

Evaluating Global Land Surface Models in CMIP5: Analysis of Ecosystem Water- and Light-Use Efficiencies and Rainfall Partitioning

LONGHUI LI,^{a,b} YINGPING WANG,^{c,d} VIVEK K. ARORA,^e DEREK EAMUS,^b HAO SHI,^b JING LI,^f LEI CHENG,^g JAMES CLEVERLY,^b T. HAJIMA,^h DUOYING JI,ⁱ C. JONES,^j M. KAWAMIYA,^h WEIPING LI,^k J. TJIPUTRA,^l A. WILTSHIRE,^j LU ZHANG,^g AND QIANG YU^b

^a *Jiangsu Centre for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing Normal University, Nanjing, China*

^b *School of Life Sciences, University of Technology Sydney, Sydney, New South Wales, Australia*

^c *CSIRO Oceans and Atmosphere, Aspendale, Victoria, Australia*

^d *South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China*

^e *Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia, Canada*

^f *State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences and Beijing Normal University, Beijing, China*

^g *CSIRO Land and Water, Canberra, Australian Capital Territory, Australia*

^h *Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan*

ⁱ *Land–Atmosphere Interaction Research Group, Beijing Normal University, Beijing, China*

^j *Met Office Hadley Centre, Exeter, United Kingdom*

^k *Beijing Climate Center, China Meteorological Administration, Beijing, China*

^l *Bjerknes Centre for Climate Research, Uni Research Climate, Bergen, Norway*

(Manuscript received 26 February 2016, in final form 5 December 2017)

ABSTRACT

Water and carbon fluxes simulated by 12 Earth system models (ESMs) that participated in phase 5 of the Coupled Model Intercomparison Project (CMIP5) over several recent decades were evaluated using three functional constraints that are derived from both model simulations, or four global datasets, and 736 site-year measurements. Three functional constraints are ecosystem water-use efficiency (WUE), light-use efficiency (LUE), and the partitioning of precipitation P into evapotranspiration (ET) and runoff based on the Budyko framework. Although values of these three constraints varied significantly with time scale and should be quite conservative if being averaged over multiple decades, the results showed that both WUE and LUE simulated by the ensemble mean of 12 ESMs were generally lower than the site measurements. Simulations by the ESMs were generally consistent with the broad pattern of energy-controlled ET under wet conditions and soil water-controlled ET under dry conditions, as described by the Budyko framework. However, the value of the parameter in the Budyko framework ω , obtained from fitting the Budyko curve to the ensemble model simulation (1.74), was larger than the best-fit value of ω to the observed data (1.28). Globally, the ensemble mean of multiple models, although performing better than any individual model simulations, still underestimated the observed WUE and LUE, and overestimated the ratio of ET to P , as a result of overestimation in ET and underestimation in gross primary production (GPP). The results suggest that future model development should focus on improving the algorithms of the partitioning of precipitation into ecosystem ET and runoff, and the coupling of water and carbon cycles for different land-use types.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-16-0177.s1>.

Corresponding author: Longhui Li, lilhcn@gmail.com

1. Introduction

Several intercomparison studies compared simulations by the Earth system models (ESMs) participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) against various

DOI: 10.1175/JCLI-D-16-0177.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy \(www.ametsoc.org/PUBSReuseLicenses\)](https://www.ametsoc.org/PUBSReuseLicenses).

observation-based in situ measurements or global datasets, including runoff, net primary productivity, and soil carbon (Alkama et al. 2013; Anav et al. 2015; Seager et al. 2013; Shao et al. 2013; Todd-Brown et al. 2013). These evaluations identified some key deficiencies in the simulated temporal dynamics and spatial patterns of specific variables or processes relating the water or carbon cycles, but none evaluated water and carbon cycles together, or the coupling of those two cycles at site and global scales.

Over the last three decades, there has been an accumulation of global flux network (FLUXNET) and observation-based gross primary production (GPP) and evapotranspiration (ET) datasets, which provide a unique opportunity to evaluate two key terrestrial processes: water and carbon cycles and their coupling, as represented in the most advanced ESMs. Such intercomparisons enable evaluation of models and highlight differences among them to further our understanding of how the carbon cycle operates and interacts with the physical climate system.

In this study, we intend to assess how the ESMs participating in CMIP5 perform against the broadscale fundamental functional constraints. Functional constraints are defined here as integrated indicators that describe integrated effects of multiple ecophysiological processes from environmental variables, and they may vary spatially but are quite conservative for a given location if averaged over multiple decades. Three functional constraints including ecosystem water-use efficiency (WUE), light-use efficiency (LUE), and the partitioning of precipitation into ET and runoff based on the Budyko framework were used to evaluate ESMs.

In ecophysiology, WUE and LUE are often used to quantify the efficiencies of different terrestrial ecosystems to use two primary resources for plant growth and ecosystem functioning: water and light (Ito and Inatomi 2012; Ruimy et al. 1999). WUE represents the interactions between carbon and water cycles at leaf, canopy, and ecosystem scales (Eamus 1991). Leaf-scale WUE and its response to meteorological variables have been relatively well understood (Medlyn et al. 2011; Wong et al. 1979). However, understanding of WUE at ecosystem or global scale is still limited (Dekker et al. 2016; Keenan et al. 2013; Medlyn et al. 2015) because many factors including soil properties, nutrient availability, and vegetation structure can play a significant role in regulating WUE (Field et al. 1995; Huntington 2008; Medlyn et al. 2015). Recent research showed that ecosystem WUE was positively related to rising atmospheric CO₂ concentration and an increase in leaf area index, but negatively impacted by a decline in vapor pressure deficits (Cheng et al. 2017). Despite multiple environmental factors influencing WUE and its temporal variations, ecosystem WUE was found to be quite conservative for a specific location when averaged over multiple decades (Gao

et al. 2014; Huxman et al. 2004; Tang et al. 2014). This robustness in ecosystem WUE is essentially determined by the intrinsic ecosystem sensitivity to water availability across precipitation regimes, independent of temporally hydroclimatic variations (Ponce-Campos et al. 2013).

LUE is used to quantify the efficiency of absorbed lights in photosynthesis or dry matter production (Monteith 1972), and has been used in ecosystem models for estimating GPP (Coops et al. 2005; Potter et al. 1993). LUE has been shown to be constant over a day in many studies (Rosati and Dejong 2003; Sims et al. 2005), but can also vary significantly at different temporal and spatial scales (Goerner et al. 2011; Turner et al. 2003; Yuan et al. 2007; Zhou et al. 2016). Phenology, temperature, vapor pressure deficit, soil water content, and diffuse radiation fraction were considered to be main factors influencing LUE (Mäkelä et al. 2008; Veroustraete et al. 2002; Yuan et al. 2007; Y. Zhang et al. 2015). However, ecosystem LUE was considered to be approximately constant over decadal time scales (Medlyn 1998; Monteith 1972; Monteith and Moss 1977), which is mainly based on the ecological optimization principle through adjusting plant nitrogen (N) content and Rubisco activity (Haxeltine and Prentice 1996).

In hydrology, the Budyko framework is a robust tool for estimating the partitioning of precipitation P into ET and runoff R_{off} as a function of aridity index (Budyko 1974; Zhang et al. 2001). Over multiannual or decadal time scales, the Budyko framework has been proven to be accurate in describing the interactions between climate and the hydrological cycle (Li et al. 2013; Yang et al. 2006; Zhang et al. 2008). In addition, the introduction of a plant-available water coefficient ω (Li et al. 2013; Yang et al. 2007; Zhang et al. 2001) in the Budyko-type equation allows the effects of vegetation on ET to be considered. Numerous studies showed the robustness of the Budyko framework in describing the joint controls of energy and water availabilities on ET, and the ratio of ET to R_{off} . The Budyko parameter ω remained considerably constant at multiple annual time scales for a given catchment (Zhang et al. 2001).

The objectives of this study are to assess the performance of the 12 ESMs for estimating WUE, LUE, and rainfall partitioning at individual sites or global scale against measurements, and to identify key sources of model errors. Results of this study will guide future model development.

2. Materials and methods

a. Earth system model output

Twelve ESMs were selected from all participating models in CMIP5 based on their availability of the

selected monthly outputs of carbon, energy, and climatic variables. These ESMs differ in their representations of terrestrial carbon cycle processes, as well as in their atmospheric and oceanic components, but share common experimental designs and output protocols. The 12 ESMs used in this study are listed in Table S1 of the supplemental material and the detailed descriptions of each ESM are available from previous studies (Shao et al. 2013; Todd-Brown et al. 2013). A number of experimental simulations were performed for CMIP5, but we only focus here on the “historical” simulation. The historical CMIP5 simulation that covers the 1850–2005 period allows evaluation of ESMs through a direct comparison with observations (Taylor et al. 2012). We selected the most recent 24 years covering the period of 1982–2005, which enables comparison with observation-based estimates from the global FLUXNET data and available global GPP and ET data products.

Monthly carbon, water, and energy fluxes including GPP, ET, downward and upward shortwave and longwave radiation, and climatic variables including precipitation, temperature, humidity, and atmospheric pressure were obtained online (http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html). All variables were aggregated to annual values and their mean annual values were used in this study.

b. FLUXNET and global GPP and ET datasets

Site-scale data of monthly GPP and ET (FLUXNET2015) were downloaded from the global FLUXNET site (<http://fluxnet.fluxdata.org/data/fluxnet2015-dataset>) that included 165 stations and 1180 site years. In the FLUXNET2015 dataset, GPP was derived from observed net ecosystem exchange (NEE) and ecosystem respiration that was estimated from nighttime NEE (Papale et al. 2006; Reichstein et al. 2005). Annual ET was estimated from observed latent heat flux, in which latent heat flux was filled by a marginal distribution sampling algorithm and was corrected for any errors in energy balance closure (Reichstein et al. 2005). Ecosystem WUE was calculated as the ratio of annual GPP to annual ET. We selected site years with both annual GPP and ET available. As a result, 736 site years (93 sites) of data were selected for analysis. Geographical distribution of the 93 sites is shown in Fig. S1 of the supplemental material.

Global spatial datasets include one set of GPP and three sets of ET. One set of GPP and one set of ET were obtained from Jung et al. (2011, hereinafter JU11). The JU11 data are essentially model products derived by combining FLUXNET observations and a model tree ensemble (MTE) approach trained with satellite-derived fraction of absorbed photosynthetically active radiation (FAPAR) and gridded mean climates. The

MTE statistical model employed by JU11 consists of a set of regression trees trained with the local GPP estimates obtained from measured NEE. The JU11 dataset represent the mean fluxes of GPP and ET during the period 1982–2008. Although JU11 GPP and ET are not direct measurements, they have been used as a “reference” for evaluating global models (Shao et al. 2013; Wang et al. 2012). To avoid the uncertainties caused by one dataset, two other independent global ET datasets were also used as “references.” The two independent ET datasets consist of the Global Land Evaporation Amsterdam Model (GLEAM; Miralles et al. 2011) and the LandFlux Eval (hereinafter EVAL; Mueller et al. 2013). The GLEAM ET dataset was generated using the Priestley and Taylor equation with a variety of satellite data including microwave-derived soil moisture, land surface temperature, and vegetation density, as well as the detailed estimation of rainfall interception loss as model inputs. The GLEAM ET data (version 3.0a) cover the period from 1980 to 2014 at $0.25^\circ \times 0.25^\circ$. The EVAL has longer data coverage (1989–2005) and is an ET synthesis product at $1^\circ \times 1^\circ$, merged from a total of 14 datasets including five diagnostic data (derived from satellite and/or in situ observations), five land surface models, and four reanalyses ET products (Mueller et al. 2013).

c. WUE, LUE, and partitioning of rainfall within the Budyko framework

For the purposes of this study, we define ecosystem WUE as GPP per unit total ET [$\text{mmol C} (\text{mol H}_2\text{O})^{-1}$],

$$\text{WUE} = \frac{\text{GPP}}{\text{ET}}. \quad (1)$$

Ecosystem LUE is defined as GPP per unit incident photosynthetically active radiation (PAR), that is,

$$\text{LUE} = \frac{\text{GPP}}{\text{PAR}}. \quad (2)$$

PAR is approximately calculated as half of downward shortwave radiation in each ESM with a conversion factor of $1 \text{ J} = 4.6 \mu\text{mol}$ for PAR. The unit of LUE is micromole of carbon per micromole of PAR.

The Budyko framework states that the mean annual water balance of a given land surface is mainly controlled by available water and atmospheric demand for water (Budyko 1974). The original Budyko framework uses net radiation to describe atmospheric demand of water and doesn't include the effect of vegetation that is expected to influence actual ET. Here we used the formulation proposed by Zhang et al. (2001) that includes the effect of vegetation on annual water balance through

the parameter ω that is related to plant-available water in soil:

$$\frac{ET}{P} = \frac{1 + \omega\phi}{1 + \omega\phi + \phi^{-1}}, \quad (3a)$$

$$\frac{R_{\text{off}}}{P} = \frac{\phi^{-1}}{1 + \omega\phi + \phi^{-1}}, \text{ or} \quad (3b)$$

$$\frac{ET}{R_{\text{off}}} = \frac{1 + \omega\phi}{\phi^{-1}}. \quad (3c)$$

The parameter ϕ is the aridity index, defined as the ratio of potential ET (PET) to P , and ω is the only model parameter that varies with vegetation type, land use, and topography, with a typical value of 2.0 for forests and 0.5 for grasses (Li et al. 2013; Zhang et al. 2001).

PET was computed using the Priestley–Taylor equation

$$PET = \alpha \frac{\Delta}{\Delta + \gamma} R_{\text{net}}, \quad (4)$$

where α is the Priestley–Taylor parameter (set to 1.26), Δ is the slope of the saturation vapor pressure–temperature relationship, and γ is the psychrometric constant. Net radiation R_{net} is computed as

$$R_{\text{net}} = r_{\text{sds}} + r_{\text{lds}} - r_{\text{sus}} - r_{\text{lus}}. \quad (5)$$

The four radiation fluxes r_{sds} , r_{lds} , r_{sus} , and r_{lus} are downward shortwave radiation, downward longwave radiation, upward shortwave, and upward longwave radiation, respectively. The four radiation fluxes and air temperature T_{air} are obtained from each ESM, and PET was computed based on Eq. (5). Combining ESM-simulated P and ET, a best-fit value of ω can be obtained by applying the Budyko framework [Eqs. (3a)–(3c)] to all ESMs.

d. CRU precipitation data

To determine if the model biases in ET and GPP were caused by biases in simulated climate (especially precipitation) or by the land model components of the ESMs themselves (i.e., in the way that ecological processes are represented), a global dataset of historical climate, the Climatic Research Unit (CRU) time series (TS) data, version 4.01 (Harris et al. 2014), was used to evaluate precipitation simulated by ESMs. The CRU gridded climate dataset is available at a monthly temporal resolution and a spatial resolution of $0.5^\circ \times 0.5^\circ$ for the period 1961–2014. To match the period of global FLUXNET data, the precipitation data from the CRU climate dataset for the period 1982–2005 were used. Although precipitation, temperature, and incoming solar radiation exhibit primary controls on GPP and ET,

simulated precipitation was considered the most important variable that differed most among ESMs.

3. Results

a. WUE

The observed multiyear mean WUE at site scale from the FLUXNET [WUE in eddy covariance (EC)] dataset varied significantly among different land-cover types. For example, the observed WUE values for OSH were significantly lower than those for ENF. As compared with the range of the observed WUE variation from 0 to $5 \text{ mmol C (mol H}_2\text{O)}^{-1}$ (Fig. 1a, where $1 \text{ mmol C (mol H}_2\text{O)}^{-1} = 2/3 \text{ g C (kg H}_2\text{O)}^{-1}$), the ESMs estimated much smaller variation of multiyear mean WUE [from 1.0 to $2.0 \text{ mmol C (mol H}_2\text{O)}^{-1}$] for those selected sites (Fig. 1a). The linear correlation of site-scale WUE estimates between FLUXNET and ensemble ESM means was poor [WUE in ESM = $0.16 \text{ WUE in EC} + 0.14$, $r^2 = 0.22$, and root-mean-square error (RMSE) = $0.86 \text{ mmol C (mol H}_2\text{O)}^{-1}$]. Overall, ensemble ESM mean underestimated WUE when the observed WUE from FLUXNET dataset was larger than $2.0 \text{ mmol C (mol H}_2\text{O)}^{-1}$.

Compared with the simulated WUE by ESMs, multiyear mean WUE from JU11 GPP and JU11 ET at site levels (WUE in JU11) had a much better linear correlation with WUE in EC [WUE in JU11 = $0.47 \text{ WUE in EC} + 1.06$, $r^2 = 0.36$, and RMSE = $0.63 \text{ mmol C (mol H}_2\text{O)}^{-1}$]. Therefore, WUE in JU11 will be used to evaluate spatial patterns of WUE from ESMs.

At the global scale, the ratio of JU11 GPP to GLEAM ET (WUE in GLEAM) and EVAL ET (WUE in EVAL) showed a spatial pattern similar to WUE in JU11 (Fig. 2). All three estimates of WUE were relatively high [$>2.5 \text{ mmol C (mol H}_2\text{O)}^{-1}$] in Europe and South America, and low [$<1.0 \text{ mmol C (mol H}_2\text{O)}^{-1}$] in central Asia and western North America. Three WUE estimates based on observations were highly correlated with each other [$r^2 = 0.69$ between WUE in JU11 and WUE in GLEAM, between WUE in JU11 and WUE in EVAL ($r^2 = 0.78$), or between WUE in GLEAM and WUE in EVAL ($r^2 = 0.85$)]. On the other hand, there are considerable variations in the multiyear mean WUEs as estimated from different ESM simulations for a given region. Particularly, GISS-E2-H simulated systematically higher WUE than the other 11 ESMs. Both GFDL-ESM2G and IPSL-CM5A-LR simulated higher WUE in the Northern Hemisphere (NH), but BCC_CSM1.1, CCSM4, and HadGEM2-CC simulated lower WUE in NH than other ESMs (see Fig. S2 in the supplemental material).

The ensemble mean of multiyear mean WUE of 12 ESMs has a spatial pattern similar to those derived

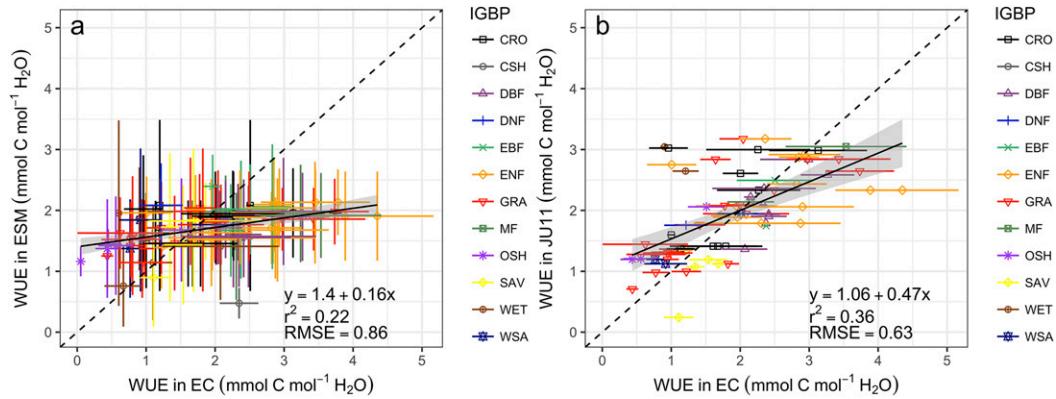


FIG. 1. Comparisons of WUE between model mean of (a) 12 ESMs or (b) **JU11** and FLUXNET measurements at site scale. IGBP land-cover types include cropland (CRO), closed shrubland (CSH), deciduous broadleaf forests (DBF), deciduous needleleaf forests (DNF), evergreen broadleaf forests (EBF), evergreen needleleaf forests (ENF), grassland (GRA), mixed forests (MF), open shrubland (OSH), savanna (SAV), wetland (WET), and woody savanna (WSA). The error bar represents one SD. [The unit of RMSE for WUE is $\text{mmol C} (\text{mol H}_2\text{O})^{-1}$.]

from the model–data products (see Fig. 2): generally low in dry areas, such as central Australia and the southern United States, and high in the tropical rain forests in Africa and South America. However, the ensemble mean of 12 ESMs underestimated the WUE in Europe and the northern temperate forest regions. Ensemble mean WUEs of 12 ESMs were significantly lower than that of **JU11** when pixel-by-pixel values were compared (WUE in ESM = 0.4 WUE in **JU11** and $r^2 = 0.78$).

Mean annual WUE from three observation-based estimates varied similarly with latitude (Fig. 3). The estimated WUE from both GLEAM and EVAL

datasets was smaller than **JU11** because of the relatively larger ET values in GLEAM and EVAL than in **JU11** (see Fig. S3 in the supplemental material). Overall, estimated WUE from three different datasets was high in the tropics, subtropics, and temperate regions, but low in the regions near 15°N , 35°N , and 30°S . At least five of the 12 ESMs (BNU-ESM, CanESM2, GISS-E2-H, IPSL-CM5A-LR, and MPI-ESM-LR) failed to reproduce this pattern of latitudinal variation in ecosystem mean annual WUE (Fig. 3). For example, in the latitudinal range of 30°S – 20°N , the estimated mean annual WUE by CanESM2 was significantly greater, up to

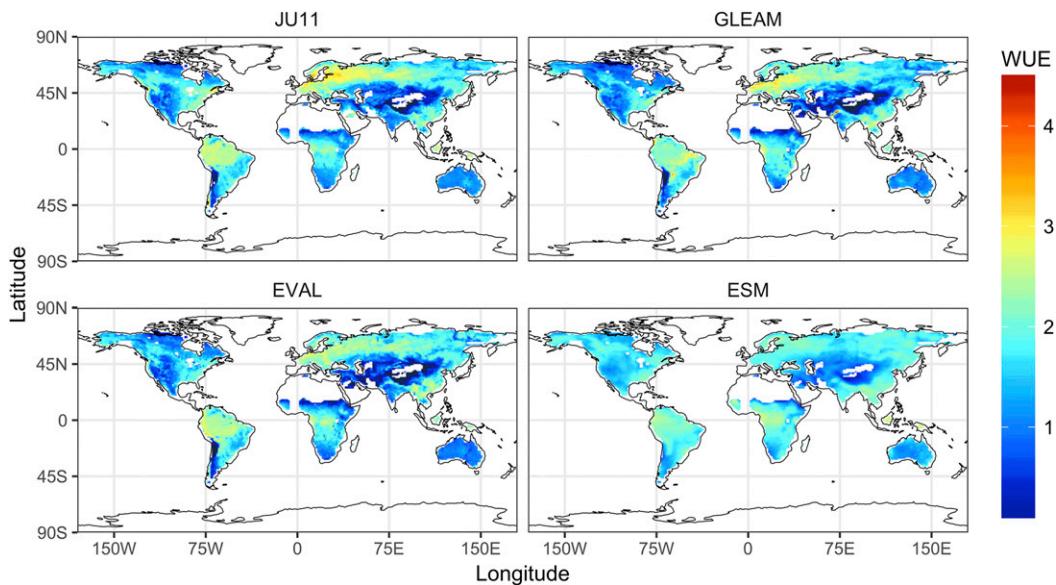


FIG. 2. Global distribution of WUE of three references (**JU11**, GLEAM, and EVAL) and model mean of 12 ESMs. [The unit of WUE is $\text{mmol C} (\text{mol H}_2\text{O})^{-1}$.]

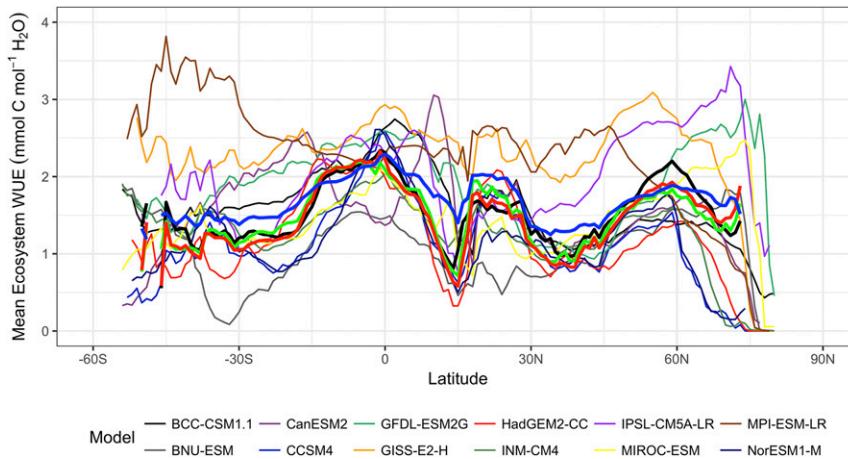


FIG. 3. Zonal profiles of WUE [$\text{mmol C (mol H}_2\text{O)}^{-1}$] for three references (thick black line for JU11, thick green for GLEAM, and thick red for EVAL), the 12 ESMs, and the mean of multiple models (thick blue line).

150% more, than the estimated WUE from the three observation-based data products. The remaining seven ESMs simulated much lower WUE across different latitudes (Fig. 3). As a result, the ensemble mean WUE averaged over 12 ESMs was comparable with the mean WUE from three observation-based estimates across different latitudes.

b. LUE

At site scale, ensemble means LUE of 12 ESMs (LUE in ESM) were significantly correlated with the observed LUE (LUE in EC) [LUE in ESM = 0.38 LUE

in EC + 0.004, $r^2 = 0.51$, and RMSE = $0.004 \mu\text{mol C (}\mu\text{mol PAR)}^{-1}$] across the selected sites (Fig. 4). The mean LUE across all the selected sites from the 12 ESMs was $0.0076 \pm 0.0019 \mu\text{mol C (}\mu\text{mol PAR)}^{-1}$, as compared with the mean of $0.0086 \pm 0.0053 \mu\text{mol C (}\mu\text{mol PAR)}^{-1}$ from the observations. Overall, the ensemble mean of 12 ESMs gave higher LUE when the observed LUE was $>0.015 \mu\text{mol C (}\mu\text{mol PAR)}^{-1}$ [$1 \mu\text{mol C (}\mu\text{mol PAR)}^{-1} = 55.2 \text{ g C MJ}^{-1}$], such as EBF, and gave lower LUE when the observed LUE was $<0.005 \mu\text{mol C (}\mu\text{mol PAR)}^{-1}$, such as OSH or GRA.

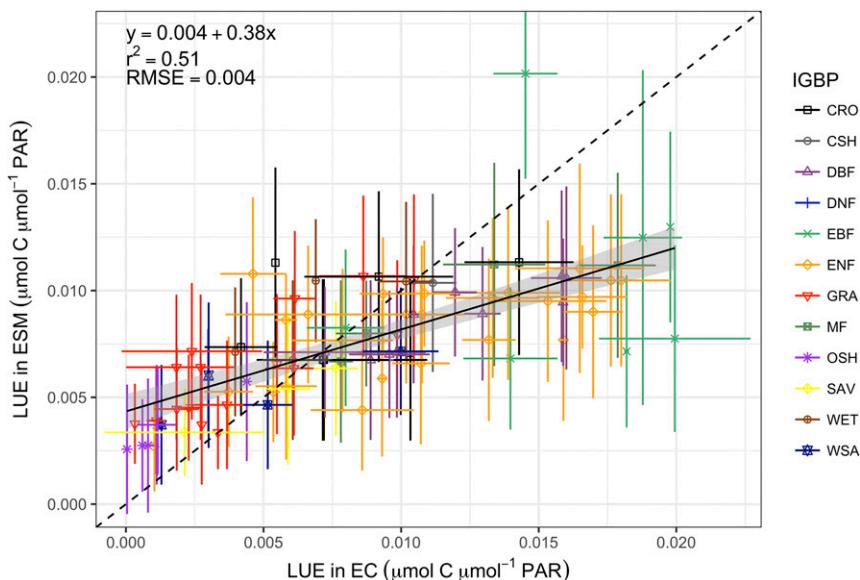


FIG. 4. Comparisons of LUE between model mean of 12 ESMs and FLUXNET measurements at site scale. Legends for IGBP land-cover types are defined in Fig. 1. [The unit of RMSE for LUE is $\mu\text{mol C (}\mu\text{mol PAR)}^{-1}$.]

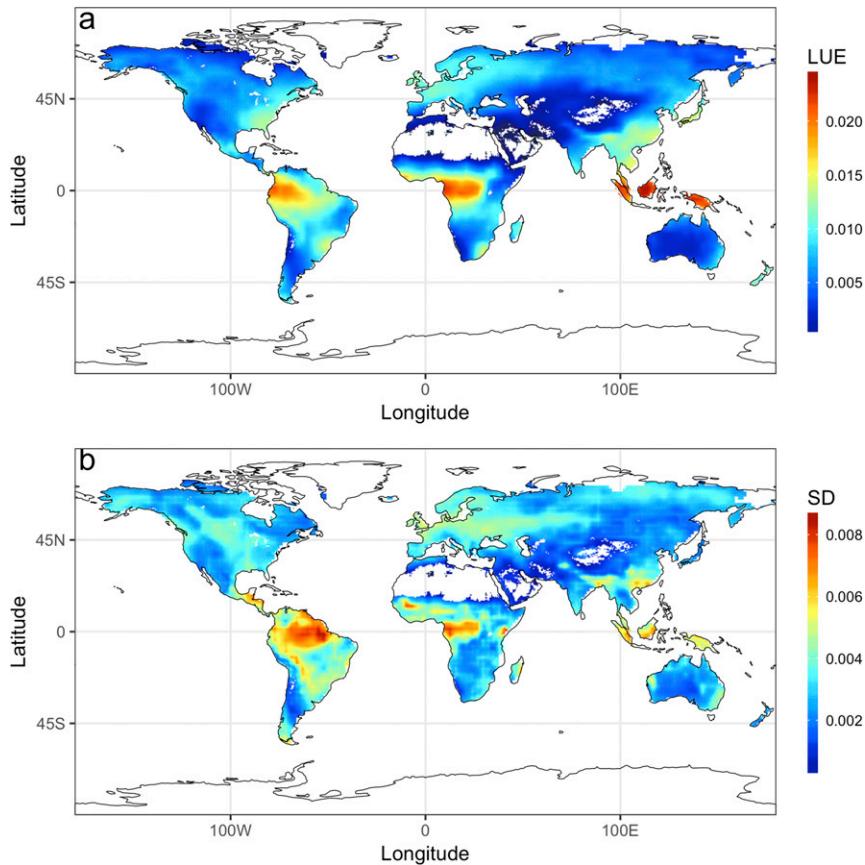


FIG. 5. Global distribution of (a) LUE model mean of 12 ESMs and (b) the SD. [The unit of RMSE for LUE is $\mu\text{mol C } (\mu\text{mol PAR})^{-1}$.]

At the global scale, the calculated LUE showed that the LUE values for the majority of land grid cells from the 12 ESMs were less than $0.02 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$ (Fig. S4 in the supplemental material). Only 7 of 12 ESMs simulated LUE $> 0.03 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$ for some land points, and none simulated LUE $> 0.04 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$. Because CCSM4 and NorESM1-M used the same land surface model, the simulated LUE by these two models were similar. BNU-ESM and MIROC-ESM simulated relatively low LUE values [$< 0.02 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$ for most land points], as compared to other ESMs.

The mean LUE of 12 ESMs over land grid cells ranged from 0.001 to $0.025 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$, with an average value of $0.007 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$ ($= 0.37 \text{ g CMJ}^{-1}$) and a standard deviation of $0.004 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$ ($= 0.23 \text{ g CMJ}^{-1}$; Fig. 5a). Larger LUE values [$> 0.015 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$] were located in the tropical EBF in South America, part of central Africa, and South Asia (Fig. 5a). Relatively low LUE [$< 0.002 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$] was mostly located in central Asia. Regions with high-ensemble mean LUE,

such as Amazonia, central Africa, and tropical Asia also showed high standard deviation (SD), which suggests a large difference in the estimated LUE among different ESMs. Furthermore, Europe and central United States also showed large variations in LUE estimates among 12 ESMs (Fig. 5b). As PAR simulated by 12 ESMs were quite similar at the global scale (see Fig. S5 in the supplemental material), differences in simulated LUE across ESMs were largely contributed by the differences in the simulated GPP by different ESMs (see Fig. S6 in the supplemental material).

To evaluate the differences in the simulated LUE by different ESMs, we constructed a probability density function of LUE by calculating the area fraction of land cells for a given LUE for each of the 12 ESMs (see Fig. 6). We grouped the 12 ESMs into three groups based on the peak values of their probability distribution functions (PDFs). BNU-ESM, CCSM4, HadGEM2, and CanESM2 all have an LUE peak around $0.001 \mu\text{mol } (\mu\text{mol})^{-1}$, MIROC-ESM, INM-CM4.0, and BCC_CSM1.1 have a peak at an LUE of about $0.005 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$, and the remaining ESMs

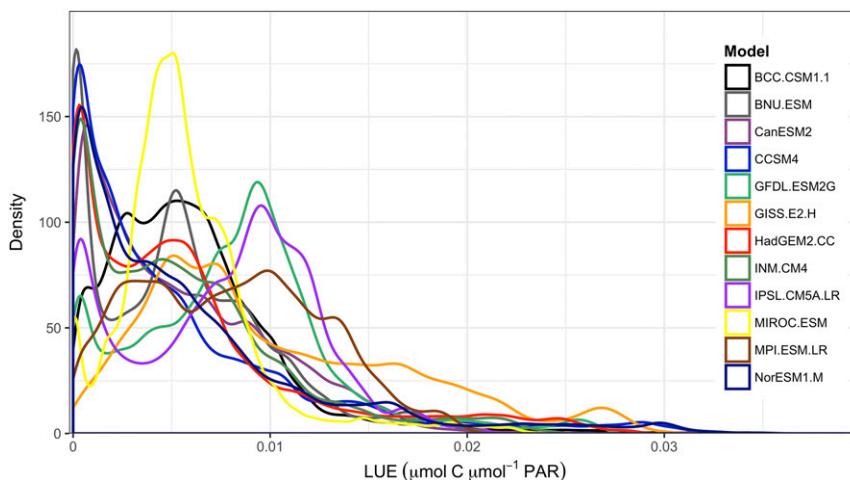


FIG. 6. PDF of LUE for the 12 ESMs. (Legends for the 12 ESMs are defined in Table 1.)

have a peak at an LUE of about $0.01 \mu\text{mol C } (\mu\text{mol PAR})^{-1}$ (see Fig. 6).

c. Budyko framework

Figure 7a showed annual ET/P [Eq. (3a)] increased nonlinearly with ϕ , as described by the Budyko curve for site-scale observations ($\omega = 1.28$, $r^2 = 0.34$; Fig. 7a) or ensemble mean of 12 ESMs ($\omega = 1.74$, $r^2 = 0.48$, Fig. 7b) at those selected sites, where ω is an empirical parameter [see Eqs. (3a)–(3c)]. In the context of Fig. 7, ω does not represent an “equivalent vegetation,” but rather how the land surface component of a given ESM partitions precipitation into evaporation and runoff. The best-fit value of ω for the ensemble mean of 12 ESMs was larger than that for the FLXUNET measurements, which indicated that the ensemble mean of 12 ESMs systematically oversimulated ET and underestimated runoff for a given value of ϕ .

Although the relationship between ET/P and aridity index from the ensemble mean of 12 ESMs is quite similar to the observed (see Fig. 7), the best-fit value of ω varied from 0.88 in BCC_CSM1.1 to 2.61 in INM-CM4.0, with a high coefficient of variance of 0.27 (mean of 1.95, and SD of 0.52) (see Table 1 and Fig. 8). The best-fit ω values for four other ESMs (BNU-ESM, CanESM2, GFDL-ESM2G, and IPSL-CM5A-LR) are lower than 2.0, and the ω values for the remaining seven ESMs all are larger than 2.0, although the correlation coefficients are all larger than 0.75. This divergence of ω in the partitioning of precipitation between ET and total runoff. For example, BCC_CSM1.1 with the smallest ω (0.88) simulated a relatively smaller proportion of precipitation into ET than runoff. In contrast, the INM-CM4.0 model with the largest ω (2.61) predicted a relatively larger proportion of ET than runoff. This also can be seen from

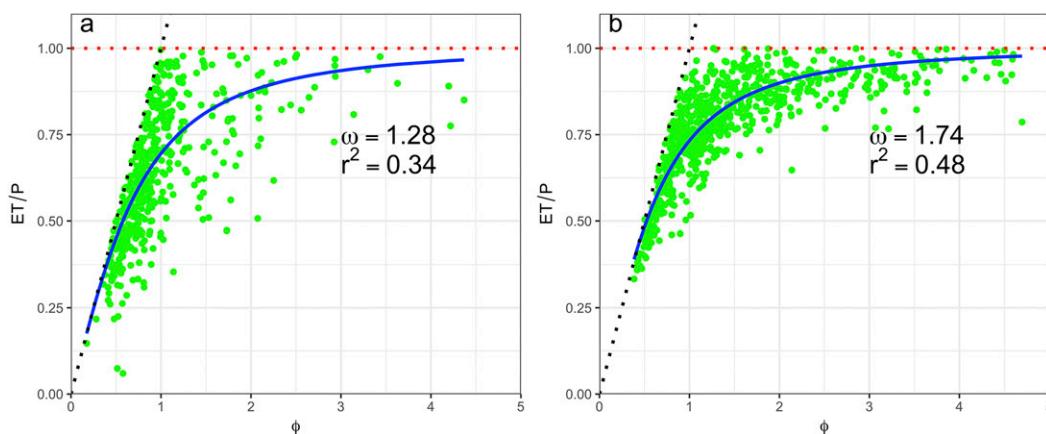


FIG. 7. The best-fit ω from the (a) FLUXNET measurements and (b) model mean of 12 ESMs, where r^2 is the squared correlation of the best-fit Budyko curve.

TABLE 1. The best-fit Budyko parameter ω from the 12 ESMs. The variable r is the correlation coefficient of the best-fit Budyko curve. (Expansions of acronyms are available online at <http://www.ametsoc.org/PubsAcronymList>.)

No.	Model	ω	r
1	BCC_CSM1.1	0.88	0.78
2	BNU-ESM	1.57	0.89
3	CanESM2	1.60	0.89
4	CCSM4	2.52	0.89
5	GFDL-ESM2G	1.65	0.87
6	GISS-E2-H	2.22	0.93
7	HadGEM2-CC	2.21	0.86
8	INM-CM4.0	2.61	0.91
9	IPSL-CM5A-LR	1.39	0.86
10	MIROC-ESM	2.07	0.92
11	MPI-ESM-LR	2.28	0.89
12	NorESM1-M	2.34	0.90
13	Model mean	1.74	0.69
14	EC observation	1.28	0.58

the ET estimates for INM-CM4.0, which had a significantly larger ET than BCC_CSM1.1 (see Fig. S7 in the supplemental material).

4. Discussion

Our study showed that the ensemble mean WUE of 12 ESMs was generally lower and less variable than the observed WUE across 736 site years. The modeled WUEs by 12 different ESMs showed a little variation

with a mean value of 1.71 and standard deviation of $0.32 \text{ mmol C (mol H}_2\text{O)}^{-1}$ for all land-cover types except CSH, as compared with the observed variation of WUE from 0 to $4 \text{ mmol C (mol H}_2\text{O)}^{-1}$ with a mean value of 1.96 and standard deviation of $0.93 \text{ mmol C (mol H}_2\text{O)}^{-1}$ [see Fig. 1 and previous reports (Beer et al. 2009; Tang et al. 2014)].

Simulated WUE by the multiple ESMs followed a similar latitudinal pattern with that of data products, and this is consistent with previous findings (Ito and Inatomi 2012; Tang et al. 2014; Xue et al. 2015). This indicates that there was strong offsetting for the simulated WUE, and some models overpredicted WUE, whereas others underestimated WUE (Mystakidis et al. 2016). Accuracy of simulated WUE was essentially dependent on accuracy in estimates of either GPP or ET or both. According to the intermodel comparisons, we found that uncertainties of simulated GPP were much larger than ET across ESMs; therefore, poor performance in simulating GPP in ESMs was the most important reason for poorly simulated WUE. In addition, recent studies found that global GPP increased more than global ET in response to the increase in atmospheric CO_2 concentration (Cheng et al. 2017), which resulted in an increase in WUE (Keenan et al. 2013). However, the temporary trend of WUE was not examined in this study.

The underlying causes for the differences in GPP and ET between ESMs and data products may come from the differences in model input (including dynamically

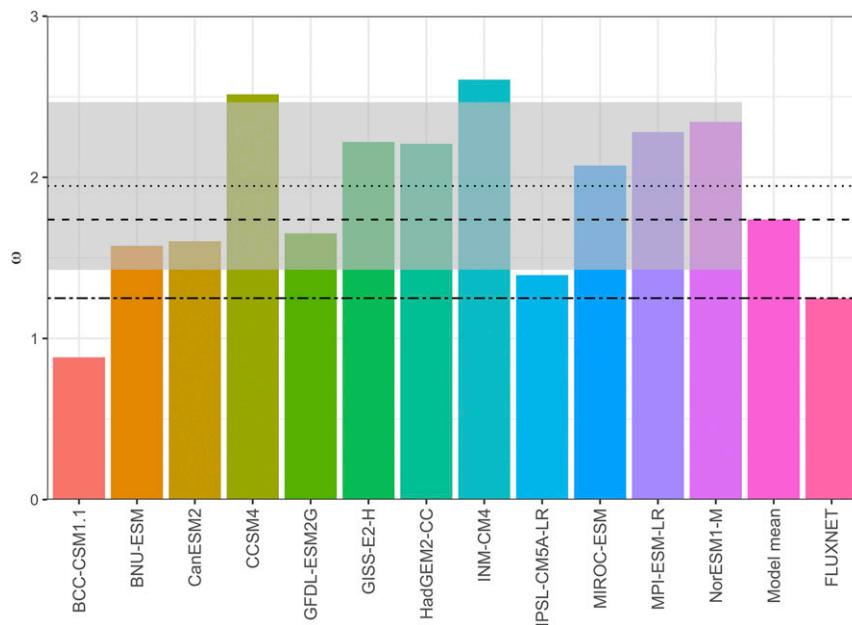


FIG. 8. Comparisons of the best-fit ω from 12 ESMs and their mean and the FLUXNET measurements. Shaded area represents one SD of the best-fit ω values from the 12 ESMs.

simulated or prescribed land cover, prescribed soil maps, and simulated climate in the models), and model structures and parameterizations (Prentice et al. 2015). Climate forcing errors were important sources of errors in regional carbon and water simulations, and the errors can be as high as 80% at site scales (Zhao et al. 2012). Here we investigated the effects of differences in model input on simulated GPP and ET. By calculating the difference between the ESM-simulated precipitation and the CRU precipitation at each grid, we found that differences in model precipitation can explain up to 31% of differences in the simulated GPP (GFDL-ESM2G, see Fig. S8 in the supplemental material) and up to 25% of differences in the simulated ET (MPI-ESM-LR, see Fig. S9 in the supplemental material). Differences in precipitation had positive linear and significant ($p < 0.001$) effects on the differences in the simulated GPP and ET for all ESMs.

Apart from the errors in climate forcing (simulated by the atmospheric component in an ESM), differences in the representation of key ecosystem processes could be an important source of differences in the simulated GPP and ET. Although global distributions of model differences for GPP or ET differed among the 12 ESMs, it was possible to identify broadscale common deficiencies across models. For example, CanESM2, BNU-ESM, and BCC_CSM1.1 significantly underestimated the GPP of the Amazon rain forests; CanESM2, BCC_CSM1.1, and GFDL-ESM2G underestimated ET over parts of the Amazon rain forest; and GISS-E2-H underestimated ET over parts of the Amazon savanna. Significant underestimations of GPP and ET in these ESMs (particularly in CanESM2) over Amazonia (see Figs. S6 and S7) can be partially explained by the underestimation of precipitation in the regions in ESMs (Shao et al. 2013). Other likely causes include inadequate representation of root functioning in the current land surface model components of ESMs, such as hydraulic redistribution (Baker et al. 2008; Oliveira et al. 2005) and the root water uptake process (Li et al. 2012), which led to underestimations of both GPP and ET. We also observed that some ESMs underestimated GPP but overestimated ET in some regions (e.g., BNU-ESM in the Amazon rain forest and over the Australian continent). Possible reasons for this include higher sensitivity of ET to precipitation than GPP in the BNU-ESM. The difference between ESM and data products can be also attributed to uncertainties within datasets themselves. For example, JU11 GPP data were essentially upscaled based on site-derived GPP but there were only a few measurement sites in the Amazon region (Oliphant 2012). For ET products, there were also large spatial differences among JU11, GLEAM, and EVAL, particularly in tropical regions and northern America (see

Fig. S3). In addition, the number of land-cover types prescribed in an LSM was not always the same across models, and associated vegetation ecophysiological parameters (e.g., maximum carboxylation rate) and state variables [e.g., leaf area index (LAI)] varied largely (Shao et al. 2013), but they were crucially important for the estimates of GPP and ET (Lu et al. 2013; Wang et al. 2007). These factors can be important sources of errors and uncertainties for the simulated GPP and ET, and hence for WUE.

We found that intermodel differences in LUE among the 12 ESMs showed quite similar spatial patterns as the differences in their GPP (see Figs. S4 and S6), because simulated incoming PAR for a given location is very similar across the 12 models (Fig. S5). This demonstrated that differences in LUE among models mainly resulted from the difference in GPP, which was overpredicted by some ESMs in some regions (e.g., GISS-E2-H and HadGEM2.CC in the Amazon forest), and LUE was overestimated as well. However, overall, most of the 12 ESMs underestimated GPP and therefore LUE globally.

It is worth noting that, apart from model structures, other factors such as different classifications of land-cover types, the definition of LUE, canopy structure, and diffuse radiation fraction may also contribute to the significant discrepancies between the model simulations and observations (Ruimy et al. 1999; Yuan et al. 2007; L.-X. Zhang et al. 2015). Because the absorbed PAR was not available for some of the 12 ESMs and the selected sites from FLUXNET, we used the incoming PAR that was calculated as half of downward shortwave radiation for calculating LUE. Therefore the influences of some factors, such as canopy LAI and PAR extinction coefficient and fraction of diffuse PAR on the absorbed PAR (Polley et al. 2011) and therefore LUE, if being defined as per unit amount of absorbed PAR basis, were not accounted for in this study. However, this study found that differences in the estimated LUE among 12 ESMs for a given location largely resulted from the differences in GPP estimates. For some tropical evergreen forests, estimated GPP by some ESMs was very high ($>5000 \text{ g C m}^{-2} \text{ yr}^{-1}$), and therefore the LUE was also very high [$>0.03 \text{ mol C (mol PAR)}^{-1}$]. Future improvement on those ESMs should focus on those regions with very high GPP and LUE values.

All of the 12 ESMs correctly captured the nonlinear dependence of ET/P , or R_{off}/P for aridity, as described within the Budyko framework. This agrees with the result by Alkama et al. (2013), who found a good agreement between the observed and simulated runoff by some CMIPs models for most of river basins globally. However, the best-fit value of the only parameter ω in the Budyko equation varied from 0.88 to 2.61 among 12

ESMs. Large variations in the best-fit ω values strongly indicated large discrepancies in the partitioning of precipitation into ET and runoff among 12 ESMs (see Fig. S10 in the supplemental material). Ensemble mean of 12 ESMs tended to perform better than most individual ESMs, as compared with the best-fit value of ω using site-scale measurements. This is consistent with results of Roderick et al. (2014) based on their analysis of CMIP3 models. The best-fit ω value (1.74) of ensemble model mean was close to the best-fit value (1.28) from FLUXNET measurements in this research but was still less than the reported value of 2.21 [converting the parameter n to ω through $\omega = n + 0.72$ (Yang et al. 2008)], which was based on an old version of global FLUXNET data (Williams et al. 2012). The differences in estimated value of ω from the FLUXNET dataset between this study and Williams et al. (2012) may result from the exclusion of sites with annual ET exceeding annual precipitation in this study. The mean values of ω at basin scale were positively correlated with long-term averaged vegetation cover (Li et al. 2013), which means that ω should be quite conservative without significant changes in land use as the scenario used in this research. It is also likely that different interpretations of land cover and land use among the 12 ESMs contributed to the large differences in ω , therefore the partitioning of rainfall into ET and runoff.

The projected future carbon and water fluxes by ESMs were reported to have large uncertainties (Cox et al. 2013). To reduce uncertainties, observation-based carbon and water flux estimates or their sensitivities to a specific climatic variable as constraints can be used to constrain the projections of ESMs (Cox et al. 2013; Mystakidis et al. 2016; Wenzel et al. 2014). Such statistical relationships between future and historical simulations or observations to constrain projections of ESMs, however, can be spurious because of errors or uncertainties in models or observations themselves (Bracegirdle and Stephenson 2013). As ecological or hydrological relationships between ecosystem functioning and climatic variables have been established on theoretical (e.g., the Budyko framework) and experimental (e.g., WUE or LUE) considerations, they offer useful alternatives in constraining the projected responses of ecosystem functioning to climate change, particularly at annual or longer temporal scales (Huxman et al. 2004; Ruimy et al. 1999).

We recommended that these functional constraints found in ecophysiological and hydrological disciplines should be used as an alternative strategy to constrain future projections of ESMs. For example, WUE, LUE, or the Budyko framework can be used in conjunction with FLUXNET measurement and/or other independent

observations (e.g., observed runoff at basin scales) to constrain ESM simulations. Corresponding mean values of parameters or constants and uncertainties (usually expressed as standard deviation) with regards to the three functional relationships can be quantified. Then we can correlate the measured parameters or constants with those derived from ESMs in present or future, and force the ESM simulations to comply with these functional relationships to reduce the uncertainties of projections of ESMs, as in the recently reported constraining approach based on WUE (Mystakidis et al. 2016). Clearly, the current research served as a prestudy of using such functional relationships to constrain ESMs.

5. Conclusions

Twelve CMIP5 ESMs were evaluated against one set of global GPP data and three sets of ET data and 736 site years of FLUXNET measurements. We assessed whether the simulated spatial variations of carbon and water fluxes by 12 ESMs were broadly consistent with our understanding using three functional constraints: WUE, LUE, and the Budyko framework. Our results showed that most CMIP5 ESMs underestimated both WUE and LUE, as results of overestimation of ET and underestimation of GPP. In addition, all 12 ESMs underestimated the spatial variation of the observed WUE among different land surface types. Although the simulated partitioning of precipitation into ET and runoff by 12 ESMs for CMIP5 is broadly consistent with the predictions by the Budyko framework, the best-fit value of the Budyko parameter ω was quite variable among the 12 ESMs. Therefore, future model development should focus on concurrently improving the algorithms of precipitation partitioning between ecosystem ET and runoff in hydrological perspective and the coupling of carbon and water cycles across different land surface types.

Acknowledgments. This work is supported by National Key R&D Program of China (2017YFA0603603) and Australian Research Council Discovery Early Career Researcher Award Project (DE120103022). UK Met Office authors were supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101) and the European Union's Horizon 2020 research and innovation programme under Grant Agreement 641816 (CRESCENDO). J. T. acknowledges RCN-funded project EVA (229771) and Bjerknesh-BIGCHANGE. We thank the CMIP5 modelers, and all simulations were downloaded online (http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html). Global flux data were provided by the global FLUXNET community (<http://fluxnet.fluxdata.org/>), and the data of Jung et al. (2011) was highly appreciated.

REFERENCES

- Alkama, R., L. Marchand, A. Ribes, and B. Decharme, 2013: Detection of global runoff changes: Results from observations and CMIP5 experiments. *Hydrol. Earth Syst. Sci.*, **17**, 2967–2979, <https://doi.org/10.5194/hess-17-2967-2013>.
- Anav, A., and Coauthors, 2015: Spatiotemporal patterns of terrestrial gross primary production: A review. *Rev. Geophys.*, **53**, 785–818, <https://doi.org/10.1002/2015RG000483>.
- Baker, I. T., L. Prihodko, A. S. Denning, M. Goulden, S. Miller, and H. R. da Rocha, 2008: Seasonal drought stress in the Amazon: Reconciling models and observations. *J. Geophys. Res.*, **113**, G00B01, <https://doi.org/10.1029/2007JG000644>.
- Beer, C., and Coauthors, 2009: Temporal and among-site variability of inherent water use efficiency at the ecosystem level. *Global Biogeochem. Cycles*, **23**, GB2018, <https://doi.org/10.1029/2008GB003233>.
- Bracegirdle, T. J., and D. B. Stephenson, 2013: On the robustness of emergent constraints used in multimodel climate change projections of Arctic warming. *J. Climate*, **26**, 669–678, <https://doi.org/10.1175/JCLI-D-12-00537.1>.
- Budyko, M. I., 1974: *Climate and Life*. International Geophysics Series, Vol. 18, Academic Press, 508 pp.
- Cheng, L., and Coauthors, 2017: Recent increases in terrestrial carbon uptake at little cost to the water cycle. *Nat. Commun.*, **8**, 110, <https://doi.org/10.1038/s41467-017-00114-5>.
- Coops, N. C., R. H. Waring, and B. E. Law, 2005: Assessing the past and future distribution and productivity of ponderosa pine in the Pacific Northwest using a process model, 3-PG. *Ecol. Modell.*, **183**, 107–124, <https://doi.org/10.1016/j.ecolmodel.2004.08.002>.
- Cox, P. M., D. Pearson, B. B. Booth, P. Friedlingstein, C. Huntingford, C. D. Jones, and C. M. Luke, 2013: Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, **494**, 341–344, <https://doi.org/10.1038/nature11882>.
- Dekker, S. C., M. Groenendijk, B. B. Booth, C. Huntingford, and P. M. Cox, 2016: Spatial and temporal variations in plant water-use efficiency inferred from tree-ring, eddy covariance and atmospheric observations. *Earth Syst. Dyn.*, **7**, 525–533, <https://doi.org/10.5194/esd-7-525-2016>.
- Eamus, D., 1991: The interaction of rising CO₂ and temperatures with water use efficiency. *Plant Cell Environ.*, **14**, 843–852, <https://doi.org/10.1111/j.1365-3040.1991.tb01447.x>.
- Field, C. B., R. B. Jackson, and H. A. Mooney, 1995: Stomatal responses to increased CO₂: Implications from the plant to the global scale. *Plant Cell Environ.*, **18**, 1214–1225, <https://doi.org/10.1111/j.1365-3040.1995.tb00630.x>.
- Gao, Y., X. Zhu, G. Yu, N. He, Q. Wang, and J. Tian, 2014: Water use efficiency threshold for terrestrial ecosystem carbon sequestration in China under afforestation. *Agric. For. Meteorol.*, **195–196**, 32–37, <https://doi.org/10.1016/j.agrformet.2014.04.010>.
- Goerner, A., M. Reichstein, E. Tomelleri, N. Hanan, S. Rambal, D. Papale, D. Dragoni, and C. Schmullius, 2011: Remote sensing of ecosystem light use efficiency with MODIS-based PRI. *Biogeosciences*, **8**, 189–202, <https://doi.org/10.5194/bg-8-189-2011>.
- Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2014: Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 dataset. *Int. J. Climatol.*, **34**, 623–642, <https://doi.org/10.1002/joc.3711>.
- Haxeltine, A., and I. C. Prentice, 1996: A general model for the light-use efficiency of primary production. *Funct. Ecol.*, **10**, 551–561, <https://doi.org/10.2307/2390165>.
- Huntington, T. G., 2008: CO₂-induced suppression of transpiration cannot explain increasing runoff. *Hydrol. Processes*, **22**, 311–314, <https://doi.org/10.1002/hyp.6925>.
- Huxman, T. E., and Coauthors, 2004: Convergence across biomes to a common rain-use efficiency. *Nature*, **429**, 651–654, <https://doi.org/10.1038/nature02561>.
- Ito, A., and M. Inatomi, 2012: Water-use efficiency of the terrestrial biosphere: A model analysis focusing on interactions between the global carbon and water cycles. *J. Hydrometeorol.*, **13**, 681–694, <https://doi.org/10.1175/JHM-D-10-05034.1>.
- Jung, M., and Coauthors, 2011: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res.*, **116**, G00J07, <https://doi.org/10.1029/2010JG001566>.
- Keenan, T. F., D. Y. Hollinger, G. Bohrer, D. Dragoni, J. W. Munger, H. P. Schmid, and A. D. Richardson, 2013: Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise. *Nature*, **499**, 324–327, <https://doi.org/10.1038/nature12291>.
- Li, D., M. Pan, Z. Cong, L. Zhang, and E. Wood, 2013: Vegetation control on water and energy balance within the Budyko framework. *Water Resour. Res.*, **49**, 969–976, <https://doi.org/10.1002/wrcr.20107>.
- Li, L., Y.-P. Wang, Q. Yu, B. Pak, D. Eamus, J. Yan, E. van Gorsel, and I. T. Baker, 2012: Improving the responses of the Australian community land surface model (CABLE) to seasonal drought. *J. Geophys. Res.*, **117**, G04002, <https://doi.org/10.1029/2012JG002038>.
- Lu, X., Y.-P. Wang, T. Ziehn, and Y. Dai, 2013: An efficient method for global parameter sensitivity analysis and its applications to the Australian community land surface model (CABLE). *Agric. For. Meteorol.*, **182–183**, 292–303, <https://doi.org/10.1016/j.agrformet.2013.04.003>.
- Mäkelä, A., and Coauthors, 2008: Developing an empirical model of stand GPP with the LUE approach: Analysis of eddy covariance data at five contrasting conifer sites in Europe. *Global Change Biol.*, **14**, 92–108, <https://doi.org/10.1111/j.1365-2486.2007.01463.x>.
- Medlyn, B. E., 1998: Physiological basis of the light use efficiency model. *Tree Physiol.*, **18**, 167–176, <https://doi.org/10.1093/treephys/18.3.167>.
- , and Coauthors, 2011: Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biol.*, **17**, 2134–2144, <https://doi.org/10.1111/j.1365-2486.2010.02375.x>.
- , and Coauthors, 2015: Using ecosystem experiments to improve vegetation models. *Nat. Climate Change*, **5**, 528–534, <https://doi.org/10.1038/nclimate2621>.
- 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, <https://doi.org/10.5194/hess-15-453-2011>.
- Monteith, J. L., 1972: Solar radiation and productivity in tropical ecosystems. *J. Appl. Ecol.*, **9**, 747–766, <https://doi.org/10.2307/2401901>.
- , and C. J. Moss, 1977: Climate and the efficiency of crop production in Britain. *Philos. Trans. Roy. Soc. London*, **281B**, 277–294, <https://doi.org/10.1098/rstb.1977.0140>.
- Mueller, B., and Coauthors, 2013: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis. *Hydrol. Earth Syst. Sci.*, **17**, 3707–3720, <https://doi.org/10.5194/hess-17-3707-2013>.

- Mystakidis, S., E. L. Davin, N. Gruber, and S. I. Seneviratne, 2016: Constraining future terrestrial carbon cycle projections using observation-based water and carbon flux estimates. *Global Change Biol.*, **22**, 2198–2215, <https://doi.org/10.1111/gcb.13217>.
- Oliphant, A. J., 2012: Terrestrial ecosystem–atmosphere exchange of CO₂, water and energy from FLUXNET: Review and meta-analysis of a global in-situ observatory. *Geogr. Compass*, **6**, 689–705, <https://doi.org/10.1111/gec3.12009>.
- Oliveira, R. S., T. E. Dawson, S. S. O. Burgess, and D. C. Nepstad, 2005: Hydraulic redistribution in three Amazonian trees. *Oecologia*, **145**, 354–363, <https://doi.org/10.1007/s00442-005-0108-2>.
- Papale, D., and Coauthors, 2006: Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: Algorithms and uncertainty estimation. *Biogeosciences*, **3**, 571–583, <https://doi.org/10.5194/bg-3-571-2006>.
- Polley, H. W., R. L. Phillips, A. B. Frank, J. A. Bradford, P. L. Sims, J. A. Morgan, and J. R. Kiniry, 2011: Variability in light-use efficiency for gross primary productivity on Great Plains grasslands. *Ecosystems*, **14**, 15–27, <https://doi.org/10.1007/s10021-010-9389-3>.
- Ponce-Campos, G. E., and Coauthors, 2013: Ecosystem resilience despite large-scale altered hydroclimatic conditions. *Nature*, **494**, 349–352, <https://doi.org/10.1038/nature11836>.
- Potter, C. S., J. T. Randerson, C. B. Field, P. A. Matson, P. M. Vitousek, H. A. Mooney, and S. A. Klooster, 1993: Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochem. Cycles*, **7**, 811–841, <https://doi.org/10.1029/93GB02725>.
- Prentice, I. C., X. Liang, B. E. Medlyn, and Y.-P. Wang, 2015: Reliable, robust and realistic: The three R's of next-generation land-surface modelling. *Atmos. Chem. Phys.*, **15**, 5987–6005, <https://doi.org/10.5194/acp-15-5987-2015>.
- Reichstein, M., and Coauthors, 2005: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm. *Global Change Biol.*, **11**, 1424–1439, <https://doi.org/10.1111/j.1365-2486.2005.001002.x>.
- Roderick, M. L., F. Sun, W. H. Lim, and G. D. Farquhar, 2014: A general framework for understanding the response of the water cycle to global warming over land and ocean. *Hydrol. Earth Syst. Sci.*, **18**, 1575–1589, <https://doi.org/10.5194/hess-18-1575-2014>.
- Rosati, A., and T. M. Dejong, 2003: Estimating photosynthetic radiation use efficiency using incident light and photosynthesis of individual leaves. *Ann. Bot.*, **91**, 869–877, <https://doi.org/10.1093/aob/mcg094>.
- Ruimy, A., and Coauthors, 1999: Comparing global models of terrestrial net primary productivity (NPP): Analysis of differences in light absorption and light-use efficiency. *Global Change Biol.*, **5**, 56–64, <https://doi.org/10.1046/j.1365-2486.1999.00007.x>.
- Seager, R., M. Ting, C. Li, N. Naik, B. Cook, J. Nakamura, and H. Liu, 2013: Projections of declining surface-water availability for the southwestern United States. *Nat. Climate Change*, **3**, 482–486, <https://doi.org/10.1038/nclimate1787>.
- Shao, P., X. Zeng, K. Sakaguchi, R. K. Monson, and X. Zeng, 2013: Terrestrial carbon cycle: Climate relations in eight CMIP5 Earth system models. *J. Climate*, **26**, 8744–8764, <https://doi.org/10.1175/JCLI-D-12-00831.1>.
- Sims, D. A., and Coauthors, 2005: Midday values of gross CO₂ flux and light use efficiency during satellite overpasses can be used to directly estimate eight-day mean flux. *Agric. For. Meteorol.*, **131**, 1–12, <https://doi.org/10.1016/j.agrformet.2005.04.006>.
- Tang, X., and Coauthors, 2014: How is water-use efficiency of terrestrial ecosystems distributed and changing on Earth? *Sci. Rep.*, **4**, 7483, <https://doi.org/10.1038/srep07483>.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An Overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Todd-Brown, K. E. O., J. T. Randerson, W. M. Post, F. M. Hoffman, C. Tarnocai, E. A. G. Schuur, and S. D. Allison, 2013: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences*, **10**, 1717–1736, <https://doi.org/10.5194/bg-10-1717-2013>.
- Turner, D. P., S. Urbanski, D. Bremer, S. C. Wofsy, T. Meyers, S. T. Gower, and M. Gregory, 2003: A cross-biome comparison of daily light use efficiency for gross primary production. *Global Change Biol.*, **9**, 383–395, <https://doi.org/10.1046/j.1365-2486.2003.00573.x>.
- Veroustraete, F., H. Sabbe, and H. Eerens, 2002: Estimation of carbon mass fluxes over Europe using the C-Fix model and Euroflux data. *Remote Sens. Environ.*, **83**, 376–399, [https://doi.org/10.1016/S0034-4257\(02\)00043-3](https://doi.org/10.1016/S0034-4257(02)00043-3).
- Wang, Y. P., D. Baldocchi, R. Leuning, E. Falge, and T. Vesala, 2007: Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight FLUXNET sites. *Global Change Biol.*, **13**, 652–670, <https://doi.org/10.1111/j.1365-2486.2006.01225.x>.
- , X. J. Lu, I. J. Wright, Y. J. Dai, P. J. Rayner, and P. B. Reich, 2012: Correlations among leaf traits provide a significant constraint on the estimate of global gross primary production. *Geophys. Res. Lett.*, **39**, L19405, <https://doi.org/10.1029/2012GL053461>.
- Wenzel, S., P. M. Cox, V. Eyring, and P. Friedlingstein, 2014: Emergent constraints on climate-carbon cycle feedbacks in the CMIP5 Earth system models. *J. Geophys. Res. Biogeosci.*, **119**, 794–807, <https://doi.org/10.1002/2013JG002591>.
- Williams, C. A., and Coauthors, 2012: Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers. *Water Resour. Res.*, **48**, W06523, <https://doi.org/10.1029/2011WR011586>.
- Wong, S. C., I. R. Cowan, and G. D. Farquhar, 1979: Stomatal conductance correlates with photosynthetic capacity. *Nature*, **282**, 424–426, <https://doi.org/10.1038/282424a0>.
- Xue, B.-L., Q. Guo, A. Otto, J. Xiao, S. Tao, and L. Li, 2015: Global patterns, trends, and drivers of water use efficiency from 2000 to 2013. *Ecosphere*, **6**, 1–18, <https://doi.org/10.1890/ES14-00416.1>.
- Yang, D., F. Sun, Z. Liu, Z. Cong, and Z. Lei, 2006: Interpreting the complementary relationship in non-humid environments based on the Budyko and Penman hypotheses. *Geophys. Res. Lett.*, **33**, L18402, <https://doi.org/10.1029/2006GL027657>.
- , —, —, —, G. Ni, and Z. Lei, 2007: Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of China using the Budyko hypothesis. *Water Resour. Res.*, **43**, W04426, <https://doi.org/10.1029/2006WR005224>.
- Yang, H., D. Yang, Z. Lei, and F. Sun, 2008: New analytical derivation of the mean annual water-energy balance equation. *Water Resour. Res.*, **44**, W03410, <https://doi.org/10.1029/2007WR006135>.
- Yuan, W., and Coauthors, 2007: Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agric. For. Meteorol.*, **143**, 189–207, <https://doi.org/10.1016/j.agrformet.2006.12.001>.
- Zhang, L., W. R. Dawes, and G. R. Walker, 2001: Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water Resour. Res.*, **37**, 701–708, <https://doi.org/10.1029/2000WR900325>.
- , N. Potter, K. Hickel, Y. Zhang, and Q. Shao, 2008: Water balance modeling over variable time scales based on

- the Budyko framework—Model development and testing. *J. Hydrol.*, **360**, 117–131, <https://doi.org/10.1016/j.jhydrol.2008.07.021>.
- Zhang, L.-X., D.-C. Zhou, J.-W. Fan, and Z.-M. Hu, 2015: Comparison of four light use efficiency models for estimating terrestrial gross primary production. *Ecol. Modell.*, **300**, 30–39, <https://doi.org/10.1016/j.ecolmodel.2015.01.001>.
- Zhang, Y., C. Song, G. Sun, L. E. Band, A. Noormets, and Q. Zhang, 2015: Understanding moisture stress on light use efficiency across terrestrial ecosystems based on global flux and remote-sensing data. *J. Geophys. Res. Biogeosci.*, **120**, 2053–2066, <https://doi.org/10.1002/2015JG003023>.
- Zhao, Y., and Coauthors, 2012: How errors on meteorological variables impact simulated ecosystem fluxes: A case study for six French sites. *Biogeosciences*, **9**, 2537–2564, <https://doi.org/10.5194/bg-9-2537-2012>.
- Zhou, Y., and Coauthors, 2016: Global parameterization and validation of a two-leaf light use efficiency model for predicting gross primary production across FLUXNET sites. *J. Geophys. Res. Biogeosci.*, **121**, 1045–1072, <https://doi.org/10.1002/2014JG002876>.