AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2013JD020864

Key Points:
- BMA-based latent heat flux estimation
- Global terrestrial LE estimation
- Integrating five process-based algorithms

Correspondence to:
Y. Yao and X. Li,
boyyunjun@163.com;
xlli@bnu.edu.cn

Citation:
Yao, Y., et al. (2014), Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations, *J. Geophys. Res. Atmos.*, *119*, 4521–4545, doi:10.1002/2013JD020864.

Received 18 SEP 2013
Accepted 27 MAR 2014
Accepted article online 31 MAR 2014
Published online 23 APR 2014

Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations

Yunjun Yao[1], Shunlin Liang[1,2], Xianglan Li[1], Yang Hong[3,4], Joshua B. Fisher[5], Nannan Zhang[6], Jiquan Chen[7], Jie Cheng[1], Shaohua Zhao[8], Xiaotong Zhang[1], Bo Jiang[1], Liang Sun[9], Kun Jia[1], Kaicun Wang[10], Yang Chen[10], Qiaozhen Mu[11], and Fei Feng[1]

[1]State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, [2]Department of Geographical Sciences, University of Maryland, College Park, Maryland, USA, [3]School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, Oklahoma, USA, [4]Department of Hydraulic Engineering and State Key Lab of Hydroscience and Engineering, Tsinghua University, Beijing, China, [5]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA, [6]School of Earth Sciences and Engineering, Hohai University, Nanjing, China, [7]Department of Environmental Sciences, University of Toledo, Toledo, Ohio, USA, [8]Environmental Satellite Center, Ministry of Environmental Protection, Beijing, China, [9]Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China, [10]State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, [11]Numerical Terradynamic Simulation Group, Department of Ecosystem and Conservation Sciences, University of Montana, Missoula, Montana, USA

Abstract Accurate estimation of the satellite-based global terrestrial latent heat flux (LE) at high spatial and temporal scales remains a major challenge. In this study, we introduce a Bayesian model averaging (BMA) method to improve satellite-based global terrestrial LE estimation by merging five process-based algorithms. These are the Moderate Resolution Imaging Spectroradiometer (MODIS) LE product algorithm, the revised remote-sensing-based Penman-Monteith LE algorithm, the Priestley-Taylor-based LE algorithm, the modified satellite-based Priestley-Taylor LE algorithm, and the semi-empirical Penman LE algorithm. We validated the BMA method using data for 2000–2009 and by comparison with a simple model averaging (SA) method and five process-based algorithms. Validation data were collected for 240 globally distributed eddy covariance tower sites provided by FLUXNET projects. The validation results demonstrate that the five process-based algorithms used have variable uncertainty and the BMA method enhances the daily LE estimates, with smaller root mean square errors (RMSEs) than the SA method and the individual algorithms driven by tower-specific meteorology and Modern Era Retrospective Analysis for Research and Applications (MERRA) meteorological data provided by the NASA Global Modeling and Assimilation Office (GMAO), respectively. The average RMSE for the BMA method driven by daily tower-specific meteorology decreased by more than 5 W/m^2 for crop and grass sites, and by more than 6 W/m^2 for forest, shrub, and savanna sites. The average coefficients of determination (R^2) increased by approximately 0.05 for most sites. To test the BMA method for regional mapping, we applied it for MODIS data and GMAO-MERRA meteorology to map annual global terrestrial LE averaged over 2001–2004 for spatial resolution of 0.05°. The BMA method provides a basis for generating a long-term global terrestrial LE product for characterizing global energy, hydrological, and carbon cycles.

1. Introduction

The terrestrial latent heat flux (LE) (a list of acronyms is given in Table A1 of Appendix A) plays an important role in exchanges of energy, water, and carbon among the terrestrial biosphere, hydrosphere, and atmosphere. It is difficult to accurately estimate terrestrial LE because the land surface is generally more heterogeneous than the sea surface and there is much uncertainty about complicated biophysical processes [*National Research Council*, 2007; *Kalma et al.*, 2008; *Liang et al.*, 2010; *Wang et al.*, 2010a, 2010b; *Jia et al.*, 2012; *Liu et al.*, 2013]. Since the 1990s, eddy covariance (EC) flux towers provided by FLUXNET projects have been used to measure LE. However, the sparse observations hamper accurate characterization of spatiotemporal LE patterns over large spatial scales [*Baldocchi et al.*, 2001; *Serreze et al.*, 2005; *Sun et al.*, 2005].

Remote sensing is considered as the most viable method for producing spatially distributed global or regional LE products because it can effectively provide temporally and spatially continuous information on

soil and vegetation variables for estimating LE, including the land surface temperature (LST), the normalized difference vegetation index (NDVI), the enhanced vegetation index (EVI), the fraction of absorbed photosynthetically active radiation (FPAR), albedo, biome type, and leaf area index (LAI) [*Los et al.*, 2000; *Liang et al.*, 2010; *Jin et al.*, 2011; *Mu et al.*, 2011; *Yao et al.*, 2013]. Currently, the existing satellite-based Global Moderate Resolution Imaging Spectroradiometer (MODIS) LE product (MOD16) [*Mu et al.*, 2011] has 1 km spatial resolution and 8 day temporal resolution, but validation results indicate that the MODIS LE product often overestimates LE at most AsiaFlux sites [*Kim et al.*, 2012]. Other existing global LE products (including satellite and reanalysis products), such as LandFlux-EVAL (merged benchmark synthesis products of evapotranspiration, ET) and the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (ERA-Interim), have high temporal resolution (daily) but rather coarse spatial resolution ($>1°$) [*Simmons et al.*, 2006; *Fisher et al.*, 2008; *Wang and Liang*, 2008; *Zhang et al.,* 2010a, 2010b; *Jiménez et al.*, 2011; *Mueller et al.*, 2011, 2013; *Wang and Dickinson*, 2012]. Therefore, there is still a need to estimate terrestrial LE accurately at high spatial resolution (e.g., 1 km) but reasonable temporal resolution (e.g., daily).

During the past 30 years, there has been much effort to develop and design algorithms for terrestrial LE estimation, as documented by a substantial body of literature[*Monteith*, 1965; *Priestley and Taylor*, 1972; *Norman et al.*, 1995; *Li et al.*, 2009; *Liang et al.*, 2010; *Katul et al.*, 2012; *Wang and Dickinson*, 2012]. In general, satellite-based LE methods can be grouped into four categories. (1) Statistical and empirical methods (SEMI) involve the development of empirical equations using vegetation parameters or LST derived from satellites and LE observations from flux tower sites [*Jackson et al.*, 1977; *Wang et al.*, 2007; *Jung et al.*, 2010; *Jin et al.*, 2011; *Yao et al.,* 2011a, 2011b]. (2) Surface energy balance (SEB) models estimate LE by calculating the evaporation fraction (EF) from satellite data or by driving a residual SEB equation with remotely sensed products and meteorological data [*Norman et al.*, 1995; *Kustas and Norman*, 1996; *Anderson et al.*, 1997; *Bastiaanssen et al.*, 1998; *Allen et al.*, 2007]. (3) Penman-Monteith (PM) and Priestley-Taylor (PT) approaches use PM and PT equations, respectively, to calculate LE [*Monteith*, 1965; *Priestley and Taylor*, 1972; *Jiang and Islam*, 2001; *Cleugh et al.*, 2007; *Mu et al.*, 2007, 2011; *Fisher et al.*, 2008; *Zhang et al.*, 2009, 2010a; *Tang et al.*, 2010; *Wang et al.,* 2010a, 2010b; *Miralles et al.*, 2011; *Yao et al.*, 2013]. (4) Data assimilation (DA) methods assimilate satellite land-surface variables (e.g., LST) into land surface models (LSMs) to improve LE predictions [*Qin et al.*, 2007; *Pipunic et al.*, 2008; *Xu et al.,* 2011a, 2011b]. Although these approaches are widely used to estimate regional or global land-surface LE, simulation results may differ substantially owing to differences in the algorithms themselves and the calibrated coefficients. For example, a global intercomparison of 12 land-surface heat flux estimates using different algorithms by *Jiménez et al.* [2011] revealed that many LE algorithms show substantial differences in partitioning LE.

Merging multiple algorithms can be effectively used to estimate terrestrial LE with higher accuracy. Previous studies showed that a traditional simple multimodel averaging method performs better than any individual model for estimating climatic and hydrologic variables [*Raftery et al.*, 1997, 2005; *Madigan et al.*, 1999; *Houghton*, 2001; *Duan and Phillips*, 2010; *Wu et al.*, 2012]. Multimodel simulation-averaged LE data sets (e.g., LE product of the Global Soil Wetness Project-2, GSWP-2) have been widely used for comparison with other LE products [*Dirmeyer et al.*, 2006; *Wang and Liang*, 2008; *Yao et al.*, 2011a]. More complicated merging methods that calculate the weightings for individual algorithms based on ground measurements have recently been designed to estimate variables for land-surface energy budgets [*Wu et al.*, 2012]. In particular, there is increasing interest in using the Bayesian model averaging (BMA) method for optimal weighting for weather and hydrology predictions [*Hoeting et al.*, 1999; *Raftery et al.*, 2005; *Duan et al.*, 2007; *Duan and Phillips*, 2010; *Wu et al.*, 2012]. In essence, BMA facilitates more accurate estimations of variables by pooling information from multiple algorithms to generate ensemble predictions similar to a weighted average of component forecasts. However, there is a lack of similar studies on improving global terrestrial LE estimates using the BMA method for understanding energy and hydrologic cycles.

In this study we used the BMA method to improve satellite-based global terrestrial LE estimation by merging five process-based LE algorithms. We had two major objectives. First, we evaluated the BMA method using long-term FLUXNET measurements between 2000 and 2009 by comparison with a simple model averaging (SA) method and five process-based algorithms. Second, we applied the BMA method to map annual global terrestrial LE averaged over 2001–2004 with spatial resolution of 0.05° using MODIS data and Modern Era Retrospective Analysis for Research and Applications (MERRA) meteorological data provided by the NASA Global Modeling and Assimilation Office (GMAO).

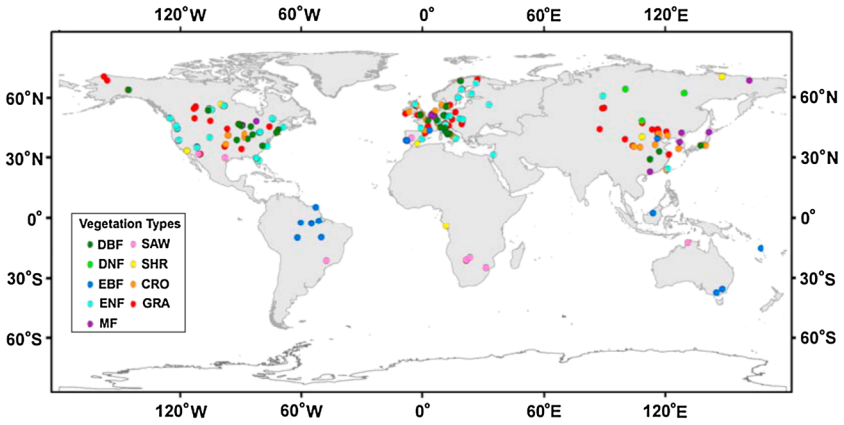

Figure 1. Locations of the 240 eddy covariance flux towers used in this study.

2. Data

2.1. Observations From Eddy Covariance Flux Towers

The BMA method, the SA method, and five traditional LE algorithms were validated and evaluated using ground-measured flux data. The data were collected at 240 EC flux tower sites and provided by AmeriFlux, AsiaFlux, LathuileFlux, the Chinese Ecosystem Research Network (CERN), the Coordinated Enhanced Observation Network of China (CEOP), the Asian Automatic Weather Station Network (ANN) Project supported by the Global Energy and Water Cycle Experiment (GEWEX) Asian Monsoon Experiment (GAME ANN), and individual principal investigators (PIs) of the FLUXNET website. These flux towers are mainly located in Europe, Asia, and North America, as well as seven flux towers in South America, five in Africa, and five in Australia (Figure 1). The climate for the flux tower locations varies from humid to dry and from tropical to frigid. The flux tower sites cover nine major global land-surface biomes: deciduous broadleaf forest (DBF; 28 sites), deciduous needleleaf forest (DNF; 6 sites), evergreen broadleaf forest (EBF; 16 sites), evergreen needleleaf forest (ENF; 64 sites), mixed forest (MF; 12 sites), savanna (SAW; 10 sites), shrubland (SHR; 14 sites), cropland (CRO; 34 sites), and grass and other types (GRA; 56 sites). These data sets include half-hourly or hourly ground-measured incident solar radiation (R_s), relative humidity (RH), air temperature (T_a), diurnal air-temperature range (DT), wind speed (WS), vapor pressure (e), sensible heat flux (H), surface net radiation (R_n), ground heat flux (G), and LE. When available, data sets were gap filled by site principal investigators (PIs) and daily data are aggregated from half-hourly or hourly data without using additional quality control [*Liu et al.*, 2011; *Jia et al.*, 2012; *Xu et al.*, 2013]. The data cover the period from 2000 to 2009 and each flux tower has at least 1 year of reliable data. All turbulent flux observations were measured by the EC method. Although this is considered a good method for measuring heat flux, it suffers from an unclosed energy problem [*Twine et al.*, 2000; *Wilson et al.*, 2002; *Jung et al.*, 2010]. Therefore, we used the method developed by *Twine et al.* [2000] to correct the LE for all flux towers according to

$$\mathrm{LE} = (R_n - G)/(\mathrm{LE}_{ori} + H_{ori}) \times \mathrm{LE}_{ori} \quad (1)$$

where LE is the corrected latent heat flux, and H_{ori} and LE_{ori} are the uncorrected sensible heat flux and latent heat flux, respectively.

2.2. Satellite and Reanalysis Inputs to LE Algorithms

To evaluate the performance of all LE algorithms in this study for all flux tower sites, the five process-based algorithms were driven by two different meteorological data sets: (1) daily point-based meteorological observations from flux towers and (2) daily GMAO-MERRA meteorological data with spatial resolution of 1/2° × 2/3°. The 8 day MODIS FPAR/LAI (MOD15A2) product [*Myneni et al.*, 2002] with 1 km spatial resolution was used to drive the LE algorithms in this study, and the daily FPAR/LAI values were temporally interpolated from the 8 day averages using linear interpolation. Similarly, the 16 day MODIS NDVI/EVI (MOD13A2) [*Huete et al.*, 2002] and MODIS albedo (MOD43B3) products [*Liang et al.,* 1999; *Schaaf et al.,* 2002] were also used to validate the LE models. We temporally filled missing or cloud-contaminated albedo, LAI, FPAR, NDVI, and EVI for

Table 1. GMAO-MERRA and MODIS Products for Mapping Global Terrestrial LE Used in This Study

Products	Short Name	Spatial Resolution	Variables Acquired
GMAO-MERRA reanalysis product	GMAO-MERRA	1/2° × 2/3°	R_s, R_n, T_a, T_{max}, T_{min}, e, RH, WS
MODIS albedo product	MOD43C1	0.05°	Albedo
MODIS NDVI/EVI product	MOD13C1	0.05°	NDVI, EVI
MODIS land cover product	MOD12C1	0.05°	Land cover
MODIS FPAR/LAI product	MOD15A2	1 km	FPAR, LAI

each 1 km MODIS pixel using the method described by *Zhao et al.* [2005]. According to this method, when 16 day albedo data (8 day LAI, 16 day NDVI/EVI) are not available, the closest reliable 16 day (8 day) values replace the original data. To match MODIS pixels, we used the method proposed by *Zhao et al.* [2005] to interpolate coarse-resolution GMAO-MERRA data to 1 km^2 MODIS pixels. Theoretically, this spatial interpolation method improves the accuracy of meteorological data for each 1 km pixel because it uses a cosine function and the four GMAO-MERRA cells surrounding a given pixel to remove sharp changes from one side of a GMAO-MERRA boundary to the other [*Zhao et al.*, 2005].

To map global terrestrial LE over 2001–2004 with 0.05° spatial resolution using the BMA method, input data for all the LE algorithms included 1/2° × 2/3° GMAO-MERRA meteorological data and MODIS data. We used the method described by *Zhao et al.* [2005] to interpolate GMAO-MERRA data to 0.05° resolution over 2001–2004. MODIS inputs included Collection 4 MODIS albedo (MOD43C1: CMG, 0.05°) [*Jin et al.*, 2003; *Salomon et al.*, 2006], Collection 5 MODIS NDVI/EVI (MOD13C1: CMG, 0.05°) [*Myneni et al.*, 2002], Collection 4 MODIS land cover (MOD12C1: CMG, 0.05°) [*Friedl et al.*, 2002], and the Collection 5 MODIS FPAR/LAI (MOD15A2) with spatial resolution of 1 km. Furthermore, the enhanced 1 km^2 LAI/FPAR has been aggregated into 0.05° data using bilinear interpolation with geographic projection. Detailed information on these inputs for mapping global terrestrial LE is summarized in Table 1.

3. Methods

3.1. Process-Based Latent Heat Flux Algorithms

We used five process-based LE algorithms in this study. For simplicity, the algorithms are denoted by their abbreviations in figure legends (e.g., Revised Remote-Sensing-Based Penman-Monteith (RRS-PM)). The algorithms (Table 2) are briefly described below.

3.1.1. MODIS LE Product Algorithm (MOD16)

The MODIS LE product algorithm (MOD16) is based on a beta version [*Mu et al.*, 2007] developed from the *Cleugh et al.* [2007] version using a PM approach [*Monteith*, 1965]. *Mu et al.* [2011] improved the beta version by (1) revising vegetation cover fraction with FPAR, (2) estimating LE as the sum of daytime and nighttime components, (3) improving calculations of aerodynamic, boundary-layer, and canopy resistance, (4) estimating the soil heat flux using available energy and simplified NDVI, (5) dividing the canopy into wet and dry components, and (6) separating moist soil surfaces from saturated wet ones. The MOD16 algorithm was evaluated at 46 EC flux tower sites and has been successfully extended to generate a MODIS global

Table 2. Summary of the Five Process-Based LE Algorithms and Forcing Input Variables

LE Algorithm	Forcing Inputs	Institute	References
MODIS LE products algorithm (MOD16)	R_n, T_a, T_{min}, e, RH, LAI, land cover, FPAR	Numerical Terradynamic Simulation Group Department of Ecosystem and Conservation Sciences, University of Montana, USA	*Mu et al.* [2011]
Revised remote-sensing-based Penman-Monteith LE algorithm (RRS-PM)	R_n, T_a, e, RH, LAI, EVI	State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, China	*Yuan et al.* [2010]
Priestley-Taylor LE algorithm of Jet Propulsion Laboratory, Caltech (PT-JPL)	R_n, T_a, T_{max}, e, RH, LAI, NDVI, FPAR	Jet Propulsion Laboratory, California Institute of Technology, USA	*Fisher et al.* [2008]
Modified satellite-based Priestley-Taylor LE algorithm (MS-PT)	R_n, T_a, T_{max}, T_{min}, NDVI	State Key Laboratory of Remote Sensing Science, Beijing Normal University, China	*Yao et al.* [2013]
Semiempirical Penman LE algorithm of the University of Maryland (UMD-SEMI)	R_s, T_a, e, WS, RH, NDVI	Department of Geography, University of Maryland, College Park, Maryland, USA	*Wang et al.* [2010a]

terrestrial LE product driven by MODIS land cover, albedo, and LAI/FPAR, and a GMAO daily meteorological reanalysis data set [*Mu et al.,* 2011].

Except for the MOD16 LE algorithm, the other four LE algorithms used in this study all neglect nighttime LE because most LE occurs during daytime. For consistency, we calculated daytime LE by using the MODIS LE algorithm with the daytime-averaged meteorological data as input.

3.1.2. Revised Remote-Sensing-Based Penman-Monteith LE Algorithm (RRS-PM)

The beta version of the MOD16 algorithm [*Mu et al.,* 2007] has a good physical basis in terms of a PM equation and constraint parameters of air temperature and vapor pressure deficit (VPD) that differ for different vegetation types. However, *Yuan et al.* [2010] found that it is possible to set invariant model parameters across different vegetation types to reduce the effects of misclassification of land cover types. Therefore, *Yuan et al.* [2010] developed a revised remote-sensing-based PM LE algorithm (RRS-PM) by revising the algorithm parameters, modifying the air temperature constraint for vegetation conductance, and improving calculation of the vegetation cover fraction using EVI. Validation for 23 EC flux tower sites in China revealed higher squared correlation coefficients (R^2) and lower root-mean-square errors (RMSEs) for the RRS-PM algorithm than for the beta version of the MOD16 algorithm [*Chen et al.*, 2014].

3.1.3. Priestley-Taylor-Based LE Algorithm (PT-JPL)

To avoid the complexity of parameterizations of both aerodynamic and surface resistance, *Priestley and Taylor* [1972] reduced the atmospheric control term in the PM equation and added an empirical factor to design a simple LE algorithm. On the basis of this algorithm, *Fisher et al.* [2008] proposed a novel PT-based LE algorithm (PT-JPL) by introducing both atmospheric (RH and VPD) and ecophysiological constraints (FPAR and LAI) without using any ground-based observed data. The PT-JPL method was validated at 16 global FLUXNET sites and the simulation-to-measurement R^2 was 0.90 (RMSE = 15.2 W/m^2 or 28%) for all sites during 2 years. It has been applied to estimate global terrestrial LE driven by the Advanced Very High Resolution Spectroradiometer (AVHRR) satellite products and the International Satellite Land Surface Climatology Project, Initiative II (ISLSCP-II) data sets [*Hall et al.*, 2006; *Jiménez et al.*, 2011].

3.1.4. Modified Satellite-Based Priestley-Taylor LE Algorithm (MS-PT)

To reduce the atmospheric inputs for the PT-JPL algorithm, *Yao et al.* [2013] used the apparent thermal inertia (ATI) derived from DT to parameterize surface soil moisture constraints and the revised linear two-source model (N95) to estimate vegetation transpiration [*Norman et al.*, 1995; *Anderson et al.*, 1997; *Yao et al.*, 2013]. This MS-PT algorithm estimates LE from four components: saturated wet soil evaporation, unsaturated wet soil evaporation, vegetation transpiration, and evaporation from vegetation interception. MS-PT is an operational satellite method for estimating global terrestrial LE because it only requires R_n, air temperature, DT, and NDVI as inputs. According to validation for 16 EC flux tower sites throughout China, the average RMSE between measured and predicted site-averaged daily LE was approximately 5 W/m^2 lower (99% confidence) for the MS-PT compared to the PT-JPL algorithm [*Yao et al.,* 2013].

3.1.5. Semiempirical Penman LE Algorithm of the University of Maryland (UMD-SEMI)

There are no satellite-based LE methods available for detecting global LE variations on a scale of several decades because of a lack of long-term satellite and ground-measured data [*Wang et al.,* 2010a, 2010b]. Therefore, *Wang et al.* [2010a] developed a semiempirical LE algorithm based on the Penman equation [*Penman*, 1948]. This UMD-SEMI method mainly considers the impact of incident solar radiation, air temperature, VPD, RH, vegetation indices, and wind speed to predict global terrestrial LE variability. It is the only method that explicitly includes wind speed, which may play an important role in annual or decadal LE variability [*Skidmore et al.*, 1969; *Wang et al.,* 2010a, 2010b; *McVicar et al.*, 2012]. Comparison of measured and predicted daily LE at 64 globally distributed flux tower sites demonstrated that the 16 day average daily LE can be reasonably predicted with an average correlation coefficient of 0.94 and average RMSE of 17 W/m^2 [*Wang et al.,* 2010a, 2010b].

3.2. Bayesian Model Averaging Method

Here we used the Bayesian model averaging (BMA) method to merge five process-based LE algorithms to estimate terrestrial LE. The predictive probability density function (PDF) for LE is a weighted average of the PDFs for the individual models, weighted by their posterior model probability [*Raftery et al.*, 1997; *Hoeting et al.*, 1999; *Duan and Phillips*, 2010; *Wu et al.*, 2012]. The BMA method can improve LE accuracy by adjusting the predictive PDF to obtain a good fit to a set of tower-based LE observations.

The BMA method considers a predictive variable y, the corresponding observation data at a given time t (y_t), and an ensemble model $F\{M_1, M_2, \ldots, M_f\}$ for variable y. In this study, y refers to predicted LE and $f = 5$. The law of total probability tells us that the predictive PDF, $p(y)$, can be expressed as

$$p(y|M_1, M_2, \ldots, M_f) = \sum_{i=1}^{f} p(y|M_i) \cdot p(M_i|y_t) \tag{2}$$

where $p(y|M_i)$ is the predictive PDF using model M_i alone and $p(M_i|y_t)$ denotes the posterior probability that model M_i is correct given the corresponding observation data. In general, $p(M_i|y_t)$ can be considered as a statistical weight u_i, which reflects how well M_i matches the observation data, and $\sum_{i=1}^{f} u_i = 1$. Thus, equation (2) can be expressed as

$$p(y|M_1, M_2, \ldots, M_f) = \sum_{i=1}^{f} u_i \cdot p(y|M_i) \tag{3}$$

It is reasonable to assume that $p(y|M_i)$ meets a normal distribution defined by a mean $\widetilde{M}_i$ and a variance ω_i^2 [*Raftery et al.*, 2005; *Duan and Phillips*, 2010; *Wu et al.*, 2012]. Assuming the parameter vector $\theta_i = \left\{\widetilde{M}_i, \omega_i^2\right\}$ and a conditional density function $h(\cdot)$ as the PDF associated with the normal distribution, this can be written as

$$p(y|M_i) = h(y|\theta_i) \tag{4}$$

Combining equations (3) and (4), we obtain

$$p(y|M_1, M_2, \ldots, M_f) = \sum_{i=1}^{f} u_i \cdot h(y|\theta_i) \tag{5}$$

The conditional expectation (E) of y is the ultimate BMA predictive LE for merging five process-based algorithms and can be expressed as

$$E(y|M_1, M_2, \ldots, M_f) = \sum_{i=1}^{f} u_i \cdot \widetilde{M}_i \tag{6}$$

where $\widetilde{M}_i$ is the estimated LE using each single algorithm. To obtain both u_i and θ_i, a log likelihood function l from the Gaussian function h on the basis of training observational data can be used. The log likelihood function can be expressed as

$$l(\theta_1, \theta_2, \ldots, \theta_f) = \sum\nolimits_{(s,t)} \log\left[\sum_{i=1}^{f} u_i \cdot h\left(y_{s,t}|\theta_i\right)\right] \tag{7}$$

where $\sum_{(s,t)}$ refers to the summation of observed LE values over all spatial points s and time points t; $y_{s,t}$ is an observed LE value at location s and time t. The BMA method can estimate the Bayesian weights u_i and parameter vectors θ_i when the log likelihood function l is maximized [*Raftery et al.*, 1997; *Duan and Phillips*, 2010]. We maximize the log likelihood function l using the expectation-maximization (EM) algorithm, which has been described elsewhere [*Raftery et al.*, 1997; *Duan and Phillips*, 2010].

To assess the terrestrial LE accuracy, we tested the performance of our BMA method using the holdout method, which is a simple type of cross validation [*Mo et al.*, 2004; *Yao et al.*, 2011a]. The data set is randomly stratified into two groups with approximately equal numbers of samples. We calibrated the weights using data from the first group and independently validated daily LE using data from the second group. We then calculated the weights using data from the second group and independently validated daily LE using data from the first group. We also used all of the data to calculate the weights for producing global terrestrial LE products.

3.3. Simple Model Averaging Method

For comparison with LE estimates based on the BMA method, we also used a simple model averaging (SA) method to merge the five process-based LE algorithms. The SA method is a simplified version of the BMA method that considers the weight for any individual algorithm as a constant ($1/f$). It can be expressed as

$$\mathrm{LE}_{\mathrm{SA}} = \frac{1}{f}\sum_{i=1}^{f} \mathrm{LE}_i \tag{8}$$

where $\mathrm{LE}_{\mathrm{SA}}$ and LE_i are terrestrial LE predicted using the SA method and each single process-based LE algorithm, respectively.

Table 3. Averaged Bias for 240 EC Flux Sites With the Same Land Cover Type[a]

		Bias				
Land Cover Type	Forcing Data	MOD16	RRS-PM	PT-JPL	MS-PT	UMD-SEMI
DBF	Tower-specific	−19.8	−19.6	18.1	16.7	8.8
	GMAO-MERRA	−21.8	−21.1	10.4	17.8	17.6
DNF	Tower-specific	−18.2	−19.1	10.8	10.3	8.7
	GMAO-MERRA	−13.5	−12.9	18.4	18.1	17.4
EBF	Tower-specific	−12.2	−17.8	22.1	18.3	9.5
	GMAO-MERRA	−19.3	−19.5	22.8	19.8	2.9
ENF	Tower-specific	−10.6	−17.4	22.4	14.2	6.7
	GMAO-MERRA	−10.8	−18.1	14.5	−1.8	5.9
MF	Tower-specific	−13.8	−13.5	10.4	12.3	5.7
	GMAO-MERRA	−2.8	−9.9	19.3	8.6	9.3
SHR	Tower-specific	−4.5	−17.4	17.3	10.1	12.9
	GMAO-MERRA	−3.8	−16.6	18.3	1.2	1.8
SAW	Tower-specific	−16.2	−18.2	19.7	12.1	15.6
	GMAO-MERRA	−22.5	−23.2	11.8	−2.8	2.9
CRO	Tower-specific	−30.2	−33.6	−4.5	−12.7	−2.3
	GMAO-MERRA	−28.7	−32.4	−1.2	−7.9	−2.4
GRA	Tower-specific	−25.2	−26.7	−1.8	−8.2	−3.6
	GMAO-MERRA	−25.4	−26.9	−1.2	−4.4	−2.5

[a]Values are in W/m^2.

3.4. Model Performance

We used Taylor diagrams to evaluate the performance of the individual LE algorithms, the SA method, and the BMA method to qualify how closely the simulated LE matched the ground observations [*Taylor*, 2001]. A Taylor diagram is a polar-style graph including the correlation coefficient (R) between simulations and observations, the centered RMSE, and the standard deviation (STD). The radial distance from the origin reflects STD, the cosine of the azimuth angle gives R, and the radial distance from the observed point is proportional to the centered RMSE difference between simulations and observations. Taylor diagrams are particularly beneficial in evaluating the relative accuracy of different complex models. We also summarized the average bias and p values for the estimated LE and those derived from tower data to evaluate the relative predictive errors for different LE models.

4. Results

4.1. Evaluation of the BMA Method for Merging Terrestrial LE Algorithms

4.1.1. Comparison of Five Process-Based LE Algorithms

At the flux tower site scale, the five process-based LE algorithms exhibit substantial differences in LE modeling. Table 3 and Figure 2 compare daily LE observed at 240 EC flux tower sites according to land cover types with estimates driven by tower meteorology and GMAO-MERRA meteorology, respectively. For DBF and DNF sites, average RMSEs are lower for the MS-PT and UMD-SEMI algorithms than for the other algorithms, and the average bias for these two algorithms is less than 17 W/m^2 for tower-driven meteorology. In contrast to the tower data as inputs, the results driven by GMAO-MERRA meteorology showed average bias of approximately 18 W/m^2 for the MS-PT and UMD-SEMI algorithms. The MOD16 algorithm has the highest average STD, and the RRS-PM algorithm has the second highest STD and the greatest bias. It is clear that the MOD16 and RRS-PM algorithms both underestimate terrestrial LE, but the PT-JPL, MS-PT, and UMD-SEMI algorithms overestimate terrestrial LE for DBF and DNF towers. This can be attributed to differences in calibrated coefficients for the different algorithms. In particular, the performance of the UMD-SEMI algorithm is strongly related to the regression coefficients because it was calibrated using the data from 64 flux tower sites (including 26 flux tower sites used in this study) over the U.S., East Asia, and Australia. Therefore, the UMD-SEMI algorithm is one of the algorithms that provide a better fit to flux tower observations.

For ENF and MF sites, the tower-driven MS-PT model has the lowest average RMSE, with average bias of 14.2 and 12.3 W/m^2, respectively. The tower-driven RRS-PM algorithm shows the second lowest average RMSE,

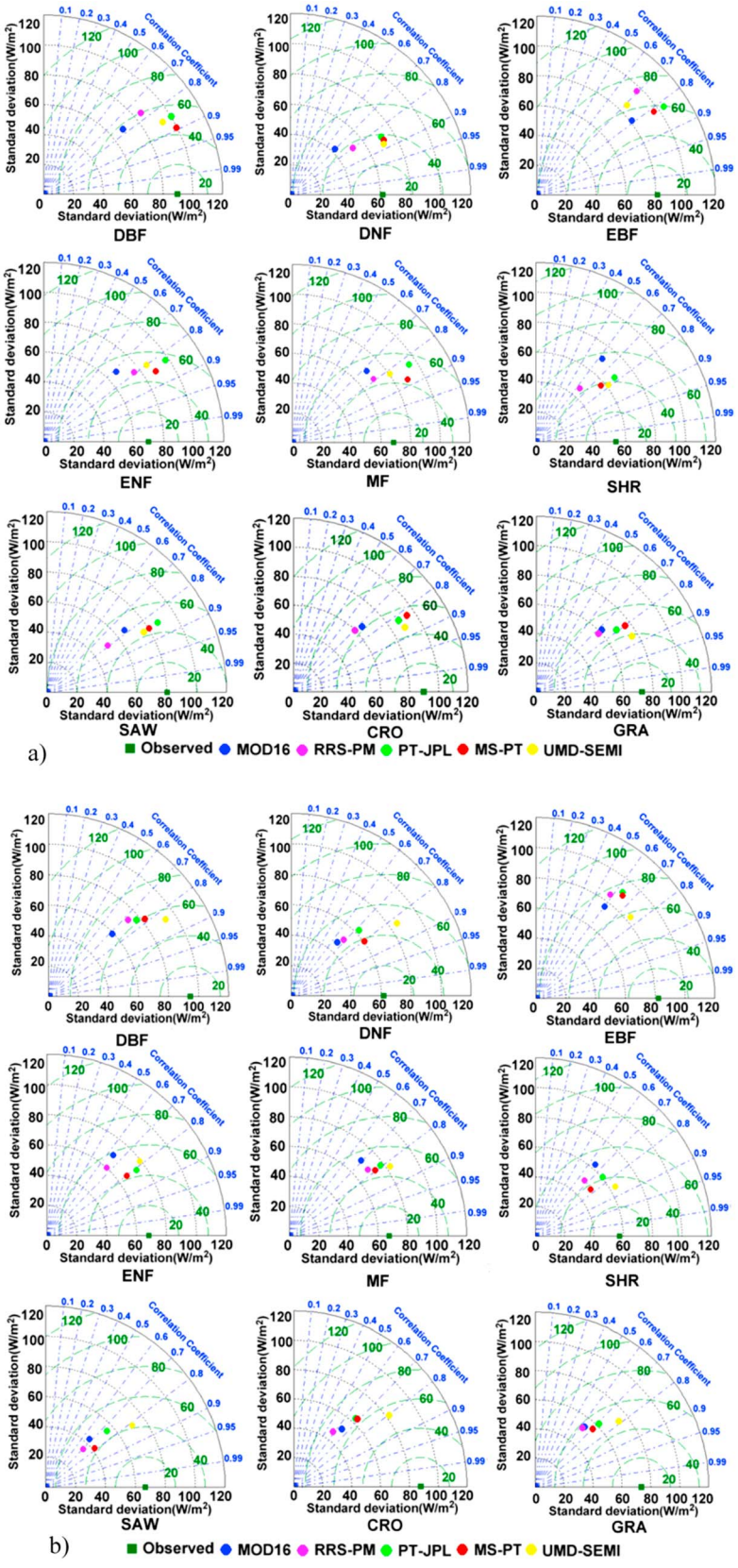

Figure 2. Taylor diagrams for daily LE observations and LE estimates using different algorithms driven by (a) tower-specific meteorology and (b) GMAO-MERRA meteorology at all 240 EC sites.

with average bias of −17.4 and −13.5 W/m^2, respectively, while the tower-driven MOD16 algorithm has the highest average RMSE. Similar conclusions are drawn for algorithms driven by GMAO-MERRA meteorology. In general, the ENF canopy conductance is half that of deciduous forests [*Eugster et al.*, 2000] and the MS-PT algorithm may capture this information by parameterization of the vegetation index [*Yao et al.*, 2013, 2014].

The PT-JPL algorithm highly overestimates terrestrial LE at EBF sites. In contrast, the MOD16, MS-PT, and UMD-SEMI algorithms show relatively lower bias. Among the tower-driven algorithms, MOD16 has the lowest average RMSE, with average bias of −12.2W/m^2 for all 16 EBF sites. However, for GMAO-MERRA meteorological data as inputs, the UMD-SEMI algorithm has the highest R^2, with an average R^2 of 0.52 ($p < 0.01$) for all 16 EBF sites. In general, most EBF sites are located in tropical and subtropical zones, and atmospheric control factors improve the LE parameterization under most humid conditions because VPD and wind speed affect LE via boundary-layer processes [*Rodriguez-Iturbe et al.*, 2007; *Zhao and Running*, 2010; *Wang and Dickinson*, 2012]. Only the UMD-SEMI algorithm, which is based on EC data calibration, considers the effects of wind speed. The MOD16 algorithm effectively considers the effects of VPD by using a look-up table for updating biome properties to adjust the VPD constraint [*Mu et al.,* 2007, 2011].

For SHR and SAW sites, although average RMSEs are higher for the MS-PT than for the UMD-SEMI algorithm, its daily average LE estimates are lower than those of the other algorithms. For instance, for all 14 SHR sites, the MS-PT algorithm has higher average RMSE (39.8 and 37.7 W/m^2) and lower average R^2 (0.60 and 0.63, $p < 0.01$) than the UMD-SEMI algorithm (RMSE: 38.2 and 36.3 W/m^2; R^2: 0.64 and 0.68, $p < 0.01$) when driven by tower and GMAO-MERRA meteorological data, respectively. However, the magnitude of average daily LE estimated using the MS-PT algorithm is much closer to tower-measured LE and the MS-PT algorithm has lower average bias (10.1 and 1.2 W/m^2) than the UMD-SEMI algorithm (12.9 and 1.8 W/m^2) when driven by tower and GMAO-MERRA meteorological data, respectively. This indicates that the MS-PT algorithm performs well under sparse vegetation conditions.

The UMD-SEMI algorithm shows lower RMSEs and R^2 for CRO and GRA sites but the differences were significantly below the confidence level of $p < 0.05$. Its average RMSE is 43.5 W/m^2 (55.1 W/m^2) for the 34 CRO sites and 38.7 W/m^2 (49.4 W/m^2) for the 56 GRA sites for tower-specific (GMAO-MERRA) meteorology. The validation results reveal average RMSE differences between tower-driven LE and GMAO-MERRA-driven LE for all five algorithms. This indicates that different algorithm parameterizations affect the accuracy of process-based terrestrial LE algorithms.

4.1.2. Cross Validation of the BMA Method

None of the individual LE algorithm provides the best LE estimate for all vegetation types. Therefore, we used the BMA method to estimate terrestrial LE by integrating five process-based LE algorithms driven by tower-specific meteorology and GMAO-MERRA meteorology, respectively. All LE algorithms chosen in this study underestimate terrestrial LE at CRO and GRA sites, and the MOD16 and RRS-PM algorithms have the highest absolute values for average bias (< −25W/m^2) for these sites. To reduce the BMA bias, the MOD16 and RRS-PM algorithms were therefore excluded from the multi-algorithm ensemble for CRO and GRA sites.

Terrestrial LE estimates calculated using the BMA method were compared with those for the SA method and the individual algorithms for each land cover type. Figure 3 compares daily LE observations and BMA estimates for the first group using the second group as training data to calibrate the weights of the five algorithms driven by tower-specific and GMAO-MERRA meteorology, respectively. The most prominent feature in Figure 3 is that the BMA method agrees best with observations that lie nearest the point marked "observed" on the *x* axis, and the BMA method has higher R^2 (95% confidence) and lower RMSE compared to the SA method and the individual algorithms driven by tower-specific (GMAO-MERRA) meteorology at all 120 EC sites with different land cover types. For the 63 forest sites using BMA driven by tower-specific (GMAO-MERRA) meteorology, the average RMSE is 42.2 W/m^2 (49.4W/m^2) and is lower than for the SA method and the individual algorithms, and the average R^2 is approximately 0.72 (0.67) ($p < 0.01$). For the 12 SHR and SAW sites using BMA driven by tower-specific (GMAO-MERRA) meteorology, the average RMSE is less than 40 W/m^2 (41 W/m^2) and the average R^2 is greater than 0.68 (0.64) ($p < 0.01$), which represents better performance than that of the SA method and the individual algorithms. For the 45 CRO and GRA sites, the average RMSE is much lower and the average R^2 is slightly higher for the BMA method compared to the SA method and the individual algorithms. Overall, the average RMSE of the BMA method for tower-specific meteorology inputs

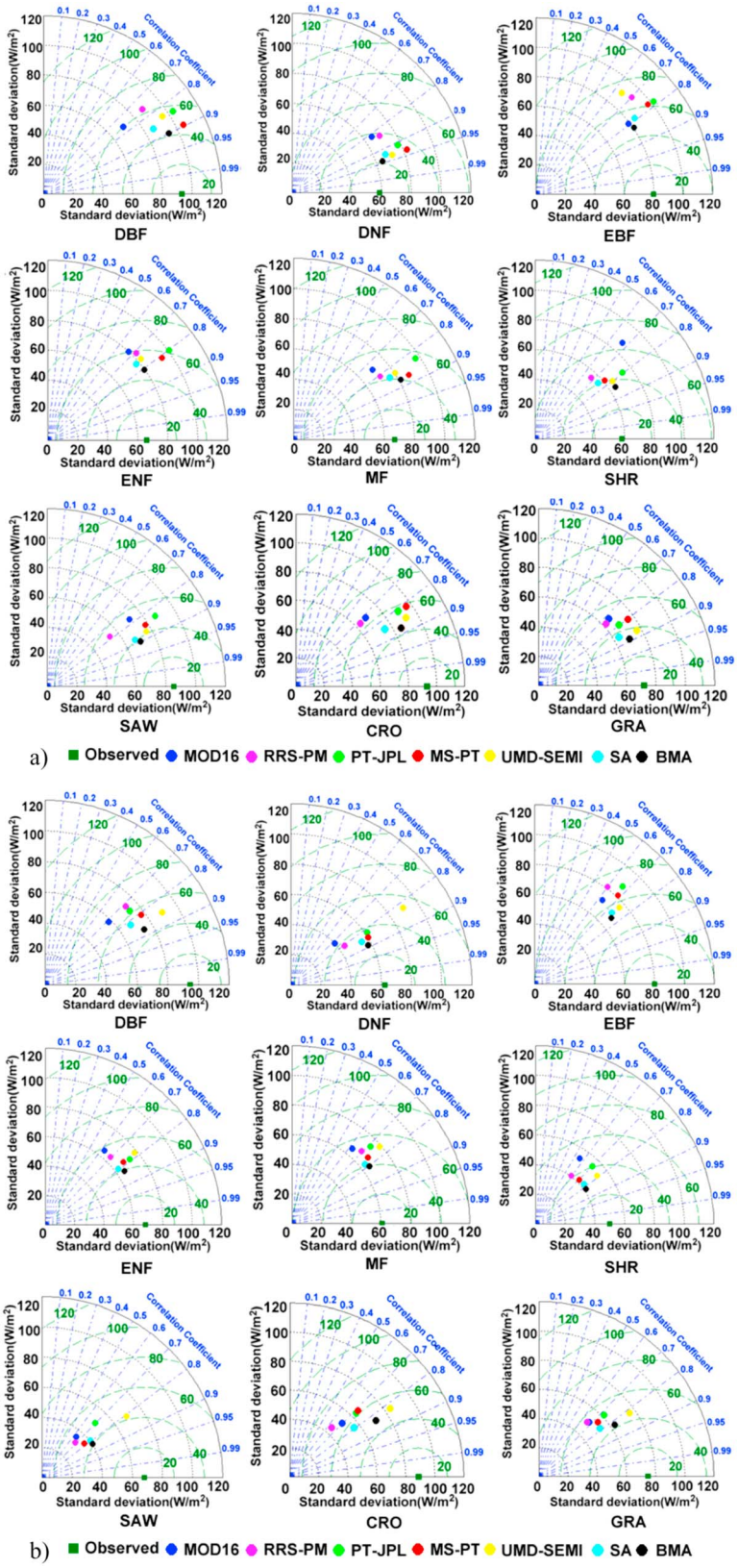

Figure 3. Taylor diagrams for daily LE observations and LE estimates using different algorithms driven by (a) tower-specific meteorology and (b) GMAO-MERRA meteorology at 120 EC sites. Simulations are for the first group and the second group was used as training data to calibrate the algorithm weights.

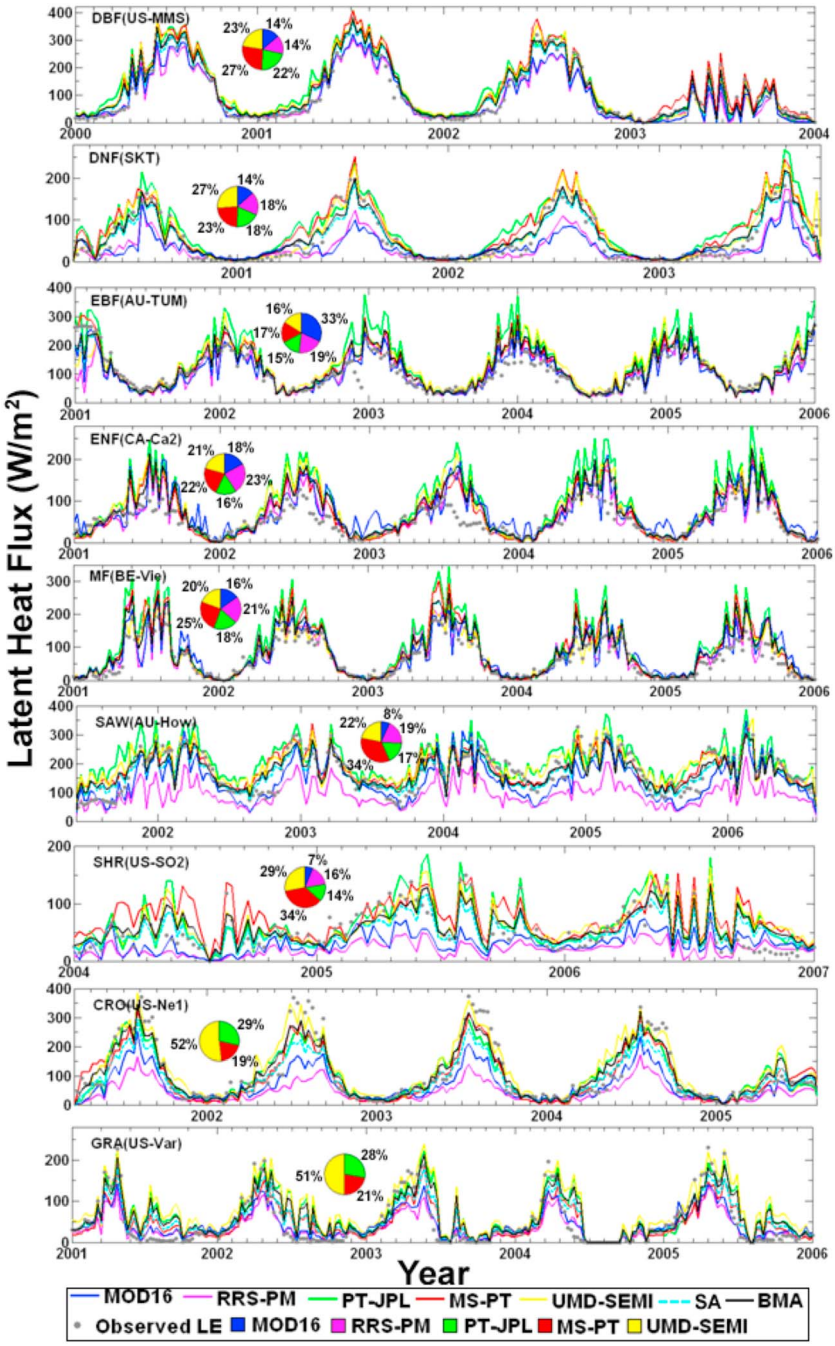

Figure 4. Example of a time series for the 8 day LE average as measured and predicted using different tower-driven algorithms for different land cover types from the first group. The pie charts show the relative contribution of each algorithm to the merged LE for the second group.

decreased by more than 5 W/m^2 for crop and grass sites and more than 6 W/m^2 for forest, shrub, and savanna sites. The average R^2 increased by approximately 0.05 at the 95% level of confidence for most sites. Figure 4 presents a time series for 8 day average LE measurements and tower-driven predictions for the first group of data for different land cover types. In comparison to the SA method and the individual algorithms, the BMA method yields seasonal LE variations that are closest to the ground-measured values. Similar conclusions can be drawn from BMA LE estimates for the second group of flux towers using the first group data to calibrate the weights (Figure 5).

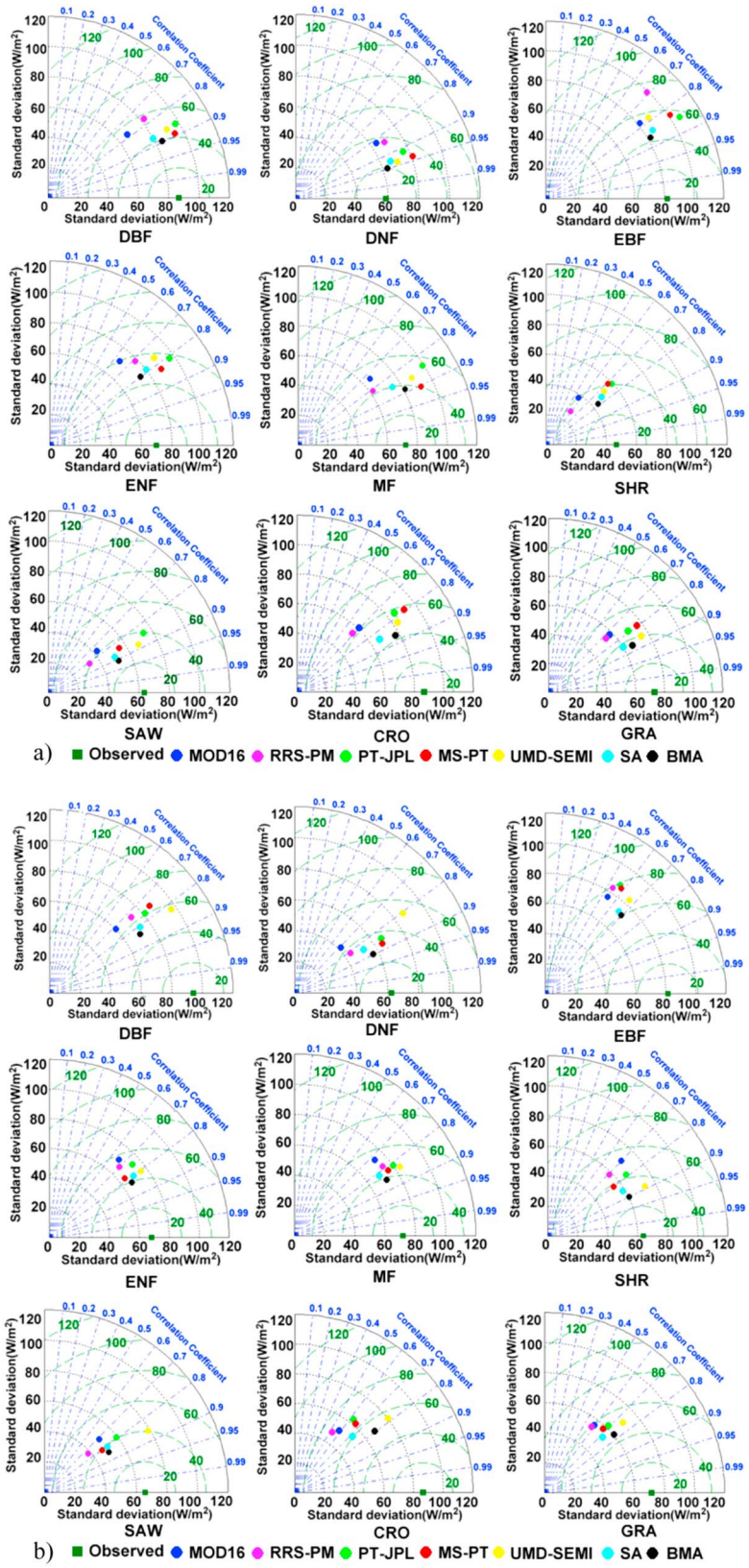

Figure 5. Taylor diagrams for daily LE observations and LE estimates using different algorithms driven by (a) tower-specific meteorology and (b) GMAO-MERRA meteorology at 120 EC sites. Simulations are for the second group and the first group was used as training data to calibrate the algorithm weights.

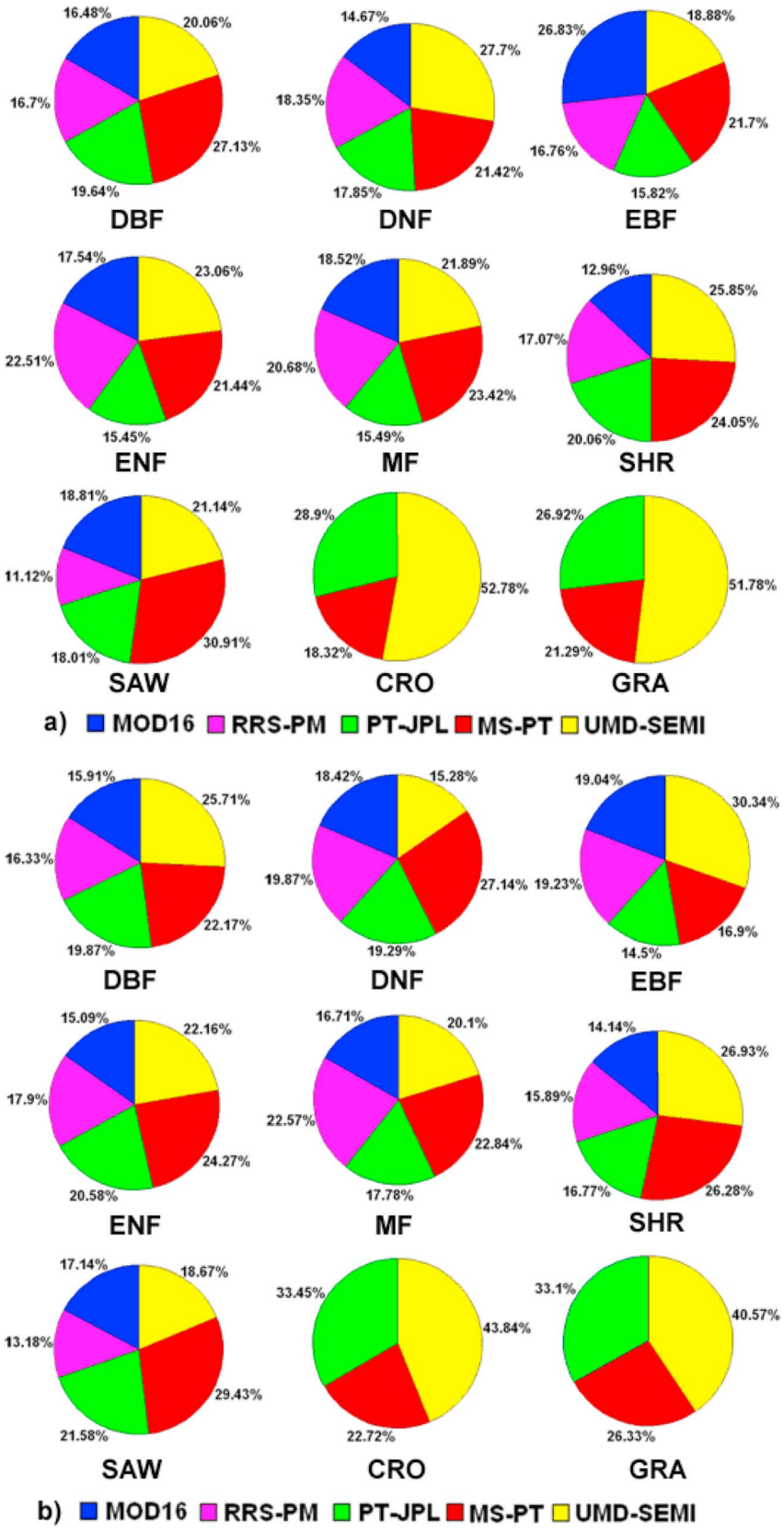

Figure 6. Weights for five process-based LE algorithms driven by (a) tower-specific meteorology and (b) GMAO-MERRA meteorology according to the BMA method for all land cover types (DBF, DNF, EBF, ENF, MF, SHR, SAW, CRO, and GRA).

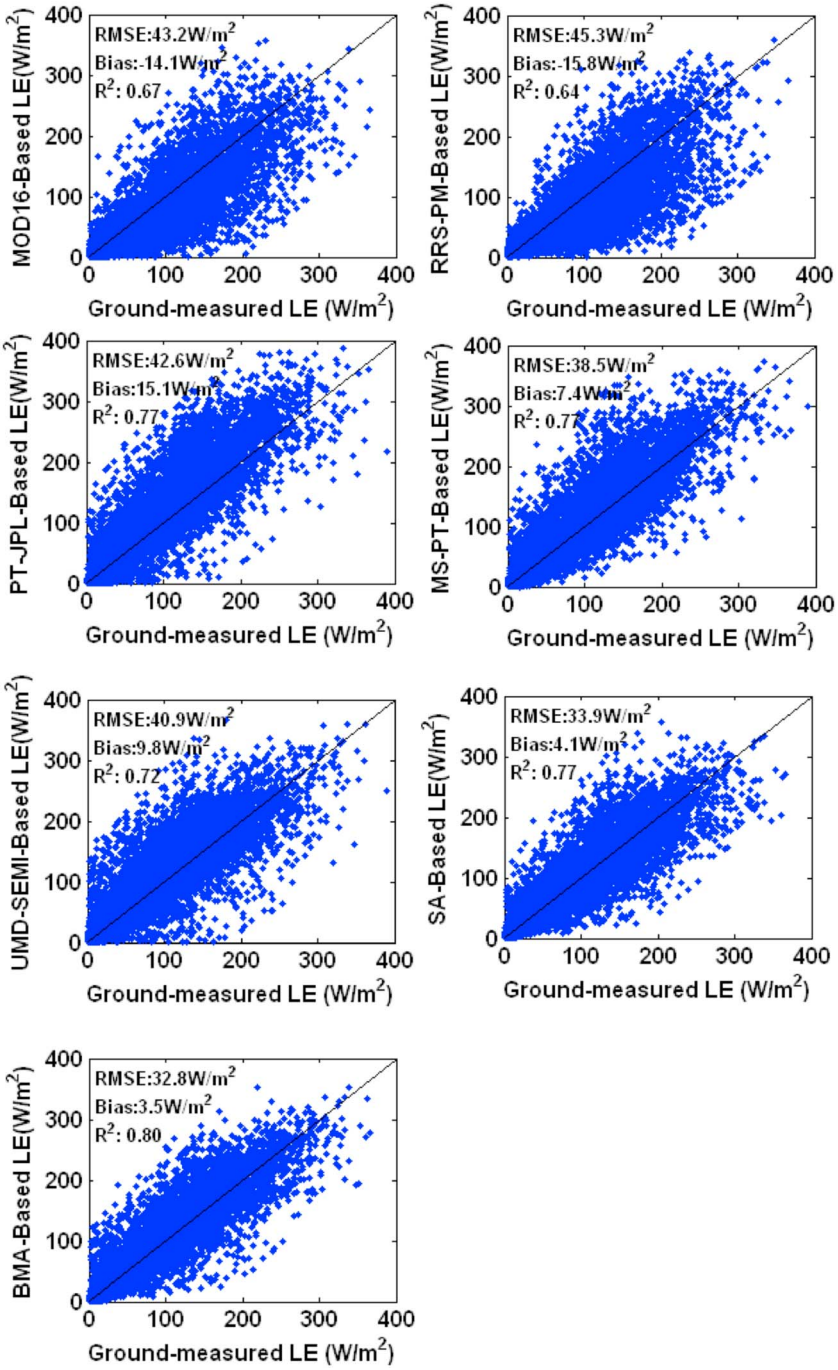

Figure 7. Comparison of monthly LE observations for all 240 flux tower sites and LE estimates using different algorithms driven by tower-specific meteorology.

To estimate global terrestrial LE according to the BMA method, all the data collected at 240 sites were considered as training data to determine weights for the individual LE algorithms for each land cover type. Figure 6 presents the weights for the five process-based LE algorithms for tower-specific and GMAO-MERRA meteorology inputs, respectively. The relative contributions vary for different land cover types. For example, MS-PT has the highest weight for DBF sites and MOD16 has the highest weight for EBF sites for tower-specific meteorology inputs. The MS-PT algorithm weight ranges from 17% to 31% for all land cover types, reflecting its low RMSE for LE estimates compared to the other algorithms. The cross validation also reveals that

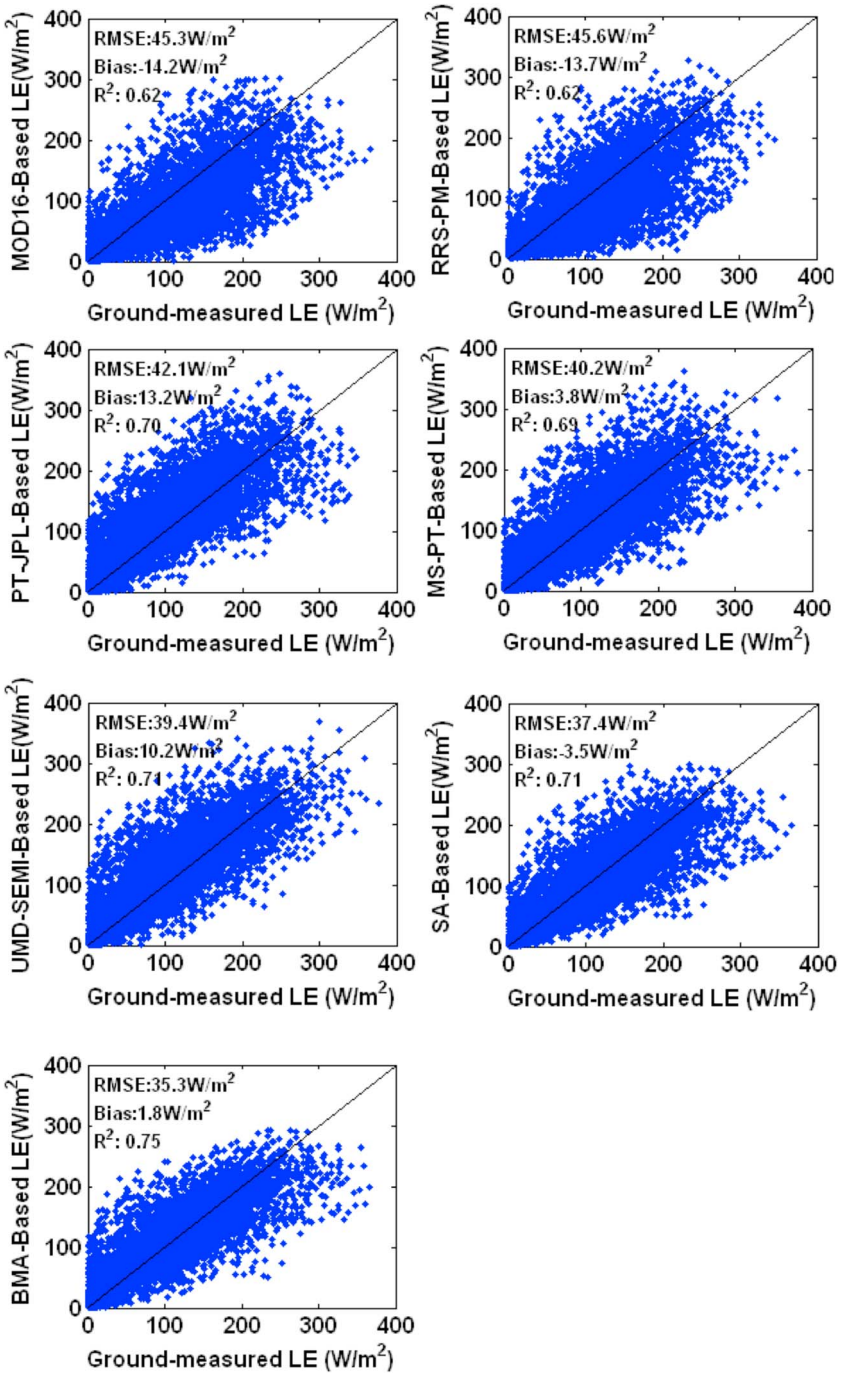

Figure 8. Comparison of monthly LE observations for all 240 flux tower sites and LE estimates using different algorithms driven by GMAO-MERRA meteorology.

MS-PT and UMD-SEMI LE estimates closely match the BMA LE estimate for most land cover types. Therefore, their contributions to the BMA LE are greater than for those of the other algorithms.

Figures 7 and 8 compare monthly LE observations at all 240 sites and LE estimates for the different algorithms driven by tower-specific and GMAO-MERRA meteorology, respectively. The results illustrate that the BMA method yields the best LE estimates, with the lowest RMSE (32.8 and 35.3 W/m^2) and highest R^2 (0.80 and 0.75) ($p < 0.01$) in comparison to the SA method and the individual algorithms for tower and GMAO-MERRA meteorological data, respectively. The error histograms for BMA are more closely centered on zero, and the

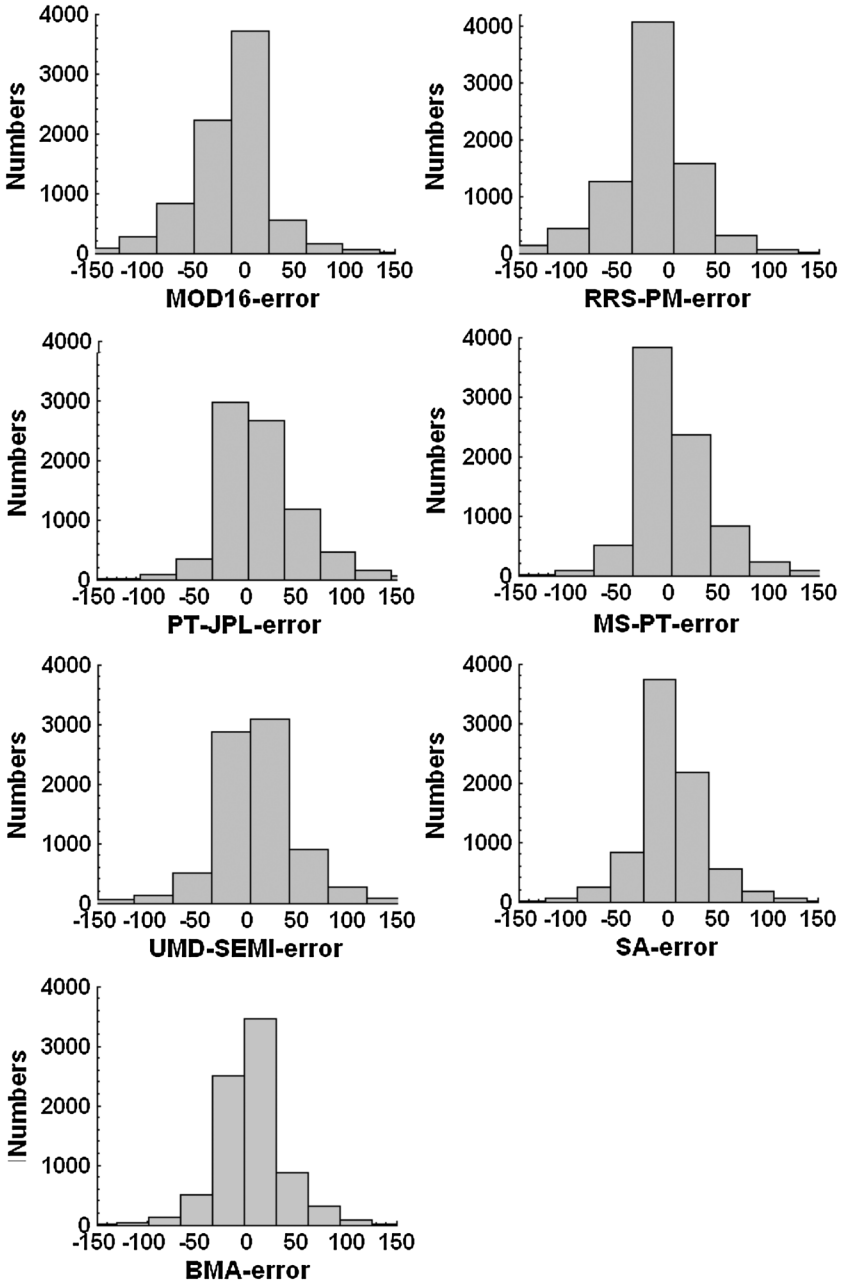

Figure 9. Error histograms for monthly LE according to five process-based algorithms, the BMA method, and the SA method driven by GMAO-MERRA meteorology for all flux towers. The unit for the *x* axis is W/m^2.

individual algorithms are more biased with respect to the tower observations (Figure 9). Previous studies showed that the terrestrial LE retrieved from remote sensing has a relative error of approximately 15–30% [*Kalma et al.*, 2008; *Wang and Dickinson*, 2012] while the relative error of the BMA method driven by eddy covariance, GMAO-MERRA, and satellite observations is about 9.7%, with an average R^2 of 0.87 ($p < 0.01$). Therefore, the BMA method can be used to substantially improve the accuracy of satellite-derived LE estimates.

4.2. Mapping of BMA-Based Global Terrestrial LE

We applied the five process-based algorithms, the SA method, and the BMA method globally for 2001–2004 at a spatial resolution of 0.05° using GMAO-MERRA meteorological data and MODIS products as described in section 2.2. Figure 10 shows maps of annual global terrestrial LE averaged over 2001–2004. Despite the general differences in

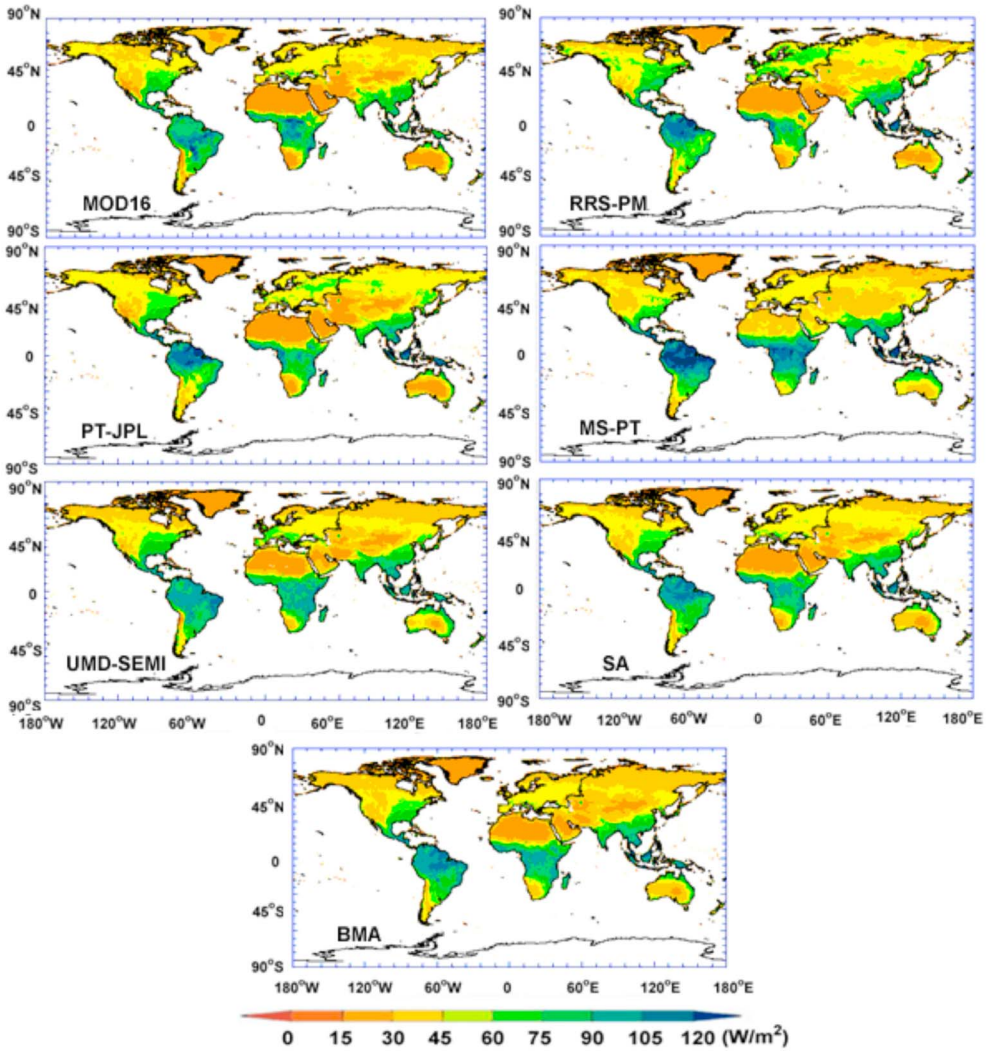

Figure 10. Maps of annual global terrestrial LE averaged for 2001–2004 at spatial resolution of 0.05° according to five process-based algorithms, the SA method, and the BMA method driven by GMAO-MERRA meteorology.

spatial LE distributions among different models, all of the models yield highest annual LE over tropical forests in Africa, South America, and Southeast Asia, whereas arid and desert regions in temperate and subtropical zones and the Arctic have the lowest annual LE owing to moisture limitations and their short growth seasons. Average annual global terrestrial LE according to the BMA method is 38.6 W/m^2, which is higher than values according to the MOD16 (35.2 W/m^2) and RRS-PM algorithms (36.4 W/m^2) and lower than values according to the SA method (39.5W/m^2) and the PT-JPL (40.3 W/m^2), MS-PT (42.1 W/m^2), and UMD-SEMI algorithms (41.5 W/m^2). According to the BMA method, EBF has the highest average LE (88.8 W/m^2), followed by SAW (65.4 W/m^2), DBF (60.1 W/m^2), CRO (45.7 W/m^2), GRA (39.4 W/m^2), MF (37.6 W/m^2), ENF (34.3 W/m^2), SHR (30.6 W/m^2), and DNF (28.9 W/m^2).

5. Discussion

5.1. Performance of the BMA Method

Cross validation for 240 globally distributed EC flux tower sites demonstrated that the BMA approach for merging five process-based LE algorithms is reliable and robust for most land cover types. Figures 3 and 5 both show that the BMA method has no significant LE bias and yields the closest LE to tower flux data relative to the SA method and the individual algorithms. However, for CRO and GRA sites, when more than two of the individual algorithms significantly underestimate LE compared to ground-measured data, BMA merging may

lead to large LE bias. When the two worst algorithms for LE estimation for CRO and GRA sites (MOD16 and RRS-PM) were excluded, the BMA performance greatly improved.

We found that the BMA method shows large interbiome differences and it performs better for DBF, SAW, and CRO sites. For example, BMA can account for 62–85% of the LE variability for 28 globally distributed DBF EC flux tower sites. Several studies have revealed that some satellite-based LE algorithms, such as the surface energy balance system (SEBS), PT-JPL, and the vita version of MOD16, can yield considerably better LE estimates for seasonal vegetation types such as DBF [*Mu et al.*, 2007; *Fisher et al.*, 2008; *Vinukollu et al.,* 2011a, 2011b; *Yebra et al.*, 2013]. These LE algorithms may exhibit strong seasonality for vegetation indices or LAI derived from remotely sensed data for accurate capture of information on seasonal changes in vegetation [*Yebra et al.*, 2013]. Thus, the BMA approach for merging these LE algorithms improves the accuracy of LE quantification. By contrast, the BMA method yields poor LE estimates for EBF sites (average RMSE 56.4 W/m^2 and average R^2 0.58 for GMAO-MERRA inputs; Figure 3). Our results indicate that the five process-based LE algorithms considered here exhibit poor LE modeling performance for EBF sites. This may be attributable in part to the fact that seasonal EBF variation is less evident when satellite-derived vegetation indices (e.g., NDVI) saturate asymptotically and signal contamination of MODIS vegetation indices by broken clouds hampers the provision of reliable vegetation information, especially over tropical forests [*Huete et al.*, 2002; *Demarty et al.*, 2007].

For merging multiple LE algorithms, the BMA method outperforms the SA method for most flux tower sites. For example, daily LE estimates for the first group of MF flux towers, the average RMSE is 38.3 W/m^2 and $R^2 = 0.76$ for the BMA approach, while the average RMSE is 40.7 W/m^2 and $R^2 = 0.70$ for the SA method (Figure 3). This may be because the weights for individual LE algorithms play an important role in algorithm fusion and the BMA method considers an ensemble distribution that has first and second processes to correct bias using ground-measured LE as training data in calculating algorithm weights whereas the SA method calculates weights by simple averaging of values without any auxiliary data. Similar conclusions have been drawn for algorithm fusion for other meteorological variables. For instance, *Raftery et al.* [2005] found that BMA yields a deterministic forecast of sea-level pressure in the Pacific Northwest, with RMSE 11% lower than for any of the ensemble members and 6% lower than for the ensemble mean. We conclude that the BMA method presented here performs better than the SA method and the individual LE algorithms in estimating terrestrial LE.

The accuracy of BMA-based LE estimates depends on the errors for EC LE observations, the LE simulation accuracy of individual algorithms, and selection of the conditional density function. EC LE observations directly determine the accuracy of BMA-based LE estimates because we consider EC LE observations as true values in calculating the weights for individual LE algorithms. However, EC observations do not correspond to true absolute LE because of ambiguous value interpretations, and the typical error for EC LE is approximately 20–50 W/m^2 [*Mahrt*, 2010; *Vickers et al.,* 2010]. An important problem regarding EC observations is an energy imbalance, whereby $LE + H < R_n - G$, and the average energy closure ratio $(LE + H)/(R_n - G)$ for more than 250 FLUXNET flux towers is approximately 0.8 [*Beer et al.*, 2010]. Although *Foken* [2006] documented several reasons for this energy imbalance and we corrected LE data measured at EC flux towers in this study, the uncertainty for energy correction arising from the closure error remains unclear [*Twine et al.*, 2000; *Shuttleworth*, 2007; *Zhang et al.*, 2009, 2010a]. Moreover, the pixel average for MODIS-based LE estimation is 1 km, whereas EC observations can only represent a small scale of several hundred meters [*Li et al.*, 2008; *McCabe and Wood*, 2006], which may lead to large differences between LE observations and the true LE.

The LE accuracy of individual algorithms has a significant impact on the accuracy of the BMA method because it is used to calculate the predictive PDF to get a good fit to EC LE observations. Several previous studies have shown that model input errors, spatial scale mismatch among different data sources, and the limitations of the algorithm itself all affect the accuracy of LE estimated by an individual algorithm [*Mu et al.*, 2007, 2011; *Vinukollu et al.,* 2011a, 2011b; *Shi and Liang*, 2013; *Yebra et al.*, 2013; *Chen et al.*, 2014], which may indirectly affect the performance of the BMA method. Validation results in a recent study revealed bias for both GMAO-MERRA data and MODIS LAI data when compared to ground measurements [*Heinsch et al.,* 2006; *Serbin et al.*, 2013]. This may be an important factor that leads to substantial bias for LE estimated using individual algorithms and BMA ensembles [*Yang et al.*, 2006; *Zhao et al.*, 2006; *Demarty et al.*, 2007; *Mu et al.*, 2011]. In addition, we used GMAO-MERRA products with spatial resolution of 1/2° × 2/3° and MODIS products with resolution of 1 km (LAI/FPAR and NDVI) or 0.05° (NDVI). In both cases, the resolution is greater than footprint for field measurements, which is usually 2–5 m [*Zhao et al.*, 2006; *Baldocchi*, 2008]. Thus, accurate meteorological and vegetation information

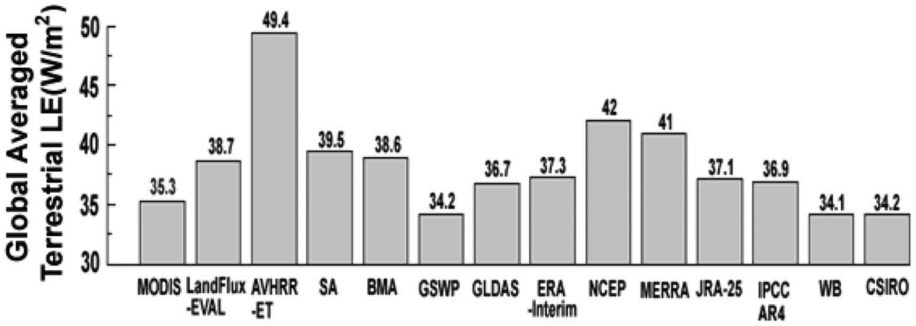

Figure 11. Mean annual LE averaged globally for the different data sets. MODIS, MODIS LE product [*Mu et al.*, 2011]; LandFlux-EVAL, merged benchmark synthesis LE product [*Mueller et al.*, 2013]; AVHRR-ET, GIMMIS-NDVI-based ET product [*Zhang et al.*, 2010a]; SA, SA-based merged LE product used in this study; BMA, BMA-based merged LE product used in this study; GSWP, LE product of the Global Soil Wetness Project [*Dirmeyer et al.*, 2006]; GLDAS, Global Land Data Assimilation System LE product [*Kumar et al.*, 2006]; ERA-Interim, Interim Reanalysis LE product of the European Centre for Medium-Range Weather Forecasts (ECMWF) [*Simmons et al.*, 2006]; NCEP, Reanalysis LE product of the National Centers for Environmental Prediction (NCEP) [*Kalnay et al.*, 1996]; MERRA, Reanalysis LE product of the Modern Era Retrospective Analysis for Research and Applications (MERRA)[*Bosilovich*, 2008]; JRA-25, LE product of the Japanese 25-year Reanalysis [*Onogi et al.*, 2007]; IPCC AR4, LE estimation according to model projections by the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) [*Mueller et al.*, 2011]; WB, ET derived from water balance equations [*Trenberth et al.*, 2007]; CSIRO, ET product of the Commonwealth Scientific and Industrial Research Organization of Australia (CSIRO) [*Zhang et al.*, 2010b].

forflux tower sites cannot be acquired using only these products because of their coarse spatial resolution [*Zhang et al.*, 2010a]. Such representation of the field measurement footprint may also lead to bias in the BMA method for surface LE. Moreover, the LE algorithms we used do not include the effects of CO_2 on LE [*Zhang et al.*, 2009; *Mu et al.*, 2011; *Yao et al.*, 2013] and this limitation will also reduce the accuracy of BMA-based LE.

The conditional density function $h(y|\theta_i)$ is widely recognized as a core regulator of the BMA method for ensembles of meteorological or hydrological variables and it determines the method performance [*Hoeting et al.*, 1999; *Raftery et al.*, 2005; *Duan and Phillips*, 2010]. Here we assumed that $h(y|\theta_i)$ is a reasonable normal density in the BMA method. In general, this works well for surface energy variables such as surface air temperature, and shortwave and longwave radiation [*Wu et al.*, 2012; *Miao et al.*, 2013; *Shi and Liang*, 2013]. However, normal densities may not apply to water variables such as precipitation and runoff because they have a positive probability and their distribution tends to be skewed [*Raftery et al.*, 2005; *Yang et al.*, 2012]. Therefore, a Gamma distribution is a suitable replacement when merging multiple precipitation products [*Yang et al.*, 2012]. In contrast, LE is a complicated variable that couples energy, hydrological, and carbon budgets, and it is difficult to accurately determine the conditional densities for the BMA method. Although we used normal densities to optimize weights for integration of predictive distributions for the five process-based LE algorithms and the cross validation confirmed good model performance, selection of the optimal conditional density function for BMA merging multiple LE algorithms remains a key topic for future research. In addition, we only used normal densities to calculate weights for the five process-based LE algorithms for different land cover types without considering weight differences for different growth seasons and climate spaces for land vegetation. However, some vegetation, such as GRA, represents a great variety in climate space, and the behavior and parameter values for LE may be slightly different for midlatitude and arctic/polar regions [*Priestley and Taylor*, 1972; *Norman et al.*, 1995; *Wang and Dickinson*, 2012]. Thus, the accuracy of the BMA method for LE estimation may be slightly lower for different growth seasons and climate spaces. The effects of growth seasons and climate spaces on the *LE* algorithm weights will be introduced in future work to improve the accuracy of long-term global terrestrial LE estimates.

5.2. BMA-Based Global Terrestrial LE Estimation

Although rigorous validation of global terrestrial LE is difficult owing to the lack of spatial continuous flux measurements for heterogeneous continental landscapes, we demonstrated the reliability of the BMA method by comparison with other hydrological and satellite models. The BMA-based estimate of annual average global terrestrial LE is 38.6 W/m^2, which is comparable to simulated LE values reported in the literature. *Trenberth et al.* [2007] estimated global land-surface ET using a residual for precipitation and runoff

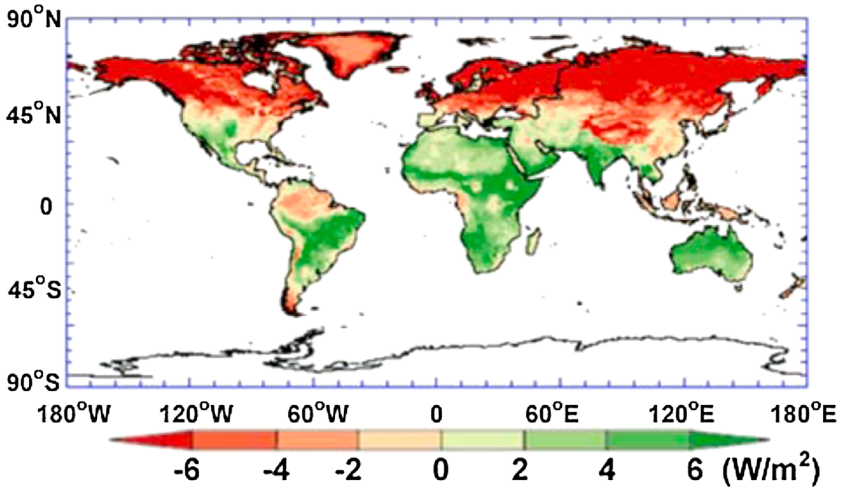

Figure 12. Spatial differences in average annual global terrestrial LE (2001–2004) between the BMA and SA methods.

and found that the volume of water evaporated annually is 72.6×10^3 km^3 (34.1 W/m^2). The annual global LE estimate using 15 models for GSWP-2 is approximately 34.2 W/m^2 [*Dirmeyer et al.*, 2006; *Wang and Dickinson*, 2012]. *Zhang et al.* [2010a] reported average annual ET of 635 ± 200 mm (49.4 ± 15.6 W/m^2) based on a PM combination equation driven by 1983–2006 satellite records. *Wang and Dickinson* [2012] compared 17 different LE data sets and inferred that the global average LE estimated from multiple models is between 1.2 mm/d (34.1 W/m^2) and 1.5 mm/d (42.7 W/m^2), with an average of 1.3mm/d (36.9 W/m^2). Figure 11 shows the mean annual values averaged globally for the different data sets and the BMA method yields a result that falls within the Wang and Dickinson range.

Spatial differences between BMA-merged LE and other LE estimates according to different algorithms are much greater than those for the global mean values. Figure 12 shows spatial differences in annual global terrestrial LE between the BMA and SA methods. Relative to the SA method, the BMA method yields higher annual global terrestrial LE in the Southern Hemisphere and lower LE estimates over almost all high-latitude regions in the Northern Hemisphere. This spatial dissimilarity is mainly caused by differences in the weights of different LE algorithms for the BMA method and SA method, since the bias-correction-based BMA method, weighted by recent performance results for the multiple models, has higher accuracy than the SA method [*Duan and Phillips*, 2010; *Wu et al.*, 2012; *Miao et al.*, 2013]. Although all the LE models used in this study have been evaluated against EC observations, there are marked spatial patterns in the accuracy for each algorithm (Figure 10). Different algorithm parameterizations have an impact on their simulation by partitioning the surface energy flux in a different manner [*Robock et al.*, 2003]. Further work is still required to compare and explain the differences between BMA-based LE and other LE products.

6. Conclusions

We described a BMA method that merges five process-based LE algorithms (MOD16, RRS-PM, PT-JPL, MS-PT, and UMD-SEMI) for estimation of global terrestrial LE. Different LE algorithms driven by tower-specific meteorology and GMAO-MERRA meteorology were evaluated using ground-based data for 2000–2009 collected from 240 flux tower sites across the world on all continents except for Antarctica. Compared to the SA method, the BMA approach yields better LE results. Performance analysis for the BMA method for merging five LE algorithms with ground-based observations and GMAO-MERRA meteorology demonstrated that the method is applicable for global terrestrial LE mapping.

The evaluation results for terrestrial LE estimation using five process-based algorithms were grouped according to nine major terrestrial biomes (DBF, DNF, EBF, ENF, MF, SAW, SHR, CRO, and GRA). The validation results for algorithms driven by tower-specific meteorology and GMAO-MERRA meteorology show that the MS-PT and UMD-SEMI algorithms yield better performance than the other algorithms for SHR, GRA, CRO,

SAW, and most of the forest types. However, for EBF sites, the tower-based MOD16 algorithm has the lowest average RMSE, with an average bias of −12.2 W/m^2 when driven by tower-specific meteorology. For cropland and grass sites, the UMD-SEMI algorithm has lower RMSEs than the other LE algorithms. The UMD-SEMI algorithm shows good performance because its regression coefficients are calibrated using data from 64 flux tower sites.

These five process-based algorithms driven by tower and GMAO-MERRA meteorological data were merged according to the BMA and SA methods using daily weights from the LE estimates and ground-based data. The cross-validation results demonstrate that the RMSEs are lower and R^2 is approximately 0.05 greater for the BMA method compared to the SA method and each individual LE algorithm. The weights for different algorithms were also calculated for different land cover types because the relative contributions of each algorithm vary. The errors for EC LE observations, the accuracy of LE simulated by a single algorithm, and selection of the conditional density function all influence the accuracy of the BMA method. Thus, it is still necessary to explore the weights for different growing seasons and climate spaces for land vegetation.

The BMA method for merging five process-based LE algorithms was applied globally using data for GMAO-MERRA, MODIS land cover, LAI/FPAR, NDVI, and albedo. The annual average global terrestrial LE for 2001–2004 estimated by the BMA method is approximately 38.6 W/m^2, which is in good agreement with other studies. Currently, this LE algorithm is being used to produce the Global Land-Surface Satellite (GLASS) LE product, which will be distributed by the Center for Global Change Data Processing and Analysis of Beijing Normal University, China.

Appendix A

Table A1 shows a list of acronyms used in this study.

Table A1. Acronyms Used in This Study

Acronym	Definition
ANN	Asian Automatic Weather Station Network
AR4	Fourth Assessment Report
ATI	Apparent Thermal Inertia
AVHRR	Advanced Very High Resolution Spectroradiometer
BMA	Bayesian Model Averaging
CEOP	Coordinated Enhanced Observation Network of China
CERN	Chinese Ecosystem Research Network
CRO	Cropland
CSIRO	Commonwealth Scientific and Industrial Research Organization of Australia
DA	Data Assimilation
DBF	Deciduous Broadleaf Forest
DNF	Deciduous Needleleaf Forest
DT	Diurnal Air-Temperature Range
e	Vapor Pressure
E	Conditional Expectation
EBF	Evergreen Broadleaf Forest
EC	Eddy Covariance
ECMWF	European Centre for Medium-Range Weather Forecasts
EF	Evaporation Fraction
ENF	Evergreen Needleleaf Forest
ERA-Interim	ERA-Interim Reanalysis
ET	Evapotranspiration
EVI	Enhanced Vegetation Index
FPAR	Photosynthetically Active Radiation
G	Ground Heat Flux
GEWEX	Global Energy and Water Cycle Experiment
GLASS	Global Land-Surface Satellite
GIMMS	Global Inventory Modeling and Mapping Studies
GLDAS	Global Land Data Assimilation System
GMAO	Global Modeling and Assimilation Office

Acknowledgments
Authors thank the three anonymous reviewers for their critical and helpful comments and suggestions. Authors also thank Shaomin Liu, Wenping Yuan, and Ziwei Xu from Beijing Normal University, China; Guangsheng Zhou from the Institute of Botany, CAS; Yan Li and Ran Liu from Xinjiang Institute of Ecology and Geography, CAS; Guoyi Zhou and Yuelin Li from South China Botanic Garden, CAS; and Bin Zhao from Fudan University, China, for providing ground-measured data. 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, and 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 U.S. 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 ANN (http://aan.suiri.tsukuba.ac.jp/), the Coordinated Enhanced Observation Project (CEOP) in arid and semiarid 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). MODIS LAI/FPAR, NDVI, albedo, and land cover satellite products were obtained online (http://reverb.echo.nasa.gov/reverb). This work was partially supported by the High-Tech Research and Development Program of China (2013AA122801), the Natural Science Fund of China (41201331, 41101310, 41101313, 41301353, and 41205104), the National Basic Research Program of China (2012CB955302), the Fundamental Research Funds for the Central Universities (2013YB34), the High Resolution Earth Observation Systems of National Science and Technology Major Projects (05-Y30B02-9001-13/15-9). J.B.F. contributed to this paper from the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Table A1. (continued)

Acronym	Definition
GRA	Grass and other types
GSWP-2	Global Soil Wetness Project-2
H	Sensible Heat Flux
H_{ori}	Uncorrected Sensible Heat Flux
IPCC	Intergovernmental Panel on Climate Change
ISLSCP-II	International Satellite Land Surface Climatology Project, Initiative II
JRA-25	Japanese 25-year Reanalysis
LAI	Leaf Area Index
LandFlux-EVAL	Merged Benchmark Synthesis Products of ET
LE	Latent Heat Flux
LE_{ori}	Uncorrected Latent Heat Flux
LSM	Land Surface Model
LST	Land Surface Temperature
MERRA	Modern Era Retrospective Analysis for Research and Applications
MF	Mixed Forest
MODIS	Moderate Resolution Imaging Spectroradiometer
MOD15A2	MODIS FPAR/LAI product
MOD13A2	MODIS NDVI/EVI
MOD43B3	MODIS albedo
NCEP	National Centers for Environmental Prediction
NDVI	Normalized Difference Vegetation Index
PDF	Probability Density Function
PI	Principal Investigator
PM	Penman-Monteith
PT	Priestley-Taylor
R^2	Squared Correlation Coefficients
RH	Relative Humidity
RMSE	Root Mean Square Error
R_n	Surface Net Radiation
R_s	Incident Solar Radiation
SA	Simple Model Averaging
SAW	Savanna
SEB	Surface Energy Balance
SEBS	Surface Energy Balance System
SEMI	Statistical and Empirical Method
SHR	Shrubland
STD	Standard Deviation
T_a	Air Temperature
T_{max}	Daily Maximum Air Temperature
T_{min}	Daily Minimum Air Temperature
VPD	Vapor Pressure Deficit
WB	Water Balance
WS	Wind Speed

References

Allen, R. G., M. Tasumi, and R. Trezza (2007), Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC) model, *J. Irrig. Drain. Eng.*, *133*(4), 380–394, doi:10.1061/(ASCE)0733-9437(2007)133:4(380).

Anderson, M. C., J. M. Norman, G. R. Diak, W. P. Kustas, and J. R. Mecikalski (1997), A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing, *Remote Sens. Environ.*, *60*(2), 195–216, doi:10.1016/S0034-4257(96)00215-5.

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, doi:10.1071/BT07151.

Baldocchi, D., et al. (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*, 2415–2434, doi:10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2.

Bastiaanssen, W. G. M., M. Menenti, R. A. Feddes, and A. A. M. Holtslag (1998), A remote sensing surface energy balance algorithm for land (SEBAL): 1. Formulation, *J. Hydrol.*, *212-213*(1–4), 198–212, doi:10.1016/S0022-1694(98)00253-4.

Beer, C., et al. (2010), Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate, *Science*, *329*(5993), 834–838, doi:10.1126/science.1184984.

Bosilovich, M. (2008), NASA's Modern Era Retrospective-Analysis for Research and Applications: Integrating Earth observations, *EarthZine*.

Chen, Y., et al. (2014), Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China, *Remote Sens. Environ.*, *140*, 279–293.

Cleugh, H. A., R. Leuning, Q. Mu, and S. W. Running (2007), Regional evaporation estimates from flux tower and MODIS satellite data, *Remote Sens. Environ.*, *106*(3), 285–304, doi:10.1016/j.rse.2006.07.007.
Demarty, J., F. Chevallier, A. D. Friend, N. Viovy, S. Piao, and P. Ciais (2007), Assimilation of global MODIS leaf area index retrievals within a terrestrial biosphere model, *Geophys. Res. Lett.*, *34*, L15402, doi:10.1029/2007GL030014.
Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki (2006), GSWP-2: Multimodel analysis and implications for our perception of the land surface, *Bull. Am. Meteorol. Soc.*, *87*, 1381–1397.
Duan, Q., and T. J. Phillips (2010), Bayesian estimation of local signal and noise in multimodel simulations of climate change, *J. Geophys. Res.*, *115*, D18123, doi:10.1029/2009JD013654.
Duan, Q., N. K. Ajami, X. Gao, and S. Sorooshian (2007), Multi-model ensemble hydrologic prediction using Bayesian model averaging, *Adv. Water Resour.*, *30*(5), 1371–1386, doi:10.1016/j.advwatres.2006.11.014.
Eugster, W., et al. (2000), Land-atmosphere energy exchange in Arctic tundra and boreal forest: Available data and feedbacks to climate, *Global Change Biol.*, *6*, 84–115, doi:10.1046/j.1365-2486.2000.06015.x.
Fisher, J., K. Tu, and D. Baldocchi (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*, 901–919.
Foken, T. (2006), 50 years of the Monin-Obukhov similarity theory, *Boundary Layer Meteorol.*, *119*(3), 431–447, doi:10.1007/s10546-006-9048-6.
Friedl, M. A., et al. (2002), Global land cover mapping from MODIS: Algorithms and early results, *Remote Sens. Environ.*, *83*(1–2), 287–302.
Hall, F., E. B. de Colstoun, G. Collatz, D. Landis, P. Dirmeyer, A. Betts, G. Huffman, L. Bounoua, and B. Meeson (2006), ISLSCP Initiative II global data sets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies, *J. Geophys. Res.*, *111*, D22S01, doi:10.1029/2006JD007366.
Heinsch, F. A., et al. (2006), Evaluation of remote sensing based terrestrial productivity from MODIS using AmeriFlux tower eddy flux network observations, *IEEE Trans. Geosci. Remote Sens.*, *44*(7), 1908–1925.
Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999), Bayesian model averaging: A tutorial, *Stat. Sci.*, *14*(4), 382–417.
Houghton, J. T. A. (Ed.) (2001), *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, 892 pp., Cambridge Univ. Press, New York.
Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira (2002), Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sens. Environ.*, *83*, 195–213.
Jackson, R., R. Reginato, and S. Idso (1977), Wheat canopy temperature: A practical tool for evaluating water requirements, *Water Resour. Res.*, *13*, 651–656.
Jia, Z., S. Liu, Z. Xu, Y. Chen, and M. Zhu (2012), Validation of remotely sensed evapotranspiration over the Hai River Basin, China, *J. Geophys. Res.*, *117*, D13113, doi:10.1029/2011JD017037.
Jiang, L., and S. Islam (2001), Estimation of surface evaporation map over Southern Great Plains using remote sensing data, *Water Resour. Res.*, *37*, 329–340, doi:10.1029/2000WR900255.
Jiménez, C., et al. (2011), Global intercomparison of 12 land surface heat flux estimates, *J. Geophys. Res.*, *116*, D02102, doi:10.1029/2010JD014545.
Jin, Y., C. B. Schaaf, F. Gao, X. Li, A. H. Strahler, W. Lucht, and S. Liang (2003), Consistency of MODIS surface BRDF/Albedo retrievals: 1. Algorithm performance, *J. Geophys. Res.*, *108*(D5), 4158, doi:10.1029/2002JD002803.
Jin, Y., J. Randerson, and M. Goulden (2011), Continental-scale net radiation and evapotranspiration estimated using MODIS satellite observations, *Remote Sens. Environ.*, *115*, 2302–2319.
Jung, M., et al. (2010), Recent decline in the global land evapotranspiration trend due to limited moisture supply, *Nature*, *467*(7318), 951–954, doi:10.1038/nature09396.
Kalma, J., T. McVicar, and M. McCabe (2008), Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data accomplished, *Surv. Geophys.*, *29*, 421–469, doi:10.1007/s10712-008-9037-z.
Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.
Katul, G., R. Oren, S. Manzoni, C. Higgins, and M. Parlange (2012), Evapotranspiration: A process driving mass transport and energy exchange in the soil-plant-atmosphere-climate system, *Rev. Geophys.*, *50*, RG3002, doi:10.1029/2011RG000366.
Kim, H., K. Hwang, Q. Mu, S. Lee, and M. Choi (2012), Validation of MODIS 16 global terrestrial evapotranspiration products in various climates and land cover types in Asia, *KSCE J. Civ. Eng.*, *16*(2), 229–238.
Kumar, S. V., et al. (2006), Land information system—An interoperable framework for high resolution land surface modeling, *Environ. Modell. Software*, *21*, 1402–1415.
Kustas, W. P., and J. M. Norman (1996), Use of remote sensing for evapotranspiration monitoring over land surfaces, *Hydrol. Sci. J.*, *41*(4), 495–516, doi:10.1080/02626669609491522.
Li, F. Q., W. P. Kustas, M. C. Anderson, J. H. Prueger, and T. L. Scott (2008), Effect of remote sensing spatial resolution on interpreting tower-based flux observations, *Remote Sens. Environ.*, *112*(2), 337–349.
Li, Z., R. Tang, Z. Wan, Y. Bi, C. Zhou, B. Tang, G. Yan, and X. Zhang (2009), A review of current methodologies for regional evapotranspiration estimation from remotely sensed data, *Sensors*, *9*(5), 3801–3853, doi:10.3390/s90503801.
Liang, S., A. H. Strahler, and C. W. Walthall (1999), Retrieval of land surface albedo from satellite observations: A simulation study, *J. Appl. Meteorol.*, *38*, 712–725.
Liang, S., K. Wang, X. Zhang, and M. Wild (2010), Review on estimation of land surface radiation and energy budgets from ground measurements, remote sensing and model simulations, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, *3*(3), 225–240, doi:10.1109/JSTARS.2010.2048556.
Liu, S., Z. Xu, W. Wang, Z. Jia, M. Zhu, J. Bai, and J. Wang (2011), A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy balance closure problem, *Hydrol. Earth Syst. Sci.*, *15*, 1291–1306.
Liu, S., Z. Xu, Z. Zhu, Z. Jia, and M. Zhu (2013), Measurements of evapotranspiration from eddy-covariance systems and large aperture scintillometers in the Hai River Basin, China, *J. Hydrol.*, *487*, 24–38.
Los, S., N. H. Pollack, M. T. Parris, G. J. Collatz, C. J. Tucker, P. J. Sellers, C. M. Malmström, R. S. DeFries, L. Bounoua, and D. A. Dazlich (2000), A global 9-yr biophysical land surface dataset from NOAA AVHRR data, *J. Hydrometeorol.*, *1*(2), 183–199.
Madigan, D., A. E. Raftery, C. T. Volinsky, and J. A. Hoeting (1999), Bayesian model averaging, *Stat. Sci.*, *14*, 382–401.
Mahrt, L. (2010), Computing turbulent fluxes near the surface: Needed improvements, *Agric. For. Meteorol.*, *150*(4), 501–509, doi:10.1016/j.agrformet.2010.01.015.
McCabe, M. F., and E. F. Wood (2006), Scale influences on the remote estimation of evapotranspiration using multiple satellite sensors, *Remote Sens. Environ.*, *105*(4), 271–285, doi:10.1016/j.rse.2006.07.006.
McVicar, T. R., et al. (2012), Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation, *J. Hydrol.*, *416–417*, 182–205.

Miao, C., Q. Duan, Q. Sun, and J. Li (2013), Evaluation and application of Bayesian multi-model estimation in temperature simulations, *Prog. Phys. Geogr.*, *37*(6), 727–744.
Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman (2011), Global land-surface evaporation estimated from satellite-based observations, *Hydrol. Earth Syst. Sci.*, *15*, 453–469.
Mo, X., Z. Lin, H. Li, and Y. Xiang (2004), Simulation of winter wheat yield and evapotranspiration with process-based model and remote sensing data in the Hebei plain [in Chinese], *Geogr. Res.*, *23*(5), 623–631.
Monteith, J. (1965), Evaporation and environment, *Symp. Soc. Exp. Biol.*, *19*, 205–224.
Mueller, B., et al. (2011), Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, *Geophys. Res. Lett.*, *38*, L06402, doi:10.1029/2010GL046230.
Mueller, B., et al. (2013), Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, *Hydrol. Earth Syst. Sci.*, *17*, 3707–3720, doi:10.5194/hess-17-3707-2013.
Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running (2007), Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, *Remote Sens. Environ.*, *111*(4), 519–536, doi:10.1016/j.rse.2007.04.015.
Mu, Q., M. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sens. Environ.*, *115*(8), 1781–1800, doi:10.1016/j.rse.2011.02.019.
Myneni, R., et al. (2002), Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, *Remote Sens. Environ.*, *83*(1–2), 214–231.
National Research Council (2007), *Earth Science and Applications From Space: National Imperatives for the Next Decade and Beyond*, 400 pp., Natl. Acad. Press, Washington, D. C.
Norman, J. M., W. P. Kustas, and K. S. Humes (1995), Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature, *Agric. For. Meteorol.*, *77*(3–4), 263–293, doi:10.1016/0168-1923(95)02265-Y.
Onogi, K., et al. (2007), The JRA-25 reanalysis, *J. Meteorol. Soc. Jpn.*, *85*, 369–432.
Penman, H. L. (1948), Natural evaporation from open water, bare soil and grass, *Proc. R. Soc. London, Ser. A*, *193*, 120–145, doi:10.1098/rspa.1948.0037.
Pipunic, R., J. Walker, and A. Western (2008), Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: A comparative synthetic study, *Remote Sens. Environ.*, *112*, 1295–1305.
Priestley, C. H. B., and R. J. Taylor (1972), On the assessment of surface heat flux and evaporation using large-scale parameters, *Mon. Weather Rev.*, *100*(2), 81–92, doi:10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2.
Qin, J., S. Liang, R. Liu, H. Zhang, and B. Hu (2007), A weak-constraint based data assimilation scheme for estimating surface turbulent fluxes, *IEEE Geosci. Remote Sens.*, *4*(4), 649–653.
Raftery, A. E., D. Madigan, and J. A. Hoeting (1997), Bayesian model averaging for linear regression models, *J. Am. Stat. Assoc.*, *92*, 179–191, doi:10.1080/01621459.1997.10473615.
Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski (2005), Using Bayesian model averaging to calibrate forecast ensembles, *Mon. Weather Rev.*, *133*, 1155–1174, doi:10.1175/MWR2906.1.
Robock, A., et al. (2003), Evaluation of the North American land data assimilation system over the southern Great Plains during the warm season, *J. Geophys. Res.*, *108*(D22), 8846, doi:10.1029/2002JD003245.
Rodriguez-Iturbe, I., P. D'Odorico, F. Laio, L. Ridolfi, and S. Tamea (2007), Challenges in humid land ecohydrology: Interactions of water table and unsaturated zone with climate, soil, and vegetation, *Water Resour. Res.*, *43*, W09301, doi:10.1029/2007WR006073.
Salomon, J., C. B. Schaaf, A. H. Strahler, F. Gao, and Y. Jin (2006), Validation of the MODIS bidirectional reflectance distribution function and albedo retrievals using combined observations from the aqua and terra platforms, *IEEE Trans. Geosci. Remote Sens.*, *44*(6), 1555–1565.
Schaaf, C., et al. (2002), First operational BRDF, albedo and nadir reflectance products from MODIS, *Remote Sens. Environ.*, *83*, 135–148.
Serbin, S., D. Ahl, and S. Gower (2013), Spatial and temporal validation of the MODIS LAI and FPAR products across a boreal forest wildfire chronosequence, *Remote Sens. Environ.*, *133*, 71–84.
Serreze, M., A. Barrett, and F. Lo (2005), Northern high-latitude precipitation as depicted by atmospheric reanalyses and satellite retrievals, *Mon. Weather Rev.*, *133*, 3407–3430.
Shi, Q., and S. Liang (2013), Characterizing the surface radiation budget over the Tibetan Plateau with ground-measured, reanalysis, and remote sensing data sets: 1. Methodology, *J. Geophys. Res. Atmos.*, *118*, 9642–9657, doi:10.1002/jgrd.50720.
Shuttleworth, W. J. (2007), Putting the "vap" into evaporation, *Hydrol. Earth Syst. Sci.*, *11*, 210–244.
Simmons, A., S. Uppala, D. Dee, and S. Kobayashi (2006), ERA-Interim: New ECMWF reanalysis products from1989 onwards, *ECMWF Newsl.*, *110*, 25–35.
Skidmore, E. L., H. S. Jacobs, and W. L. Powers (1969), Potential evapotranspiration as influenced by wind, *Agron. J.*, *61*, 543–546.
Sun, G., S. McNulty, D. Amatya, Y. Liang, and R. Kolka (2005), Regional annual water yield from forest lands and its response to potential deforestation across the southeastern United States, *J. Hydrol.*, *308*, 258–268.
Tang, R., Z.-L. Li, and B. Tang (2010), An application of the Ts-VI triangle method with enhanced edges determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: Implementation and validation, *Remote Sens. Environ.*, *114*(3), 540–551, doi:10.1016/j.rse.2009.10.012.
Taylor, K. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*, 7183–7192, doi:10.1029/2000JD900719.
Trenberth, K. E., L. Smith, T. T. Qian, A. Dai, and J. Fasullo (2007), Estimates of the global water budget and its annual cycle using observational and model data, *J. Hydrometeorol.*, *8*, 758–769.
Twine, T. E., W. P. Kustas, J. M. Norman, D. R. Cook, P. R. Houser, T. P. Meyers, J. H. Prueger, P. J. Starks, and M. L. Wesely (2000), Correcting eddy-covariance flux underestimates over a grassland, *Agric. For. Meteorol.*, *103*(3), 279–300, doi:10.1016/S0168-1923(00)00123-4.
Vickers, D., M. Gockede, and B. E. Law (2010), Uncertainty estimates for 1-h averaged turbulence fluxes of carbon dioxide, latent heat and sensible heat, *Tellus, Ser. B*, *62*(2), 87–99, doi:10.1111/j.1600-0889.2009.00449.x.
Vinukollu, R. K., E. F. Wood, C. R. Ferguson, and J. B. Fisher (2011a), Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches, *Remote Sens. Environ.*, *115*, 801–823.
Vinukollu, R. K., R. Meynadier, J. Sheffield, and E. F. Wood (2011b), Multi-model, multi-sensor estimates of global evapotranspiration: Climatology, uncertainties and trends, *Hydrol. Processes*, *25*, 3993–4010, doi:10.1002/hyp.8393.
Wang, K., and R. Dickinson (2012), A review of global terrestrial evapotranspiration: Observation, modeling, climatology and climatic variability, *Rev. Geophys.*, *50*, RG2005, doi:10.1029/2011RG000373.
Wang, K., and S. Liang (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*, 712–727, doi:10.1175/2007JHM911.1.

Wang, K., P. Wang, Z. Q. Li, M. Cribb, and M. Sparrow (2007), A simple method to estimate actual evapotranspiration from a combination of net radiation, vegetation index, and temperature, *J. Geophys. Res.*, *112*, D15107, doi:10.1029/2006JD008351.

Wang, K., R. Dickinson, M. Wild, and S. Liang (2010a), Evidence for decadal variation in global terrestrial evapotranspiration between 1982 and 2002. Part 1: Model development, *J. Geophys. Res.*, *115*, D20112, doi:10.1029/2009JD013671.

Wang, K., R. Dickinson, M. Wild, and S. Liang (2010b), Evidence for decadal variation in global terrestrial evapotranspiration between 1982 and 2002. Part 2: Results, *J. Geophys. Res.*, *115*, D20113, doi:10.1029/2010JD013847.

Wilson, K., et al. (2002), Energy balance closure at FLUXNET sites, *Agric. For. Meteorol.*, *113*(1–4), 223–243, doi:10.1016/S0168-1923(02)00109-0.

Wu, H., X. Zhang, S. Liang, H. Yang, and G. Zhou (2012), Estimation of clear-sky land surface longwave radiation from MODIS data products by merging multiple models, *J. Geophys. Res.*, *117*, D22107, doi:10.1029/2012JD017567.

Xu, T., S. Liang, and S. Liu (2011a), Estimating turbulent fluxes through assimilation of geostationary operational environmental satellites data using ensemble Kalman filter, *J. Geophys. Res.*, *116*, D09109, doi:10.1029/2010JD015150.

Xu, T., S. Liu, S. Liang, and J. Qin (2011b), Improving predictions of water and heat fluxes by assimilating MODIS land surface temperature products into common land model, *J. Hydrometeorol.*, *12*(2), 227–244.

Xu, Z., S. Liu, X. Li, S. Shi, J. Wang, Z. Zhu, T. Xu, W. Wang, and M. Ma (2013), Intercomparison of surface energy flux measurement systems used during the HiWATER-MUSOEXE, *J. Geophys. Res. Atmos.*, *118*, 13,140–13,157, doi:10.1002/2013JD020260.

Yang, C., Z. Yan, and Y. Shao (2012), Probabilistic precipitation forecasting based on ensemble output using generalized additive models and Bayesian model averaging, *Acta Meteorol. Sin.*, *26*(1), 1–12, doi:10.1007/s13351-012-0101-8.

Yang, F., M. A. White, A. R. Michaelis, K. Ichii, H. Hashimoto, P. Votava, A.-X. Zhu, and R. R. Nemani (2006), Prediction of continental-scale evapotranspiration by combining MODIS and AmeriFlux data through support vector machine, *IEEE Trans. Geosci. Remote Sens.*, *44*, 3452–3461, doi:10.1109/TGRS.2006.876297.

Yao, Y., et al. (2013), MODIS-driven estimation of terrestrial latent heat flux in China based on a modified Priestley-Taylor algorithm, *Agric. For. Meteorol.*, *171–172*, 187–202, doi:10.1016/j.agrformet.2012.11.016.

Yao, Y., S. Liang, S. Zhao, Y. Zhang, Q. Qin, J. Cheng, K. Jia, X. Xie, N. Zhang, and M. Liu (2014), Validation and application of the modified satellite-based Priestley-Taylor algorithm for mapping terrestrial evapotranspiration, *Remote Sens.*, *6*, 880–904, doi:10.3390/rs6010880.

Yao, Y., S. Liang, Q. Qin, K. Wang, and S. Zhao (2011a), Monitoring global land surface drought based on a hybrid evapotranspiration model, *Int. J. Appl. Earth Obs.*, *13*, 447–457, doi:10.1016/j.jag.2010.09.009.

Yao, Y., Q. Qin, A. Ghulam, S. Liu, S. Zhao, Z. Xu, and H. Dong (2011b), Simple method to determine the Priestley–Taylor parameter for evapotranspiration estimation using Albedo-VI triangular space from MODIS data, *J. Appl. Remote Sens.*, *5*, 053,505, doi:10.1117/1.3557817.

Yebra, M., A. Van Dijk, R. Leuning, A. Huete, and J. P. Guerschman (2013), Evaluation of optical remote sensing to estimate actual evapotranspiration and canopy conductance, *Remote Sens. Environ.*, *129*, 250–261.

Yuan, W., et al. (2010), Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data, *Remote Sens. Environ.*, *114*, 1416–1431.

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

Zhang, K., J. S. Kimball, R. R. Nemani, and S. W. Running (2010a), A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006, *Water Resour. Res.*, *46*, W09522, doi:10.1029/2009WR008800.

Zhang, Y., R. Leuning, L. B. Hutley, J. Beringer, I. McHugh, and J. P. Walker (2010b), Using long-term water balances to parameterize surface conductances and calculate evaporation at 0.05° spatial resolution, *Water Resour. Res.*, *46*, W05512, doi:10.1029/2009WR008716.

Zhao, M., and S. W. Running (2010), Drought-induced reduction in global terrestrial net primary production from 2000 through 2009, *Science*, *329*(5994), 940–943, doi:10.1126/science.1192666.

Zhao, M., F. A. Heinsch, R. Nemani, and S. W. Running (2005), Improvements of the MODIS terrestrial gross and net primary production global data set, *Remote Sens. Environ.*, *95*, 164–176.

Zhao, M., S. W. Running, and R. R. Nemani (2006), Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses, *J. Geophys. Res.*, *111*, G01002, doi:10.1029/2004JG000004.