Journal of Hydrology 610 (2022) 127990

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

The Global LAnd Surface Satellite (GLASS) evapotranspiration product Version 5.0: Algorithm development and preliminary validation

Zijing Xie [a], Yunjun Yao [a,*], Xiaotong Zhang [a,*], Shunlin Liang [b], Joshua B. Fisher [c], Jiquan Chen [d], Kun Jia [a], Ke Shang [a], Junming Yang [a], Ruiyang Yu [a], Xiaozheng Guo [a], Lu Liu [a], Jing Ning [a], Lilin Zhang [e]

[a] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[b] Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA
[c] Schmid College of Science and Technology, Chapman University, 1 University Drive, Orange, CA 92866, USA
[d] Department of Geography, Environment, and Spatial Sciences, Michigan State University, East Lansing, MI 48823, USA
[e] Faculty of Geo-Information and Earth Observation (ITC), University of Twente, 7500 AE Enschede, The Netherlands

ARTICLE INFO

This manuscript was handled by Marco Borga, Editor-in-Chief, with the assistance of Hanbo Yang, Associate Editor

Keywords:
GLASS
Evapotranspiration
Deep neural networks
Machine learning
Bayesian model averaging

ABSTRACT

An accurate estimation of spatially and temporally continuous global terrestrial evapotranspiration (ET) is essential in the assessment of surface energy, water and carbon cycles. The Global LAnd Surface Satellite (GLASS) ET product Version 4.0 (v4.0) based on the Bayesian model averaging (BMA) method was generated to estimate global terrestrial ET. However, certain uncertainty for the GLASS ET product v4.0 limits its application. In this study, we introduced the deep neural networks (DNN) merging framework to improve terrestrial ET estimation for GLASS ET product Version 5.0 (v5.0) generation by integrating five satellite-derived ET products [Moderate Resolution Imaging Spectroradiometer (MODIS) ET product (MOD16), Shuttleworth–Wallace dual-source ET product (SW), Priestley–Taylor-based ET product (PT-JPL), modified satellite-based Priestley–Taylor ET product (MS-PT) and simple hybrid ET product (SIM)]. We compared the performance of DNN method against other merging methods, including GLASS ET algorithm v4.0 (BMA), the gradient boosting regression tree (GBRT) method and the random forest (RF) method, based on 195 global eddy covariance (EC) flux towers covering observations from 2000 through 2015. Validations indicated that the DNN had the highest accuracy among four merging methods across different land cover types, yielding the highest average determination coefficients (R^2, 0.62), root-mean-squared-error (RMSE, 24.1 W/m^2) and Kling–Gupta efficiency (KGE, 0.77) with a of 99% confidence interval. Compared with GLASS ET algorithm v4.0, the DNN improved on the R^2 by approximately 7% ($p < 0.01$) and the KGE by 10%. Based on the DNN, we then generated 8-day GLASS ET product v5.0 globally with a 1 km spatial resolution from 2001 to 2015 driven by GLASS vegetation and surface net radiation (R_n) datasets and Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA2) datasets. This global terrestrial ET product provides a valuable dataset for monitoring regional and global water resources and environmental changes.

1. Introduction

Evapotranspiration (ET), the loss of water from the Earth's surface to the atmosphere from soil evaporation, water bodies evaporation, canopy interception evaporation and vegetation transpiration, is a major component of the surface energy budget and water cycle (Fisher et al., 2017; Ma et al., 2021). For the terrestrial energy budget, surface latent heat flux (LE) accompanied by the ET process accounts for approximately 50% of the surface net radiation, using up more than half of the total solar energy absorbed by land surfaces and helping cool the land surface as an energy flux (Murphy et al., 2009; Trenberth et al., 2009). Additionally, for the water cycle, ET via energy flux exchanges returns approximately 65% of precipitation on the land surface back to the atmosphere at the global scale annually by consuming an enormous amount of heat (Oki and Kanae, 2006) and thereby is an important constraint on water availability at the land surface (Gleeson et al.,

* Corresponding authors.
E-mail addresses: boyyunjun@163.com (Y. Yao), xtngzhang@bnu.edu.cn (X. Zhang).

https://doi.org/10.1016/j.jhydrol.2022.127990
Received 9 March 2022; Received in revised form 15 May 2022; Accepted 23 May 2022
Available online 28 May 2022

2020). Accurate estimation of ET at the global scale provides important scientific information that is valuable to high-interest research fields, such as global environmental changes, water resource management, and sustainable agricultural development (Condon et al., 2020; Helbig et al., 2020; Liu et al., 2022; Ma and Szilagyi, 2019; Pascolini-Campbell et al., 2021).

ET can be measured by in-situ eddy covariance (EC) flux tower networks (e.g., FLUXNET), but these site-based measurements are spatially sparse and are available only over limited time periods (Baldocchi et al., 2001). Fortunately, satellites have provided temporally regular and spatially continuous ET estimates from fields to global scales because they can successfully retrieve land surface and atmospheric variables that are linked to ET variability (Brust et al., 2021; Fisher et al., 2020). In the past three decades, many satellite-derived ET products have been available (Table 1), including the Global Moderate Resolution Imaging Spectroradiometer (MODIS) ET products (MOD16) (500 m, 8-day) (Mu et al., 2007; Mu et al., 2011); the Breathing Earth System Simulator (BESS) ET product (1 km, 8-day) (Jiang and Ryu, 2016); the Global Land Evaporation Amsterdam Model (GLEAM) v3 ET product (0.25°, daily) (Martens et al., 2017); the Penman–Monteith–Leuning (PML)_V2 global ET product (0.05°, 8-day) (Zhang et al., 2019; Zhang et al., 2016); the ETMonitor ET product (1 km, daily) (Hu and Jia, 2015), the FLUXCOM ET product (0.0833°, 8-day), and the Global LAnd Surface Satellite (GLASS) ET product Version 4.0 (v4.0) (0.05°/1 km, 8-day) (Liang et al., 2021; Yao et al., 2014). However, except for FLUXCOM and GLASS ET product v4.0, these products are generated using individual models that will lead to low performance in ET estimation (Shang et al., 2021). Previous studies have shown that the multi-model merging method performs better than any individual model for estimating terrestrial ET (Aires, 2014; Jimenez et al., 2018; Jung et al., 2019; Mueller et al., 2013). For example, the FLUXCOM ET product (0.0833°, 8-day) employs an ensemble of multiple machine learning approaches (Jung et al., 2019), but their accuracy remains unclear because theirs errors are inherent in the underlying EC measurements and the spatially biased distribution of FLUXNET sites (Ma et al., 2020). To obtain ET estimations with high accuracy, more studies on merging ET products are still needed.

As an established multi-model ensemble ET product, GLASS ET product v4.0 (1 km, 8-day) was produced based on the Bayesian model averaging (BMA) method by merging five satellite-based physical ET models (Yao et al., 2014), including the MOD16 model (Mu et al., 2011), the Revised Remote-Sensing-based Penman–Monteith (RRS-PM) ET model (Yuan et al., 2010), the Priestley–Taylor-based (PT-JPL) ET model (Fisher et al., 2008), the Modified Satellite-based Priestley–Taylor ET (MS-PT) model (Yao et al., 2013) and the semi-empirical Penman ET model of the University of Maryland (UMD-SEMI) (Wang et al., 2010). Subsequently, GLASS ET product v4.0 has been evaluated and validated, and the preliminary results indicate that it has relatively higher quality and accuracy than some existing satellite-derived ET products (Liang et al., 2021; Song et al., 2018; Yang et al., 2022). For example, Yang et al. (2022) used GLASS ET product v4.0 to investigate the impacts of various climate and vegetation factors on ET variation in northwest China because its values are closest to the ground-observations. Besides, GLASS ET product v4.0 was used in related studies on topics such as climate change and vegetation dynamics (Song et al., 2018; Shang et al., 2021). For instance, Song et al. (2018) compared five major satellite ET products over the Heihe River Basin (HRB) in China and found that GLASS ET product v4.0 was slightly more accurate than many other products because the BMA method yields a more reliable output in the ensemble than any individual model. However, the BMA method is a linear combination of each single ET model, which may perform worse than many nonlinear multi-model ensemble approaches (e.g., machine learning methods) (Duan et al., 2007; Yao et al., 2017b). For example, Bai et al. (2021) found that the machine learning method is superior to the BMA method in merging individual ET models for cropland ET estimates across a wide range of environmental conditions. Currently, with the development of deep learning methods, there is a new opportunity for updating GLASS ET product v4.0 using deep learning methods to replace the BMA method by merging individual ET products (Ball et al., 2017; Yann et al., 2015; Yuan et al., 2020).

Recently, deep neural networks (DNN) has been widely used to upscale ET from site to regional scales by relating ground-measured ET to satellite-derived variables and other meteorological data because DNN has multiple layers of parameterized differentiable nonlinear modules trained by backpropagation to tackle a wide range of problems with the large accumulation of satellite and ground observation data (Elbeltagi et al., 2020; Saggi and Jain, 2019; Shang et al., 2021). For example, Cui et al. (2021) developed a new combined model coupling DNN and the two-source energy balance (TSEB) model (TSEB_DNN) over the HRB in China to generate spatiotemporally continuous ET. The TSEB_DNN model was consistent with the in-situ measurements and had overall correlation coefficient (R) of 0.88, root-mean-square-error (RMSE) of 0.88 mm/day and bias of 0.37 mm/day. However, there is a lack of studies on improving global terrestrial ET estimation using the DNN method by merging multiple satellite-derived ET products.

In the study, we proposed a DNN-merging framework to generate GLASS ET product v5.0 by merging five satellite-derived ET products. Our specific objectives are to (1) evaluate GLASS ET algorithm v5.0 that merges five satellite-derived ET products driven by the individual model using long-term FLUXNET EC data from 2000 through 2015; (2) compare the performances of DNN method against other merging methods (GLASS ET algorithm v4.0, Gradient boosting regression tree and Random forest) using EC observations at the site scale; and (3) generate the 8-day GLASS ET product v5.0 with a 1 km spatial resolution during 2001–2015 driven by GLASS vegetation and surface net radiation (R_n) datasets and Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA2) datasets.

2. Data

2.1. Eddy covariance observations

Ground-measured EC data were used to evaluated the GLASS ET algorithm v5.0 and other three merging methods. The data from FLUXNET2015 database were collected from 195 EC flux tower sites (Fig. 1) distributed mainly in North America, Europe and Asia, with 5 EC flux tower sites in Africa, 4 EC flux tower sites in Australia and 4 EC flux tower sites in South America (Pastorello et al., 2020). These flux tower sites covered ten International Geosphere-Biosphere Program (IGBP) land cover types for the 2000–2015 period: cropland (CRO, 29 flux

Table 1
Summary of satellite-based ET products.

Product	Domain	Resolution	Time span	Algorithm
MOD16	Global	500-m/8-day	2001–2021	Penman–Monteith algorithm
BESS	Global	1 km/8-day	2000–2015	Simplified process-based model
GLEAM v3	Global	0.25°/daily	1980–2020	Priestley–Taylor algorithm
PML_V2	Global	0.05°/8-day	2002–2019	Penman–Monteith–Leuning model
ETMonitor	Global	1 km/daily	2009–2011	Shuttleworth–Wallace Dual-Source model
FLUXCOM	Global	0.0833°/8-day	2000–2015	Machine learning method
GLASS v4.0	Global	(0.05°)1 km/8-day	(1981) 2000–2020	Bayesian model averaging method

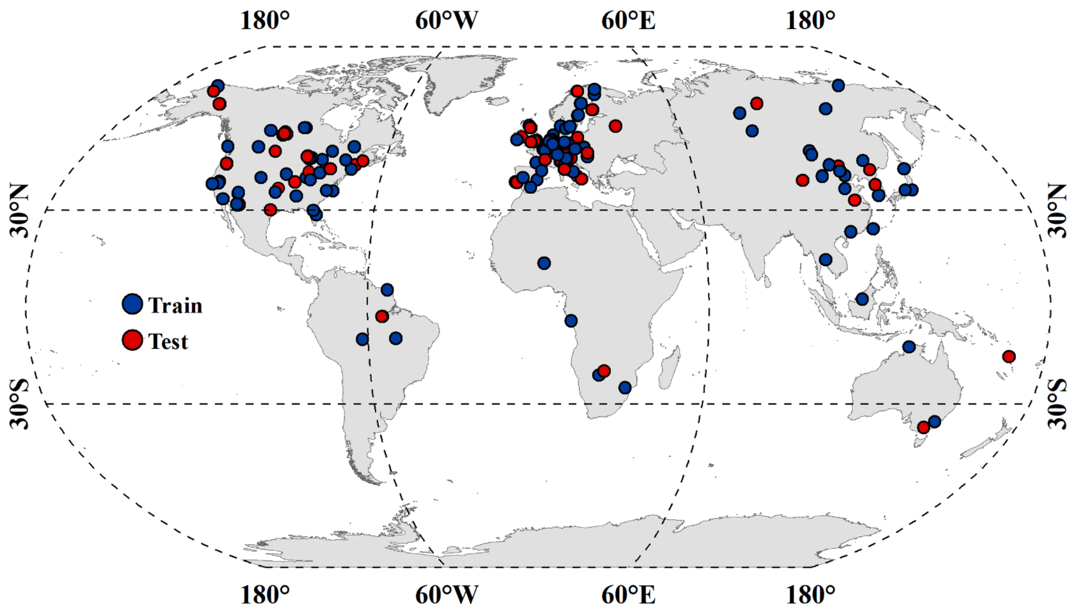

Fig. 1. Distribution of the 195 EC sites in the study. "Train" represents the training sites, and "Test" represents the validation sites.

tower sites), deciduous broadleaf forest (DBF, 25 flux tower sites), deciduous needleleaf forest (DNF, 4 flux tower sites), evergreen broadleaf forest (EBF, 12 flux tower sites), evergreen needleleaf forest (ENF, 50 flux tower sites), grassland (GRA, 36 flux tower sites), mixed forest (MIF, 9 flux tower sites), savanna (SAW, 8 flux tower sites), shrubland (SHR, 12 flux tower sites) and wetland (WET, 10 flux tower sites). We randomly divided the 195 EC flux tower sites into training and independent validation groups: 132 sites (~70%) for training and the remaining 63 sites (~30%) for independent validation, each representing major global land cover types.

Observations from the EC towers include surface net radiation (R_n), soil heat flux (G), sensible heat flux (H), air temperature (T_a), diurnal air temperature range (DT), vapor pressure (e), shortwave radiation (R_s), relative humidity (RH), wind spread (WS), soil moisture (SM) and ET. The hourly or half-hourly ET, H and the corresponding meteorological variables were subsequently aggregated into daily means using the gap-filling method proposed by Reichstein et al. (2005), which utilizes both the co-variation of fluxes with meteorological variables and the temporal autocorrelation of fluxes. If more than 25% of the data were missing on a given day, the values of that day were considered missing. Otherwise, daily values were obtained by multiplying the average hourly rate by 24 (hours) (Yao et al., 2015). The data covered the period of 2000–2015, and each flux tower site had at least one year of data. Although the EC method is considered a good method to measure ET, it suffers an energy imbalance where the measured available energy (R_n − G) is greater than the sum of the measured ET and H (Foken, 2008). Therefore, we used the Bowen ratio closure method proposed by Twine et al. (2000) to correct ET as follows:

$$ET_{cor} = ET \times \frac{(R - G)}{(H + ET)} \quad (1)$$

where ET_{cor} is the corrected evapotranspiration.

2.2. Satellite-based ET products

GLASS ET product v5.0 was generated based on the DNN method, which merges five traditional satellite-based ET products (MOD16, SW, PT-JPL, MS-PT and SIM) with a daily temporal resolution and 1 km spatial resolution (Table 2). The inputs for each ET model include (1) GLASS vegetation and radiation datasets: fraction of absorbed photosynthetically active radiation (FPAR), leaf area index (LAI), normalized difference vegetation index (NDVI) products with a 1 km spatial resolution and 8-day temporal resolution, and the daily R_n with a 0.05° spatial resolution (Jiang et al., 2015; Liang et al., 2021; Xiao et al., 2016; Xiao et al., 2017; Xiao et al., 2014); and (2) MERRA2 datasets: T_a, DT, e, RH, SM and WS. The spatial resolution of the MERRA2 meteorological data is 1/2° X 2/3°. The daily FPAR, LAI, and NDVI values were temporally interpolated from the 8-day averages using the linear interpolation method, and all coarse resolution MERRA2 data and GLASS R_n data were resampled spatially to 1 km using the bilinear interpolation method. In addition, we used the MOD16 (Mu et al., 2011) and FLUXCOM ET products (Jung et al., 2019) as comparison to evaluate the performance of the DNN for ET estimation. Information of FLUXCOM ET product is summarized in supplementary material.

(1) MOD16 ET product.

The MOD16 ET product was calculated based on the Penman–Monteith equation, and the forcing data included daily satellite vegetation and daily surface meteorological data. The original beta version proposed by Mu et al. (2007) combined land cover, albedo and reanalysis data and improved Cleugh's algorithm (Cleugh et al., 2007) to produce the MODIS-based ET product at the global scale. Subsequently,

Table 2
Summary of the five satellite-based ET products generated in this study for 2001–2015.

Product	Resolution	Time span	Algorithm	Forcing Input of the LE product	
				MERRA2	GLASS
MOD16	1 km/8-day	2001–2015	Penman–Monteith algorithm	T_a, T_{min}, e, RH	FPAR, LAI, R_n
SW	1 km/8-day	2001–2015	Shuttleworth–Wallace dual-source model	RH, T_a, e, SM, WS	FPAR, LAI, NDVI, R_n
PT-JPL	1 km/8-day	2001–2015	Priestley–Taylor algorithm	T_a, T_{max}, e, RH	FPAR, LAI, NDVI, R_n
MS-PT	1 km/8-day	2001–2015	Priestley–Taylor algorithm	T_a, DT	NDVI, R_n
SIM	1 km/8-day	2001–2015	Simple hybrid algorithm	T_a, T_{max}, T_{min}	NDVI, R_n

Mu et al. (2011) improved the algorithm by utilizing the evaporation from canopy interception, which was not considered in the former algorithm. The latest improvements by Mu et al. (2011) include (1) simplifying the calculation of the vegetation cover fraction; (2) calculating ET as the sum of daytime and nighttime components; (3) adding the soil heat flux calculation; (4) improving estimates of the stomatal conductance, aerodynamic resistance and boundary layer resistance; (5) separating the dry canopy surface from the wet surface; and (6) dividing the soil surface into a saturated wet surface and a moist surface (Mu et al., 2011). In this study, considering the latest MOD16A2 Version 6 global ET product is an 8-day composite dataset and it doesn't include urban and barren areas, we calculated daily ET based on the MOD16 algorithm driven by daily meteorological and satellite data. The difference between the MOD16 we produced and the original product is that our estimated MOD16 ET includes the ET values of non-vegetated areas. We produced the daily MOD16 ET product with a 1 km spatial resolution for the period from 2001 to 2015.

(2) SW ET product.

SW ET product was calculated using the Shuttleworth–Wallace Dual-Source (SW) model proposed by Shuttleworth and Wallace (1985), which accounted separately for the energy balance for vegetation and soil components of a soil–vegetation canopy unity (Hu and Jia, 2015; Yao et al., 2017a). It is a one-dimensional model to describe the energy partition of sparse crops. Theoretical improvement of this model introduces a combination equation that describes evaporation in terms of controlling resistances associated with the plants, soil and water. The equation provides a simple but physically explainable illustration of the transition between bare substrate and a closed canopy (Shuttleworth and Wallace, 1985). Inputs of the model include RH, T_a, WS, LAI, SM and e from the MERRA2 datasets and R_n from the GLASS datasets. We produced the daily SW ET product with a 1 km spatial resolution for the period from 2001 to 2015.

(3) PT-JPL ET product.

PT-JPL ET product was derived from the Priestley–Taylor algorithm of the Jet Propulsion Laboratory (PT-JPL). The PT-JPL algorithm was based on the original work of Priestley and Taylor (1972) and a novel improvement by Fisher et al. (2008). The simplified version of the Penman–Monteith model proposed by Fisher et al. (2008) overcomes the calculation of aerodynamic and surface resistance and thereby requires no inputs of any ground-based observations. The model requires no site calibration, tuning or spin-ups and is applied on a per-pixel basis. The PT-JPL algorithm was validated by EC observations from 16 FLUXNET sites, and the validation results indicated that the R^2 was 0.90 (RMSE = 16 mm/month or 28%) across 2 years (Fisher et al. (2008)). The input variables include RH, R_n, T_a, WS, LAI, and e from the MERRA2 datasets and NDVI and FPAR from the GLASS datasets. We produced the daily PT-JPL ET product with a 1 km spatial resolution for the period from 2001 to 2015.

(4) MS-PT ET product.

MS-PT ET product was produced based on the Modified Satellite-based Priestley–Taylor (PT-JPL) algorithm developed by Yao et al. (2013). To circumvent the difficulty of the satellite-based estimation of RH and vapor pressure deficit (VPD) for the PT-JPL algorithm, the MS-PT algorithm used the Apparent Thermal Inertia (ATI) derived from the temperature (T_a, or land surface temperature, LST) change over time to replace RH and VPD for calculating the SM constraint. Similar to PT-JPL, the MS-PT algorithm avoids the computational complexities of aerodynamic resistance parameters. The MS-PT algorithm only needs four variables: R_n, T_a, DT and NDVI. The MS-PT algorithm was evaluated based on ground observations from 16 flux tower sites in China, and the results showed that compared with the PT-JPL model, the MS-PT algorithm increased the coefficient of determination (R^2) by approximately 10% and slightly reduced the root-mean-square error (RMSE) and mean biases. We produced the daily MS-PT ET product with a 1 km spatial resolution for the period from 2001 to 2015.

(5) SIM ET product.

SIM ET product was estimated based on a simple hybrid (SIM) algorithm proposed by Wang and Liang (2008). In a previous study, Wang et al. (2007) used a simple and relatively accurate algorithm that combines R_n, NDVI, and T_a. However, the former algorithm fails to effectively consider the influence of SM. Previous studies have found that SM has a potentially important effect on ET and that DT is also a good indicator to characterize SM (Detto et al., 2006; Krishnan et al., 2006). Therefore, Wang and Liang (2008) developed SIM algorithm using DT to replace SM to further improve ET parameterization. The SIM algorithm was validated based on long-term ground measurements from the Atmospheric Radiation Measurement (ARM) and AmeriFlux sites, and the results indicated that the SIM algorithm improved the accuracy of ET estimation relative to the beta algorithm of Wang et al. (2007). We produced the daily SIM ET product with a 1 km spatial resolution for the period from 2001 to 2015.

3. Methods

3.1. GLASS ET algorithm v5.0

GLASS ET algorithm v5.0 is a deep neural networks (DNN)-merging method that is a mathematical model analogous to human neurons, including input layers, hidden layers and output layers (Krizhevsky et al., 2012). Each layer of the DNN contains several neurons that are fully connected between layers. For instance, any neuron in Layer i is connected to all other neuron in Layer $i+1$. The connections between adjacent neurons are linear. Due to the fully connected relationships between multiple hidden layers, the DNN is able to simulate nonlinear feature relationships to maximize the target information extraction from limited data (Srivastava et al., 2014). The DNN has powerful nonlinear fitting capabilities and can approximate nonlinear continuous functions with arbitrary accuracy. The DNN can automatically extract training rules from data and has strong feature extraction capabilities. In addition, it is significantly more efficient in processing high-dimensional datasets than machine learning methods (Hinton and Salakhutdinov, 2006).

In this study, the DNN is applied to generate GLASS ET product v5.0 by merging five satellite-based ET products. We set five ET products (MOD16, SW, PT-JPL, MS-PT and SIM) as the input layer and set the estimated ET as the output layer (Fig. 2). Multiple hidden layers exist between the input and output layers, and the nodes within are fully connected. For each node, the output (b_i) is calculated as:

$$b_i = \sum_{n=1}^{m} W_{ni} \times a_n \tag{2}$$

where W_{ni} is the connection weight of the nth neuron in the hidden layer and the ith output layer neuron, and a_n is the output of the nth neuron in the hidden layer.

The availability of outputs from nodes is impacted by the activation function. We applied the ReLU as the activation function:

$$\sigma(b_i) = \begin{cases} \beta_i - \theta_i, if\ (\beta_i - \theta_i) \geqslant 0 \\ 0,\ if\ (\beta_i - \theta_i) < 0 \end{cases} \tag{3}$$

where σ is the activation function of the neural network, β_i is the output received by the ith neuron in the output layer and θ_i is the threshold of the jth neuron in the output layer.

For each node, the output of b_i is linear. The ReLU function introduces the nonlinear feature into connections among nodes while simultaneously avoiding the problems of vanishing gradients and exploding gradients, thus guaranteeing the accuracy of DNN fitting. The learning process of the DNN is used to adjust the connection weight between neurons and the threshold of each neuron based on the results of the training.

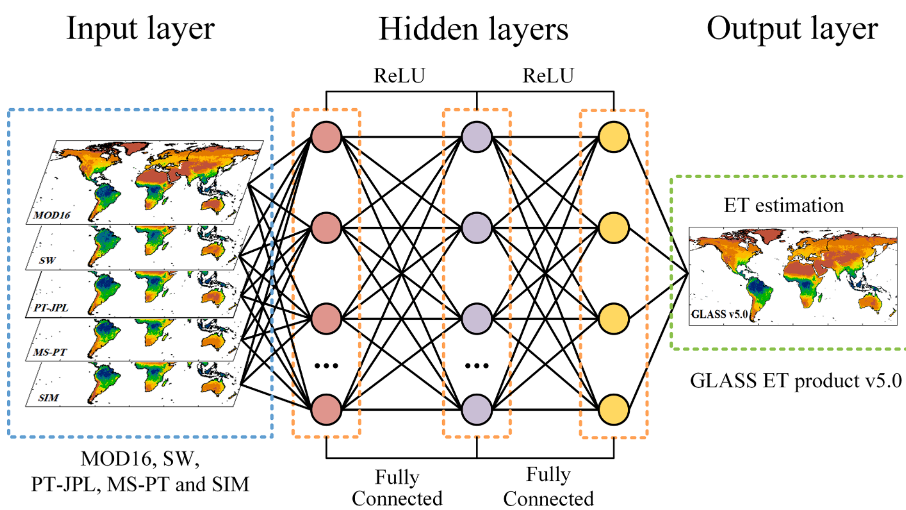

Fig. 2. The structure of the DNN used in this study.

3.2. Other merging algorithms

(1) GLASS ET algorithm v4.0.

GLASS ET algorithm v4.0 is a Bayesian model averaging (BMA)-merging method – a statistical probability theory-based method for synthesizing model results proposed by Raftery et al. (1997). The BMA establishes a probability density function (PDF) for the variables using the weighted average of the individual models' forecasted probability distributions after correction for bias, weighted by the posterior probability distribution of the corresponding model. GLASS ET product v4.0 used the BMA method to merge five individual ET products for terrestrial ET estimation (Yao et al., 2014). The PDF for ET weights each model by its posterior model probabilities and uses average weights of the PDFs for each model, which may reduce the uncertainties in individual models and thus improve ET accuracy (Chen et al., 2015).

(2) Gradient boosting regression tree.

Gradient boosting regression tree (GBRT) proposed by Friedman (2001) is an iterative decision tree algorithm formed by combining models from individual decision trees. The core idea of GBRT is to use the negative gradient of the loss function in the current model as an approximation to the residuals by minimizing the loss function, and each subtree is formed by iterating over the residuals of all previous subtrees. In each iteration, GBRT fits the residuals of a randomly sampled subset rather than the original training data. GBRT has better predictive capacity and stability than a single decision tree algorithm (Kotsiantis, 2013).

(3) Random forest.

Random forest (RF) is a machine learning algorithm that integrates multiple decision trees proposed by Breiman (2001), with the result relying on the judgment of each decision tree. For each subtree, random sampling is applied, and for each sample, random features are selected to build a subtree, and the optimal set of subtrees is iteratively selected to build the RF. The RF subsets perform well in both classification and regression and generate excellent generalization results using randomly selected subtree samples and features. The method has shown superiority in reducing variance and preventing overfitting as an ensemble of multiple subtrees (Elith et al., 2008).

3.3. Model evaluation metrics

To evaluate the performance of different algorithms, the correlation coefficient of determination (R^2), root-mean-square error (RMSE), bias and Kling–Gupta efficiency (KGE) were adopted. The KGE is a comprehensive assesment for model performance (Gupta et al., 2009) incorporating the mean value ratio (β), relative variability ratio (α) and correlation (r):

$$KGE = 1 - \sqrt{(\beta - 1)^2 + (\alpha - 1)^2 + (r - 1)^2} \tag{4}$$

$$\beta = \frac{\mu_{est}}{\mu_{obs}} \tag{5}$$

$$\alpha = \frac{\sigma_{est}}{\sigma_{obs}} \tag{6}$$

where μ_{est} and μ_{obs} are the mean values of estimations and observations, respectively; σ_{est} and σ_{obs} are the standard deviation of estimations and observations, respectively. In ideal circumstances with no simulation error, the values are $\beta = \alpha = r = 1$; thus, the optimal value of KGE is 1.

3.4. Experimental setup

To build models, we extracted ET observations from EC flux tower sites and the corresponding ET values from five satellite-based products. ET observations from 195 flux tower sites were collected as target variables, and ET values extracted from five ET products (MOD16, SW, PT-JPL, MS-PT and SIM) were used as predictor variables (Fig. 3). To train and validate the models, all 195 EC flux tower sites were separated into two independent subsets for model training (132 flux tower sites) and validation (63 flux tower sites), each representing major global land cover types. The optimal parameters that provided the highest correlation coefficient were chosen in the training data and then were used to estimate ET.

We built the DNN, BMA, RF and GBRT based on Sklearn, Keras and TensorFlow modules on the Python platform. To determine the optimal parameter for each method, we applied the GridSearchCV module. The GridSearchCV method tunes parameters by trying every possibility through loop traversal among all parameter combinations and thereby selecting the optimal ones. For DNN models built with Keras and TensorFlow modules, the major parameters of selection are activation and density. We tested the second-layer to the fifth-layer which fully connect the DNN model, implementing the merging of the satellite-derived ET products and ground-based measurements. For each hidden layer, we used nodes (2, 4, 8, 16, 32, 64, 128) for random combinations. To reduce the risk of overfitting, we restricted the hidden layers to three (16, 8, and 4) and adopted a 'rectified linear units' (ReLU) activation function for the hidden layer. We used the Adam optimizer and the mean squared error (MSE) loss function to compile the DNN model. The training terminal was set as 1000 echoes, and the RMSE was used as the cost function. Acquisition of the optimal parameters of the model not only improves the accuracy of model simulation but also shortens model

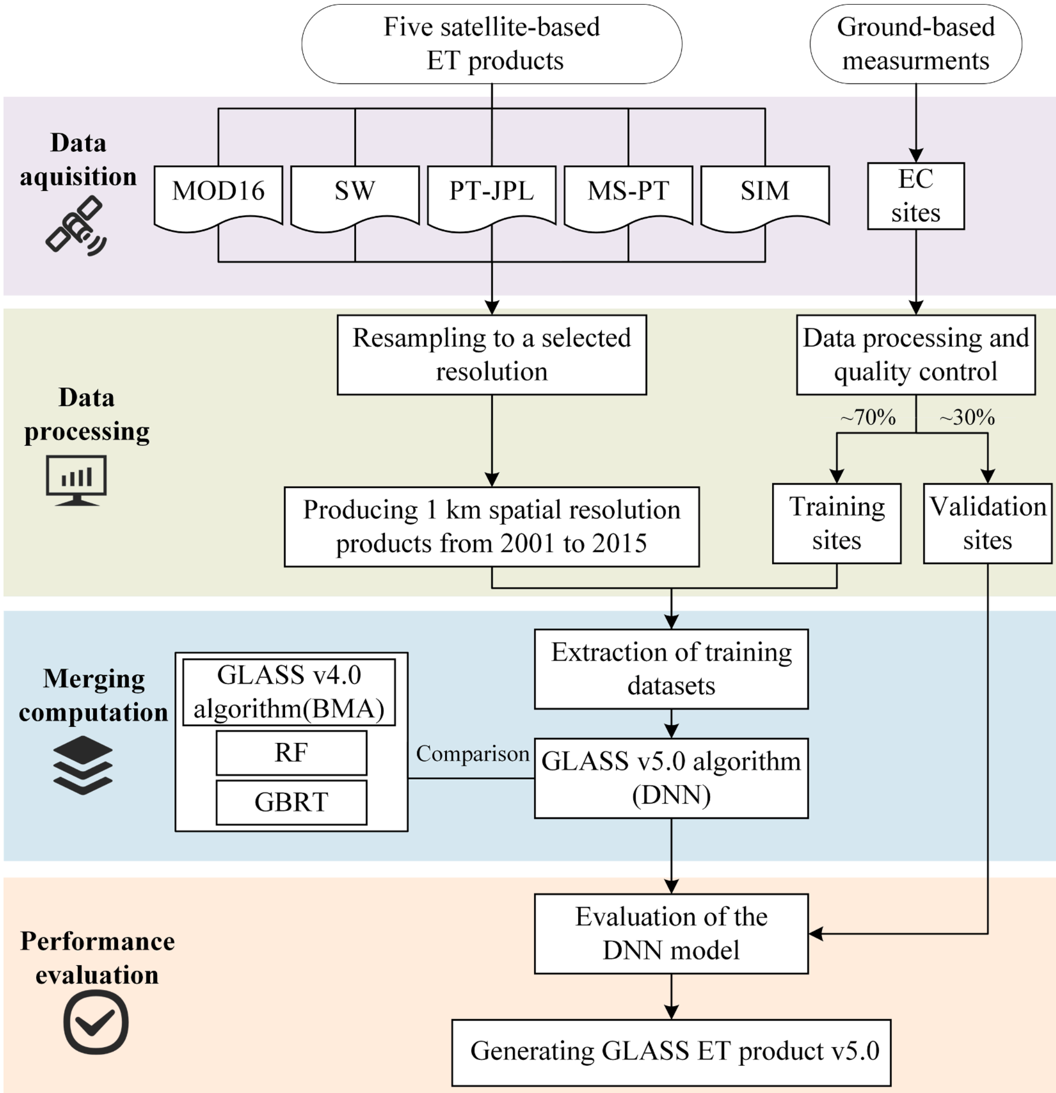

Fig. 3. The flowchart of the study.

running time.

4. Results

4.1. Evaluation of GLASS ET algorithm v5.0

4.1.1. Model development based on the training EC observations

We trained GLASS ET algorithm v5.0 (DNN), GLASS ET algorithm v4.0 (BMA), GBRT and RF using the EC data collected from 132 training flux tower sites for merging ET products. It's clear that the estimated daily ET from GLASS ET algorithm v4.0 is much closer to the EC observation than that of the other individual satellite-based ET products (supplementary material). Fig. 4 shows the comparison between the daily ET observations and the estimated ET from different methods at the 132 training sites. Regarding the parameters of R^2, RMSE and bias, GLASS ET algorithm v5.0, GBRT and RF yield rather similar results based on R^2 (0.68, 0.67 and 0.68, respectively), RMSE (22.2, 22.3 and 22.1 W/m^2, respectively) and bias (−0.2, −0.1 and −0.1 W/m^2, respectively), and preformed slightly better than the results of GLASS Algorithm v4.0 with a R^2 value of 0.64, RMSE of 23.5 W/m^2 and bias of 0.4 W/m^2. However, in terms of KGE, GLASS ET algorithm v5.0 performed better than with GBRT and RF. Compared with GLASS ET v4.0, GLASS ET v5.0 increased R^2 and KGE by 6% and 11%, respectively. GLASS ET v4.0 tended to underestimate ET when ET had high values, while GLASS ET v5.0 can partly improve the issue. Moreover, GLASS ET v5.0 shows a relatively symmetrical bias distribution and is thus more likely to provide unbiased ET estimation on average.

Fig. 5 compares statistics (R^2, RMSE, Bias and KGE) of different merging methods (GBRT, RF, GLASS ET v4.0 and GLASS ET v5.0) at 132 training flux tower sites across different land cover types. It is clear that for CRO, all four merging methods present the lowest performance with the lowest KGE (0.55–0.68) and R^2 (0.55–0.63, $p < 0.01$), highest RMSE (28.1–30.1 W/m^2) and negative bias (−10.9 to −8.6 W/m^2). In addition, all merging methods yield a relatively lower performance for EBF, GRA and WET. In contrast, all merging methods for MIF produce the best performance with the highest KGE (0.82–0.88) and R^2 (0.80–0.81, $p <$ 0.01) and lowest RMSE (16.4–17.3 W/m^2). Overall, for all land cover types, GLASS ET v5.0 outperforms GBRT, RF and GLASS ET v4.0, with favorable R^2 values ranging from 0.62 to 0.81 ($p < 0.01$), KGE ranging from 0.68 to 0.88, RMSE ranging from 16.4 to 28.1 W/m^2 and bias ranging from −10.1 to 4.2 W/m^2. Therefore, we can conclude that GLASS ET v5.0 is the best method among all four merging methods used in this study.

.

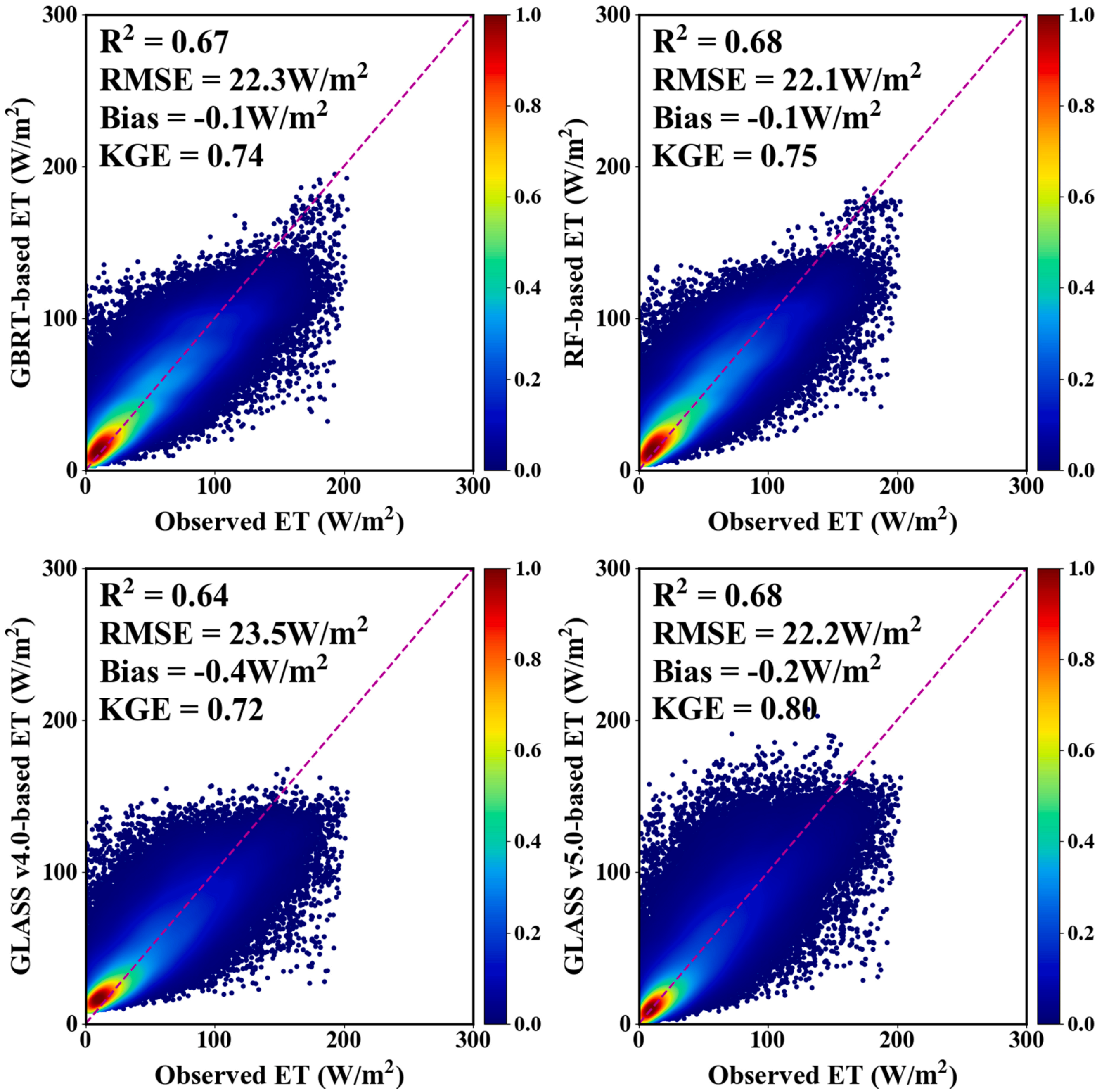

Fig. 4. Comparison of the estimated ET from different merging methods and ground measurements at 132 training flux tower sites for all land cover types.

4.1.2. Model evaluation based on the validation EC observations

We evaluated GLASS ET v5.0 (DNN), GLASS ET v4.0 (BMA), GBRT and RF using EC observations collected from 63 validation flux tower sites. Fig. 6 presents the comparison of the estimated ET from different merging methods and ground-observed ET at 63 validation flux tower sites for all land cover types. Although GLASS ET v4.0 has an acceptable accuracy based on R^2 (0.58, $p < 0.01$) and RMSE (24.3 W/m^2), it performs the worst in terms of KGE (0.70) by comparison with other merging methods. In contrast, GLASS ET v5.0 yields the best performance in terms of R^2 (0.62, $p < 0.01$), RMSE (24.1 W/m^2) and KGE (0.77), followed by RF and GBRT. Compared with GLASS ET algorithm v4.0, GLASS ET algorithm v5.0 improved R^2 and KGE by 7% ($p < 0.01$) and 10%, respectively. Similar to the conclusions for the training data, GLASS ET v4.0 still tends to underestimate ET when high ET occurs, while GLASS ET v5.0 yields unbiased ET estimation.

Fig. 7 displays the statistics (R^2, RMSE, Bias and KGE) of the comparison among different merging methods (GLASS ET v5.0, GLASS ET v4.0, GBRT and RF) at 63 validation flux tower sites for different land cover types. For EBF, all merging models illustrate the worst performance with the lowest KGE (0.52–0.64) and R^2 (0.39–0.46, $p < 0.01$), and the largest RMSE (27.1–30.2 W/m^2). In contrast, all merging methods for SHR have the highest LE accuracy with the highest KGE (0.81–0.85) and R^2 (0.72–0.79, $p < 0.01$) and the lowest RMSE (12.4–15.5 W/m^2). In addition, GLASS ET algorithm v5.0, along with other merging methods, also demonstrates satisfactory LE performance [KGE of 0.69–0.79, R^2 of 0.49–0.78 ($p < 0.01$) and RMSE of 20.5–26.0 W/m^2] for the DBF and MIF land cover types. Overall, although the merging models might generate good LE estimations for different land cover types, GLASS ET algorithm v5.0 demonstrates the best performance with the largest KGE (0.57–0.85) and R^2 (0.46–0.79, $p < 0.01$)

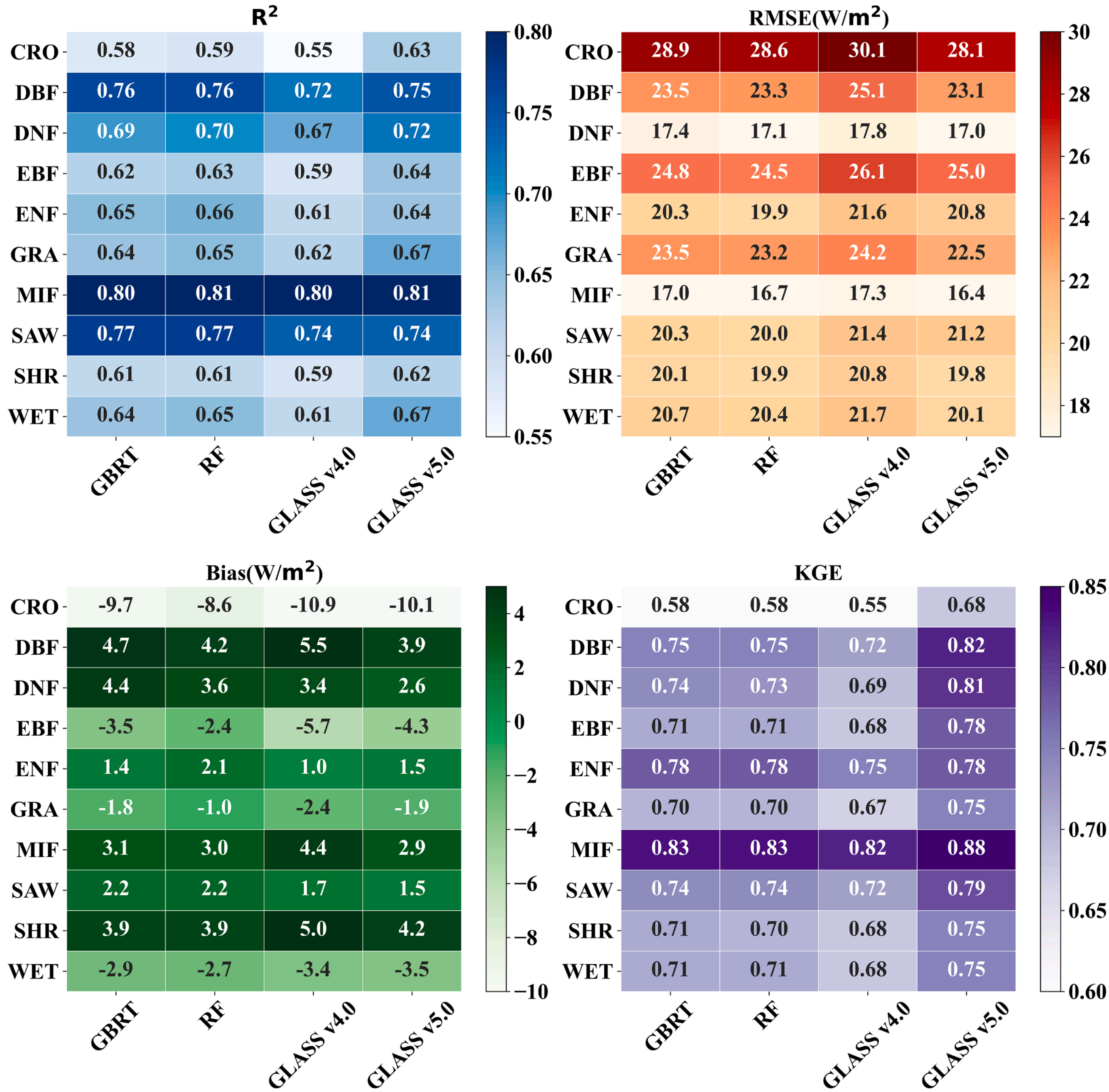

Fig. 5. Diagrams of the statistics (R^2, RMSE, Bias and KGE) of the comparison between different merging methods (GBRT, RF, GLASS ET v4.0 and GLASS ET v5.0) at 132 training flux tower sites for different land cover types.

and with the smallest RMSE (12.4–27.2 W/m^2) and bias (−3.2 to 11.6 W/m^2), respectively. Compared with GLASS ET algorithm v4.0, the performance of GLASS ET algorithm v5.0 shows a large improvement among all land cover types. The most obvious improvement occurs in the DBF and WET land cover types; the KGE values are enhanced over 5%, and the R^2 values are enhanced more than 12%. Similarly, for CRO and EBF, the GLASS ET algorithm v5.0 enhanced the R^2 values more than 11%.

Fig. 8 shows a time series for 8-day EC observations and the estimated ET from multiple algorithms (or products) for ten typical land cover types. In comparison to the individual LE products, GLASS v5.0-derived ET estimates capture more accurate seasonal ET variations that were closest to the ET observations for multiple land cover types. Overall, the error of the estimated ET based on GLASS is smaller than the error from the individual ET products. Therefore, the GLASS ET algorithm v5.0 can effectively capture the ET seasonal variance and is reliable for acquiring long-term ET products.

4.2. Global summary and spatial pattern

We generated GLASS ET product v5.0 with a 1 km spatial resolution and 8-day temporal resolution. Fig. 9 displays the spatial pattern of average annual terrestrial ET (2001–2015) for GLASS ET product v5.0, GLASS ET product v4.0, MOD16, SW, PT-JPL, MS-PT and SIM ET product over the globe. For GLASS ET v5.0, the highest ET occurs in equatorial regions, typically including the rainforests of Amazon, Ituri and Harapan, and the lowest ET estimates occur in high-latitude regions. Compared with the other six ET products, GLASS ET v5.0 has a similar spatial pattern during 2001–2015. GLASS ET v5.0 yields a global average annual terrestrial ET (excluding Antarctica) of 41.2 W/m^2,

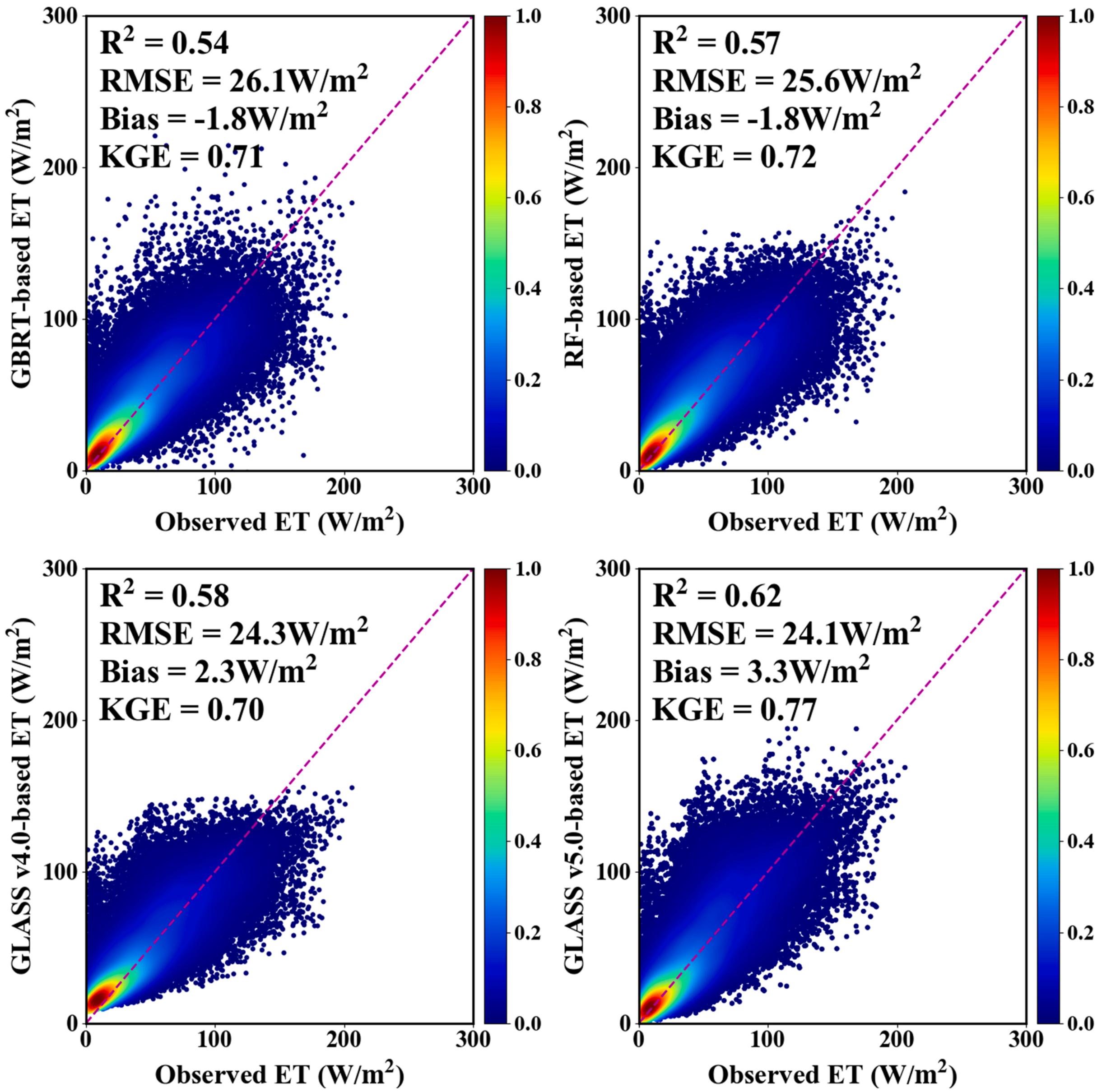

Fig. 6. Comparison of the estimated ET from different merging methods (GBRT, RF, GLASS ET algorithm v4.0 and GLASS ET algorithm v5.0) and ground-observed ET at 63 validation flux tower sites for all land cover types.

which is lower than the ET values (42.1 W/m^2) of GLASS ET v4.0, but it falls within the range of 35–41 W/m^2 documented by previous studies (Ma et al., 2021; Wang and Dickinson, 2012).

Fig. 10 shows the spatial difference in the average annual global terrestrial ET (2001–2015) between GLASS ET v5.0 and the other six ET products. GLASS ET v5.0 yields significantly lower ET values than GLASS ET v4.0 around tropical areas and higher ET values in North Africa and a portion of Australia. In these regions, the differences in ET estimates are within $\pm$8 W/m^2. In general, GLASS ET v5.0 yields small spatial differences when compared with PT-JPL and SIM. However, relative to SW, GLASS ET v5.0 yields a large spatial difference in South America and South Africa. When compared with MS-PT, GLASS ET v5.0 has lower ET values for most regions, including North Africa, the Arabian Peninsula and northwestern Australia. This spatial dissimilarity is possibly caused by differences in the structures of different ET models.

5. Discussion

5.1. The performance of GLASS ET v5.0

5.1.1. The ability of GLASS ET v5.0 to estimate ET

We applied the DNN to produce GLASS ET product v5.0 by merging five individual satellite products and found that the DNN preserves spatiotemporal consistency and produces reliable and robust estimation for most cover types. Independent validations for 63 EC flux tower sites demonstrated that GLASS ET v5.0 showed the best performance with the highest KGE and R^2 and the lowest RMSE and bias when compared with the other merging methods for most land cover types (Fig. 6 and Fig. 7). We found that the GLASS ET v5.0 method showed large differences among biomes and performed better for DBF [KGE of 0.79, R^2 of 0.78 ($p < 0.01$) and RMSE of 20.9 W/m^2] and SHR [KGE of 0.85, R^2 of 0.79 ($p < 0.01$) and RMSE of 12.4 W/m^2] flux tower sites when compared with the other land cover types. This might be explained by the fact that GLASS ET v5.0 uses vegetation indices or LAI to successfully capture the seasonal ET variations for these vegetation types (Ershadi et al., 2014)

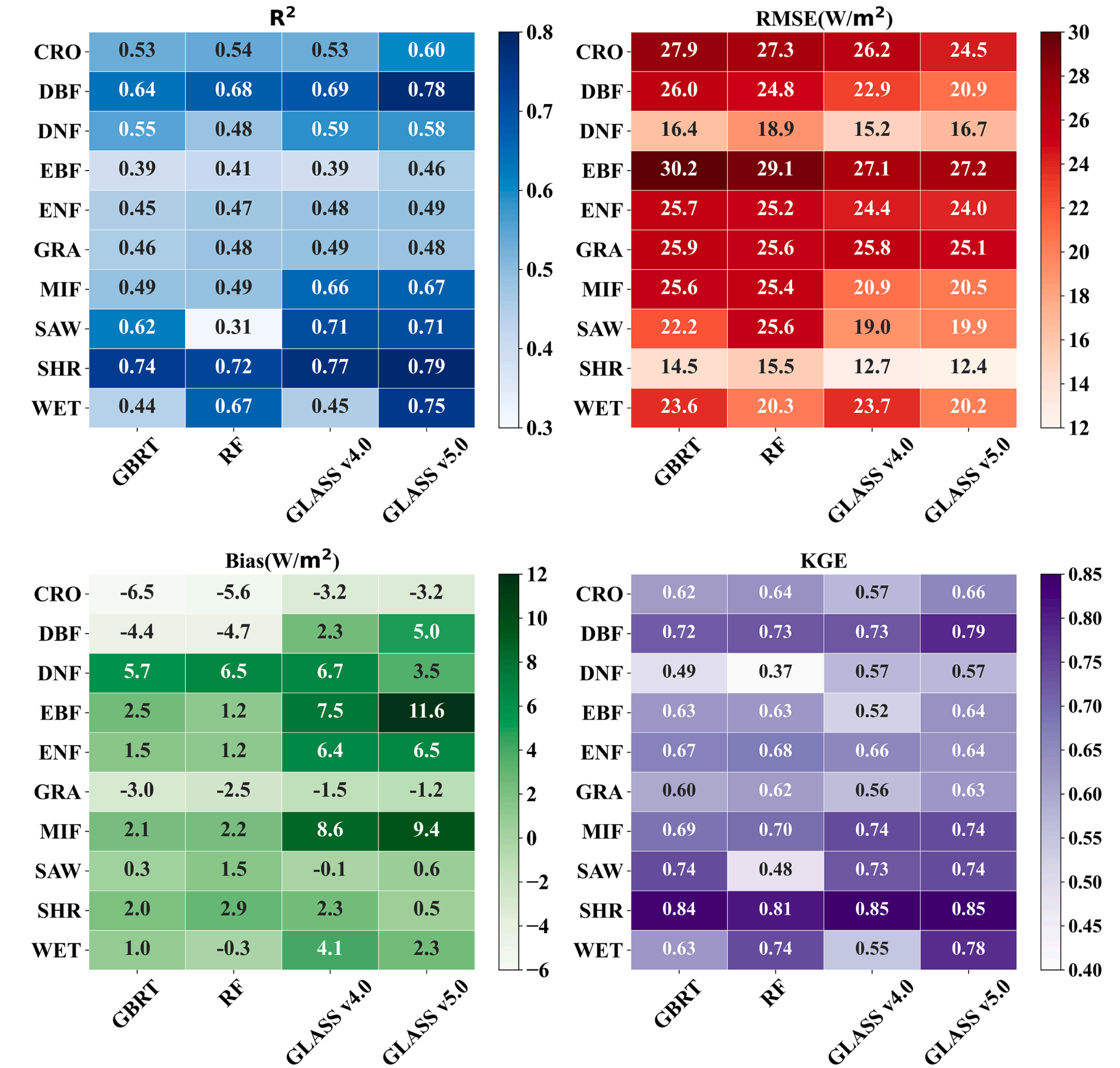

Fig. 7. Diagrams of the statistics (R^2, RMSE, Bias and KGE) of the comparison between different merging methods (GBRT, RF GLASS ET v4.0 and GLASS ET v5.0) at 63 validation flux tower sites for different land cover types.

because the strong seasonality of satellite vegetation variables (NDVI, FPAR and LAI) could reflect accurate information on seasonal changes in vegetation (Yebra et al., 2013). Several previous studies have documented that some satellite-based ET algorithms, i.e., the surface energy balance system (SEBS), PT-JPL, and the beta version of MOD16, could generate more accurate ET estimates for vegetation with significant seasonal variation (e.g., DBF) (Mu et al., 2007; Vinukollu et al., 2011a; Vinukollu et al., 2011b). Therefore, by merging these individual ET products that can provide a reasonable seasonality of ET variability, GLASS ET v5.0 improves the accuracy of ET estimation. In contrast, for some evergreen forests, such as the EBF [KGE of 0.64, R^2 of 0.46 ($p <$ 0.01) and RMSE of 27.2 W/m^2] and ENF [KGE of 0.64, R^2 of 0.49 ($p <$ 0.01) and RMSE of 24.0 W/m^2], GLASS ET v5.0 yields a relatively poorer estimate. This may be attributable in part to the fact that seasonal EBF variation is less evident (Yebra et al., 2013). GLASS ET v5.0 tends to underestimate ET for CRO and GRA flux tower sites while it tends to overestimate ET for forest flux tower sites. This might be mainly caused by the limitations of the individual ET products. For example, the MS-PT ET product was produced by a simplified PT method that tends to underestimate ET for CRO and GRA flux tower sites (Yao et al., 2013), which may make a large contribution to the uncertainties of DNN-based estimations. Despite its underestimation or overestimation, GLASS ET v5.0 has relatively higher accuracy compared with other individual products.

A series of validations using EC observations also confirmed the improvement of GLASS ET v5.0 over GLASS ET v4.0. Generally, ET underestimation for high values was found for GLASS ET v4.0 but not for GLASS ET v5.0 (Fig. 6 and Fig. 7). This might partially be attributed to the fact that the DNN can fit nonlinear input data well owing to multilayer learning, while BMA cannot (Duan et al., 2007). Moreover, the

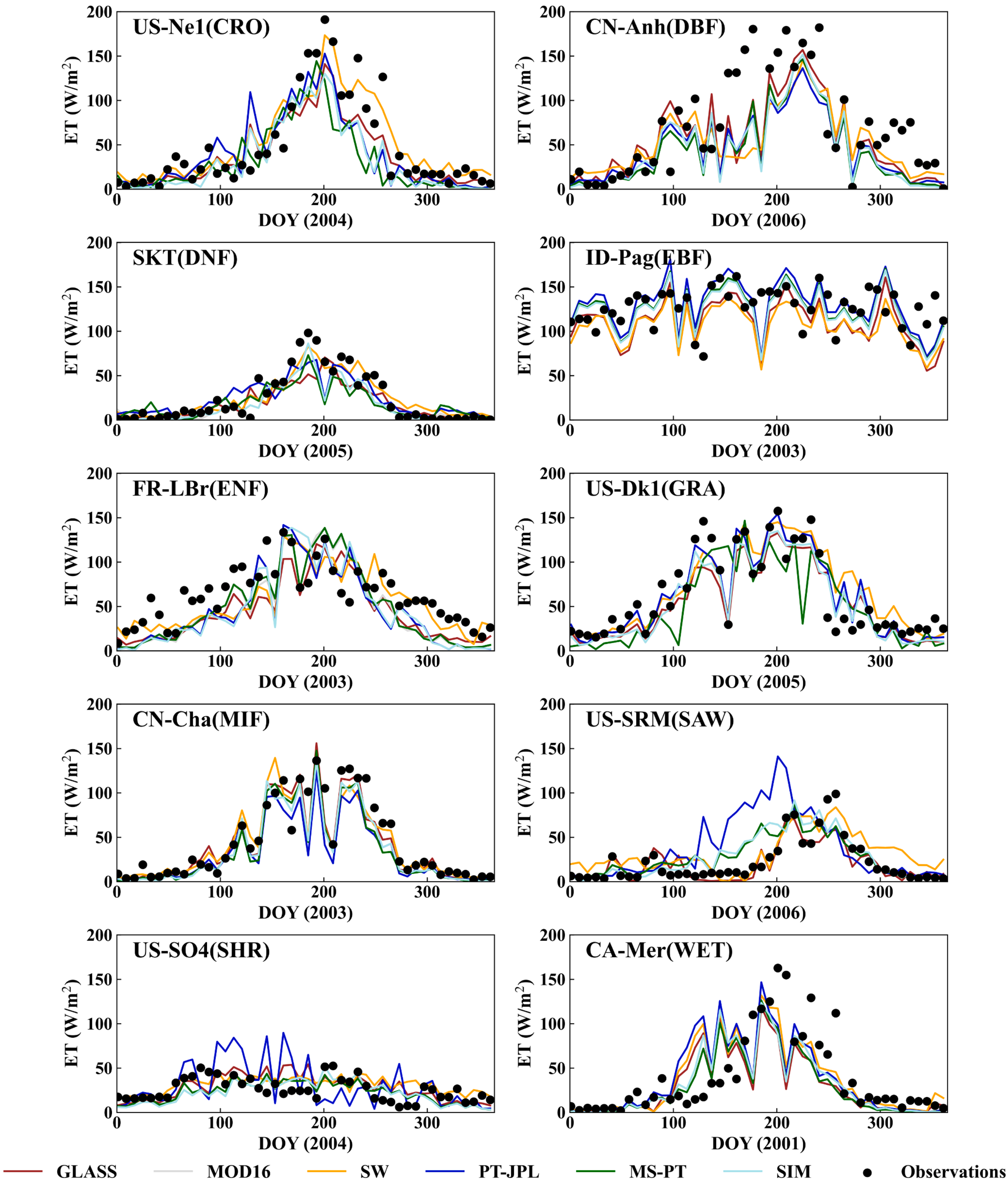

Fig. 8. Examples of the 8-day ET average as observed and estimated using the different models for the different land cover types.

DNN is superior to the BMA for transforming input data and reaching a depth of tens/hundreds of multiple layers, saving much time and effort (Rumelhart et al., 1986; Yoshua, 2009).

To investigate the difference between the merged ET and the upscaled ET from the flux tower site to the global scale, we also estimated global terrestrial ET using the DNN method driven by the forcing data of the five individual products as input variables. To compare the results of upscaling and fusion, we used forcing data from the same group of sites to upscale ET. Fig. 10 shows that the same satellite-derived data sources and ground observations and integration approach provide overall comparable performance in ET estimation with upscaling ET (KGE of 0.77 and 0.75 for integration and upscaling of ET, respectively). This is supported by a study by Shang et al. (2021) who compared the results of a DNN trained model and upscaling using forcing data of five

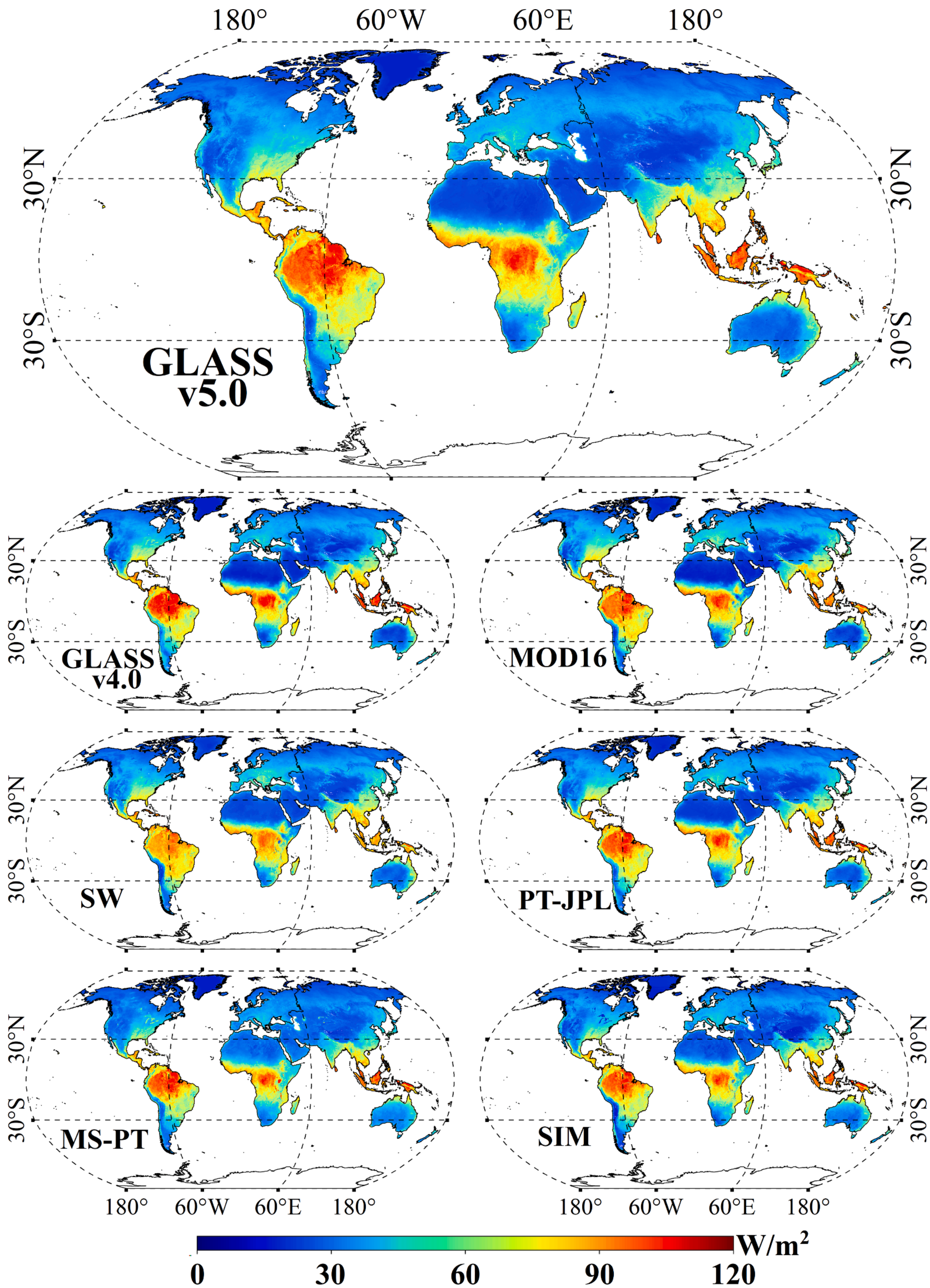

Fig. 9. Spatial pattern of average annual terrestrial ET during 2001–2015 for different ET products.

products over the Heihe River Basin in China and found that merging ET produced consistent results with upscaling at the regional scale.

However, the integration approach shows significant improvement in estimation accuracy in the extreme cases. Specifically, the KGE of the merged ET in the case of NDVI ≤ 0.15 arrives at 0.50, while the KGE of the upscaled ET is 0.40 (Fig. 11). This may be explained by the fact that the ET products used in the merging procedure are constrained by process-based models, which are capable of generating more reliable results under extreme circumstances, while the upscaling approach is confined to the training sample and fails to deal with extreme data and hence introduces uncertainty (Chen et al., 2014; Zhao et al., 2019). Therefore, the integration approach we chose can preserve model physics to maintain strong extrapolation capacity in extreme cases and thereby outperform the upscaling approach.

5.1.2. The uncertainties of GLASS ET v5.0

The uncertainties of GLASS ET v5.0 mainly include the errors of the individual ET products, the bias of EC observations, the spatial scale mismatch between flux tower sites and satellite pixels, and the algorithm structure of the DNN (Fisher et al., 2020; Mu et al., 2011;

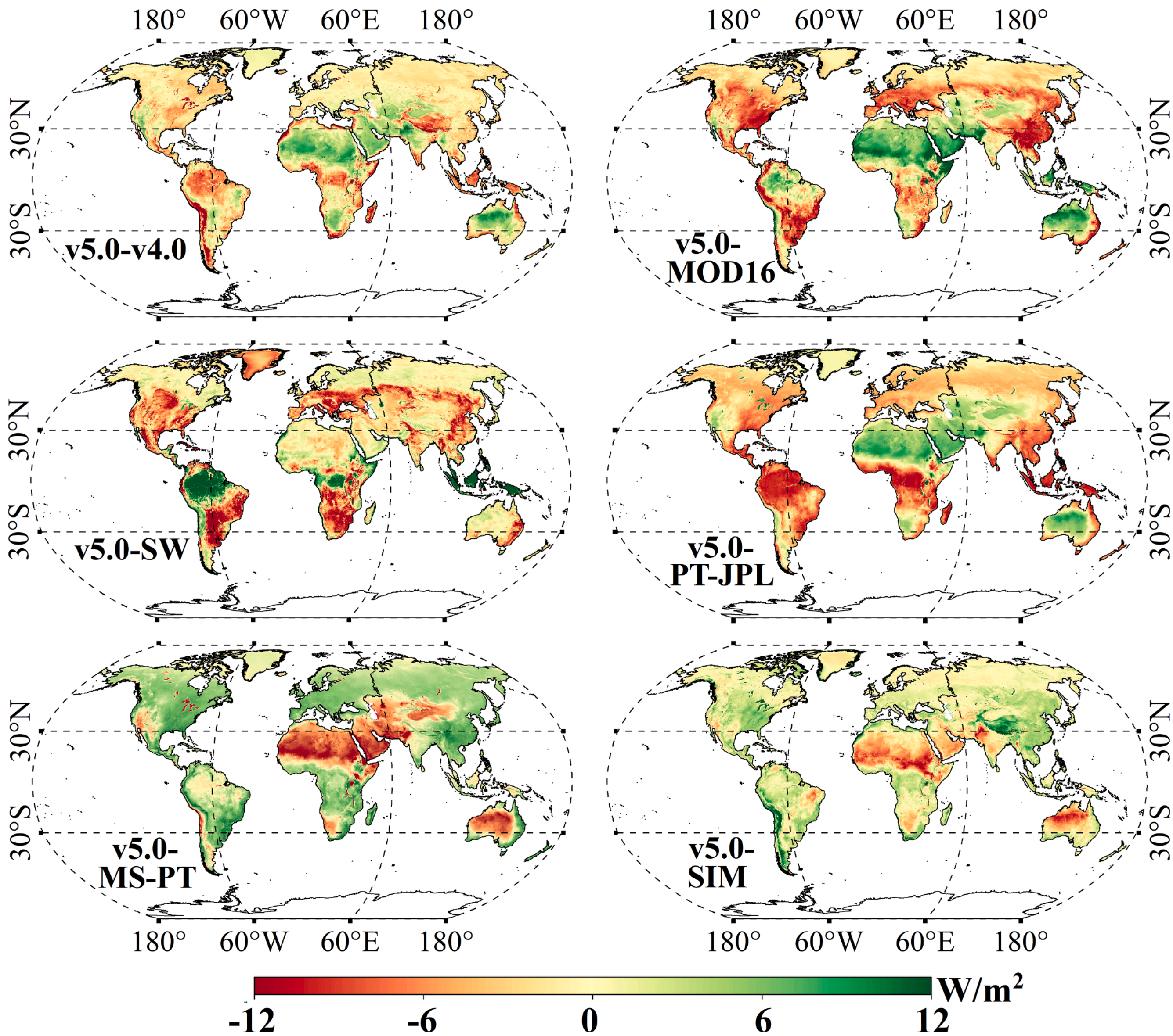

Fig. 10. Spatial differences in the average annual global terrestrial ET (2001–2015) between GLASS ET v5.0 and the other six ET products.

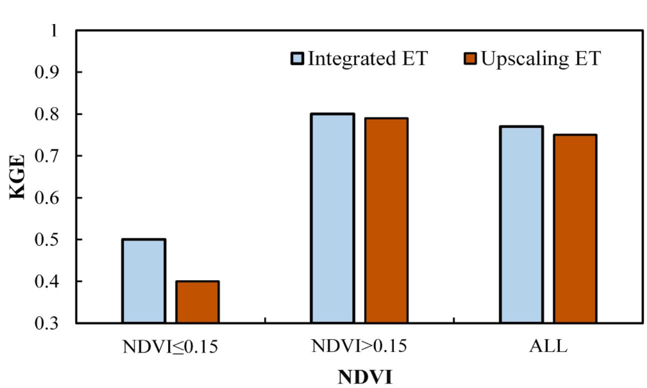

Fig. 11. Comparison of KGE of the upscaled LE and integration of LE estimates by the DNN method under different NDVI cases.

Polhamus et al., 2013). The errors of the individual ET products are mainly inherited through the errors in forcing inputs (MERRA2 and satellite-derived vegetation variables) and the ET model structures. Previous studies showed that MERRA2 meteorological data also contain a large bias when compared to in-situ measurements, and no single reanalysis dataset is superior to others in terms of meteorological variables to estimate land surface energy budgets (Badgley et al., 2015; Ferguson et al., 2010). We used MERRA2 products with spatial resolution of 1/2° X 2/3° that were resampled spatially into 1 km using the bilinear interpolation method. However, the resolution of MRERRA2 is greater than footprint for field measurements (Baldocchi, 2008). Thus, accurate meteorological information for flux tower sites cannot be acquired due to their coarse spatial resolution and errors in the bilinear interpolation method (Zhang et al., 2010). Additionally, there also exist approximately 15–30% errors in the vegetation parameters (e.g., LAI, NDVI, or FPAR) retrieved from satellite-based observations (Ganguly et al., 2012; Kalma et al., 2008). Moreover, different ET model structures lead to a 20% bias in ET estimation (Polhamus et al., 2013). Therefore, the errors of inputs and model structures for the individual ET products contribute to large uncertainties in GLASS ET v5.0. Although EC observations are relatively accurate for ET acquisition, biases of approximately 5%–25% still exist (Foken, 2008). Importantly, the EC method suffers from an energy imbalance, and Foken (2008) reported that the EC method cannot observe large eddies, which will cause $H + LE \neq R_n -$

G. Although we corrected the energy imbalance using a method developed by Twine et al. (2000), the uncertainties in EC observations remain unclear (Wang and Dickinson, 2012; Wilson et al., 2002). Additionally, the gap filling from half-hour intervals to daily periods will also introduce a 5% bias into the daily ET values (Barcza et al., 2009; Falge et al., 2001).

The spatial mismatch between the flux tower footprint and satellite pixels could also introduce large uncertainties in the merged ET. The footprint of the flux tower site is several hundred meters (Barcza et al., 2009), and the spatial resolution of GLASS pixels is 1 km. Inaccurate representations of the EC footprint may lead to large errors for merging ET (Wang et al., 2019). The DNN generally depends on the representation of training data. The lack of a sufficient amount of labeled data could lead to the inferior performance of the DNN (Bengio et al., 2013). Therefore, the uncertainties could be inherited through errors from training samples. Furthermore, DNN requires many hyperparameters as input and careful tuning to deliver a favorable learning performance (Yann et al., 2015). On the other hand, too many interfering factors could definitely introduce great uncertainties into the model construction process (Chollet, 2017).

5.2. Comparison with the MOD16A2 and FLUXCOM ET products

To evaluate the accuracy and spatial consistency of GLASS ET product v5.0 (8-day, 1 km) for different land cover types, we compared GLASS ET product v5.0 with the 8-day MOD16A2 ET product that was resampled into 1 km from 500 m using the bilinear method and 8-day FLUXCOM ET product resampled into 1 km from 0.0833° using the bilinear method. Fig. 12 shows the validation results of GLASS ET product v5.0, MOD16A2 ET product and FLUXCOM ET product based on EC observations at 63 validation flux tower sites. It is clear that GLASS ET v5.0 is superior to MOD16A2 with increasing R^2 by 23% ($p < 0.01$), and KGE by 19% and reducing RMSE by 17%. This supports the conclusion that the ET estimated by merging multiple ET products using machine learning outperforms the ET estimated using the individual ET model (Shang et al., 2021). MOD16A2 was produced by the improved Penman–Monteith equation, which includes a complicated ET process. Thus, the errors in input data, including MODIS and MERRA2 reanalysis data, and the error transmission through the Penman–Monteith equation could reduce the accuracy of MOD16A2 (Ruhoff et al., 2013). For example, the classification accuracy of the MODIS land cover type product (MCD12Q1) is approximately 75% (Myneni et al., 1997; Wang et al., 2004), and misclassification of MCD12Q1 could also cause the large bias of MOD16A2. Additionally, the same biophysical parameters for different phonologies in the MOD16 algorithm might ignore the actual ET variations and reduce the accuracy of MOD16A2 (Turner et al., 2003). Additionally, GLASS ET v5.0 outperforms FLUXCOM with increasing R^2 by 6% ($p < 0.01$), and KGE by 7% and reducing RMSE by 2%. Validation results illustrated that the machine learning-based ET products (e.g., GLASS ET v5.0 and FLUXCOM ET) are superior to the process-based ET products (e.g., MOD16) at site scale when the training sample is sufficiently representative, which may be attributed to the fact that both GLASS and FLUXCOM ET products make full use of prior knowledge of training samples partially inherited from EC observations (Shang et al., 2021). Moreover, GLASS ET v5.0 outperforms FLUXCOM in all dimensions because GLASS ET v5.0 not only preserves model physics to effectively simulate ET under extreme cases but has a strong DNN structure.

We also found that the MOD16A2 ET product yields ET values below 45 W/m^2 for most regions of the world (Fig. 9). GLASS ET v5.0 uses the DNN to yield reliable global terrestrial ET, which are approximately consistent with those of MOD16A2 and FLUXCOM with acceptable differences within ±20 W/m^2 for most regions (Fig. 13). However, relative to MOD16A2, GLASS ET v5.0 underestimates ET in northern South America and South Asia and overestimates ET in Africa and Australia. Previous studies also revealed that MOD16A2 generally underestimated ET in arid and sparsely vegetated areas (Brust et al., 2021; Khan et al., 2018; Marshall et al., 2020). Compared to FLUXCOM, GLASS ET v5.0 underestimates ET in Central and South Africa and part of South America and overestimates ET in North Africa, West Asia and Australia. Previous studies reported FLUXCOM estimated with higher evaporation levels than GLDAS especially in Africa, while GLDAS performed stably in Africa (Staal et al., 2020). This may indicate that GLASS ET v5.0 has a more reliable spatial distribution of ET than MOD16A2 and FLUXCOM.

5.3. Advantages and limitations of GLASS ET algorithm v5.0

Compared with GLASS ET algorithm v4.0 and the individual ET products, GLASS ET algorithm v5.0 (DNN) has two distinct advantages. On the one hand, GLASS ET algorithm v5.0 improves global terrestrial ET estimation by merging five individual ET products and EC observations because it can accurately approximate the complicated nonlinear relationship between observed EC and the individual ET estimates using multilayer learning, which helps capture the potential association between different variables for spatiotemporal fusion of remote sensing data (Ngiam et al., 2011; Yuan et al., 2020). Several studies have also shown that merging algorithms can avoid intrinsic errors generated by a single model structure caused by overconfidence and significant bias, and applying machine learning methods to integrate multiple models can lead to higher ET estimate accuracy (Parrish et al., 2012; Zhu et al., 2016). On the other hand, GLASS ET algorithm v5.0 preserves some physical mechanisms of individual algorithms for producing the individual ET products and can better simulate ET under extreme conditions than pure machine learning. Substantial studies have reported that

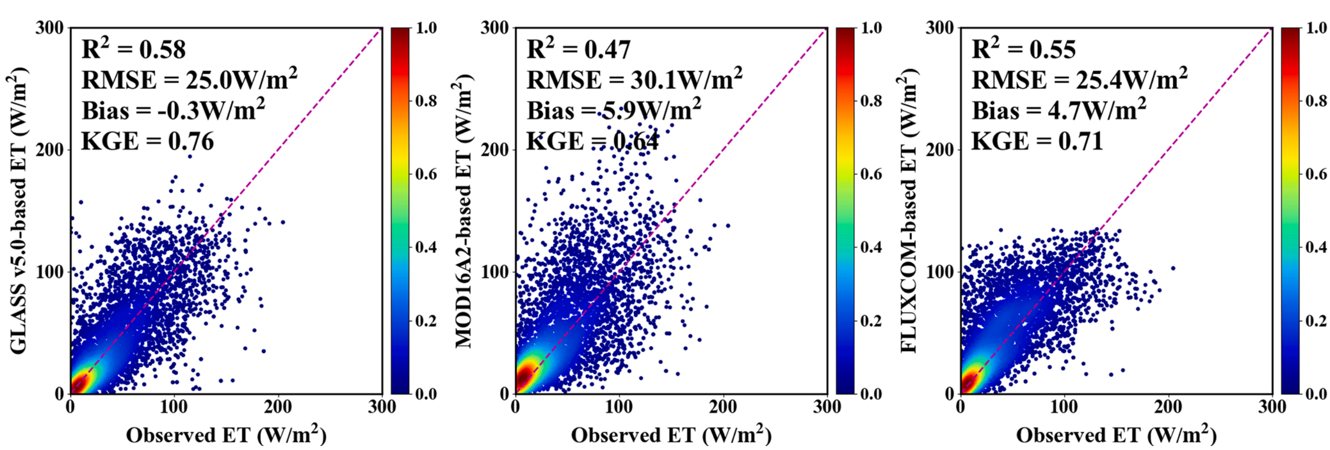

Fig. 12. Validation of GLASS ET product v5.0, MOD16A2 ET product and FLUXCOM ET product based on EC observations at 63 validation flux tower sites.

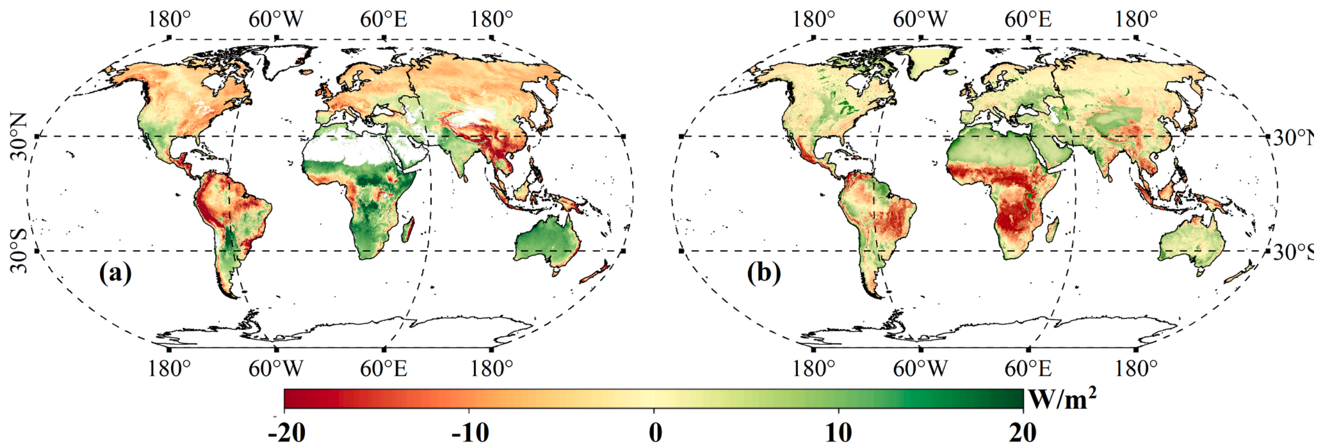

Fig. 13. Map of the spatial difference in the annual average (a) between GLASS ET product v5.0 and MOD16A2 ET product, and (b) between GLASS ET product and FLUXCOM ET product over the globe from 2001 to 2015.

fusion models can merge the information provided by different sources to improve algorithm performance as well as provide a better physical interpretation of the results to enhance the understanding of ET processes (Hu et al., 2021; Salcedo-Sanz et al., 2020).

Similar to other machine learning methods, GLASS ET algorithm v5.0 also has two limitations. First, the DNN has a complex neural network model that requires a long time (~10 s for 1000 samples) to complete the generation of global ET products (Chollet, 2017). In general, DNN always requires many hyperparameters, with the performance of DNN depending largely on their parameter tuning, which will reduce the efficiency of the generation of ET products (Hinton and Salakhutdinov, 2006). Second, the DNN relies on sufficient training samples adequate forcing inputs to enhance model performance (Hinton et al., 2012). If the samples are not sufficient for some land cover types, the DNN could result in large biases of ET estimation. Furthermore, as different sources of ET products introduce different meteorological forcing into the model, the possible inclusion of unnecessary interfering factors and almost infinite configurational combinations might lead to the model construction process more uncertain (Zhou and Feng, 2017).

6. Conclusions

We introduced a DNN-merging framework to produce GLASS ET v5.0 by merging five satellite-derived ET products and ground observations at 195 globally distributed EC sites from 2000 through 2015. Compared with the existing GLASS ET v4.0, GLASS ET v5.0 has relatively lower uncertainty and spatial consistency. The main results of this study are summarized as follows:

(1) GLASS ET algorithm v5.0 outperformed other merging models, including GLASS ET algorithm v4.0 (BMA method), and two machine learning methods (RF and GBRT), which was superior to five individual satellite-derived ET products.

(2) The EC validation results illustrated that GLASS ET product v5.0 outperformed the MOD16A2 and FLUXCOM ET products and demonstrated improvement in spatiotemporal prediction accuracy.

(3) Compared with the upscaling of ET from site to global, the integration approach not only has comparable accuracy but also has enhanced performance for extreme cases because it can preserve model physics and maintain extrapolation capacity in extreme conditions.

CRediT authorship contribution statement

Zijing Xie: Software, Data curation, Validation, Writing – original draft. **Yunjun Yao:** Resources, Conceptualization, Data curation, Funding acquisition, Writing – review & editing. **Xiaotong Zhang:** Resources, Conceptualization, Data curation, Funding acquisition, Writing – review & editing. **Shunlin Liang:** Project administration, Investigation, Writing – review & editing. **Joshua B. Fisher:** Formal analysis, Writing – review & editing. **Jiquan Chen:** Formal analysis, Writing – review & editing. **Kun Jia:** Formal analysis, Writing – review & editing. **Ke Shang:** Software, Methodology, Investigation. **Junming Yang:** Software, Methodology, Investigation. **Ruiyang Yu:** Software, Methodology, Investigation. **Xiaozheng Guo:** Software, Methodology, Investigation. **Lu Liu:** Software, Investigation, Visualization. **Jing Ning:** Software, Investigation, Visualization. **Lilin Zhang:** Software, Investigation, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Natural Science Fund of China (No. 42171310 and No. 42192581). Acknowledgement for the data support from “National Earth System Science Data Center, National Science & Technology Infrastructure of China. (http://www.geodata.cn)”. This work used eddy covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (U.S. Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont,ChinaFlux, Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia, USCCC. We acknowledge the financial support to the eddy covariance data harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval, Environment Canada and US Department of Energy and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California – Berkeley and the University of Virginia. Other ground-measured data were obtained from the GAME AAN (http://aan.suiri.tsukuba.ac.jp/), the Coordinated Enhanced Observation Project (CEOP) in arid and semi-arid regions of northern China (http://observation.tea.ac.cn/), and the water experiments of Environmental and Ecological Science Data Center for West China (http://westdc.westgis.ac.cn/water).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2022.127990.

References

Aires, F., 2014. Combining datasets of satellite-retrieved products. Part I: Methodology and water budget closure. J. Hydrometeorol. 15 (4), 1677–1691.
Badgley, G., Fisher, J.B., Jiménez, C., Tu, K.P., Vinukollu, R., 2015. On uncertainty in global terrestrial evapotranspiration estimates from choice of input forcing datasets. J. Hydrometeorol. 16 (4), 1449–1455.
Bai, Y., Zhang, S., Bhattarai, N., Mallick, K., Liu, Q.i., Tang, L., Im, J., Guo, L.i., Zhang, J., 2021. On the use of machine learning based ensemble approaches to improve evapotranspiration estimates from croplands across a wide environmental gradient. Agric. For. Meteorol. 298–299, 108308.
Baldocchi, D., 2008. 'Breathing' of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust. J. Bot. 56 (1), 1–26.
Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C.h., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82 (11), 2415–2434.
Ball, J.E., Anderson, D.T., Chan, C.S., 2017. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. J. Appl. Remote. Sens. 11 (4), 042609.
Barcza, Z., Kern, A., Haszpra, L., Kljun, N., 2009. Spatial representativeness of tall tower eddy covariance measurements using remote sensing and footprint analysis. Agric. For. Meteorol. 149 (5), 795–807.
Bengio, Y., Courville, A., Vincent, P., 2013. Representation learning: a review and new perspectives. IEEE Trans. Pattern. Anal. Mach. Intell. 35 (8), 1798–1828.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Brust, C., Kimball, J.S., Maneta, M.P., Jencso, K., He, M., Reichle, R.H., 2021. Using SMAP Level-4 soil moisture to constrain MOD16 evapotranspiration over the contiguous USA. Remote Sens. Environ. 255, 112277.
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., Yuan, W., 2014. Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China. Remote Sens. Environ. 140, 279–293.
Chen, Y., Yuan, W., Xia, J., Fisher, J.B., Dong, W., Zhang, X., Liang, S., Ye, A., Cai, W., Feng, J., 2015. Using Bayesian model averaging to estimate terrestrial evapotranspiration in China. J. Hydrol. 528, 537–549.
Chollet, F.O., 2017. Deep Learning with Python. Manning Publications.
Cleugh, H.A., Leuning, R., Mu, Q., Running, S.W., 2007. Regional evaporation estimates from flux tower and MODIS satellite data. Remote Sens. Environ. 106 (3), 285–304.
Condon, L.E., Atchley, A.L., Maxwell, R.M., 2020. Evapotranspiration depletes groundwater under warming over the contiguous United States. Nat. Commun. 11 (1), 1–8.
Cui, Y., Song, L., Fan, W., 2021. Generation of spatio-temporally continuous evapotranspiration and its components by coupling a two-source energy balance model and a deep neural network over the Heihe River Basin. J. Hydrol. 597, 126176.
Detto, M., Montaldo, N., Albertson, J.D., Mancini, M., Katul, G., 2006. Soil moisture and vegetation controls on evapotranspiration in a heterogeneous Mediterranean ecosystem on Sardinia. Italy. Water Resour. Res. 42 (8).
Duan, Q., Ajami, N.K., Gao, X., Sorooshian, S., 2007. Multi-model ensemble hydrologic prediction using Bayesian model averaging. Adv. Water. Resour. 30 (5), 1371–1386.
Elbeltagi, A., Deng, J., Wang, K., Malik, A., Maroufpoor, S., 2020. Modeling long-term dynamics of crop evapotranspiration using deep learning in a semi-arid environment. Agric. Water. Manag. 241, 106334.
Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77 (4), 802–813.
Ershadi, A., Mccabe, M.F., Evans, J.P., Chaney, N.W., Wood, E.F., 2014. Multi-site evaluation of terrestrial evaporation models using FLUXNET data. Agric. For. Meteorol. 187, 46–61.
Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Ta Lai, C., Law, B. E., Meyers, T., Moncrieff, J., Moors, E., William Munger, J., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Gap filling strategies for long term energy flux data sets. Agric. For. Meteorol. 107 (1), 71–77.
Ferguson, C.R., Sheffield, J., Wood, E.F., Gao, H., 2010. Quantifying uncertainty in a remote sensing-based estimate of evapotranspiration over continental USA. Int. J. Remote Sens. 31 (14), 3821–3865.
Fisher, J.B., Lee, B., Purdy, A.J., Halverson, G.H., Dohlen, M.B., Cawse-Nicholson, K., Wang, A., Anderson, R.G., Aragon, B., Arain, M.A., Baldocchi, D.D., Baker, J.M., Barral, H., Bernacchi, C.J., Bernhofer, C., Biraud, S.C., Bohrer, G., Brunsell, N., Cappelaere, B., Castro-Contreras, S., Chun, J., Conrad, B.J., Cremonese, E., Demarty, J., Desai, A.R., De Ligne, A., Foltýnová, L., Goulden, M.L., Griffis, T.J., Grünwald, T., Johnson, M.S., Kang, M., Kelbe, D., Kowalska, N., Lim, J.-H., Maïnassara, I., McCabe, M.F., Missik, J.E.C., Mohanty, B.P., Moore, C.E., Morillas, L., Morrison, R., Munger, J.W., Posse, G., Richardson, A.D., Russell, E.S., Ryu, Y., Sanchez-Azofeifa, A., Schmidt, M., Schwartz, E., Sharp, I., Šigut, L., Tang, Y., Hulley, G., Anderson, M., Hain, C., French, A., Wood, E., Hook, S., 2020. ECOSTRESS: NASA's next generation mission to measure evapotranspiration from the international space station. Water Resour. Res. 56 (4).
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., Wood, E.F., 2017. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Water Resour. Res. 53 (4), 2618–2626.
Fisher, J.B., Tu, K.P., Baldocchi, D.D., 2008. Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. Remote Sens. Environ. 112 (3), 901–919.
Foken, T., 2008. The energy balance closure problem: an overview. Ecol. Appl. 18 (6), 1351–1367.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.
Ganguly, S., Nemani, R.R., Zhang, G., Hashimoto, H., Milesi, C., Michaelis, A., Wang, W., Votava, P., Samanta, A., Melton, F., Dungan, J.L., Vermote, E., Gao, F., Knyazikhin, Y., Myneni, R.B., 2012. Generating global Leaf Area Index from Landsat: Algorithm formulation and demonstration. Remote Sens. Environ. 122, 185–202.
Gleeson, T., Wang-Erlandsson, L., Porkka, M., Zipper, S.C., Jaramillo, F., Gerten, D., Fetzer, I., Cornell, S.E., Piemontese, L., Gordon, L.J., Rockstrom, J., Oki, T., Sivapalan, M., Wada, Y., Brauman, K.A., Florke, M., Bierkens, M.F.P., Lehner, B., Keys, P., Kummu, M., Wagener, T., Dadson, S., Troy, T.J., Steffen, W., Falkenmark, M., Famiglietti, J.S., 2020. Illuminating water cycle modifications and Earth system resilience in the Anthropocene. Water Resour. Res. 56(4), e2019WR024957.
Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91.
Helbig, M., Waddington, J.M., Alekseychik, P., Amiro, B.D., Aurela, M., Barr, A.G., Black, T.A., Blanken, P.D., Carey, S.K., Chen, J., Chi, J., Desai, A.R., Dunn, A., Euskirchen, E.S., Flanagan, L.B., Forbrich, I., Friborg, T., Grelle, A., Harder, S., Heliasz, M., Humphreys, E.R., Ikawa, H., Isabelle, P.-E., Iwata, H., Jassal, R., Korkiakoski, M., Kurbatova, J., Kutzbach, L., Lindroth, A., Löfvenius, M.O., Lohila, A., Mammarella, I., Marsh, P., Maximov, T., Melton, J.R., Moore, P.A., Nadeau, D.F., Nicholls, E.M., Nilsson, M.B., Ohta, T., Peichl, M., Petrone, R.M., Petrov, R., Prokushkin, A., Quinton, W.L., Reed, D.E., Roulet, N.T., Runkle, B.R.K., Sonnentag, O., Strachan, I.B., Taillardat, P., Tuittila, E.-S., Tuovinen, J.-P., Turner, J., Ueyama, M., Varlagin, A., Wilmking, M., Wofsy, S.C., Zyrianov, V., 2020. Increasing contribution of peatlands to boreal evapotranspiration in a warming climate. Nat. Clim. Chang. 10 (6), 555–560.
Hinton, G., Deng, L.i., Yu, D., Dahl, G., Mohamed, A.-R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T., Kingsbury, B., 2012. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process. Mag. 29 (6), 82–97.
Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Science 313 (5786), 504–507.
Hu, G.C., Jia, L., 2015. Monitoring of evapotranspiration in a semi-arid inland river basin by combining microwave and optical remote sensing observations. Remote Sens. 7 (3), 3056–3087.
Hu, X., Shi, L., Lin, G., Lin, L., 2021. Comparison of physical-based, data-driven and hybrid modeling approaches for evapotranspiration estimation. J. Hydrol. 601, 126592.
Jiang, B., Zhang, Y., Liang, S.L., Wohlfahrt, G., Arain, A., Cescatti, A., Georgiadis, T., Jia, K., Kiely, G., Lund, M., Montagnani, L., Magliulo, V., Ortiz, P.S., Oechel, W., Vaccari, F.P., Yao, Y.J., Zhang, X.T., 2015. Empirical estimation of daytime net radiation from shortwave radiation and ancillary information. Agric. For. Meteorol. 211, 23–36.
Jiang, C., Ryu, Y., 2016. Multi-scale evaluation of global gross primary productivity and evapotranspiration products derived from Breathing Earth System Simulator (BESS). Remote Sens. Environ. 186, 528–547.
Jimenez, C., Martens, B., Miralles, D.M., Fisher, J.B., Beck, H.E., Fernandez-Prieto, D., 2018. Exploring the merging of the global land evaporation WACMOS-ET products based on local tower measurements. Hydrol. Earth Syst. Sci. 22 (8), 4513–4533.
Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., Reichstein, M., 2019. The FLUXCOM ensemble of global land-atmosphere energy fluxes. Sci. Data 6 (1), 1–14.
Kalma, J.D., McVicar, T.R., McCabe, M.F., 2008. Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. Surv. Geophys. 29 (4), 421–469.
Khan, M.S., Liaqat, U.W., Baik, J., Choi, M., 2018. Stand-alone uncertainty characterization of GLEAM, GLDAS and MOD16 evapotranspiration products using an extended triple collocation approach. Agric. For. Meteorol. 252, 256–268.
Kotsiantis, S.B., 2013. Decision trees: a recent overview. Artif. Intell. Rev. 39 (4), 261–283.
Krishnan, P., Black, T.A., Grant, N.J., Barr, A.G., Hogg, E.T., Jassal, R.S., Morgenstern, K., 2006. Impact of changing soil moisture distribution on net ecosystem productivity of a boreal aspen forest during and following drought. Agric. For. Meteorol. 139 (3–4), 208–223.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems – Volume 1. Curran Associates Inc, Lake Tahoe, pp. 1097–1105.

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., Zhao, X., Zhou, J., 2021. The Global Land Surface Satellite (GLASS) product suite. Bull. Am. Meteorol. Soc. 102 (2), E323–E337.

Liu, Y.B., Qiu, G.Y., Zhang, H.S., Yang, Y.H., Zhang, Y.S., Wang, Q., Zhao, W.Z., Jia, L., Ji, X.B., Xiong, Y.J., Yan, C.H., Ma, N., Han, S.M., Cui, Y.F., 2022. Shifting from homogeneous to heterogeneous surfaces in estimating terrestrial evapotranspiration: Review and perspectives. Sci. China Earth Sci. 1–18.

Ma, N., Szilagyi, J., 2019. The CR of evaporation: a calibration-free diagnostic and benchmarking tool for large-scale terrestrial evapotranspiration modeling. Water Resour. Res. 55 (8), 7246–7274.

Ma, N., Szilagyi, J., Jozsa, J., 2020. Benchmarking large-scale evapotranspiration estimates: a perspective from a calibration-free complementary relationship approach and FLUXCOM. J. Hydrol. 590, 125221.

Ma, N., Szilagyi, J., Zhang, Y.Q., 2021. Calibration-free complementary relationship estimates terrestrial evapotranspiration globally. Water Resour. Res. 57(9), e2021WR029691.

Marshall, M., Tu, K., Andreo, V., 2020. On parameterizing soil evaporation in a direct remote sensing model of ET: PT-JPL. Water Resour. Res. 56(5), e2019WR026290.

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., Verhoest, N.E.C., 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev. 10 (5), 1903–1925.

Mu, Q., Heinsch, F.A., Zhao, M., Running, S.W., 2007. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sens. Environ. 111 (4), 519–536.

Mu, Q., Zhao, M., Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sens. Environ. 115 (8), 1781–1800.

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., Seneviratne, S.I., 2013. Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis. Hydrol. Earth Syst. Sci. 17 (10), 3707–3720.

Murphy, D.M., Solomon, S., Portmann, R.W., Rosenlof, K.H., Forster, P.M., Wong, T., 2009. An observationally based energy balance for the Earth since 1950. J. Geophys. Res. 114 (D17).

Myneni, R.B., Keeling, C.D., Tucker, C.J., Asrar, G., Nemani, R.R., 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. Nature 386 (6626), 698–702.

Ngiam, J., Khosla, A., Kim, M., Nam, J., Ng, A.Y., 2011. Multimodal Deep Learning. In ICML.

Oki, T., Kanae, S., 2006. Global hydrological cycles and world water resources. Science 313 (5790), 1068–1072.

Parrish, M.A., Moradkhani, H., DeChant, C.M., 2012. Toward reduction of model uncertainty: Integration of Bayesian model averaging and data assimilation. Water Resour. Res. 48 (3).

Pascolini-Campbell, M., Reager, J.T., Chandanpurkar, H.A., Rodell, M., 2021. A 10 per cent increase in global land evapotranspiration from 2003 to 2019. Nature 593 (7860), 543–547.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., et al., 2020. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci Data 7, 225.

Polhamus, A., Fisher, J.B., Tu, K.P., 2013. What controls the error structure in evapotranspiration models? Agric. For. Meteorol. 169, 12–24.

Priestley, C., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large scale parameters. Mon. Weather Rev. 100 (2), 81–92.

Raftery, A.E., Madigan, D., Hoeting, J.A., 1997. Bayesian model averaging for linear regression models. J. Am. Stat. Assoc. 92 (437), 179–191.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., Valentini, R., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Change Biol. 11 (9), 1424–1439.

Ruhoff, A.L., Paz, A.R., Aragao, L.E.O.C., Mu, Q., Malhi, Y., Collischonn, W., Rocha, H.R., Running, S.W., 2013. Assessment of the MODIS global evapotranspiration algorithm using eddy covariance measurements and hydrological modelling in the Rio Grande basin. Hydrol. Sci. J. 58 (8), 1658–1676.

Rumelhart, D., Hinton, G.E., Williams, R.J., 1986. Learning representations by back propagating errors. Nature 323 (6088), 533–536.

Saggi, M.K., Jain, S., 2019. Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. Comput. Electron. Agric. 156, 387–398.

Salcedo-Sanz, S., Ghamisi, P., Piles, M., Werner, M., Cuadra, L., Moreno-Martínez, A., Izquierdo-Verdiguier, E., Muñoz-Marí, J., Mosavi, A., Camps-Valls, G., 2020. Machine learning information fusion in Earth observation: a comprehensive review of methods, applications and data sources. Inf. Fusion 63, 256–272.

Shang, K., Yao, Y., Liang, S., Zhang, Y., Fisher, J.B., Chen, J., Liu, S., Xu, Z., Zhang, Y., Jia, K., Zhang, X., Yang, J., Bei, X., Guo, X., Yu, R., Xie, Z., Zhang, L., 2021. DNN-MET: A deep neural networks method to integrate satellite-derived evapotranspiration products, eddy covariance observations and ancillary information. Agric. For. Meteorol. 308-309, 108582.

Shuttleworth, W.J., Wallace, J.S., 1985. Evaporation from sparse crops-an energy combination theory. Q. J. R. Meteorol. Soc. 111 (469), 839–855.

Song, L., Liu, S., Kustas, W.P., Nieto, H., Sun, L., Xu, Z., Skaggs, T.H., Yang, Y., Ma, M., Xu, T., Tang, X., Li, Q., 2018. Monitoring and validating spatially and temporally continuous daily evaporation and transpiration at river basin scale. Remote Sens. Environ. 219, 72–88.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15 (1), 1929–1958.

Staal, A., Fetzer, I., Wang-Erlandsson, L., Bosmans, J.H.C., Dekker, S.C., van Nes, E.H., Rockstrom, J., Tuinenburg, O.A., 2020. Hysteresis of tropical forests in the 21st century. Nat. Commun. 11 (1), 1–8.

Trenberth, K.E., Fasullo, J.T., Kiehl, J., 2009. Earth's global energy budget. Bull. Am. Meteorol. Soc. 90 (3), 311–323.

Turner, D.P., Urbanski, S., Bremer, D., Wofsy, S.C., Meyers, T., Gower, S.T., Gregory, M., 2003. A cross-biome comparison of daily light use efficiency for gross primary production. Glob. Chang Biol. 9 (3), 383–395.

Twine, T.E., Kustas, W.P., Norman, J.M., Cook, D.R., Houser, P.R., Meyers, T.P., Prueger, J.H., Starks, P.J., Wesely, M.L., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agric. For. Meteorol. 103 (3), 279–300.

Vinukollu, R.K., Meynadier, R., Sheffield, J., Wood, E.F., 2011a. Multi-model, multi-sensor estimates of global evapotranspiration: climatology, uncertainties and trends. Hydrol. Process. 25 (26), 3993–4010.

Vinukollu, R.K., Wood, E.F., Ferguson, C.R., Fisher, J.B., 2011b. Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches. Remote Sens. Environ. 115 (3), 801–823.

Wang, K., Dickinson, R.E., 2012. A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. Rev. Geophys. 50 (2).

Wang, K., Dickinson, R.E., Wild, M., Liang, S., 2010. Evidence for decadal variation in global terrestrial evapotranspiration between 1982 and 2002: 1. Model development. J. Geophys. Res. 115 (D20).

Wang, K., Liang, S., 2008. An improved method for estimating global evapotranspiration based on satellite determination of surface net radiation, vegetation index, temperature, and soil moisture. J. Hydrometeorol. 9 (4), 712–727.

Wang, K., Wang, P., Li, Z., Cribb, M., Sparrow, M., 2007. A simple method to estimate actual evapotranspiration from a combination of net radiation, vegetation index, and temperature. J. Geophys. Res. 112 (D15).

Wang, Y., Li, R., Min, Q., Fu, Y., Wang, Y.u., Zhong, L., Fu, Y., 2019. A three-source satellite algorithm for retrieving all-sky evapotranspiration rate using combined optical and microwave vegetation index at twenty AsiaFlux sites. Remote Sens. Environ. 235, 111463.

Wang, Y., Woodcock, C.E., Buermann, W., Stenberg, P., Voipio, P., Smolander, H., Häme, T., Tian, Y., Hu, J., Knyazikhin, Y., Myneni, R.B., 2004. Evaluation of the MODIS LAI algorithm at a coniferous forest site in Finland. Remote Sens. Environ. 91 (1), 114–127.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B.E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., Verma, S., 2002. Energy balance closure at FLUXNET sites. Agric. For. Meteorol. 113 (1–4), 223–243.

Xiao, Z., Liang, S., Wang, J., Xiang, Y., Zhao, X., Song, J., 2016. Long-time-series global land surface satellite leaf area index product derived from MODIS and AVHRR surface reflectance. IEEE Trans Geosci. Remote Sens. 54 (9), 5301–5318.

Xiao, Z., Liang, S., Tian, X.D., Jia, K., Yao, Y., Jiang, B., 2017. Reconstruction of long-term temporally continuous NDVI and surface reflectance from AVHRR Data. IEEE J. Sel. Top Appl. Earth Obs. Remote Sens. 10 (2), 5551–5568.

Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L., Song, J., 2014. Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance. IEEE Trans Geosci. Remote Sens. 52 (1), 209–223.

Yang, L., Feng, Q.i., Zhu, M., Wang, L., Alizadeh, M.R., Adamowski, J.F., Wen, X., Yin, Z., 2022. Variation in actual evapotranspiration and its ties to climate change and vegetation dynamics in northwest China. J. Hydrol. 607, 127533.

Yann, L., Yoshua, B., Geoffrey, H., 2015. Deep learning. Nature 521 (7553), 436–444.

Yao, Y., Liang, S., Cheng, J., Liu, S., Fisher, J.B., Zhang, X., Jia, K., Zhao, X., Qin, Q., Zhao, B., Han, S., Zhou, G., Zhou, G., Li, Y., Zhao, S., 2013. MODIS-driven estimation of terrestrial latent heat flux in China based on a modified Priestley-Taylor algorithm. Agric. For. Meteorol. 171, 187–202.

Yao, Y., Liang, S., Li, X., Chen, J., Liu, S., Jia, K., Zhang, X., Xiao, Z., Fisher, J.B., Mu, Q., Pan, M., Liu, M., Cheng, J., Jiang, B., Xie, X., Grünwald, T., Bernhofer, C., Roupsard, O., 2017a. Improving global terrestrial evapotranspiration estimation using support vector machine by integrating three process-based algorithms. Agric. For. Meteorol. 242, 55–74.

Yao, Y., Liang, S., Li, X., Chen, J., Wang, K., Jia, K., Cheng, J., Jiang, B.o., Fisher, J.B., Mu, Q., Grünwald, T., Bernhofer, C., Roupsard, O., 2015. A satellite-based hybrid algorithm to determine the Priestley-Taylor parameter for global terrestrial latent heat flux estimation across multiple biomes. Remote Sens. Environ. 165, 216–233.

Yao, Y., Liang, S., Li, X., Hong, Y., Fisher, J.B., Zhang, N., Chen, J., Cheng, J., Zhao, S., Zhang, X., Jiang, B., Sun, L., Jia, K., Wang, K., Chen, Y., Mu, Q., Feng, F., 2014. Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations. J. Geophys. Res. Atmos. 119 (8), 4521–4545.

Yao, Y., Liang, S., Li, X., Zhang, Y., Chen, J., Jia, K., Zhang, X., Fisher, J.B., Wang, X., Zhang, L., Xu, J., Shao, C., Posse, G., Li, Y., Magliulo, V., Varlagin, A., Moors, E.J., Boike, J., Macfarlane, C., Kato, T., Buchmann, N., Billesbach, D.P., Beringer, J., Wolf, S., Papuga, S.A., Wohlfahrt, G., Montagnani, L., Ardö, J., Paul-Limoges, E., Emmel, C., Hörtnagl, L., Sachs, T., Gruening, C., Gioli, B., López-Ballesteros, A.,

Steinbrecher, R., Gielen, B., 2017b. Estimation of high-resolution terrestrial evapotranspiration from Landsat data using a simple Taylor skill fusion method. J. Hydrol. 553, 508–526.
Yebra, M., Dijk, A.V., Leuning, R., Huete, A., Guerschman, J.P., 2013. Evaluation of optical remote sensing to estimate actual evapotranspiration and canopy conductance. Remote Sens. Environ. 129, 250–261.
Yoshua, B., 2009. Learning Deep Architectures for AI. Now Publishers Inc.
Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., Zhang, L., 2020. Deep learning in environmental remote sensing: achievements and challenges. Remote Sens. Environ. 241, 111716.
Yuan, W., Liu, S., Yu, G., Bonnefond, J.-M., Chen, J., Davis, K., Desai, A.R., Goldstein, A. H., Gianelle, D., Rossi, F., Suyker, A.E., Verma, S.B., 2010. Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data. Remote Sens. Environ. 114 (7), 1416–1431.
Zhang, K., Kimball, J.S., Nemani, R.R., Running, S.W., 2010. A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006. Water Resour. Res. 46 (9).
Zhang, Y., Kong, D., Gan, R., Chiew, F.H.S., McVicar, T.R., Zhang, Q., Yang, Y., 2019. Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. Remote Sens. Environ. 222, 165–182.
Zhang, Y., Peña-Arancibia, J.L., McVicar, T.R., Chiew, F.H.S., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang, Y., Liu, Y.Y., Miralles, D.G., Pan, M., 2016. Multi-decadal trends in global terrestrial evapotranspiration and its components. Sci. Rep. 6 (1), 1–12.
Zhao, W., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S., Wen, Y., Lin, C., Li, X., Qiu, G. Y., 2019. Physics-constrained machine learning of evapotranspiration. Geophys. Res. Lett. 46 (24), 14496–14507.
Zhou, Z., Feng, J., 2017. Deep Forest: Towards an Alternative to Deep Neural Networks. In: international joint conference on artificial intelligence, pp. 3553–3559.
Zhu, G., Li, X., Zhang, K., Ding, Z., Han, T., Ma, J., Huang, C., He, J., Ma, T., 2016. Multi-model ensemble prediction of terrestrial evapotranspiration across north China using Bayesian model averaging. Hydrol. Process. 30 (16), 2861–2879.