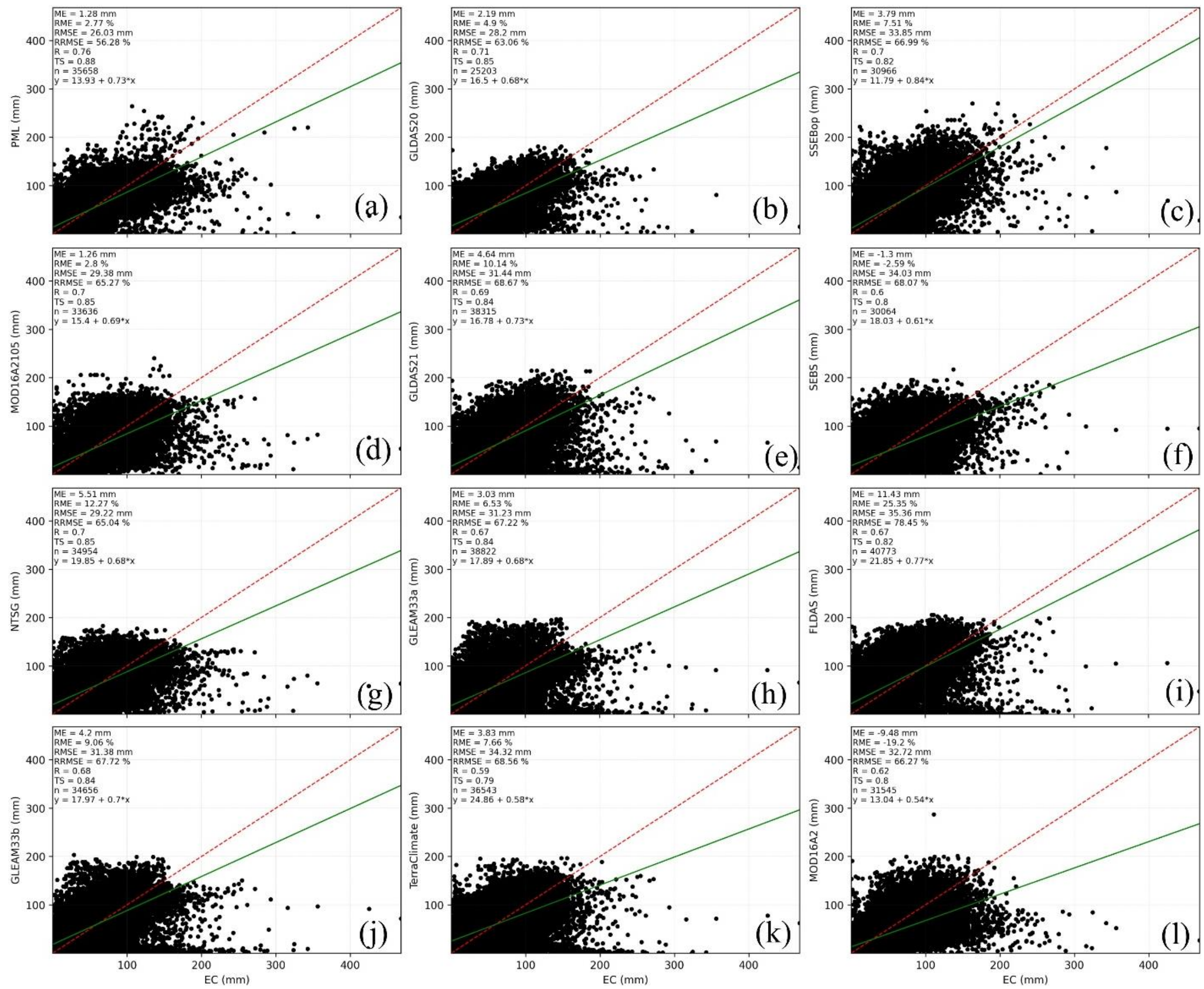

Figure ~~3~~4. Monthly ET products (PML: (**a**); GLDAS20: (**b**); SSEBop: (**c**); MOD16A2105: (**d**); GLDAS21: (**e**); SEBS: (**f**); NTSG: (**g**); GLEAM33a: (**h**); FLDAS: (**i**); GLEAM33b: (**j**); TerraClimate: (**k**); MOD16A2: (**l**)) against flux EC ET aggregated for all sites.

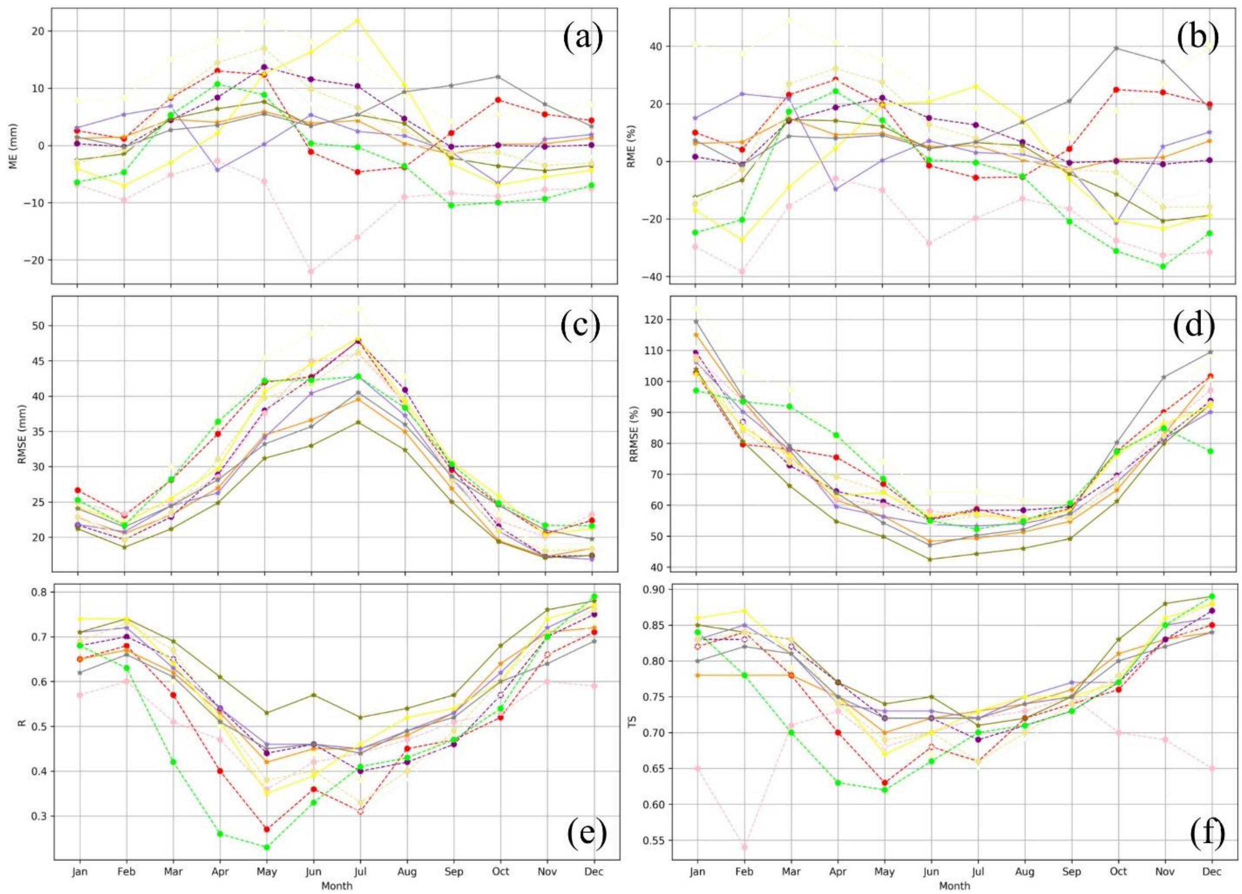

Figure ~~4~~5. Monthly validation metrics (ME (mm): (**a**); RME (%): (**b**); RMSE (mm): (**~~d~~c**); RRMSE (%): (**d**); R: (**e**); TS: (**f**)) of ET products against flux EC ET for all sites (legend as Figure ~~2k~~3k).

4.1.3. Validation by all sites' annual ET

Figure ~~5~~6 shows all ET products overestimate the observed ET with two exceptions~~,~~; SEBS and MOD16A2. In all environmental conditions, PML has the highest R (TS) and the lowest ME (RME) and RMSE (RRMSE). Figures ~~3~~4 and ~~5~~6 indicate the obvious error metrics of annual scale performances that are consistent with those that come from the monthly time step. The lowest and highest absolute values of ME (RME) for monthly ET exist in MOD16A2105 (SEBS) and FLDAS, respectively, while those for annual ET exist in PML and FLDAS, respectively. Furthermore, PML yields the largest R and TS values for monthly and annual ET, but the minimum values of R and TS were registered with TerraClimate and MOD16A2 for monthly and annual ET, respectively. This result may be attributed to the aggregation of monthly ET into annual values.

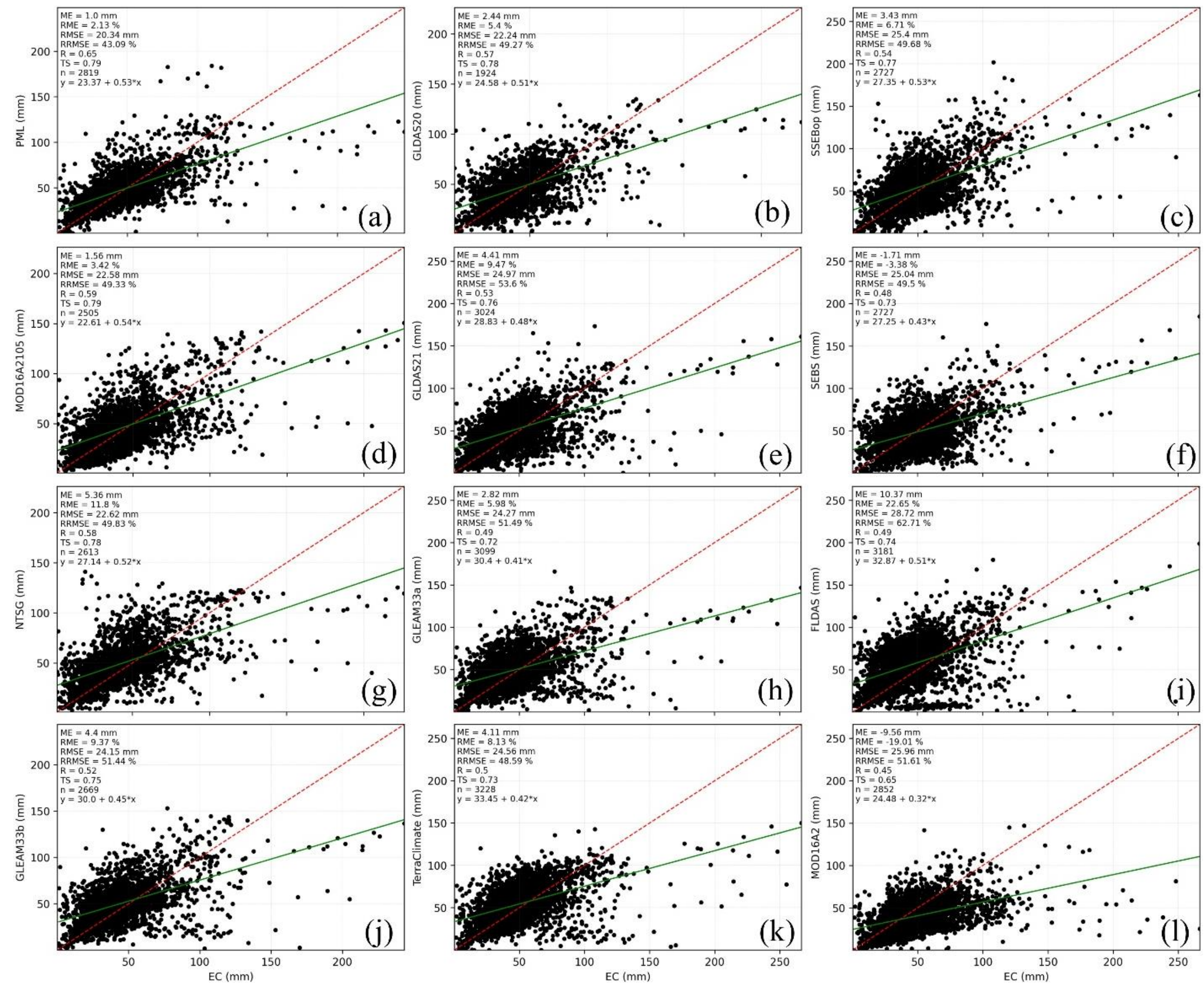

Figure ~~5~~6. Annually ET products ~~(PML: (**a**); GLDAS20: (**b**); SSEBop: (**c**); MOD16A2105: (**d**); GLDAS21: (**e**); SEBS: (**f**); NTSG: (**g**); GLEAM33a: (**h**); FLDAS: (**i**); GLEAM33b: (**j**); TerraClimate: (**k**); MOD16A2: (**l**))~~ against flux EC ET aggregated for all sites~~.~~ (subplot label as in Figure 4).

4.1.4. Validation by land cover types

Figures ~~6~~7 and ~~7~~8 show that, according to the ME (RME) sign, except for some ET products over croplands (i.e., MOD16A2, SEBS, MOD16A2105, and PML), grasslands (i.e., MOD16A2, SEBS, MOD16A2105, GLDAS20, and PML), forests (MOD16A2), and barren land and permanent snow and ice (i.e., MOD16A2105, MOD16A2, FLDAS, and GLDAS20), which underestimate the flux EC ET, the other ET products overestimate. For water bodies, MOD16A2105, GLEAM33b, GLDAS20, and FLDAS overestimate, while the other products produce underestimates. Over croplands, grasslands, and forests, PML is the best product for R (TS) and RMSE (RRMSE). Additionally, it has the highest TS over water bodies. SSEBop, GLEAM33a, SEBS, NTSG, and GLDAS20 obtained the desired ME (RME) over croplands, grasslands, forests, water bodies, and barren land and permanent snow and ice, respectively. GLEAM33a also represents the highest R (TS) with the lowest RRMSE, while GLDAS20 has the smallest RMSE over barren land and permanent snow and ice. In addition, GLDAS20 has the lowest RMSE, while SSEBop has the highest R and lowest RRMSE over water bodies, see Table 5 (level one: 15–19).

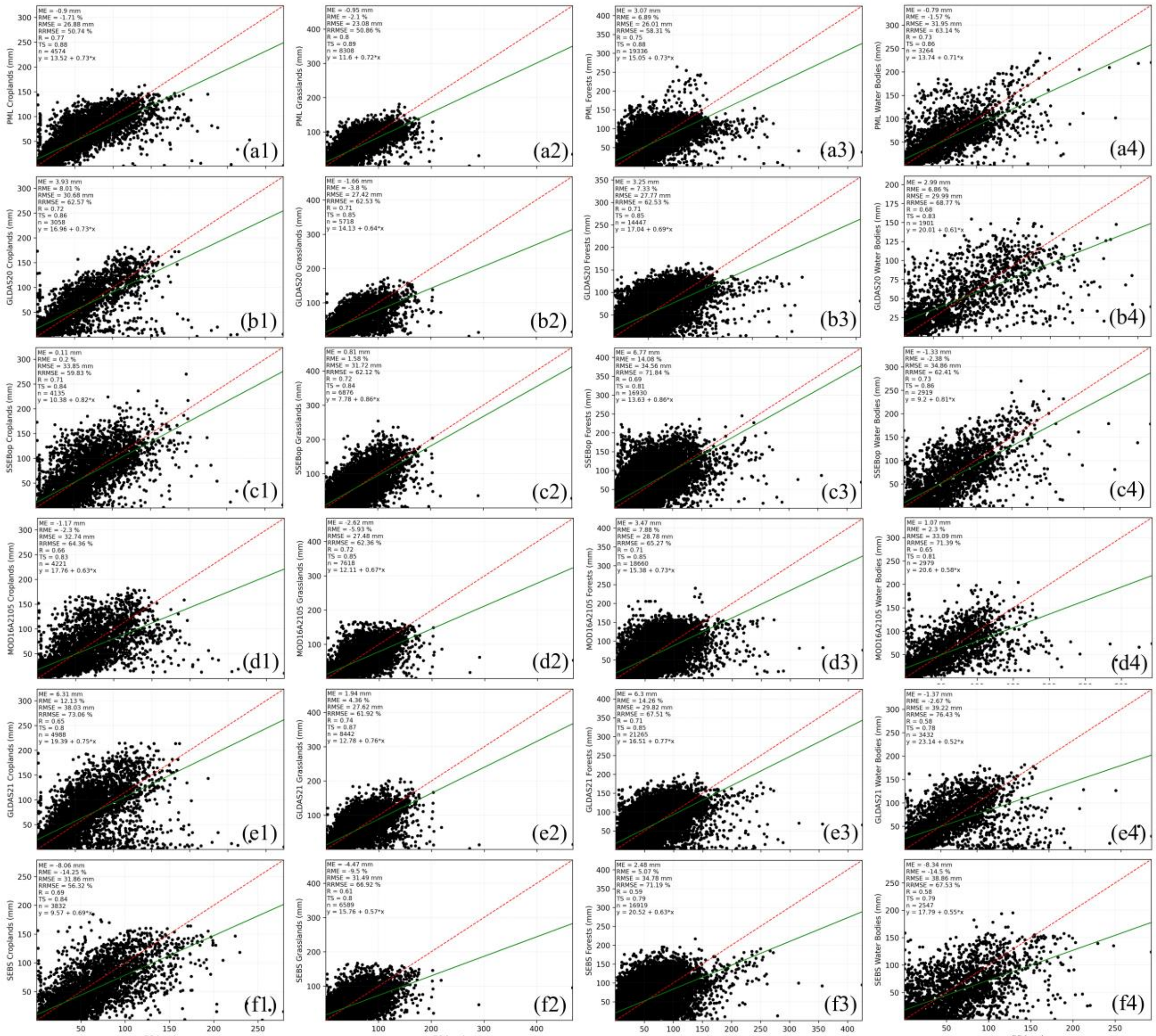

Figure ~~6~~7. Monthly ET products (PML: (**a**); GLDAS20: (**b**); SSEBop: (**c**); MOD16A2105: (**d**); GLDAS21: (**e**); SEBS: (**f**)) against flux EC ET aggregated for all sites for each land cover type (croplands: (**1**); grasslands: (**2**); frosts: (**3**); water bodies: (**4**)).

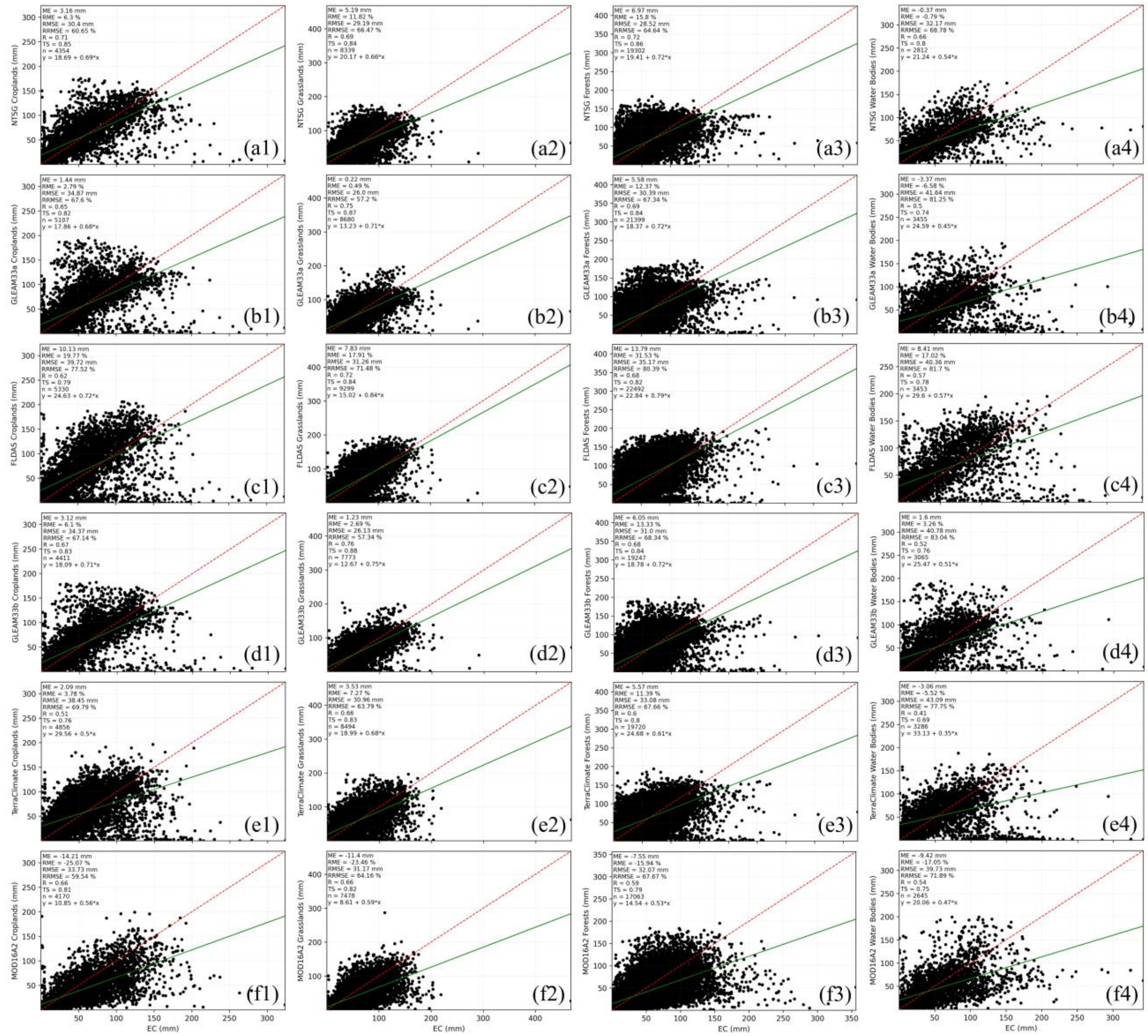

Figure ~~7~~8. Monthly ET products (NTSG: (**a**); GLEAM33a: (**b**); FLDAS: (**c**); GLEAM33b: (**d**); TerraClimate: (**e**); MOD16A2: (**f**)) against flux EC ET aggregated for all sites for each land cover type (croplands: (**1**); grasslands: (**2**); frosts: (**3**); water bodies: (**4**)).

4.1.5. Validation by climate classes

Figures ~~8~~9 and ~~9~~10 show that SEBS, PML, NTSG, and SSEBop in arid areas and PML, NTSG, and SSEBop in semiarid areas overestimate values, while MOD16A2 and SEBS in dry sub-humid areas and MOD16A2, SEBS, and PML in humid areas underestimate values; for each aridity index class, other products were the opposite. Over humid areas, PML represents the highest agreement and accurate dataset compared to the flux EC ET. Furthermore, it had the highest R (TS) in the arid and semiarid areas and the smallest RMSE (RRMSE) in semiarid areas. GLDAS20 yielded the largest R (TS) with the smallest RMSE (RRMSE) in dry-sub-humid regions; over these regions, MOD16A2105 presented the best ME (RME). FLDAS has two contributions, with the smallest ME (RME) and RMSE (RRMSE) in semiarid and arid areas, respectively, while GLDAS21 has only one point over arid areas where the best ME (RME) is found, see Table 5 (level one: 20–23).

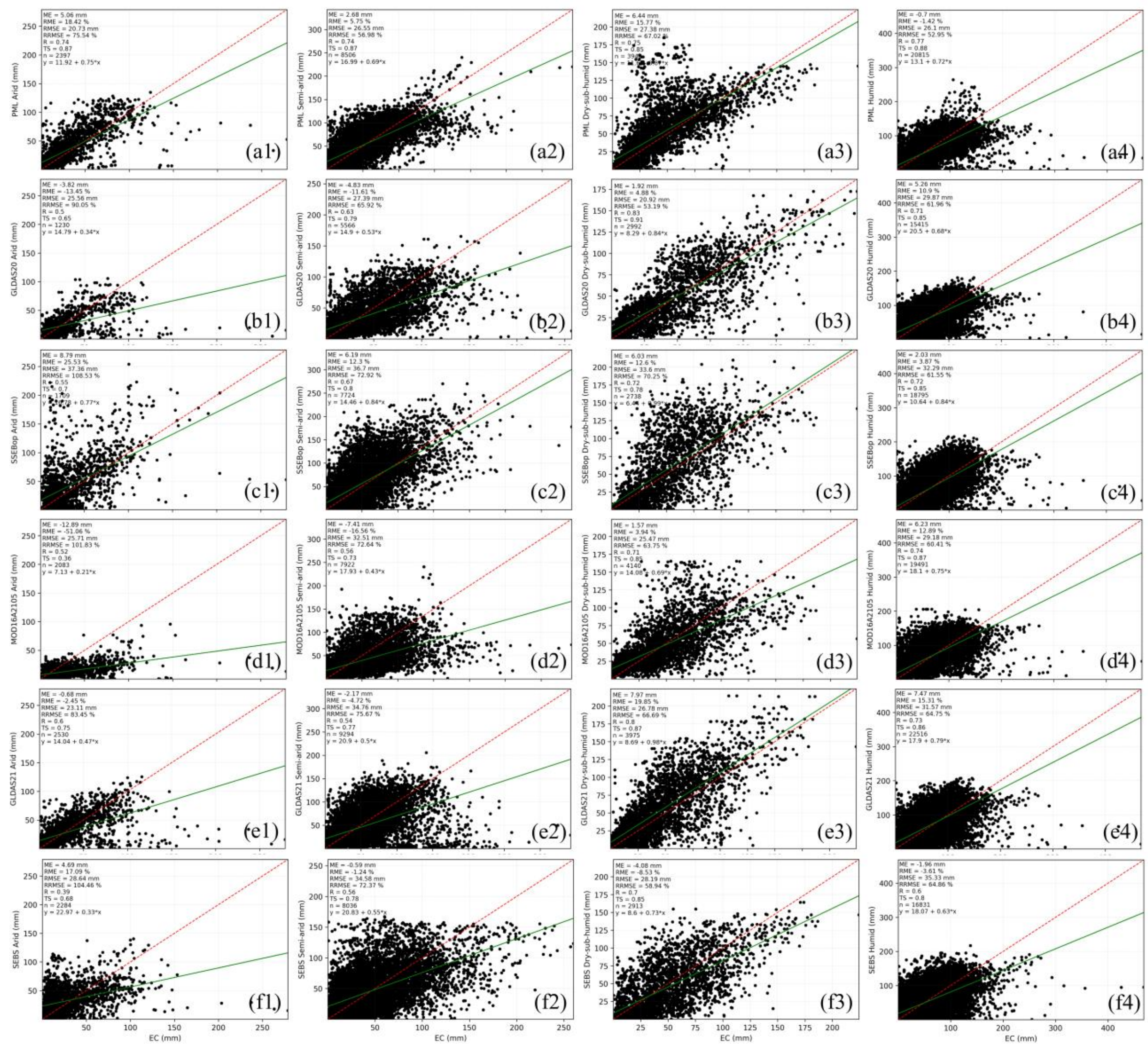

Figure 89. Monthly ET products (PML: (**a**); GLDAS20: (**b**); SSEBop: (**c**); MOD16A2105: (**d**); GLDAS21: (**e**); SEBS: (**f**)) against flux EC ET aggregated for all sites for each climate class (arid: (**1**); semiarid: (**2**); dry-sub-humid: (**3**); humid: (**4**)).

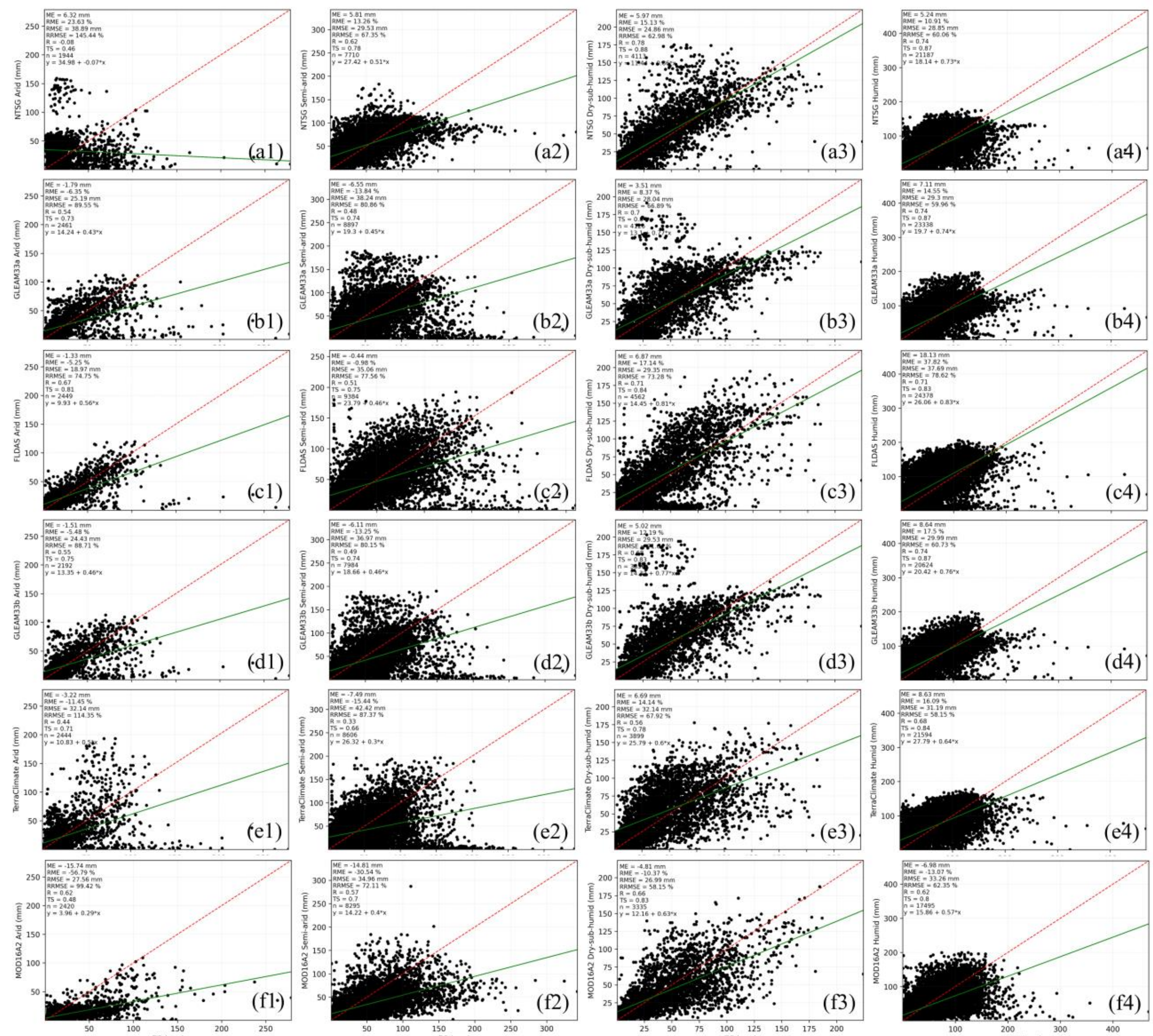

Figure ~~9~~10. Monthly ET products (NTSG: (**a**); GLEAM33a: (**b**); FLDAS: (**c**); GLEAM33b: (**d**); TerraClimate: (**e**); MOD16A2: (**f**)) against flux EC ET aggregated for all sites for each climate class (arid: (**1**); semiarid: (**2**); dry-sub-humid: (**3**); humid: (**4**)).

4.1.6. Validation by elevation levels

Figures ~~10~~11 and ~~11~~12 show that MOD16A2 and SEBS over elevation levels <500 and MOD16A2 and MOD16A2105 over elevation levels from 500 m to 1500 underestimate the values, while the other ET products overestimate the values; additionally, at elevations >1500, only SSEBop and NTSG overestimate the values. The ET product agreed best with the desired RMSE (RRMSE) in the PML product. Moreover, it yielded the best ME (RME) at elevations <500 m. The preferred ME (RME) over elevations 500 m to 1500 m and elevations > 500 m was obtained using SEBS and FLADS, respectively, see Table 5 (level one: 24–26).

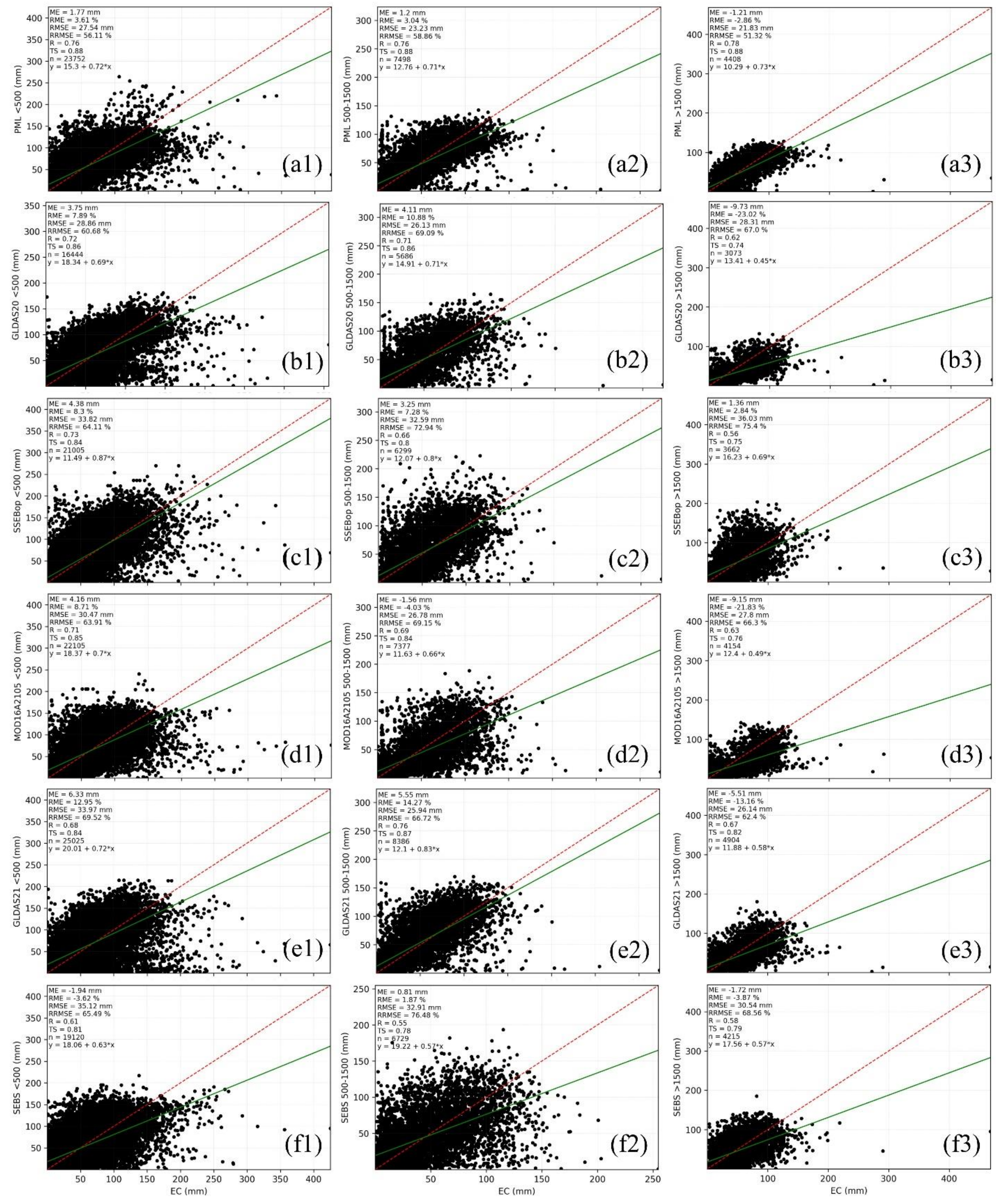

Figure ~~10~~11. Monthly ET products (PML: (**a**); GLDAS20: (**b**); SSEBop: (**c**); MOD16A2105: (**d**); GLDAS21: (**e**); SEBS: (**f**)) against flux EC ET aggregated for all sites for each elevation level (<500 m: (**1**); 500 m–1500 m: (**2**); >1500 m: (**3**)).

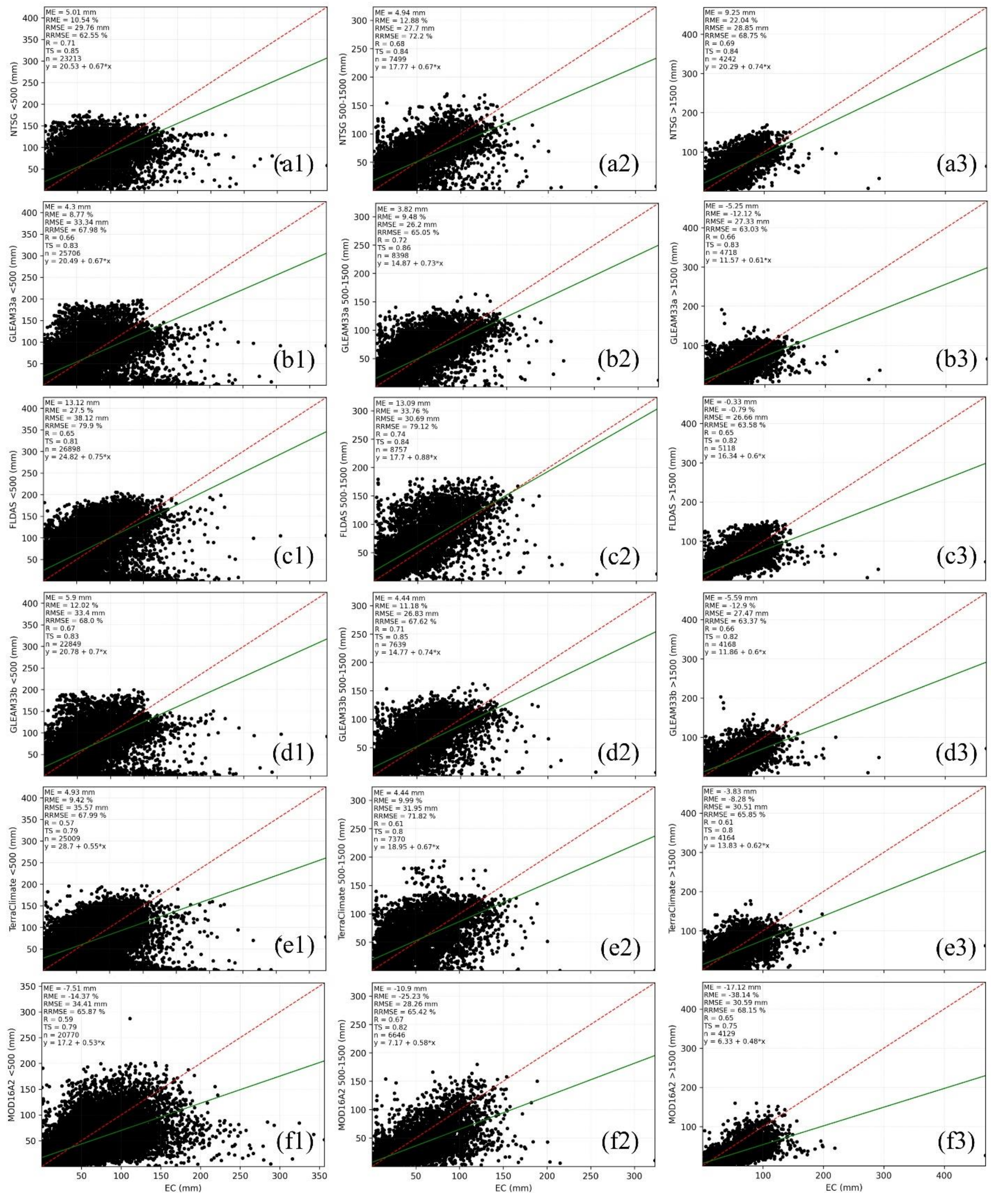

Figure ~~11~~12. Monthly ET products (NTSG: (**a**); GLEAM33a: (**b**); FLDAS: (**c**); GLEAM33b: (**d**); TerraClimate: (**e**); MOD16A2: (**f**)) against flux EC ET aggregated for all sites for each elevation level (<500 m: (**1**); 500 m–1500 m: (**2**); >1500 m: (**3**)).

4.2. Ensemble ET product

4.2.1. Ensemble steps

Table 5 provides levels one and two validation metrics of all ET products for monthly (01), annual (02), interannual (January-December: 03–14), land cover types (croplands, grasslands, forests, water bodies, others: 15–19), climatic classes (arid, semiarid, dry sub-humid, humid: 20–23), and elevation levels (<500 m, 500 m-1500 m, >1500 m: 24–26). Each cell represents one of the validation levels (01–26) and the best-performing ET product based on the selected validation metric, see Sect. 3.

Table 6 shows that, according to ~~first~~ the occurrence of ET products in level ~~accuracies~~one, PML, GLDAS20, and SEBS represent the first three best-performing ET products, while according to the ~~second~~occurrence of ET products in level two GLDAS20, PML, and MOD16A2105, and according to the total ~~of the first and second~~occurrence in levels one and two, PML, GLDAS20, and SSEBop are the best, respectively. For example, PML yielded the best validation ~~indices~~metrics (the lowest ME, RME, RMSE, and RRMSE as well as the highest R and TS) over 83 (53%) and 24 (15%) cells in levels one and two, respectively; thus, the total count was 107 (34%) cells. Accordingly, the three best-performing ET products over most of the all conditions are MPL followed by GLDAS20 (~~first~~ level one: 10 (6%); ~~second~~ level two: 37 (24%); total: 37 (15%)) and SSEBop (~~first~~ level one: 12 (8%); ~~second~~ level two: 15 (10%); total: 27 (9%)).

Since the three best-performing ET products differ in their spatial resolution and algorithms, we introduced an ensemble mean product at a 1000 m × 1000 m spatial resolution that spans from 2003 to 2017 (15 years) and relies on remotely sensed models (PML and SSEBop). It should be noted that although SEBS has one point more than SSEBop ~~in the first~~on level one, it has 7 fewer points than SSEBop in ~~the second~~ level two (5%). In addition, SSEBop has a higher spatial resolution than that of SEBS. In the same manner, SSEBop and MOD16A2105 have the same performance in terms of total count (27 (9%)), but SSEBop is higher by 5 points in ~~the first~~ level one.

Obviously, from Table 7, the ensemble ET products cannot perform highly across all regions, and it had a total count of 50%, followed by PML (44%). Looking to the ensemble mean from Table 7 compared to PML from Table 6, the total count increased from 34% to 50% (+16%), indicating that the ensemble mean, which created from PML and SSEBop, enhanced PML performance across all conditions by 16% and PML itself still has the best performance by 44%.

To introduce an ensemble product before 2003, firstly, PML and SSEBop were ignored, and the same steps were repeated. Table 8 shows that the best-performing products are GLDAS20, MOD16A2105, and NTSG in terms of the total count. Since the last two products are based on remote sensing, they were selected to create the ensemble product before 2003 at a 1000 m × 1000 m spatial resolution. Although GLDAS20 agreed well over 42% and had the lowest maximum ME among all datasets (9.73 mm), NTSG was selected to provide the ET estimates before 2000 because it had a higher spatial resolution, so it could capture more spatial details than GLDAS20.

Table 9 shows that the ensemble ET for 2001 and 2002 performed better than the original ET products, with values of 62%, 38%, and 50% for level one, level two and the total, respectively. For the periods before 2001, NTSG can be used from 1982 to 2001 or GLDAS20 can be used instead. Hence, remotely sensed-based long-term ensemble

ET can be synthesized from PML and SSEBop between 2003 and 2017, MOD16A2105 and NTSG between 2001 and 2002. SSEBop can be used after 2018, while before 2000, NTSG can be used.

Table 5. Levels one and two validation metrics of ~~all~~the 12 ET products for monthly (**01**), annually (**02)** interannual (January-December: **03-14**), land cover types (croplands, grasslands, forests, water bodies, others: **15-19**), climatic classes (arid, semiarid, dry sub-humid, humid: **20-23**), and elevation levels (<500 m, 500 m-1500 m, >1500 m: **24-26**), cells ~~colour~~color as Table **6**.

Level	~~Indicator~~Metrics	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
~~Level one~~	ME																										
	RME																										
	RMSE																										
One	RRMSE																										
	R																										
	TS																										
~~Level two~~	ME																										
	RME																										
	RMSE																										
Two	RRMSE																										
	R																										
	TS																										

Table 6. The ~~count, percent and~~ occurrence of the ~~total count and percent of levels one and two of all~~12 ET products ~~performance~~based on Table 5.

ET products	Occurrence in level 1		Occurrence in level 2		Total	
	~~Level 1~~ count	~~Level 1 count (%)~~%	~~Level 2~~ count	~~Level 2 count (%)~~%	~~Total~~ count	%~~Total count (%)~~
PML	83	53	24	15	107	34
GLDAS20	10	6	37	24	47	15
SSEBop	12	8	15	10	27	9
MOD16A2105	7	4	20	13	27	9
GLDAS21	14	9	11	7	25	8
SEBS	13	8	8	5	21	7
NTSG	4	3	16	10	20	6
GLEAM33a	5	3	6	4	11	4
FLDAS	6	4	4	3	10	3
GLEAM33b	1	1	6	4	7	2
TerraClimate	1	1	6	4	7	2
MOD16A2	0	0	3	2	3	1

Table 7. The ~~count, percent and the total count and percent of levels one and two~~occurrence of PML and SSEBop products and their ensemble mean ~~for the period~~during 2003~~-~~ and 2017.

ET products	Occurrence in level 1		Occurrence in level 2		Total	
	~~Level 1~~ count	~~Level 1 count (%)~~%	~~Level 2~~ count	~~Level 2 count (%)~~%	~~Total~~ count	%~~Total count (%)~~
Mean	43	28	113	72	156	50
PML	103	66	33	21	136	44
SSEBop	10	6	10	6	20	6

Table 8. The ~~count, percent and the total count and percent of levels one and two~~occurrence of all ET products ~~performance~~ except PML and SSEBop products.

ET products	Occurrence in level 1		Occurrence in level 2		Total	
	~~Level 1~~ count	~~Level 1 count (%)~~%	~~Level 2~~ count	~~Level 2 count (%)~~%	~~Total~~ count	~~Total count (%)~~%
GLDAS20	42	27	27	17	69	22
MOD16A2105	28	18	28	18	56	18
NTSG	14	9	35	22	49	16
GLDAS21	23	15	14	9	37	12
SEBS	21	13	7	4	28	9
GLEAM33a	8	5	16	10	24	8
GLEAM33b	6	4	15	10	21	7
FLDAS	9	6	5	3	14	4
TerraClimate	3	2	5	3	8	3
MOD16A2	2	1	4	3	6	2

Table 9. The ~~count, percent and the total count and percent of levels one and two~~occurrence of NTSG and MOD16A2105 products and their ensemble mean ~~for~~during 2001 and 2002.

ET products	Occurrence in level 1		Occurrence in level 2		Total	
	~~Level 1~~ count	~~Level 1 count (%)~~%	~~Level 2~~ count	~~Level 2 count (%)~~%	~~Total~~ count	~~Total count (%)~~%
Mean	96	62	59	38	155	50
NTSG	19	12	68	44	87	28
MOD16A2105	41	26	29	19	70	22

4.2.2 Contribution of ET datasets to the synthesized ET

The synthesized ET dataset was created at a 1000 m × 1000 m spatial resolution from 1982 to 2019 based on remotely sensed ET products. PML, SSEBop, MOD16A2105, and NTSG were augmented together to create the new dataset. Since SSEBop and MOD16A2105 have a 1000 m × 1000 m spatial resolution, PML was upscaled and NTSG was downscaled by pixel average and nearest neighbor resampling techniques in GEE, respectively. The synthesized ET was fully contributed by SSEBop for the years 2018 and 2019 and by NTSG from 1982 to 2000, while for the years 2001 and 2002, it was contributed by the simple mean of MOD16A2105 and NTSG. Finally, between 2003 and 2017, the value represents the simple mean of PML and SSEBop.

Since the synthesized ET performance was governed by each ET product(s) for the corresponding year from 1994 to 2019 (25 years), where the ET EC fluxes were available, most of the performance comes from PML and SSEBop for the 15 years from 2003 to 2017 (60%), from MOD16A2105 and NTSG for 2 years (2001 and 2002; 8%), from SSEBop for individual values in years 2018 and 2019 (8%), and from NTSG for 7 years (24%) from 1994 to 2000.

4.2.3. Synthesized global ET product

Figure ~~12~~13 shows, looking to July, except over barren land, permanent snow and ice, and arid areas (not shown), the maximum value of the synthesized ET lies between SSEBop, which yields the largest ET during all months, and PML. Hence, the long-term monthly synthesized ET performance is affected by PML and SSEBop more than by NTSG and MOD16A2105, as mentioned in Sect. 4.2.2.

Table 10 provides the average monthly and annual synthesized ET (mm month^{-1}), land cover types, aridity index classes, and elevation levels (mm year^{-1}). The average annual ET from 1982–2019 is 567 mm year^{-1}. July represents the maximum synthesized ET (Fig. ~~12~~13). Table 10 also provides average annual ET for land cover types calculated from flux sites. Across land cover types, croplands are higher than forests, followed by grassland, where the average synthesized ET was 597, 548, and 542 for croplands, forests, and grasslands, respectively. Low synthesized ET values across arid areas (average = 392 mm year^{-1}) can be attributed to low vegetation cover. It should be noted that Table 10 does not represent the perfect calculation of ET over each Land cover class because the total number of fluxes for each class was not distributed well; for instance, in the arid areas, there were 35 (5%) fluxes, while in the humid area, there were 361 (56%) fluxes.

Figure ~~13~~ 14 shows the decadal (1982–1989, 1990–1999, 2000–2009, and 2010–2019) and long-term (1982–2019) average synthesized ET maps worldwide, except for Antarctica. Regarding the spatial distribution, the higher ET is shown in Malaysia, Singapore, and Indonesia and the northern part of South America. During the first and second decades, the synthesized ET is based on the NTSG product; thus, the same spatial distribution was observed. Although PML and SSEBop mainly contribute the synthesized ET between 2003 and 2017, there is little difference in their spatial distributions, where higher ET can be observed during 2010–2019 over the northern parts of South America.

Table 11 shows statistics of the maps provided in Fig. ~~13~~ 14 for all continents except Antarctica. The standard deviation is higher over Africa followed by Oceania and Asia. The mean values of the synthesized ET is sequenced from South America followed by Oceania and Africa. The maximum value of the synthesized ET is recorded over Asia followed Africa and Australia. The total ETs are 29.1%, 21.7%, 19.9%, 16.7%, 7.9%, 4.2%, and 0.5% for Asia, South America, Africa, North America, Europe, Australia, and Oceania, respectively.

4.2.4 Validation of the synthesized ET

Figures ~~14~~15–~~17~~ 18 show that the synthesized ET agreed well with the observed data, where the R (TS) ranged between 0.70 (0.85) and 0.78 (0.89), except at the annual time step (Fig. ~~14b~~15b) and over barren land and permanent snow and ice (not shown), where R (TS) was 0.65 (0.81) and 0.68 (0.80), respectively. Based on the ME sign, the value was underestimated only over water bodies. The magnitude of ME (RME) ranged between 0.54 mm (1.05%) and 6.76 mm (16.62%), while the RMSE (RRMSE) ranged from 20.95 mm (45.22%) to 30.12 mm (59.61%). Looking at the regression line equation, with no exceptions, the synthesized ET overestimated the flux EC ET at lower ET values and underestimated the flux EC ET at higher ET values. As mentioned above, even the long-term synthesized ET cannot perform best across all comparison levels (Tables 12 and 13).

During the periods 2018–2019 and before 2001, the synthesized ET performance came from the original datasets of SSEBop and NTSG, respectively. The ensemble mean has a total count of 50% over the periods 2003–2017 and 2001–2002 compared to the original datasets, indicating that it can perform better than other ET products over half of all comparison levels, see Tables 7 and 9.

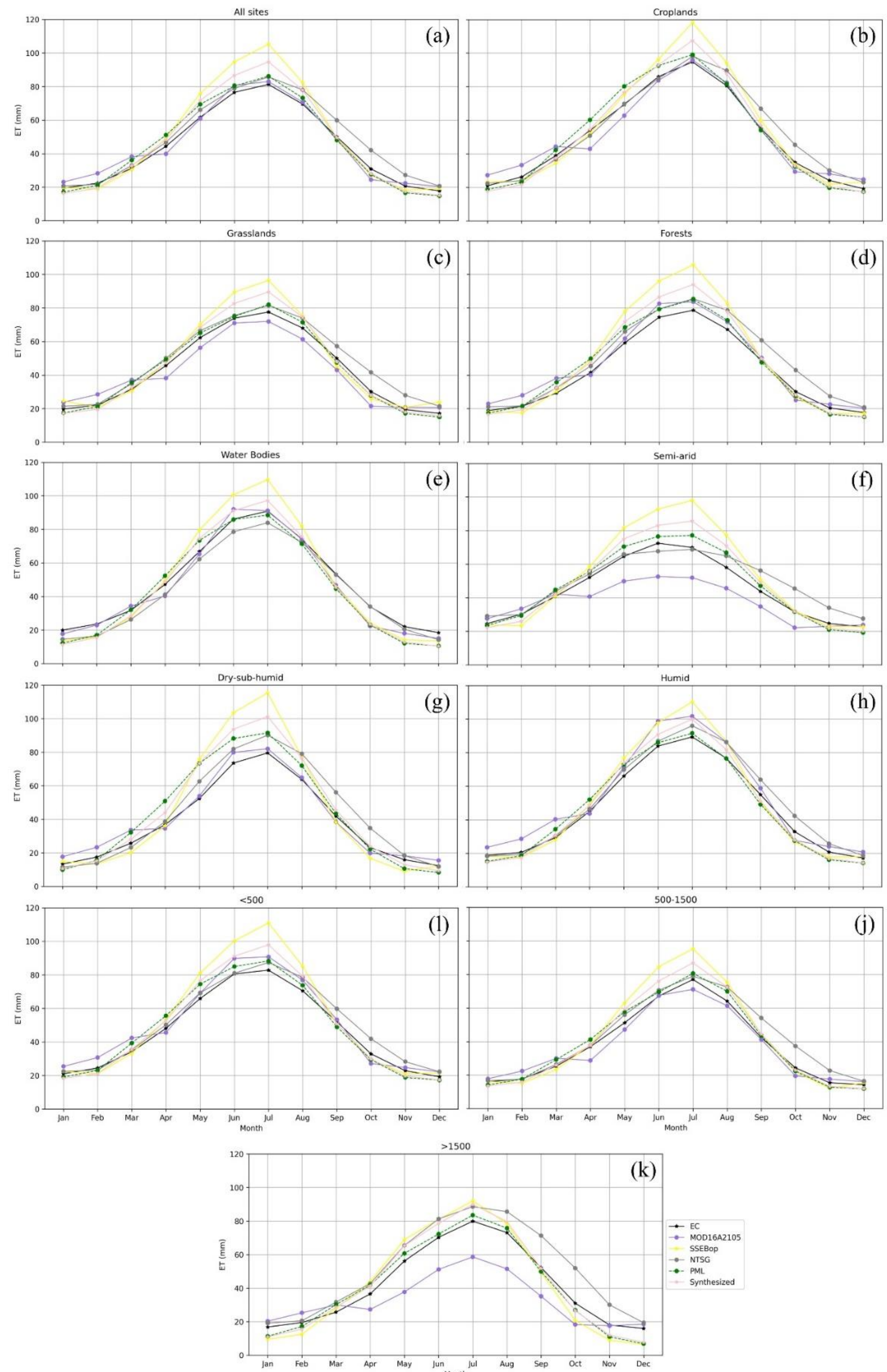

Figure 1~~2~~3. ~~Monthly average synthesized ET and the original products over all flux sites (**a**), land cover types (croplands: (**b**); grasslands: (**c**); forests: (**d**); water bodies: (**e**)), climate classes (semiarid: (**f**); dry sub-humid: (**g**); humid: (**h**)), and elevation levels (<500 m: (**l**), 500 m-1500 m: (**j**), and >1500m: (**k**))~~ Monthly average flux EC ET, MOD16A2105, SSEBop, NTSG, PML and the synthesized ET (subplot label as in Figure 3)~~.~~

Table 10. The average decadal synthesized ET of monthly (mm month^{-1}) and land cover types, aridity index classes and elevation levels (mm year^{-1}).

Level	1982–1989	1990–1999	2000–2009	2010–2019	1982–2019
January	43.22	44.10	44.94	45.99	44.56
February	39.73	41.14	42.83	42.09	41.45
March	44.83	45.09	43.73	42.93	44.15
April	45.84	46.04	39.32	38.57	42.44
May	52.86	53.36	47.13	46.61	49.99
June	56.15	57.31	53.98	54.00	55.36
July	60.83	61.80	57.06	56.99	59.17
August	58.02	58.77	51.25	50.25	54.57
September	49.99	50.15	44.10	42.79	46.76
October	46.76	46.91	38.53	38.77	42.74
November	42.55	42.45	41.52	42.29	42.20
December	42.66	43.58	42.92	44.43	43.40
Annual	583	591	547	546	567
Croplands	597	619	595	577	597
Grasslands	526	546	539	557	542
Forests	541	561	544	546	548
Water bodies	499	517	519	534	517
Others	280	288	230	195	248
Arid	400	405	366	398	392
Semiarid	519	538	528	541	532
Dry sub-humid	479	498	498	511	497
Humid	577	600	582	577	583
Elevation <500m	551	570	570	579	568
Elevation 500 m – 1500 m	498	519	484	484	496
Elevation >1500 m	557	583	506	471	529

Note: Monthly and annual estimates have based on synthesized ET raster layers averaged over a decade. Land cover types, aridity index classes and elevation levels estimates have based on annual synthesized ET values extracted over all flux sites.

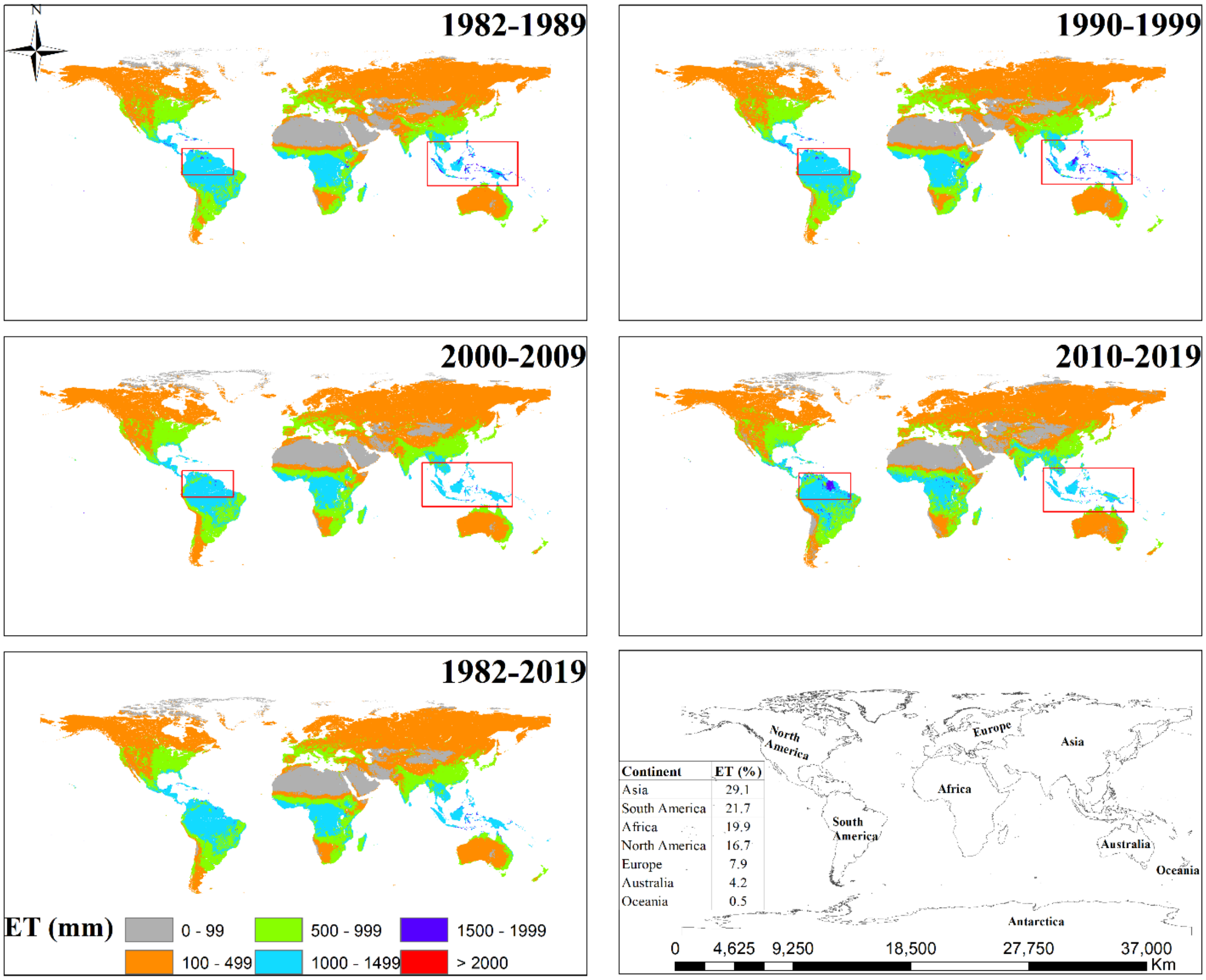

Figure ~~13~~14. Decadal and long-term synthesized ET, the last plot shows continental-scale used to create Table ~~13~~11 accompanied by the percent of ET over each continent for the periods 1982–2019 except Antarctica. Use the following link of the GEE application to preview these maps: https://elnashar.users.earthengine.app/view/synthesizedet/

Table 11. Statistics of the decadal and long-term synthesized ET (mm).

Period	Continent	Minimum	Maximum	Mean	Standard Deviation	Sum
	Africa	0	3588	541	540	17091316777
	Asia	0	3979	377	392	25075224084
	Australia	0	4076	445	275	3812181627
1982-1989	Europe	0	2934	403	189	6902627799
	North America	0	3818	413	331	14682344407
	Oceania	111	2155	903	392	431987028
	South America	4	3585	1002	364	18968179507
	Global	0	4076	583	355	86963861230
	Africa	0	3673	555	545	17552175432
	Asia	0	4054	387	398	25755440497
	Australia	0	4240	438	281	3748291789
1990-1999	Europe	0	2825	424	203	7260038441
	North America	0	3742	423	338	15051753185
	Oceania	111	2176	892	394	426754913
	South America	8	3409	1015	363	19218216796
	Global	0	4240	591	360	89012671053
	Africa	0	4326	538	504	17073575117
	Asia	0	4794	393	377	26457856410
	Australia	0	4804	397	260	3417383567
2000-2009	Europe	0	4108	399	165	7119724411
	North America	0	3915	333	310	15229417841
	Oceania	0	3349	811	398	425095485
	South America	0	3975	960	411	18312021115
	Global	0	4804	547	346	88035073946
	Africa	0	4892	556	530	17631809454
	Asia	0	6167	398	401	26760551956
	Australia	0	4692	425	271	3658944492
2010-2019	Europe	0	3866	384	165	6834742252
	North America	0	4366	338	320	15454707917
	Oceania	0	3387	766	417	391231772
	South America	0	4452	953	453	18166326886
	Global	0	6167	546	365	88898314729
	Africa	0	4892	548	530	17337219195
	Asia	0	6167	389	392	26012268237
	Australia	0	4804	426	272	3659200369
1982-2019	Europe	0	4108	402	180	7029283226
	North America	0	4366	377	325	15104555837
	Oceania	0	3387	843	400	418767300
	South America	0	4452	983	398	18666186076
	Global	0	6167	567	357	88227480239

Table 12. Same as Table 5 but~~Levels one and two validation metrics of all ET products (except~~ MOD16A2~~) and~~ replaced by the synthesized ET ~~for monthly (**01**), annually (**02**) interannual (January-December: **03-14**), land cover types (croplands, grasslands, forests, water Bodies, others: **15-19**), climatic classes (arid, semiarid, dry sub-humid, humid: **20-23**), and elevation levels (<500 m, 500 m–1500 m, >1500 m: **24-26**),~~ (cells colour as Table 13~~.~~).

Levels ~~Level~~	~~Indicator~~Metrics	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
~~Level one~~	ME																										
	RME																										
	RMSE																										
One	RRMSE																										
	R																										
	TS																										
~~Level two~~	ME																										
	RME																										
	RMSE																										
Two	RRMSE																										
	R																										
	TS																										

Note: MOD16A2 ignored according to Sec. 4.1.

Table 13. ~~The count, percent and the total count and percent of levels one and two of all ET products (except MOD16A2) and~~ Same as Table 6 but MOD16A2 replaced by the synthesized ET ~~performance~~and based on Table 12.

ET products	Occurrence in level 1		Occurrence in level 2		Total	
	~~Level 1~~ count	~~Level 1 count (%)~~%	~~Level 2~~ count	~~Level 2 count (%)~~%	~~Total~~ count	~~Total count (%)~~%
PML	66	42	33	21	99	32
Synthesized	26	17	57	37	83	27
GLDAS20	12	8	12	8	24	8
GLDAS21	12	8	7	4	19	6
SEBS	12	8	7	4	19	6
MOD16A2105	6	4	12	8	18	6
SSEBop	8	5	8	5	16	5
NTSG	2	1	8	5	10	3
FLDAS	6	4	2	1	8	3
GLEAM33a	5	3	3	2	8	3
TerraClimate	1	1	4	3	5	2
GLEAM33b	0	0	3	2	3	1

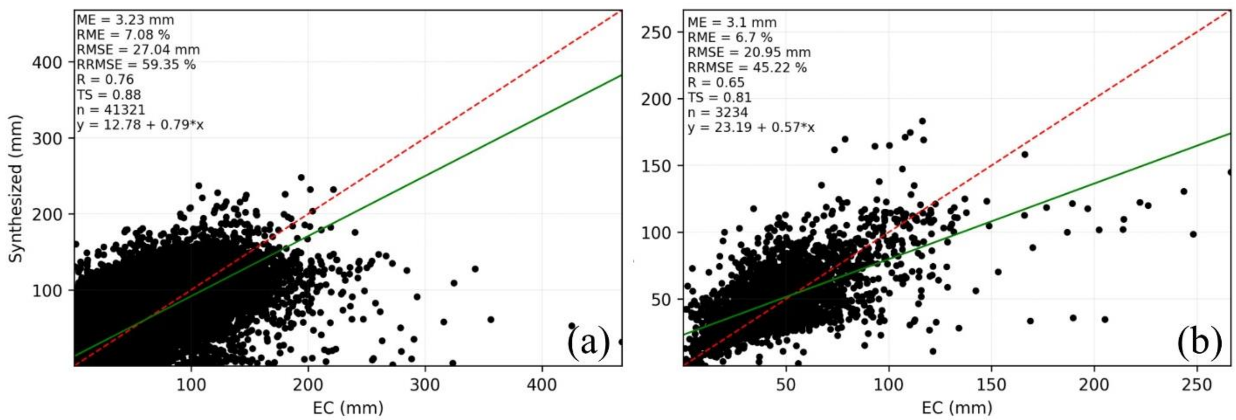

Figure ~~14~~15. Monthly (**a**) and annually (**b**) synthesized ET against flux EC ET aggregated for all sites.

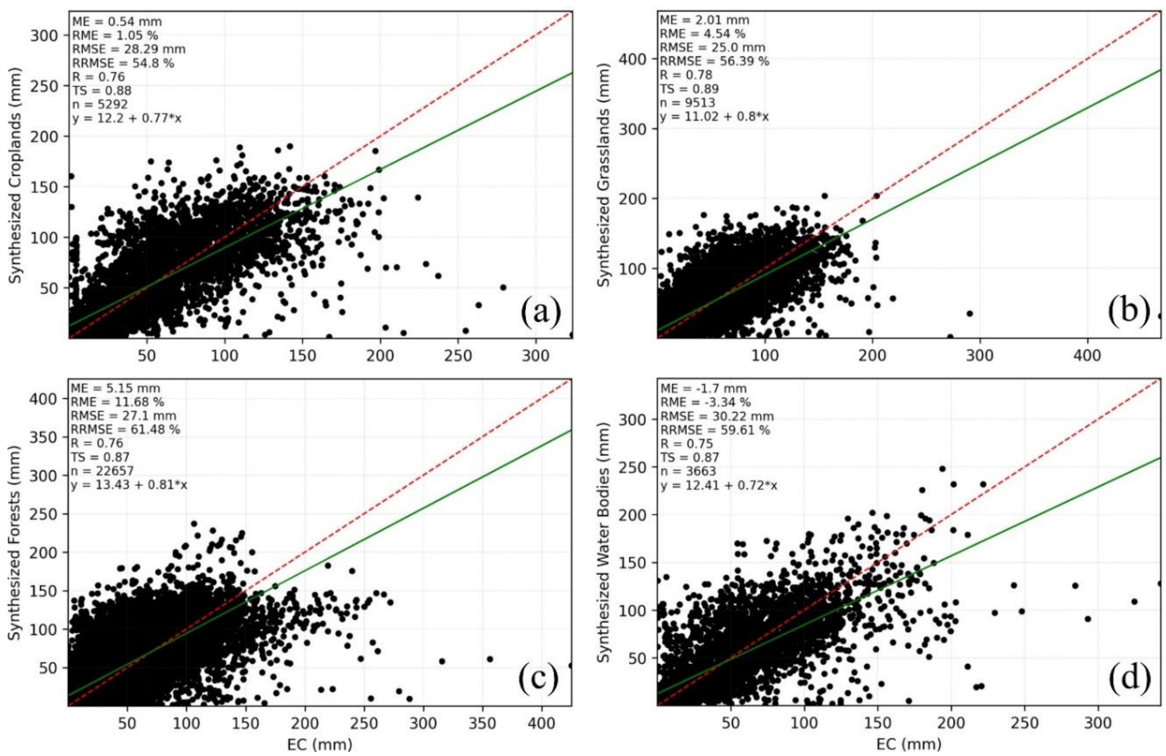

Figure ~~15~~16. Monthly synthesized ET against flux EC ET aggregated for all sites for each land cover type (croplands: (**a**); grasslands: (**b**); forest: (**c**); water bodies: (**d**)).

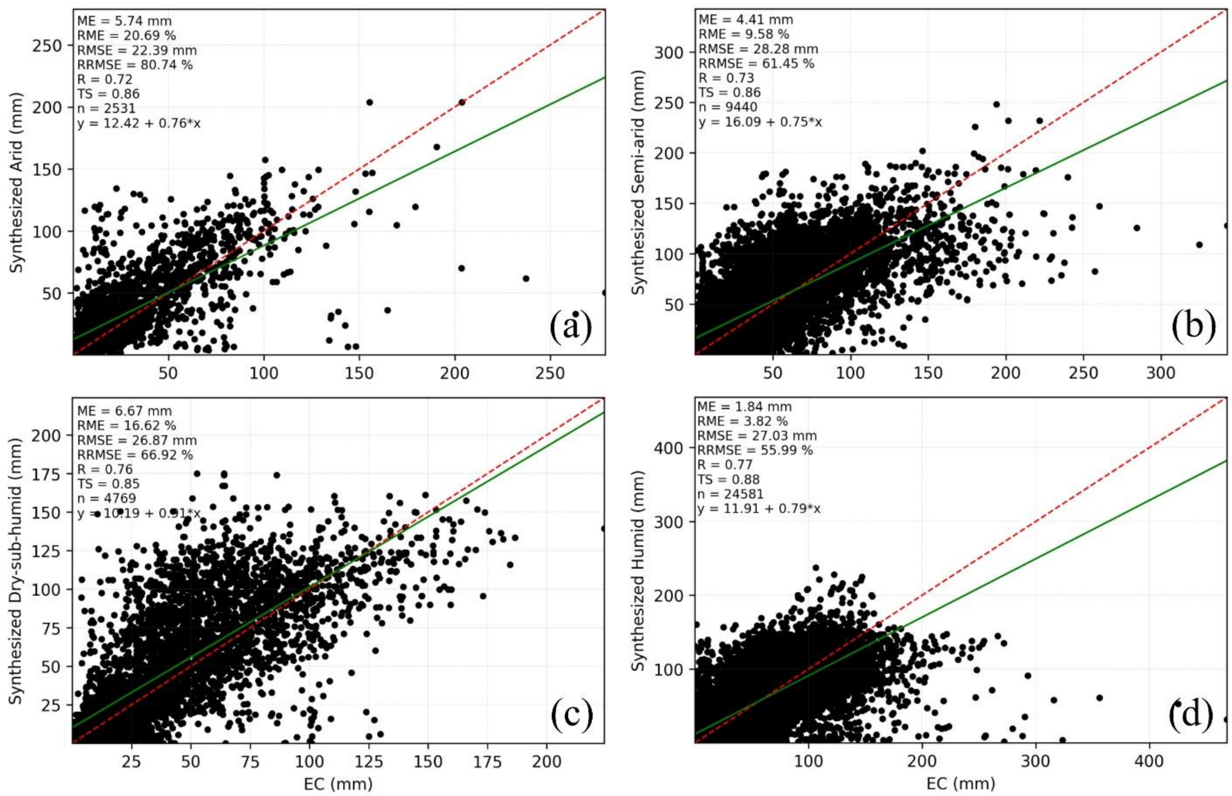

Figure ~~16~~17. Monthly synthesized ET against flux EC ET aggregated for all sites for each climate class (arid: (**a**); semiarid: (**b**); dry-sub-humid: (**c**); humid: (**d**)).

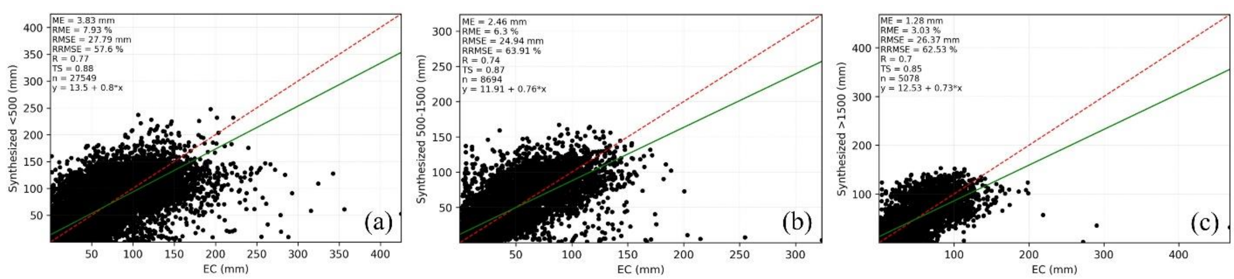

Figure ~~17~~18. Monthly synthesized ET against flux EC ET aggregated for all sites for each elevation level (<500 m: (**a**); 500 m–1500 m: (**b**); >1500 m: (**d**)).

Figure ~~18~~19 presents a monthly comparison between the synthesized ET with the country-based ET products over China and the United States as well as over the African continent. In general, the synthesized ET returned higher agreement (R and TS) and accuracy (RMSE) with the flux EC ET than did the other ET products (CR, SSEBop, and FAO WaPOR). Moreover, it has lower biases over the United States and the African continent.

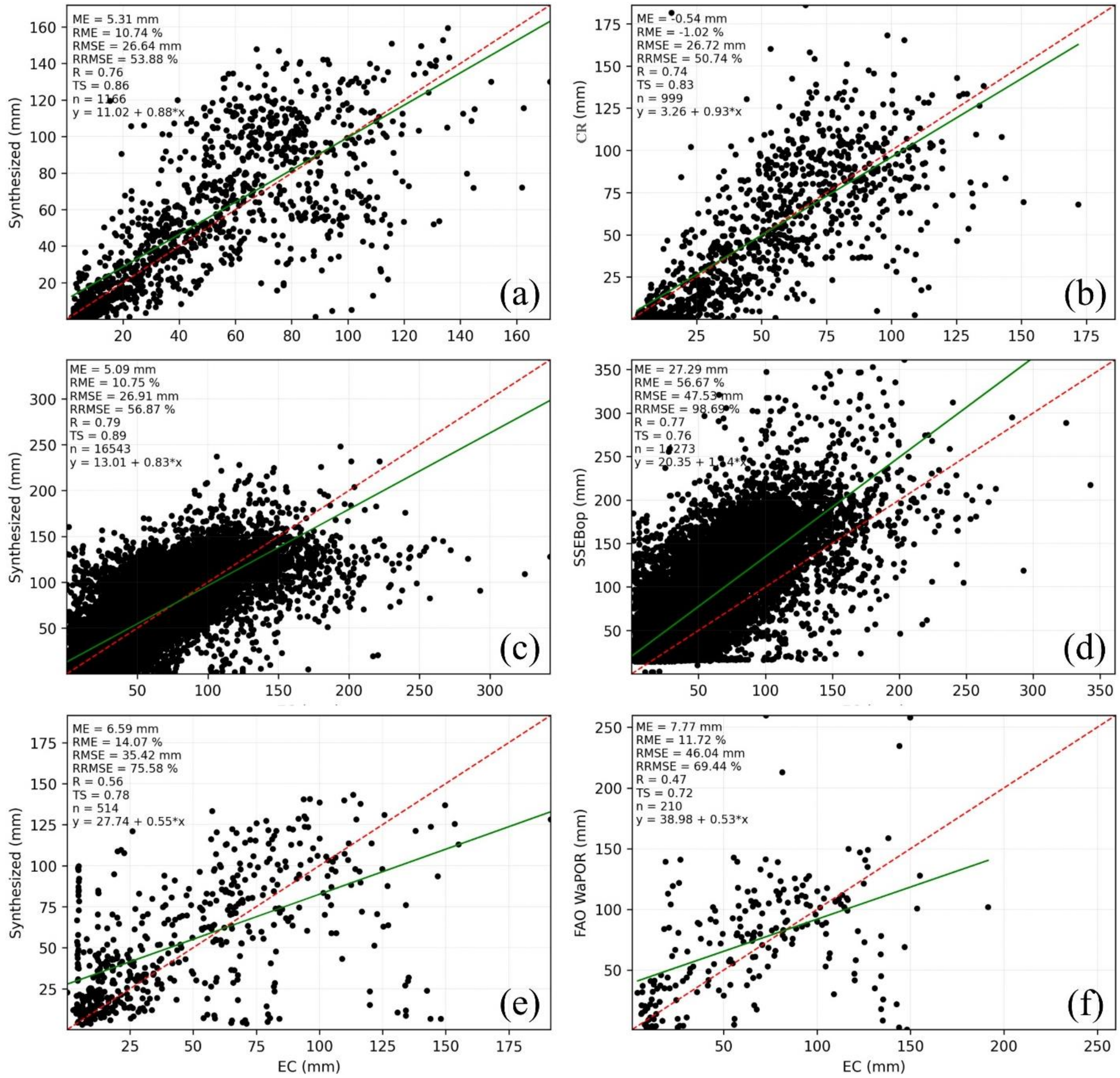

Figure ~~18~~19. Monthly comparison between the synthesized ET (**a**, **c** and **e**) and CR (**b**), SSEBop (**d**), and FAO WaPOR (**f**) ET products against flux EC ET aggregated for all sites over China (**a** and **b**), the USA (**c** and **d**) and the African continent (**e** and **f**).

5. Discussion

Since global land ET plays a paramount role in the hydrological cycle, its accurate estimation is essential for further studies. Although there are many global ET products that have been derived from remote sensing models, land surface models, and hydrological models, they differ in their algorithms, parameterization, and temporal span, and none of these products can be used for a long time with a reasonable spatial resolution and lower uncertainty. In this study, we ensemble the best-performing, currently available global ET products at a reasonable spatial resolution (kilometer) as one consistent global ET dataset covering a long temporal period. Users can use this dataset assuredly without looking at other datasets and performing additional assessments.

We used a high-quality dataset of global flux towers as a site-pixel-level validation for certain global ET products (Leuning et al., 2008;Zhang et al., 2010;Ershadi et al., 2014;Michel et al., 2016) to assess them and select the best products to create a synthesized ET covering a long temporal period. For that, a matrix of 6 validation criteria and 26 comparison levels was created, and then levels one and two of the validation metrics were used to select the best-performing products. Finally, by the simple mean of the products that performed best over the different periods, the synthesized ET was created.

Among all global ET products investigated in this study, the products that performed best are PML, GLDAS20, SSEBop, MOD16A2105, GLDAS21, SEBS, and NTSG (Table 6). From the perspective of all comparison levels, the performance of these products varied, and no single product performed well across all land surface types and conditions (Vinukollu et al., 2011a;Li et al., 2018). The PML represents the ET product with the highest agreement, with lower ME (RME) and RMSE (RRMSE) values, followed by the synthesized ET (Tables 12 and 13); however, it should be noted that PML estimates span a 15-yr period, while the synthesized ET presents longer estimates from 1982 to 2019 (38 years).

The main advantage of the new dataset is that, for the first time, a synthesized remotely sensed ET product with a reasonable spatial resolution and lower long-term uncertainties has been provided, where the maximum absolute ME (RME) and RMSE (RRMSE) values are 13.94 mm (17.13%) and 38.61 mm (47.45%), respectively. Furthermore, it agreed well (R > 0.70) in 62% of all comparison levels (Table 14). This dataset can provide ensemble ET estimates for all land cover types, where MOD16A2105 does not provide ET estimates over water bodies and desert areas other products are. Moreover, a comparison among the synthesized ET against CR, SSEBop, and FAO WaPOR ET products over China, the United States, and the African continent proved that the synthesized ET outperformed these products in terms of a higher agreement, higher accuracies and lower biases. Hence, the synthesized ET can play an essential role, especially for regional and global scale studies, over a long time (1892–2019).

Table 14. Percentage of R more than 0.70 and the maximum absolute value of ME (mm), RME (%) RMSE (mm), and RRMSE (%) across all comparisons levels (01–26) of the highly preformed ET products and the synthesized ET.

ET products	R>0.7 (%)	ME	RME	RMSE	RRMSE
PML	65	7.64	12.22	36.28	44.30
Synthesized	62	13.94	17.13	38.61	47.45
GLDAS20	42	9.73	23.02	39.53	49.32
SSEBop	42	21.82	26.07	48.14	57.50
MOD16A2105	42	12.89	51.06	42.78	53.27
GLDAS21	35	13.69	22.07	47.84	58.32
NTSG	23	14.46	86.35	40.50	50.26

The synthesized ET used SSEBop ET for the years 2018 and 2019 and NTSG from 1982 to 2000 because NTSG is the only remotely sensed global ET product available and has a good spatial resolution compared to GLDAS20. It is the simple mean of MOD16A2105 and NTSG for the years 2001 and 2002 and the simple mean of PML and SSEBop between 2003 and 2017 (see Tables 7 and 9).

Because the ET was synthesized during the first and second decades as well as the year 2000 based on resampled NTSG to a 1 km spatial resolution to be comparable with other products, future improvements may be focused on statistical downscaling of NTSG during this period. Moreover, since different datasets were selected due to ~~enhance~~data availability, also future improvements may be focused on the ~~product proposed in this paper~~adjustment of the ensemble means particularly for long-term pixel-based studies.

6. Data availability

All data used in this study are freely available; see Sect. 2 and Appendix A. The synthesized ET is available in https://doi.org/10.7910/DVN/ZGOUED (Elnashar et al., 2020) and as GEE application from the following link: https://elnashar.users.earthengine.app/view/synthesizedet. In addition, it can be accessed ~~-~~in the GEE JavaScript editor (the updated link embedded in the GEE application interface). Through this application, the user can query and display as well as download the synthesized ET. It should be noted that SSEBop and NTSG datasets are not available in Earth Engine so they were uploaded as assets in GEE for this purpose.

7. Conclusion

In the current study, a site-pixel-level validation was conducted for certain global ET products across a variety of land surface types and conditions to select the best performing ET products and then produce a global long-term synthesized ET dataset. To apply a comprehensive evaluation from different perspectives, land cover types, climate and elevations were classified into five, four, and three classes, respectively. According to six comprehensive validation criteria, the evaluated ET products ranked based on the lowest error metrics and highest accuracy and consistency over different classification levels to choose the ensemble members over different times.

~~Concerning the study investigation, PML, GLDAS20, SSEBop, MOD16A2105, GLDAS21, SEBS, and NTSG were ET products that performed best.~~ The average annual ET from 1982–2019 is 567 mm $year^{-1}$. Although no product performed ~~best~~better in terms of all selected validation criteria in all classification levels, ~~the~~ PML, GLDAS20, SSEBop, MOD16A2105, GLDAS21, SEBS, and NTSG are the sequence of their performances. The synthesized ET ~~produced~~ from PML, SSEBop, MOD16A2105 and NTSG ~~had high~~agreed with the flux EC ET with R-values higher than 0.70, a maximum ME (RME) of 13.94 mm (17.13%) and a maximum RMSE (RRMSE) of 38.61 mm (47.45%) over 62% of all comparisons levels, as remotely sensed based ET product spanning from 1982 to 2019 with highest agreements~~ and~~, accuracies ~~with low~~and lower biases over most of the land surface types and conditions. ~~In addition, this study provides ET estimates from 1982 to 2019 and for all land cover types. Furthermore, it performed~~It performs well when compared with country-based and continental ET products over China, the United States and the African continent. However, the further synthesis of local ET products is encouraged if regional ET products are available.

The results from this study provide a better understanding of the high performing ET products in each land cover type, elevation level and climate region as well as a monthly, annual and interannual time steps. Hence, this study provides an ET product that can be used to improve the quality of ET at regional and global levels and, consequently, can be used to improve agriculture, water resource management, and climate change studies.

Author Contribution: Abdelrazek Elnashar was responsible for experimental designing, manuscript preparation, and data processing and presentation. Linjiang Wang, Dr. Weiwei Zhu, and Dr. Hongwei Zeng contributed to data processing. Prof. Dr. Bingfang Wu contributed to conceptual designing, reviewing of the manuscript, funding acquisition, and project administration.

Acknowledgments: This research was financially supported by the National Key Research and Development Program of China (Grant No. 2016YFA0600303), the National Natural Scientific Foundations of China (grant numbers: 41991232) and the Key Research Program of Frontier Sciences, CAS (grant numbers: QYZDY-SSW-DQC014).

Conflicts of Interest: The authors declare that they have no conflict of interest.

Appendix A

A summary of ET datasets used in this research is presented here. It should be noted that except for SSEBop, SEBS, NTSG ET, and GLEAM, which are downloaded from their providers, other datasets are available in Earth Engine Data Catalog through the following link https://developers.google.com/earth-engine/datasets/catalog/. Each dataset in GEE has Earth Engine Snippet as following:

MOD16A2 ET V6: ee.ImageCollection("MODIS/006/MOD16A2")
MOD16A2 ET V105: ee.ImageCollection("MODIS/NTSG/MOD16A2/105")
PML ET: ee.ImageCollection("CAS/IGSNRR/PML/V2")
GLDAS ET V20: ee.ImageCollection("NASA/GLDAS/V20/NOAH/G025/T3H")
GLDAS ET V021: ee.ImageCollection("NASA/GLDAS/V021/NOAH/G025/T3H")
FLADS ET: ee.ImageCollection("NASA/FLDAS/NOAH01/C/GL/M/V001")
TerraClimate ET: ee.ImageCollection("IDAHO_EPSCOR/TERRACLIMATE")

MOD16 ET

The Moderate Resolution Imaging Spectroradiometer (MODIS) Global Evapotranspiration Project (MOD16A2) estimates terrestrial ET as the sum of evaporation and plant transpiration. MOD16A2 ET uses the Penman-Monteith model, which includes MODIS remotely sensed data (e.g., vegetation, surface albedo, and land cover classification) and daily meteorological reanalysis. There are two products of MOD16A2 ET (V6 and V105) with an 8-day temporal resolution, but they differ in their spatial resolution and temporal coverage (Mu et al., 2011;Mu et al., 2014b). V6 spans from 2001 until now with a 500 m × 500 m spatial resolution and is provided by NASA LP DAAC at the USGS EROS Center; it can be downloaded from https://doi.org/10.5067/MODIS/MOD16A2.006/. V105 estimates span the period from 2001 to 2014 with a 1000 m × 1000 m spatial resolution and are provided by the Numerical Terradynamic Simulation Group (NTSG) at the University of Montana in conjunction with the NASA Earth Observing System (Mu et al., 2014a).

PML ET

The Penman-Monteith Leuning (PML) ET product partitions ET into three components: plant transpiration, soil evaporation, and intercepted rainfall by the canopy as well as water evaporation. PML data span from 2002 to 2017 with a 500 m × 500 m spatial resolution and an 8-day temporal resolution (Zhang et al., 2019).

SSEBop

The operational Simplified Surface Energy Balance (SSEBop) model is based on the Simplified Surface Energy Balance (SSEB) approach with a unique parameterization for operational applications. Using a thermal index approach, it combines ET fractions generated from remotely sensed MODIS land surface temperature, acquired every 10 days, with reference ET from global weather datasets. The SSEBop uses predefined, seasonally dynamic, boundary conditions that are unique to each pixel for the hot and cold reference points (Senay et al., 2007;Senay et al., 2011;Senay et al., 2013;Senay et al., 2020). SSEBop estimates are from 2003 with a 0.0096°×0.0096° (≈1 km) spatial resolution and a monthly temporal resolution. Data were provided by The Early Warning and Environmental Monitoring Program via the United States Geological Survey and can be downloaded from the following website https://earlywarning.usgs.gov/.

SEBS

The Surface Energy Balance System (SEBS) is an approach designed to estimate ET from the evaporative fraction using satellite remote sensing augmented with meteorological data at corresponding scales (Su, 2002). MODIS-LST and the Normalized Difference Vegetation Index (NDVI), GLASS-LAI, GLAS global forest height, GlobAlbedo, and ERA-Interim meteorological data have been used in these ET calculations with the revised SEBS algorithm (Chen et al., 2013;Chen et al., 2014a;Chen et al., 2019). SEBS is available during the period from 2000 to 2017 with a 5 km × 5 km spatial resolution and monthly temporal resolution. It is copyrighted by the Institute of Tibetan Plateau Research, Chinese Academy of Sciences and is available at http://en.tpedatabase.cn/.

NTSG ET

The Numerical Terradynamic Simulation Group (NTSG) ET data are based on an algorithm that estimates transpiration from the canopy and evaporation from soil using a modified Penman-Monteith model and evaporation from open water using a Priestley-Taylor model. These algorithms were applied globally using the Advanced Very High-Resolution Radiometer (AVHRR) Global Inventory Modeling and Mapping Studies (GIMMS) NDVI, NCEP/NCAR Reanalysis daily surface meteorology, and NASA/GEWEX Surface Radiation Budget Release-3.0 solar radiation inputs (Zhang et al., 2009;Zhang et al., 2010). NTSG estimates cover a period from 1982 to 2013 at a spatial resolution of 8 km × 8 km and a monthly temporal resolution. It is produced by NTSG at the University of Montana and can be retrieved from http://files.ntsg.umt.edu/.

GLEAM

The Global Land Evaporation Amsterdam Model (GLEAM) is physically based on an algorithm that estimate ET components separately (i.e., transpiration, interception loss, bare soil evaporation, snow sublimation, and open-water evaporation). The potential evaporation is estimated by the Priestley and Taylor equation based on observations of surface net radiation and near-surface air temperature and is then converted into actual evaporation based on the evaporative (soil) stress factor. The soil stress factor is based on microwave vegetation optical depth and simulated root-zone soil moisture calculated from a multilayer water balance model. Separately, interception loss is calculated based on vegetation and rainfall observations. There are two datasets available for GLEAM (i.e., v3.3a, and v3.3b) that differ only in their forcing and temporal coverage. v3.3a spans from 1980 to 2018 and relies on reanalysis radiation and air temperature, a combination of gauge-based, reanalysis and satellite-based precipitation, and satellite-based vegetation optical depth, while v3.3b spans from 2003 to 2018, and its forcing factors are the same as v3.3a except for radiation and air temperature, which are based on remotely sensed data. Both v3.3a and v3.3b estimates are provided at a monthly temporal resolution and a 0.25°×0.25° (≈25 km) spatial resolution (Miralles et al., 2011b;Miralles et al., 2011a;Martens et al., 2017).

GLDAS ET

The Global Land Data Assimilation System (GLDAS) generates optimal fields of land surface states and fluxes using advanced land surface modeling and data assimilation techniques by ingesting satellite and ground-based observational data products. GLDAS Version 2 has two components (GLDAS-2.0 and GLDAS-2.1) with a 0.25°×0.25° (≈25 km) spatial resolution and a 3-hour temporal resolution. GLDAS-2.0 is reprocessed with the updated Princeton Global Meteorological Forcing Dataset and upgraded Land Information System Version 7. The model simulation was initialized from January 1, 1948, to December 31, 2010, using soil moisture and other state fields from the LSM climatology for that day of the year. The simulation used the common GLDAS datasets for land cover (MCD12Q1), land-water mask (MOD44W), and soil texture and elevation (GTOPO30). The GLDAS-2.1 simulation started on January 1, 2000, and lasted until December 31, 2019, using the conditions from the GLDAS-2.0 simulation. This simulation was forced with the National Oceanic and Atmospheric Administration (NOAA)/Global Data Assimilation System (GDAS) atmospheric analysis, disaggregated Global Precipitation Climatology Project (GPCP) precipitation, and Air Force Weather Agency's AGRicultural METeorological modeling system (AGRMET) radiation. The MODIS-based land surface parameters were used in the current GLDAS-2.x products, while the AVHRR base parameters were used in previous GLDAS-2 products before October 2012 (Rodell et al., 2004).

FLDAS ET

The Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System (FLDAS) dataset uses remotely sensed and reanalysis inputs to drive land surface models. It includes information on many climate-related variables, including evapotranspiration, moisture content, humidity, average soil temperature, and total precipitation rate. For forcing data, this FLDAS dataset uses a combination of the new version of Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) data and Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), a quasi-global rainfall dataset designed for seasonal drought monitoring and trend analysis (McNally et al., 2017). FLDAS is provided at a 0.1°×0.1° (≈10 km) spatial resolution and monthly temporal resolution during the period 1982–2019.

TerraClimate ET

TerraClimate ET is estimated based on a monthly one-dimensional soil water balance for global terrestrial surfaces, which incorporates evapotranspiration, precipitation, temperature, and interpolated plant extractable soil water capacity. The water balance model is very simple and does not account for heterogeneity in vegetation types or their physiological responses to changing environmental conditions (Abatzoglou et al., 2018). TerraClimate estimates are provided at a monthly temporal resolution from 1958 to 2018 and 0.041°×0.041° (≈5 km) grid cells.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C.: TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015, Scientific data, 5, 170191, https://doi.org/10.1038/sdata.2017.191, 2018.

Almusaed, A.: Evapotranspiration and Environmental Benefits, in: Biophilic and Bioclimatic Architecture: Analytical Therapy for the Next Generation of Passive Sustainable Architecture, edited by: Almusaed, A., Springer London, London, 167-171, 2011.

Andam-Akorful, S. A., Ferreira, V. G., Awange, J. L., Forootan, E., and He, X. F.: Multi-model and multi-sensor estimations of evapotranspiration over the Volta Basin, West Africa, International Journal of Climatology, 35, 3132-3145, https://doi.org/10.1002/joc.4198, 2015.

Arnell, N. W.: Climate change and global water resources, Global Environmental Change, 9, S31-S49, https://doi.org/10.1016/S0959-3780(99)00017-5, 1999.

Arnell, N. W., and Lloyd-Hughes, B.: The global-scale impacts of climate change on water resources and flooding under new climate and socio-economic scenarios, Climatic Change, 122, 127-140, https://doi.org/10.1007/s10584-013-0948-4, 2014.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies, Bulletin of the American Meteorological Society, 96, 69-83, https://doi.org/10.1175/bams-d-13-00068.1, 2015.

Badgley, G., Fisher, J. B., Jiménez, C., Tu, K. P., and Vinukollu, R.: On Uncertainty in Global Terrestrial Evapotranspiration Estimates from Choice of Input Forcing Datasets, Journal of Hydrometeorology, 16, 1449-1455, https://doi.org/10.1175/jhm-d-14-0040.1, 2015.

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.

Bastiaanssen, W. G. M., Karimi, P., Rebelo, L.-M., Duan, Z., Senay, G., Muthuwatte, L., and Smakhtin, V.: Earth observation based assessment of the water production and water consumption of Nile basin agro-ecosystems, Remote Sensing, 6, 10306-10334, https://doi.org/10.3390/rs61110306, 2014.

Bhattarai, N., Mallick, K., Stuart, J., Vishwakarma, B. D., Niraula, R., Sen, S., and Jain, M.: An automated multi-model evapotranspiration mapping framework using remotely sensed and reanalysis data, Remote Sensing of Environment, 229, 69-92, https://doi.org/10.1016/j.rse.2019.04.026, 2019.

Chen, X., Su, Z., Ma, Y., Yang, K., Wen, J., and Zhang, Y.: An Improvement of Roughness Height Parameterization of the Surface Energy Balance System (SEBS) over the Tibetan Plateau, Journal of Applied Meteorology and Climatology, 52, 607-622, https://doi.org/10.1175/jamc-d-12-056.1, 2013.

Chen, X., Su, Z., Ma, Y., Liu, S., Yu, Q., and Xu, Z.: Development of a 10-year (2001–2010) 0.1° data set of land-surface energy balance for mainland China, Atmospheric Chemistry and Physics, 14, 13097-13117, https://doi.org/10.5194/acp-14-13097-2014, 2014a.

Chen, X., Massman, W. J., and Su, Z.: A Column Canopy-Air Turbulent Diffusion Method for Different Canopy Structures, Journal of Geophysical Research: Atmospheres, 124, 488-506, https://doi.org/10.1029/2018jd028883, 2019.

Chen, Y., Xia, J., Liang, S., Feng, J., Fisher, J. B., Li, X., Li, X., Liu, S., Ma, Z., Miyata, A., Mu, Q., Sun, L., Tang, J., Wang, K., Wen, J., Xue, Y., Yu, G., Zha, T., Zhang, L., Zhang, Q., Zhao, T., Zhao, L., and Yuan, W.: Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China, Remote Sensing of Environment, 140, 279-293, https://doi.org/10.1016/j.rse.2013.08.045, 2014b.

Danielson, J. J., and Gesch, D. B.: Global multi-resolution terrain elevation data 2010, 2011-1073, 2011.

Degefu, D. M., Weijun, H., Zaiyi, L., Liang, Y., Zhengwei, H., and Min, A.: Mapping Monthly Water Scarcity in Global Transboundary Basins at Country-Basin Mesh Based Spatial Resolution, Scientific Reports, 8, 2144-2144, https://doi.org/10.1038/s41598-018-20032-w, 2018.

Elnashar, A., Wang, L., Wu, B., Zhu, W., and Zeng, H.: Synthesis of Global Actual Evapotranspiration from 1982 to 2019, V1, Harvard Dataverse, https://doi.org/10.7910/DVN/ZGOUED, 2020.

Ershadi, A., McCabe, M. F., Evans, J. P., Chaney, N. W., and Wood, E. F.: Multi-site evaluation of terrestrial evaporation models using FLUXNET data, Agricultural and Forest Meteorology, 187, 46-61, https://doi.org/10.1016/j.agrformet.2013.11.008, 2014.

FAO: WaPOR Database Methodology: Level 1 Data using remote sensing in support of solutions to reduce agricultural water productivity gaps, Technical Report, FAO, Rome, 2018.

FAO: WaPOR V2 Database Methodology. Remote Sensing for Water Productivity Technical Report: Methodology Series. Rome, FAO., 2020.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, Reviews of Geophysics, 45, https://doi.org/10.1029/2005rg000183, 2007.

Ferguson, P. R., and Veizer, J.: Coupling of water and carbon fluxes via the terrestrial biosphere and its significance to the Earth's climate system, Journal of Geophysical Research: Atmospheres, 112, http://dx.doi.org/10.1029/2007jd008431, 2007.

Fisher, J. B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., McCabe, M. F., Hook, S., Baldocchi, D., Townsend, P. A., Kilic, A., Tu, K., Miralles, D. D., Perret, J., Lagouarde, J.-P., Waliser, D., Purdy, A. J., French, A., Schimel, D., Famiglietti, J. S., Stephens, G., and Wood, E. F.: The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources, Water Resources Research, 53, 2618-2626, https://doi.org/10.1002/2016wr020175, 2017.

Foken, T.: The energy balance closure problem: An overview, Ecological Applications, 18, 1351-1367, https://doi.org/10.1890/06-0922.1, 2008.

Foken, T., Aubinet, M., and Leuning, R.: The Eddy Covariance Method, in: Eddy Covariance: A Practical Guide to Measurement and Data Analysis, edited by: Aubinet, M., Vesala, T., and Papale, D., Springer Netherlands, Dordrecht, 1-19, 2012.

Forootan, E., Khaki, M., Schumacher, M., Wulfmeyer, V., Mehrnegar, N., van Dijk, A. I. J. M., Brocca, L., Farzaneh, S., Akinluyi, F., Ramillien, G., Shum, C. K., Awange, J., and Mostafaie, A.: Understanding the global hydrological droughts of 2003–2016 and their relationships with teleconnections, Science of The Total Environment, 650, 2587-2604, https://doi.org/10.1016/j.scitotenv.2018.09.231, 2019.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., and Hoell, A.: The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes, Scientific Data, 2, 150066, https://doi.org/10.1038/sdata.2015.66, 2015.

Gentine, P., Green, J. K., Guérin, M., Humphrey, V., Seneviratne, S. I., Zhang, Y., and Zhou, S.: Coupling between the terrestrial carbon and water cycles-a review, Environmental Research Letters, 14, 083003, http://dx.doi.org/10.1088/1748-9326/ab22d6, 2019.

Ghilain, N., Arboleda, A., and Gellens-Meulenberghs, F.: Evapotranspiration modelling at large scale using near-real time MSG SEVIRI derived data, Hydrology and Earth System Sciences, 15, 771-786, https://doi.org/10.5194/hess-15-771-2011, 2011.

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., Stacke, T., Tessler, Z. D., Wada, Y., and Wisser, D.: Global water resources affected by human interventions and climate change, Proceedings of the National Academy of Sciences, 111, 3251, https://doi.org/10.1073/pnas.1222475110, 2014.

Helgason, W., and Pomeroy, J.: Problems Closing the Energy Balance over a Homogeneous Snow Cover during Midwinter, Journal of Hydrometeorology, 13, 557-572, https://doi.org/10.1175/JHM-D-11-0135.1, 2012.

Henderson-Sellers, B.: A new formula for latent heat of vaporization of water as a function of temperature, Quarterly Journal of the Royal Meteorological Society, 110, 1186-1190, https://doi.org/10.1002/qj.49711046626, 1984.

Hofste, R. W.: Comparative analysis among near-operational evapotranspiration products for the Nile basin based on earth observations first steps towards an ensemble product, M.Sc. Thes, Delft University of Technology, the Netherlands, 2014.

Hu, G., Jia, L., and Menenti, M.: Comparison of MOD16 and LSA-SAF MSG evapotranspiration products over Europe for 2011, Remote Sensing of Environment, 156, 510-526, https://doi.org/10.1016/j.rse.2014.10.017, 2015.

Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., and Schneider, U.: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset, Bulletin of the American Meteorological Society, 78, 5-20, https://doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2, 1997.

Jia, Z., Liu, S., Xu, Z., Chen, Y., and Zhu, M.: Validation of remotely sensed evapotranspiration over the Hai River Basin, China, Journal of Geophysical Research, 17, 1-21, https://doi.org/10.1029/2011JD017037, 2012.

Jiménez, C., Prigent, C., and Aires, F.: Toward an estimation of global land surface heat fluxes from multisatellite observations, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/10.1029/2008jd011392, 2009.

Kim, H. W., Hwang, K., Mu, Q., Lee, S. O., and Choi, M.: Validation of MODIS 16 global terrestrial evapotranspiration products in various climates and land cover types in Asia, Journal of Civil Engineering, 16, 229-238, https://doi.org/10.1007/s12205-012-0006-1, 2012.

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, https://dx.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, 1692, https://doi.org/10.3390/rs10111692, 2018.

Liu, S. M., Xu, Z. W., Zhu, Z. L., Jia, Z. Z., and Zhu, M. J.: 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.

Lu, Y., Cai, H., Jiang, T., Sun, S., Wang, Y., Zhao, J., Yu, X., and Sun, J.: Assessment of global drought propensity and its impacts on agricultural water use in future climate scenarios, Agricultural and Forest Meteorology, 278, 107623, https://doi.org/10.1016/j.agrformet.2019.107623, 2019.

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, 2019.

Majozi, N., Mannaerts, C., Ramoelo, A., Mathieu, R., Mudau, A., and Verhoef, W.: An intercomparison of satellite-based daily evapotranspiration estimates under different eco-climatic regions in South Africa, Remote Sensing, 9, 307, https://doi.org/10.3390/rs9040307, 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.

McCabe, M. F., Ershadi, A., Jimenez, C., Miralles, D. G., Michel, D., and Wood, E. F.: The GEWEX LandFlux project: evaluation of model evaporation using tower-based and globally gridded forcing data, Geoscientific Model Development, 9, 283-305, https://doi.org/10.5194/gmd-9-283-2016, 2016.

McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., Funk, C., Peters-Lidard, C. D., and Verdin, J. P.: A land data assimilation system for sub-Saharan Africa food and water security applications, Scientific Data, 4, 170012, https://doi.org/10.1038/sdata.2017.12, 2017.

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., Martens, B., McCabe, M. F., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project – Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms, Hydrology and Earth System Sciences, 20, 803-822, https://doi.org/10.5194/hess-20-803-2016, 2016.

Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., and 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-15-967-2011, 2011a.

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.

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.

Mu, Q., Zhao, M., and Steven, W.: Running and Numerical Terradynamic Simulation Group: MODIS Global Terrestrial Evapotranspiration (ET) Product MOD16A2 Collection 5, 2014a.

Mu, Q., Zhao, M., and Steven , W.: MODIS Global Terrestrial Evapotranspiration (ET) Product MOD16A2 Collection 5, 2014b.

Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan, F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y., and Seneviratne, S. I.: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, Hydrology and Earth System Sciences, 17, 3707-3720, https://doi.org/10.5194/hess-17-3707-2013, 2013.

Munia, H., Guillaume, J. H. A., Mirumachi, N., Porkka, M., Wada, Y., and Kummu, M.: Water stress in global transboundary river basins: significance of upstream water use on downstream stress, Environmental Research Letters, 11, 014002, https://dx.doi.org/10.1088/1748-9326/11/1/014002, 2016.

Naumann, G., Alfieri, L., Wyser, K., Mentaschi, L., Betts, R. A., Carrao, H., Spinoni, J., Vogt, J., and Feyen, L.: Global Changes in Drought Conditions Under Different Levels of Warming, Geophysical Research Letters, 45, 3285-3296, https://doi.org/10.1002/2017gl076521, 2018.

Oki, T., and Kanae, S.: Global hydrological cycles and world water resources, Science, 313, 1068-1072, https://doi.org/10.1126/science.1128845, 2006.

Pan, S., Tian, H., Dangal, S. R. S., Yang, Q., Yang, J., Lu, C., Tao, B., Ren, W., and Ouyang, Z.: Responses of global terrestrial evapotranspiration to climate change and increasing atmospheric CO2 in the 21st century, Earth's Future, 3, 15-35, https://doi.org/10.1002/2014ef000263, 2015.

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

Revelli, R., and Porporato, A.: Ecohydrological model for the quantification of ecosystem services provided by urban street trees, Urban Ecosystems, 21, 489-504, https://dx.doi.org/10.1007/s11252-018-0741-2, 2018.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85, 381-394, https://doi.org/10.1175/bams-85-3-381, 2004.

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., Adler, R., Bosilovich, M. G., Clayson, C. A., Chambers, D., Clark, E., Fetzer, E. J., Gao, X., Gu, G., Hilburn, K., Huffman, G. J., Lettenmaier, D. P., Liu, W. T., Robertson, F. R., Schlosser, C. A., Sheffield, J., and Wood, E. F.: The Observed State of the Water Cycle in the Early Twenty-First Century, Journal of Climate, 28, 8289-8318, https://doi.org/10.1175/jcli-d-14-00555.1, 2015.

Savoca, M. E., Senay, G. B., Maupin, M. A., Kenny, J. F., and Perry, C. A.: Actual evapotranspiration modeling using the operational Simplified Surface Energy Balance (SSEBop) approach, U.S. Geological Survey Scientific Investigations Report 2013-126, 16 p, 2013.

Schaffrath, D., and Bernhofer, C.: Variability and distribution of spatial evapotranspiration in semi arid Inner Mongolian grasslands from 2002 to 2011, SpringerPlus, 2, 547, https://doi.org/10.1186/2193-1801-2-547, 2013.

Scheff, J., and Frierson, D. M. W.: Scaling Potential Evapotranspiration with Greenhouse Warming, Journal of Climate, 27, 1539-1558, https://doi.org/10.1175/jcli-d-13-00233.1, 2014.

Scott, C. A., Silva-Ochoa, P., Florencio-Cruz, V., and Wester, P.: Competition for Water in the Lerma-Chapala Basin, in: The Lerma-Chapala Watershed: Evaluation and Management, edited by: Hansen, A. M., and van Afferden, M., Springer US, Boston, MA, 291-323, 2001.

Senay, G., Budde, M., Verdin, J., and Melesse, A.: A coupled remote sensing and simplified surface energy balance approach to estimate actual evapotranspiration from irrigated fields, Sensors, 7, 979-1000, https://doi.org/10.3390/s7060979, 2007.

Senay, G. B., Budde, M. E., and Verdin, J. P.: Enhancing the Simplified Surface Energy Balance (SSEB) approach for estimating landscape ET: Validation with the METRIC model, Agricultural Water Management, 98, 606-618, https://doi.org/10.1016/j.agwat.2010.10.014, 2011.

Senay, G. B., Bohms, S., and Verdin, J. P.: Remote sensing of evapotranspiration for operational drought monitoring using principles of water and energy balance, in, USGS Staff - Published Research. 979, 2012.

Senay, G. B., Bohms, S., Singh, R. K., Gowda, P. H., Velpuri, N. M., Alemu, H., and Verdin, J. P.: Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach, Journal of the American Water Resources Association, 49, 577-591, https://doi.org/10.1111/jawr.12057, 2013.

Senay, G. B., and Kagone, S.: Daily SSEBop Evapotranspiration: U. S. Geological Survey Data Release, https://doi.org/10.5066/P9L2YMV, 2019.

Senay, G. B., Kagone, S., and Velpuri, N. M.: Operational Global Actual Evapotranspiration: Development, Evaluation, and Dissemination, Sensors, 20, 1915, https://doi.org/10.3390/s20071915, 2020.

Sheffield, J., Wood, E. F., and Roderick, M. L.: Little change in global drought over the past 60 years, Nature, 491, 435-438, https://doi.org/10.1038/nature11575, 2012.

Sörensson, A. A., and Ruscica, R. C.: Intercomparison and Uncertainty Assessment of Nine Evapotranspiration Estimates Over South America, Water Resources Research, 54, 2891-2908, https://doi.org/10.1002/2017wr021682, 2018.

Spinoni, J., Barbosa, P., De Jager, A., McCormick, N., Naumann, G., Vogt, J. V., Magni, D., Masante, D., and Mazzeschi, M.: A new global database of meteorological drought events from 1951 to 2016, Journal of Hydrology: Regional Studies, 22, 100593, https://doi.org/10.1016/j.ejrh.2019.100593, 2019.

Su, Z.: The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes, Hydrology and Earth System Sciences, 6, 85-99, https://doi.org/10.5194/hess-6-85-2002, 2002.

Tang, R., Shao, K., Li, Z.-L., Wu, H., Tang, B.-H., Zhou, G., and Zhang, L.: Multiscale validation of the 8-day MOD16 evapotranspiration product using flux data collected in China, Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8, 1478-1486, https://doi.org/10.1109/JSTARS.2015.2420105, 2015.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research, 106, 7183-7192, https://doi.org/10.1029/2000JD900719, 2001.

Trambauer, P., Dutra, E., Maskey, S., Werner, M., Pappenberger, F., van Beek, L. P. H., and Uhlenbrook, S.: Comparison of different evaporation estimates over the African continent, Hydrology and Earth System Sciences, 18, 193-212, https://doi.org/10.5194/hess-18-193-2014, 2014.

Trenberth, K. E., Smith, L., Qian, T., Dai, A., and Fasullo, J.: Estimates of the Global Water Budget and Its Annual Cycle Using Observational and Model Data, Journal of Hydrometeorology, 8, 758-769, https://doi.org/10.1175/jhm600.1, 2007.

UNEP: World atlas of desertification, United Nations Environment Programme, 1997.

Velpuri, N. M., Senay, G. B., Singh, R. K., Bohms, S., and Verdin, J. P.: A comprehensive evaluation of two MODIS evapotranspiration products over the conterminous United States: Using point and gridded FLUXNET and water balance ET, Remote Sensing of Environment, 139, 35-49, https://doi.org/10.1016/j.rse.2013.07.013, 2013.

Vinukollu, R. K., Meynadier, R., Sheffield, J., and Wood, E. F.: Multi-model, multi-sensor estimates of global evapotranspiration: climatology, uncertainties and trends, Hydrological Processes, 25, 3993-4010, https://dx.doi.org/10.1002/hyp.8393, 2011a.

Vinukollu, R. K., Wood, E. F., Ferguson, C. R., and Fisher, J. B.: Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches, Remote Sensing of Environment, 115, 801-823, https://dx.doi.org/10.1016/j.rse.2010.11.006, 2011b.

Walter, I. A., Allen, R. G., Elliott, R., Jensen, M. E., Itenfisu, D., Mecham, B., Howell, T. A., Snyder, R., Brown, P., Echings, S., Spofford, T., Hattendorf, M., Cuenca, R. H., Wright, J. L., and Martin, D.: ASCE's Standardized Reference Evapotranspiration Equation, in: Watershed Management and Operations Management 2000, 1-11, 2001.

Wang, K., and Dickinson, R. E.: A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, Reviews of Geophysics, 50, https://doi.org/10.1029/2011rg000373, 2012.

Waring, R. H., and Running, S. W.: CHAPTER 10 - Advances in Eddy-Flux Analyses, Remote Sensing, and Evidence of Climate Change, in: Forest Ecosystems, Third Edition ed., edited by: Waring, R. H., and Running, S. W., Academic Press, San Diego, 317-344, 2007a.

Waring, R. H., and Running, S. W.: Chapter 2 - Water Cycle, in: Forest Ecosystems, Third Edition ed., edited by: Waring, R. H., and Running, S. W., Academic Press, San Diego, 19-57, 2007b.

Weerasinghe, I., Bastiaanssen, W., Mul, M., Jia, L., and van Griensven, A.: Can we trust remote sensing evapotranspiration products over Africa?, Hydrology and Earth System Sciences, 24, 1565-1586, https://doi.org/10.5194/hess-24-1565-2020, 2020.

Wu, B., Tian, F., Zhang, M., Zeng, H., and Zeng, Y.: Cloud services with big data provide a solution for monitoring and tracking sustainable development goals, Geography and Sustainability, 1, 25-32, https://doi.org/10.1016/j.geosus.2020.03.006, 2020.

Xu, T., Guo, Z., Xia, Y., Ferreira, V. G., Liu, S., Wang, K., Yao, Y., Zhang, X., and Zhao, C.: Evaluation of twelve evapotranspiration products from machine learning, remote sensing and land surface models over conterminous United States, Journal of Hydrology, 578, 124105, https://doi.org/10.1016/j.jhydrol.2019.124105, 2019.

Yamamoto, M. K., and Shige, S.: Implementation of an orographic/nonorographic rainfall classification scheme in the GSMaP algorithm for microwave radiometers, Atmospheric Research, 163, 36-47, https://doi.org/10.1016/j.atmosres.2014.07.024, 2015.

Yang, H., Luo, P., Wang, J., Mou, C., Mo, L., Wang, Z., Fu, Y., Lin, H., Yang, Y., and Bhatta, L. D.: Ecosystem Evapotranspiration as a Response to Climate and Vegetation Coverage Changes in Northwest Yunnan, China, PLOS ONE, 10, e0134795, https://doi.org/10.1371/journal.pone.0134795, 2015.

Yang, X., Yong, B., Ren, L., Zhang, Y., and Long, D.: Multi-scale validation of GLEAM evapotranspiration products over China via ChinaFLUX ET measurements, International Journal of Remote Sensing, 38, 5688-5709, https://doi.org/10.1080/01431161.2017.1346400, 2017.

Yang, Z., Zhang, Q., and Hao, X.: Evapotranspiration trend and its relationship with precipitation over the loess plateau during the last three decades, Advances in Meteorology, 1-10, https://doi.org/10.1155/2016/6809749, 2016.

Zhang, K., Kimball, J. S., Mu, Q., Jones, L. A., Goetz, S. J., and Running, S. W.: Satellite based analysis of northern ET trends and associated changes in the regional water balance from 1983 to 2005, Journal of Hydrology, 379, 92-110, https://doi.org/10.1016/j.jhydrol.2009.09.047, 2009.

Zhang, K., Kimball, J. S., Nemani, R. R., and Running, S. W.: A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006, Water Resources Research, 46, https://doi.org/10.1029/2009wr008800, 2010.

Zhang, K., Kimball, J. S., and Running, S. W.: A review of remote sensing based actual evapotranspiration estimation, Wiley Interdisciplinary Reviews: Water, 3, 834-853, https://doi.org/10.1002/wat2.1168, 2016.

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

Zhong, Y., Zhong, M., Mao, Y., and Ji, B.: Evaluation of Evapotranspiration for Exorheic Catchments of China during the GRACE Era: From a Water Balance Perspective, Remote Sensing, 12, 511, https://dx.doi.org/10.3390/rs12030511, 2020.

Zomer, R. J., Trabucco, A., Bossio, D. A., and Verchot, L. V.: Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation, Agriculture, Ecosystems & Environment, 126, 67-80, https://doi.org/10.1016/j.agee.2008.01.014, 2008.