Global Change Biology
celebrating 20 years
Global Change Biology (2014), doi: 10.1111/gcb.12473

Climate-driven uncertainties in modeling terrestrial energy and water fluxes: a site-level to global-scale analysis

RAHUL BARMAN, ATUL K. JAIN and MIAOLING LIANG[1]
Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Abstract

We used a land surface model constrained using data from flux tower sites, to analyze the biases in ecosystem energy and water fluxes arising due to the use of meteorological reanalysis datasets. Following site-level model calibration encompassing major vegetation types from the tropics to the northern high-latitudes, we repeated the site and global simulations using two reanalysis datasets: the NCEP/NCAR and the CRUNCEP. In comparison with the model simulations using observed meteorology from sites, the reanalysis-driven simulations produced several systematic biases in net radiation (R_n), latent heat (*LE*), and sensible heat (*H*) fluxes. These include: (i) persistently positive tropical/subtropical biases in R_n using the NCEP/NCAR, and gradually transitioning to negative R_n biases in the higher latitudes; (ii) large positive *H* biases in the tropics/subtropics using the NCEP/NCAR; (iii) negative *LE* biases using the NCEP/NCAR above 40°N; (iv) high tropical *LE* using the CRUNCEP in comparison with observationally derived global estimates; and (v) flux-partitioning biases from canopy and ground components. Across vegetation types, we investigated the role of the meteorological drivers (shortwave and longwave radiation, atmospheric humidity, temperature, precipitation) and their seasonal biases in controlling these reanalysis-driven uncertainties. At the global scale, our site-level analysis explains several model-data differences in the *LE* and *H* fluxes when compared with observationally derived global estimates of these fluxes. Using our results, we discuss the implications of site-level model calibration on subsequent regional/global applications to study energy and hydrological processes. The flux-partitioning biases presented in this study have potential implications on the couplings among terrestrial carbon, energy, and water fluxes, and for the calibration of land–atmosphere parameterizations that are dependent on *LE*/*H* partitioning.

Keywords: Biosphere, biogeophysics, energy cycle, FLUXNET, ISAM, land surface model, latent heat, sensible heat, uncertainties

Received 9 May 2013 and accepted 18 July 2013

Introduction

Uncertainties in energy and water fluxes in land surface models (LSMs) arise from the choice of schemes used to represent land surface and boundary layer processes (Sellers *et al.*, 1997; Overgaard *et al.*, 2006; Dickinson, 2011), and from the meteorological inputs (Santanello *et al.*, 2009). Quantification of such modeling uncertainties is becoming increasingly important to successfully study the implications of climate change on the terrestrial energy and hydrological cycles (Trenberth *et al.*, 2007, 2009). With the availability of multi-year eddy covariance data from FLUXNET (Baldocchi *et al.*, 2001), the parameterizations in LSMs can be constrained leading to improved estimates of terrestrial latent heat (*LE*) and sensible heat (*H*) fluxes at the site-level (e.g., Stöckli *et al.*, 2008; Blyth *et al.*, 2010, 2011). However, uncertainties in these fluxes continue to remain high even in the recent model-data intercomparison assessments (Jiménez *et al.*, 2011; Mueller *et al.*, 2011), calling for continued model evaluation and improvement. It remains important to systematically investigate the causes of these uncertainties, because discrepancies in simulated terrestrial *LE* and *H* fluxes can strongly influence the simulated climate in climate models through the land–atmosphere interactions (Ban-Weiss *et al.*, 2011).

While the impacts of specific model schemes/parameterizations on energy and water fluxes have been documented in many studies, fewer studies have focused on the meteorology-driven uncertainties. For example, to the best of our knowledge, only one study (Mu *et al.*, 2012) has recently documented the impacts of meteorological uncertainties on evapotranspiration biases (in remote sensing applications). Similar uncertainties can arise in the regional/global applications of LSMs, associated with biases in the meteorological reanalyses with

[1]Present address: Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA

Correspondence: Rahul Barman, tel. +217 372 7134, fax +217 372 1752, e-mail: account1.rahul@gmail.com; Atul K. Jain, tel. +217 333 2128, fax +217 372 1752, e-mail: jain1@illinois.edu

respect to weather station data. Significant biases in atmospheric reanalysis datasets have been acknowledged and documented in existing literature, with varying degrees of accuracy across meteorological variables (Janowiak *et al.*, 1998; Fekete *et al.*, 2004; Zhao *et al.*, 2006; Simmons *et al.*, 2007). In this context, previous studies reported large sensitivity of modeled carbon fluxes at regional/global scales due to various reanalysis datasets (Zhao *et al.*, 2006; Jung *et al.*, 2007; Barman *et al.*, 2013). However, the associated impacts on energy and water fluxes have not yet been documented. Because the calibration of energy/water fluxes in a LSM usually takes place utilizing ancillary meteorological data from the sites, subsequent applications of a 'calibrated' LSM using reanalysis datasets are likely to produce errors in the computed fluxes. The extent of such flux biases resulting directly due to site-level model calibration has also not been presented in literature.

Assessing the meteorology-driven (interchangeably referred to as 'climate-driven' in this study) *LE* and *H* biases is difficult at the global scale, because it is not feasible to obtain concurrent observations of meteorological and flux variables at a global scale (e.g., Henderson-Sellers *et al.*, 2003). A simple alternative may be to study the biases directly at the flux tower sites, by comparing the modeled *LE* and *H* fluxes obtained using observed meteorology vs. that using reanalysis datasets. Using this approach, we previously explored the impacts of climate-driven uncertainties on gross primary production (GPP) using one particular LSM, the Integrated Science Assessment Model (ISAM) (Barman *et al.*, 2013). There, we first optimized ISAM using meteorology and eddy covariance data from 25 FLUXNET sites, and subsequently applied the model to quantify the GPP biases using two reanalysis datasets: CRUNCEP and NCEP/NCAR. In the present study, we extend the aforementioned analysis for the *LE* and *H* fluxes.

Specifically, here we address the following questions: (i) what are the key biotic controls influencing *LE* and *H* fluxes in the calibrated model for various ecosystems; (ii) what are the flux biases using two different reanalysis datasets (CRUNCEP and NCEP/NCAR) directly at the flux tower sites; and (iii) what are the impacts on the global estimates of the corresponding fluxes. We also present the corresponding implications on the partitioning of *LE* and *H* fluxes – which can strongly influence the dynamics and thermodynamics of atmospheric circulation and biosphere–climate feedbacks (Lawrence *et al.*, 2007). Additionally, we analyzed the biases in partitioning of *LE* into canopy evapotranspiration and soil evaporation – which affects GPP and hydro-climatology (Lawrence & Chase, 2009). Along with our previous study (Barman *et al.*, 2013), this work presents a consistently integrated analysis of climate-driven biases in carbon, energy, and water fluxes using the same modeling framework.

Materials and methods

Energy/water cycle components in ISAM

Integrated Science Assessment Model computes terrestrial energy, water, and momentum fluxes at half-hourly to hourly time steps, integrated with prognostic carbon and nitrogen cycles (Jain *et al.*, 2009; Yang *et al.*, 2009; Barman *et al.*, 2013). The boundary layer turbulent processes are described based on the Monin–Obukhov Similarity Theory (MOST) approach (see Wang & Dickinson (2012) for method review). Latent heat transfer to atmosphere is resolved using canopy transpiration, ground evaporation, and canopy dew evaporation; sensible heat is partitioned into ground and canopy components (e.g., Sellers *et al.*, 1996; Oleson *et al.*, 2008). For the computation of soil evaporation, the model includes resistance to moisture transfer from soil, root, and litter components (Sakaguchi & Zeng, 2009). Surface albedo is resolved into ground albedo (function of soil color and wetness), exposed vegetation albedo (function of leaf orientation, leaf/stem reflectivity and transmissivity, and ground albedo), and snow albedo (schemes from Dai *et al.* 2004).

Simulated hydrology processes in ISAM include canopy interception and throughfall of precipitation, infiltration, redistribution of soil water within the soil column, surface, and subsurface runoffs – all adapted based on Oleson *et al.* (2008). The vertical soil column (ca. 50 m) is represented by 15 layers (adapted from Lawrence *et al.* 2008), consisting of 11 hydrologically active top layers up to total depth of ca. 6 m, and four hydrology inactive bedrock layers below. Originally, ISAM contained 10 hydrologically active layers up to a total depth of 3.5 (similar to Lawrence *et al.* 2008); however, for this study, we increased the hydrological zone to ca. 6 m to incorporate deep tropical roots implemented in ISAM (Barman *et al.*, 2013). Snow is discretized into a maximum of five layers, and dynamics include: snow accumulation, various modes of compaction (from snow weight, aging, melting, and winds, Schaefer *et al.*, 2009), depth hoar formation (Schaefer *et al.*, 2009), and water transfer across the snow layers.

The soil thermal and hydrological properties in ISAM vary with depth, depending on soil liquid and ice water contents, soil texture, and organic carbon profiles (Lawrence & Slater, 2008). In ISAM, the soil organic carbon dataset is taken from Harmonized World Soil Database (HWSD) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012; Todd-Brown *et al.*, 2013).

Data

FLUXNET Data. Table 1 lists the 25-flux tower sites used in this study, grouped into various plant functional types (PFTs). These represent major vegetation types in the global land surface: tropical broadleaf evergreen tree (Trop.BET) and broadleaf deciduous tree

Table 1 FLUXNET sites used in this study*; the annual budgets of *LE* and *H* were compiled from FLUXNET data, or from published studies based on the FLUXNET data

				LE (W m^{-2})		*H* (W m^{-2})	
				FLUXNET	ISAM**	FLUXNET	ISAM**
Site Code	Site Name	Years	EBR†	Mean ± Uncertainty		Mean ± Uncertainty	
Tropical Broadleaf Evergreen Tree (Trop.BET)							
LBA-Km34‡	Manaus KM34	2002–3004	0.82	106 ± 16	100	36 ± 13	31
LBA-Km67‡	Santarem KM83	2003–2004	0.84	104 ± 15	78	23 ± 12	19
LBA-Km83‡	Santarem KM67	2001–2003	1.04	103 ± 15	105	24 ± 12	21
LBA-Rja§	Reserva Jaru	2000–2001	0.77	106 ± 16	102	35 ± 13	41
Tropical Deciduous Evergreen Tree (Trop.BDT)							
LBA-Ban	Bananal Island	2004–2004	1.01	106 ± 16	106	29 ± 13	31
Temperate Broadleaf Deciduous Tree (Temp.BDT)							
CA-Oas	South OldAspen	1997–2004	0.84	30 ± 10	31	28 ± 14	29
US-Syv	Sylvania Wilderness	2002–2004	NA¶	NA‖	35	NA‖	42
US-WCr	Willow Creek	1999–2004	NA¶	NA‖	27	NA‖	38
Needleleaf Evergreen Tree (NET)							
CA-Gro	Groundhog River	2004–2004	0.96	40 ± 12	31	35 ± 13	42
CA-Obs	South OldBlackSpruce	2000–2004	0.85	26 ± 10	28	46 ± 14	41
CA-Ojp	South OldJackPine	2000–2003	0.87	20 ± 9	29	45 ± 14	41
CA-Qfo	East OldSpruce	2004–2004	0.85	32 ± 10	30	42 ± 13	34
US-Me3	Metolius 2nd YoungPine	2004–2004	0.63	NA‖	27	NA‖	55
US-NR1	Niwot Ridge	1999–2004	0.97	48 ± 13	48	49 ± 19	59
Savanna							
US-Ton	Tonzi Ranch	2002–2004	0.94	NA‖	29	NA‖	75
LBA-Pdg‡	Reserva Pe-de-Gigante	2001–2003	0.75	95 ± 7	75	38 ± 10	40
Grass							
CA-Let	Lethbridge	1999–2004	0.67	31 ± 6	22	39 ± 10	44
US-Shd	Shidler Tallgrass Prairie	1998–1999	NA¶	NA‖	70	NA‖	36
US-Var	Vaira Ranch	2001–2004	0.90	NA‖	32	NA‖	83
Shrub							
CA-Mer	Mer Bleue	1999–2004	0.80	NA‖	33	NA‖	48
US-Los	Lost Creek	2001–2004	0.81	NA‖	22	NA‖	45
US-SO2	Sky Oaks Old	1999–2004	1.03	NA‖	40	NA‖	62
Tundra							
US-Atq	Atqasuk	2004–2004	0.67	NA‖	11	NA‖	25
US-Brw	Barrow	2001–2001	0.82	NA‖	8	NA‖	17
Pasture							
LBA-Fns§	Fazenda Nossa Senhora	2000–2001	0.77	73 ± 6	81	47 ± 10	49

*Additional detail of these sites pertaining to site description, and references/methods are available in Barman *et al*. (2013).

†EBR, Energy Balance Ratio (see Wilson *et al*., 2002), here calculated using 3-hourly FLUXNET data for *LE*, *H* and R_n according to formula: $\mathrm{EBR}=\frac{\sum LE+H}{\sum R_n}$, where Σ is summation over the timescale for flux correction (see Data S2).

‡Annual *LE*, *H*, and EBR were calculated based on data from Da Rocha *et al*. (2009) (see Table S2 for calculation details).

§Annual *LE* and *H* were calculated based on data from Randow *et al*. (2004), and EBR from Hasler & Avissar (2007).

¶NA (not available) as net-radiation data was not available for the calculation of EBR.

‖NA (not available) as percentage of missing data exceeded ca. 40% for the years used in this study; hence we did not list the FLUXNET Mean ± Uncertainty (Table S1).

**Model simulated GPP using a model version *ISAM-FLUXNET*.

The annual budgets from the published studies shown here may not correspond to the site-years used in this study. The ISAM estimated fluxes are from the *ISAM-FLUXNET* simulations.

(Trop.BDT), temperate broadleaf deciduous tree (Temp.BDT), needleleaf evergreen tree (NET), savanna, grass, shrub, tundra, and pasture. For this study, we used the half-hourly/hourly u*-corrected *LE*, *H*, and net radiation (R_n) data from the FLUXNET database to calibrate/evaluate the model estimated energy fluxes

(see Data S1 for details). Pronounced gaps in data were present in most site measurements (e.g., Table S1 – for the North American sites), consistent with those reported in previous published studies (Falge *et al.*, 2001). For the present analysis, we calculated the annual estimates of *LE* and *H* at a site if the cumulative missing data at the site was <40% of the entire study years (see Table S1). This is a more relaxed threshold value than used in some other previously published studies. For example, Law *et al.* (2002) used 25% yr^{-1} threshold value to correct the *LE* and *H* fluxes. We converted the half-hourly/hourly data to daily, and subsequently filled the daily data gaps using linear interpolation, before aggregating to monthly/annual estimates. For sites not satisfying the missing threshold, we used the flux-corrected annual estimates of *LE* and *H* from published literature, whenever available (Table 1). This was usually a case for the South American (LBA) sites, and hence we used flux-corrected *LE* and *H* at all LBA sites based on published literature (Table S2). At all sites, we performed flux correction to enforce energy balance closure (Twine *et al.*, 2000; Jung *et al.*, 2011) on a monthly basis (details are provided in Data S2). The resulting mean annual estimates of *LE* and *H*, as used for in this analysis, are listed in Table 1.

We also calculated the random flux errors stochastically at half-hourly/hourly timescales based on Hollinger & Richardson (2005), and subsequently aggregated them to annual timescales (±Uncertainty in Table 1) for the purposes of our model-data comparison. However, we note that the annual random uncertainties are most likely to be lower than these estimates, primarily due to compensation from positive and negative uncertainties (Richardson *et al.*, 2006); hence, our ±Uncertainty estimates in Table 1 should only be representative of the maximum theoretical bounds in random measurement uncertainties.

Forcing data and model experimental setup. We performed three off-line simulations at every flux tower site: (i) *ISAM-FLUXNET* – using the observed site-level meteorology; (ii) *ISAM-NCEP* – using the NCEP/NCAR reanalysis (Qian *et al.*, 2006); and (iii) *ISAM-CRUNCEP* – using the recently available CRUNCEP reanalysis (Viovy & Ciais, 2009; Wei *et al.*, 2013). For details on boundary data, model spin-up procedure, the readers are referred to our companion study (Barman *et al.*, 2013) where we used identical experimental setups.

We first calibrated and evaluated the model using the *ISAM-FLUXNET* simulations. Subsequently, we computed the site-level biases (Δ) in *LE*, *H* and R_n (=*LE* + *H*) fluxes in the reanalysis-driven simulations (*ISAM-NCEP* and *ISAM-CRUNCEP*) with respect to the *ISAM-FLUXNET* simulations. Similarly, for site-level biases in any input/output variable, we consistently computed the biases with respect to the corresponding *ISAM-FLUXNET* variables. In all subsequent discussions, the use of the 'Δ' notation in the prefix of variable names indicates site-level biases.

ISAM calibration, and evaluation of LE and H

Utilizing the observed site-level meteorology and eddy covariance data, we used the 'trial and error' approach to tune several PFT-specific parameters in ISAM, to concurrently optimize the modeled GPP, *LE*, and *H*. An analogous model calibration approach has also been used in other studies using the ISAM framework (El-Masri *et al.*, 2013; Song *et al.*, 2013). During model calibration of each PFT, our goal was to optimize the overall model performance across sites within the PFT. For the overall calibration approach, including key parameters influencing GPP (e.g., $V_{cmax25}^{opt} f(N)$, plant rooting depths, etc.; Barman *et al.* 2013. Here, we summarize the choice of other model parameters related to *LE* and *H* (such as stomatal conductance parameters, leaf/stem optical properties), and the evaluation of modeled *LE* and *H* fluxes.

The stomatal conductance in the model (g_s, μmol m^{-2} s^{-1}) is parameterized using two tunable parameters, the stomatal conductance slope (*m*, dimensionless) and the stomatal conductance intercept (*b*, μmol m^{-2} s^{-1}), based on the following equation (Collatz *et al.*, 1991; Sellers *et al.*, 1996; Dai *et al.*, 2004):

$$g_s = m\frac{A_n}{C_s/P_{atm}} \times \frac{e_s}{e_i} + b\beta_t$$

where A_n is net photosynthesis (μmol m^{-2} s^{-1}), C_s the CO_2 partial pressure at the leaf surface (Pa), P_{atm} the atmospheric pressure (Pa), e_s the vapor pressure at leaf surface (Pa), e_i the saturation vapor pressure inside the leaf (Pa), and β_t a soil water availability factor between 0 and 1 (dimensionless, 1 implies no soil moisture stress on photosynthesis, while 0 implies no available water to plant roots). Generally, it was sufficient to use the values of *m* and the *b* directly from existing literature ($m = 9$ for C_3 and 4 for C_4; $b = 0.01$ mol-CO_2 m^{-2} s^{-1} for C_3, and 0.04 mol-CO_2 m^{-2} s^{-1}) (Dai *et al.*, 2004; Bonan *et al.*, 2011). However, for certain PFTs (Trop.BET/BDT, C_3 grass), these values produced lower than observed *LE* in the model. For these PFTs, we slightly increased their values to *m* ca. 10 and *b* ca. 0.02–0.03 mol-CO_2 m^{-2} s^{-1}, resulting in increased *LE* in the model to better match with the *LE* data (and also improving *LE*/*H* partitioning).

For PFT optical properties, the model requires the values of reflectance and transmittance for leaves and

stems of each PFT – separately for the visible and near infrared bands. While these are not tunable parameters, they need to be constrained in the model based on data. The original ISAM used the reflectance/transmittance values from the published CoLM model (Dai *et al*., 2004; Chen *et al*., 2011). There, the optical properties of grass/crop type PFTs were obtained from Dorman & Sellers (1989), and other modeling studies reported errors in simulated grass/crop albedos using these estimates (e.g., Lawrence *et al*., 2011). Hence, following Lawrence *et al*. (2011), we replaced these values for the grass type PFTs (e.g., grass, tundra, pasture) in the calibrated model based on improved field data from Asner *et al*. (1998).

In the calibrated model, the annual *H* estimates were within the random uncertainty range at all the sites with available flux data (15 out of the 25 sites; see Table 1), and the *LE* was generally within the uncertainty range as well (except at LBA-Km67, LBA-Pdg, CA-Let, and LBA-Fns). Below, we briefly describe the potential causes for *LE* disagreements at the aforementioned sites where the model was outside the uncertainty range.

At the LBA-Km67 Trop.BET site, there was a large underestimation in modeled annual *LE* (ISAM: 78 W m^{-2}, FLUXNET: 104 ± 15 W m^{-2}). Based on other site-level observational studies that have shown energy-limited *LE* response for the Trop.BET/BDT (Hasler & Avissar, 2007; Juárez *et al*., 2007), this underestimation could be largely attributed to the lower total radiation inputs at this site in comparison with the adjacent Trop.BET/BDT sites. For example, in comparison with LBA-Km83 that is in the same model 0.5 × 0.5 grid, the mean annual shortwave radiation (*Srad*) at this site was lower by 30 W m^{-2}, while the downwelling longwave radiation (LW_{down}) was same at both sites (Table S3). Hence, the total $Srad + LW_{down}$ at this site was lower by ca. 30 W m^{-2}, strongly contributing to the lower modeled *LE*. Noting that there may be problems in the radiation data at this site, the ratio of *LE*/(*LE* + *H*) may be a better statistic for model evaluation at this site, which compared favorably between the model and the data (ISAM: 0.80, FLUXNET mean: 0.82).

The tropical savanna site (LBA-Pdg) also underestimated the annual averaged *LE* (ISAM: 75 W m^{-2}, FLUXNET: 95 ± 7 W m^{-2}). This was partly due to the underestimation of simulated net radiation in the model (R_n; ISAM: 115 W m^{-2}, FLUXNET: 133 W m^{-2}). Because R_n depends on the full suite of energy processes (including the incoming and outgoing energy balance, soil temperature, vegetation/ground albedos, etc.), it is generally difficult to sufficiently diagnose/constrain R_n in the model. We tested with different estimates of savanna reflectance/transmittance from Asner (1998), and different root depth/profile values for savanna (from Schenk & Jackson 2002); however, these did not provide the required sensitivity to correct the existing bias at this site. In addition, the use of static rooting profiles in ISAM was also most likely to limit the model's ability to correctly simulate the soil water stress (especially important for herbaceous PFTs; Arora & Boer, 2003). Due to the availability of only the LBA-Pdg for model calibration/evaluation of tropical savanna, we could not further evaluate these modeling shortcomings. Furthermore, this site also had a relatively poor energy balance ratio (EBR) in the flux tower measurements (see Table 1), reducing the confidence in model-data comparison.

At the midlatitude CA-Let grass site, the underestimation of *LE* (ISAM: 22 W m^{-2}, FLUXNET: 31 ± 6 W m^{-2}) could be explained by the upper bound of annual precipitation (*Precip*) in the meteorology data (*Precip* = 0.28 m m^{-2} yr^{-1}, see Table S3), resulting in *ET*/*Precip* = 0.987 (where *ET* = evapotranspiration). Finally, at the single available pasture site (LBA-Fns), the *LE* was slightly overestimated – though the differences were not substantial (ISAM: 81 W m^{-2}, FLUXNET: 73 ± 6 W m^{-2}). Overall, due to the limited number of non-tree/herbaceous sites for model calibration (e.g., grass, savanna, pasture), more sites will be beneficial to reduce the existing uncertainties in the model.

Results

Site-level reanalysis-driven uncertainties in R_n, LE, and H

Uncertainties in annual estimates of R_n. To explain the reanalysis-driven modeled annual Δ*LE* and Δ*H*, we first analyzed the net-radiation biases ($\Delta R_n = \Delta LE + \Delta H$) at each site, along with biases in several input meteorology variables (Fig. 1). With respect to the input meteorological drivers, two primary variables controlling the R_n are the *Srad* and LW_{down}. In the tropics (i.e. LBA sites), the ΔR_n using the NCEP/NCAR dataset was very high – primarily driven by the consistently $+\Delta LW_{down}$ and $+\Delta Srad$ (Fig. 1a, d–e). Indeed, the NCEP/NCAR driven $+\Delta LW_{down}$ at the LBA sites were the largest among all the analyzed sites (ca. 20–50 W m^{-2}, Fig. 1e). In contrast to the tropical Trop.BET/BDT, the NCEP/NCAR driven ΔR_n were progressively negative at the mid- and high-latitude sites. For the Temp.BDT and NET, mean annual Δ*Srad* in the NCEP/NCAR were usually small, and consequently the $-\Delta R_n$ was driven by the $-\Delta LW_{down}$ (typically around ca. −15 W m^{-2}, see Fig. 1e). Such negative ΔLW_{down} were also present in

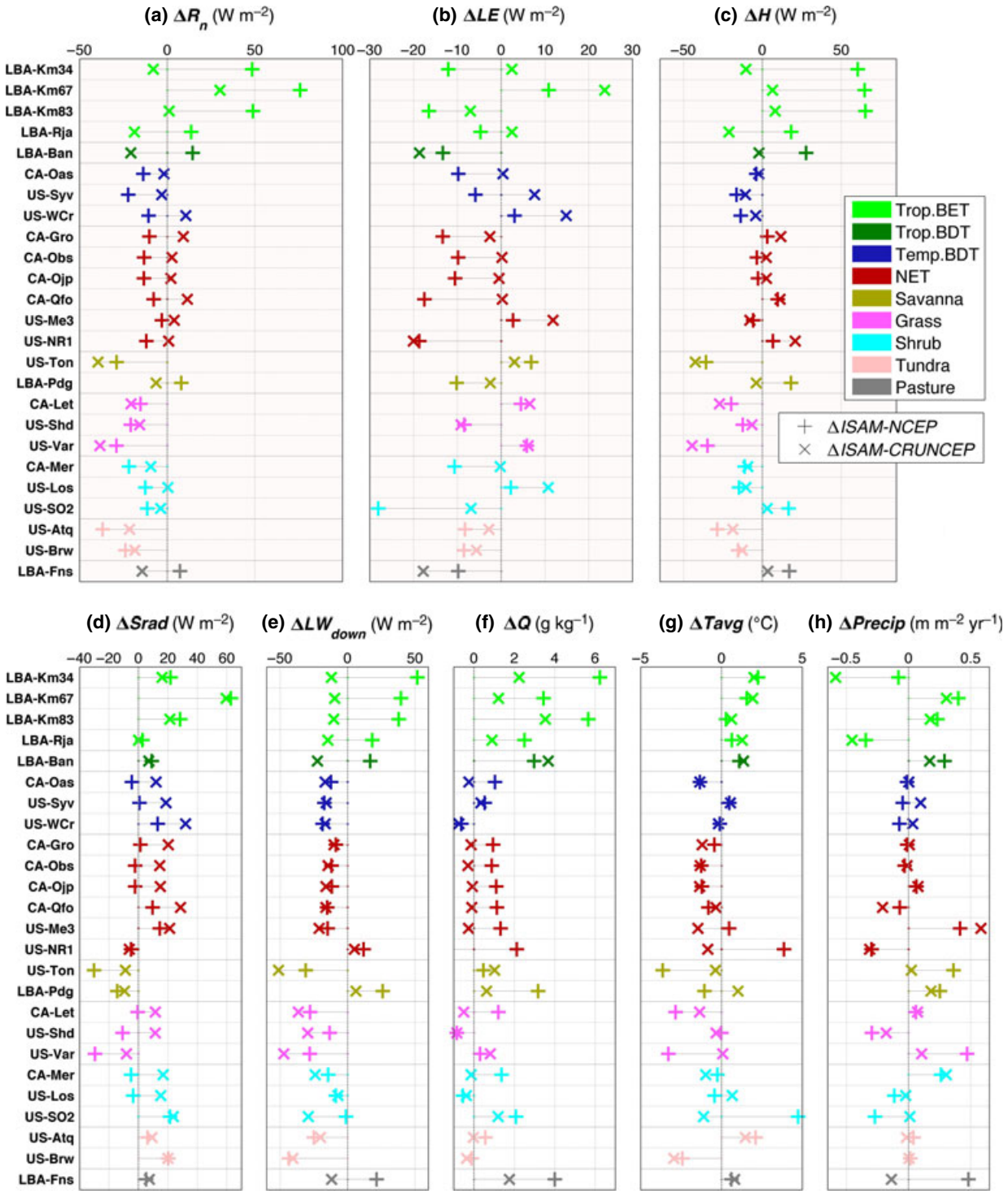

Fig. 1 Site-level mean annual biases (Δ) in (a) net radiation ($\Delta R_n = \Delta LE + \Delta H$), (b) latent heat ($\Delta LE$), and (c) sensible heat (ΔH) in the *ISAM-CRUNCEP*, *ISAM-NCEP* simulations. (d–h) Mean annual biases in input meteorology variables: (d) Δ*Srad*, (e) ΔLW_{down}, (f) Δ*Q*, (g) Δ*Tavg*, and (h) Δ*Precip*. All the biases were calculated with respect to the *ISAM-FLUXNET* counterpart.

the CRUNCEP; however, the annual total energy input were compensated by the +Δ*Srad* in the CRUNCEP (Fig. 1d). As a result, at most of the mid/high-latitude sites, the mean annual ΔR_n using the CRUNCEP was in better agreement with the respective *ISAM-FLUXNET* simulations.

For the nontree sites (savanna, grass, shrub, tundra, pasture), all except the NCEP/NCAR driven LBA

simulations produced a negative mean annual ΔR_n (Fig. 1a). For these cases with negative R_n, the reanalyses ΔLW_{down} were also negative. In many instances, the mean annual ΔR_n at the mid/high-latitude nontree sites were more negative than those simulated at the Temp.BDT and NET sites; for such cases, the corresponding input ΔLW_{down} were also among the most negative (e.g., US-Ton and US-Var, using the CRUNCEP data). At the high-latitude tundra sites in Alaska, the mean annual ΔR_n were significantly negative in both the datasets, driven by the corresponding $-\Delta LW_{down}$ which offset the $+\Delta Srad$. Overall, the patterns of ΔR_n as seen in the site-level analysis were also present in the global ISAM simulations using the two reanalysis datasets (further discussed in the section 'Global uncertainties in modeled R_n, *LE* and *H*).

Uncertainties in annual estimates of LE and H. Among the Trop.BET/BDT sites, even though the mean annual ΔR_n were positive in most instances (especially large using the NCEP/NCAR), the ΔLE were either negative or moderately positive (Fig. 1a–b). In this context, here we should also mention that the corresponding mean annual ΔGPP were positive at all the Trop.BET/BDT sites (average of ca. 0.45 kgC m^{-2} yr^{-1}) (Barman *et al.*, 2013). Hence, given the positive input energy anomaly (i.e. $\Delta R_n > 0$) together with ΔGPP > 0, the negative ΔLE suggest that the $+\Delta Q$ played a dominant role in suppressing the *LE* (by notably decreasing the atmospheric dryness). Consequently, this also means increased ecosystem water use efficiency in the corresponding reanalysis simulations. A notable characteristic of the NCEP/NCAR data is the high $+\Delta Q$ at the LBA sites (Fig. 1f). Further analysis showed that this was also true for most of the vegetated land surface, as shown by the consistently larger *Q* in the NCEP/NCAR than in the CRUNCEP dataset (Fig. S1e). For example, at two of the LBA Trop.BET sites (LBA-Km34 and LBA-Km83), the mean annual $+\Delta Q$ in the NCEP/NCAR data were among the highest of all the analyzed sites; there, the resulting ΔLE were also largely negative. Of all the LBA sites, only at LBA-Ban the ΔQ was relatively larger in the CRUNCEP, and the corresponding ΔLE using the CRUNCEP was more negative. These results highlight the importance of ΔQ in the reanalysis data for determining the direction of ΔLE in the tropics. Nonetheless, because the annual *LE response* at the tropics is known to be strongly dependent on annual R_n (e.g., Hasler & Avissar, 2007; Juárez *et al.*, 2007; Costa *et al.*, 2010), the resulting ΔLE can be positive given a sufficiently large $+\Delta R_n$ in the reanalysis data. This was the case for the LBA-Km67 site, where the impact of $+\Delta Q$ was more than offset by the increased radiation inputs in the NCEP/NCAR and the CRUNCEP datasets (ΔR_n ca. 75 and 30 W m^{-2}, respectively; $+\Delta Srad$ ca. 60 W m^{-2} in both the datasets), resulting in predominantly positive mean annual ΔLE of ca. 10 and 25 W m^{-2}, respectively.

Driven by the large $+\Delta R_n$ and $-\Delta LE$ (from $+\Delta Q$) in the NCEP/NCAR, the corresponding $+\Delta H$ was significantly high at all the Trop.BET/BDT sites. On average, the NCEP/NCAR driven ΔH for the Trop.BET (4 sites) was 52 W m^{-2}, an error of 186% of the corresponding *ISAM-FLUXNET H* (Table 2). At the single Trop.BDT site (LBA-Ban), the corresponding ΔH and% errors were 28 W m^{-2} and 89%, respectively. In comparison, the overall ΔH using the CRUNCEP was much smaller for these PFTs (Table 2).

Due to the $-\Delta R_n$ and $+\Delta Q$ using the NCEP/NCAR, the associated ΔLE were systematically negative at the NET sites (−20 W m^{-2} to −10 W m^{-2}) (Fig. 1b). An exception to this was the US-Me3 site, where the large

Table 2 Site-averaged annual latent heat (*LE*) and sensible heat (*H*) estimates, grouped by PFTs. ΔCRUNCEP = *ISAM-CRUNCEP – ISAM-FLUXNET*, ΔNCEP = *ISAM-NCEP – ISAM-FLUXNET*. Values in parenthesis are percentage differences with respect to *ISAM-FLUXNET* simulations

PFT	Number of sites	*LE* (W m^{-2})			*H* (W m^{-2})		
		ISAM	ΔCRUNCEP	ΔNCEP	ISAM	ΔCRUNCEP	ΔNCEP
Trop.BET	4	96	5 (6)	−6 (−6)	28	−4 (−15)	52 (186)
Trop.BDT	1	106	−19 (−18)	−13 (−12)	31	−2 (−7)	28 (89)
Temp.BDT	3	31	8 (24)	−4 (−13)	37	−6 (−16)	−11 (−31)
NET	6	32	−2 (−6)	−11 (−35)	45	7 (15)	1 (3)
Savanna	2	57	1 (2)	−0 (−0)	57	−24 (−42)	−10 (−17)
Grass	3	41	1 (3)	1 (2)	54	−26 (−48)	−22 (−41)
Shrub	3	32	1 (4)	−12 (−38)	52	−6 (−11)	−3 (−6)
Tundra	2	9	−4 (−46)	−8 (−90)	21	−16 (−74)	−22 (−103)
Pasture	1	81	−18 (−22)	−10 (−12)	49	3 (7)	17 (35)

+Δ*Precip* in both the reanalysis datasets produced +Δ*LE*. This shows that besides the energy inputs, the incoming water from precipitation can also be a limiting factor to the ET (or *LE*) in these ecosystems. Overall, averaged over the six NET sites, the Δ*LE* of −11 W m^{-2} using the NCEP/NCAR amounted to an error of −35% in comparison to the *LE* from *ISAM-FLUXNET* (Table 2). Across the individual Temp.BDT sites, the Δ*LE* using the NCEP/NCAR was negative to moderately positive (−10 ca. 2.5 W m^{-2}), while the mean annual Δ*LE* using the CRUNCEP were approximately 0–15 W m^{-2}. Based on Fig. 1a, a consistent feature at the Temp.BDT and NET sites was that the mean annual *LE* using the CRUNCEP were systematically greater than those using the NCEP/NCAR, i.e. $LE_{\mathrm{CRUNCEP}} > LE_{\mathrm{NCEP/NCAR}}$. These in turn, could be explained by the energy biases in the reanalysis simulations, i.e. $R_{\mathrm{n\ CRUNCEP}} > R_{\mathrm{n\ NCEP/NCAR}}$ (Fig. 1a).

For the nontree sites/PFTs, the reanalysis-driven Δ*LE* exhibited lesser consistent patterns than the tree/forest PFTs (Fig. 1b). This can partly be explained by the stronger dependence of *LE* on *Precip* in these ecosystems, which tend to be very spatially heterogeneous (Fig. 1e). In Barman *et al.* (2013), we showed that the annual ΔGPP for these ecosystems was strongly controlled by factors affecting ecosystem water stress, e.g., *Tavg*, *Q*, and *Precip*. Correspondingly, these factors generally controlled the modeled Δ*LE* as well – attributable to the stomatal coupling between GPP and transpiration. In addition, the influence of input energy over the modeled *LE* (or equivalently the *ET*) was also apparent at several of the nontree sites (similar to the tree PFTs). For example, while the *ET* increased with R_n, for monthly R_n > ca. 150 W m^{-2} some PFTs such as grass and savanna showed radiation-induced heat stress in the model – thereby progressively decreasing the *ET* with further increases in R_n (Fig. S2). Because these ecosystems are generally prevalent in drier environments, the instances of very large $+\Delta R_n$ can therefore potentially produce a –Δ*LE* using the reanalysis data (as opposed to the tree/forest PFTs).

Following the negative annual $-\Delta R_n$ at most of the mid/high-latitude nontree sites, the corresponding mean annual Δ*H* were also largely negative (Fig. 1a, c; Table 2). In this context, the only exception was the US-SO2 shrub site, where the *LE* (and the GPP) was strongly reduced in the reanalysis datasets due to very low *Precip*. To maintain the energy balance ($\Delta R_{n\ =}\ \Delta LE + \Delta H$), this resulted in the +Δ*H* using the reanalysis datasets. At the two high-latitude tundra sites (US-Atq, US-Brw), the mean annual ΔR_n were significantly negative in both the datasets, and the Δ*LE* and Δ*H* were both negative.

Uncertainties in partitioning of annual LE into canopy and ground components. Across all sites/PFTs, we analyzed the relative role of ΔSrad and $\Delta LW_{\mathrm{down}}$ on Δ*LE* by separately plotting the ground and canopy components of *LE* (i.e. LE_{ground} – ground evaporation, and LE_{veg} – canopy evapotranspiration). In the *ISAM-FLUXNET* simulations, based on the LAI, the LE_{ground}/LE increased nonlinearly from 0.1 for Trop.BET to 0.6 for NET, and reaching a maximum of ca. 0.8 for the extremely high-latitude tundra sites (Fig. 2a). In ISAM, LE_{veg} consists of two components – the canopy transpiration that is coupled with GPP (through stomatal conductance), and the leaf evaporation of dew. While energy input from *Srad* directly influences the total *LE*, LW_{down} only impacts the evaporation components (ground evaporation, leaf evaporation) due to the lack of any photosynthetically active radiation contained in it (hence LW_{down} does not influence either GPP or transpiration). Hence, these results indicate the increasing importance of $\Delta LW_{\mathrm{down}}$ toward Δ*LE* for PFTs with lower LAI where the ground is increasingly exposed (i.e. midlatitude forests with strong LAI seasonality, nontree PFTs). Consequently, at most of the mid- to high-latitude sites, the reanalysis datasets produced negative annual $\Delta LE_{\mathrm{ground}}$ (Fig. 2b), by virtue of the $-\Delta LW_{\mathrm{down}}$. Note that, because the nontree PFTs usually have lower LAI than tree PFTs, the role of Δ*Srad* toward $\Delta LE_{\mathrm{ground}}$ (and $\Delta H_{\mathrm{ground}}$) also becomes relatively more important, because the exposed ground can absorb a larger fraction of the *Srad*, thereby increasing the ground heat intake (data not shown). Also, as especially evident from the results, the LE_{ground}/LE at the mid- and high-latitude sites may be much lower in the *ISAM-NCEP* simulations than in the *ISAM-FLUXNET* counterparts; these are driven by the lower atmospheric energy input at these sites using the NCEP. Therefore, the regression fits for LE_{ground}/LE vs. LAI (Fig. 2a) may be different for the three sets of simulations performed in this study, though there is always a decrease in LE_{ground}/LE with increasing LAI.

Uncertainties in seasonality of LE and H. We investigated the daily patterns of Δ*LE* and Δ*H* because of their implications on the seasonality of the fluxes themselves. Here, we illustrate the results for the forest/tree PFTs (Fig. 3; for reference, daily patterns of Δ*LE* and Δ*H* for the nontree sites are also shown in Figs S3 and S4). The results show several points, as follows. (i) For the Trop.-BET/BDT, daily instances of high +Δ*Q* produced –Δ*LE*, and daily instances of high $+\Delta LW_{\mathrm{down}}$ produced high +Δ*H* (supplementing our analysis at annual timescales). (ii) The daily climatological biases in Δ*LE* and Δ*H* in the mid/high-latitude Temp.BDT and NET exhibited strong seasonality ranging from negative to positive

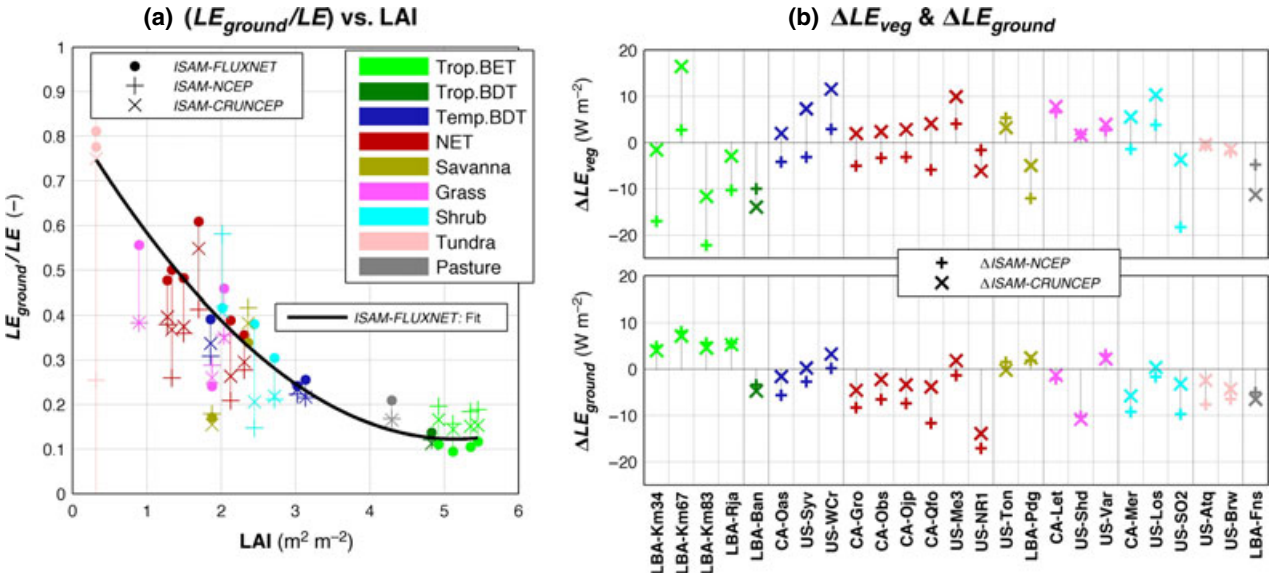

Fig. 2 (a) Ratio of annual ground evaporation to total evapotranspiration (LE_{ground}/LE) plotted vs. mean annual LAI, at individual sites. Also shown is a line of best fit (using a quadratic polynomial) across all sites for the *ISAM-FLUXNET* simulation. (b) Mean annual biases in latent heat components from the canopy (ΔLE_{veg}, top panel) and from ground (ΔLE_{ground}, bottom panel). All the biases (Δ) were calculated with respect to the *ISAM-FLUXNET* counterpart.

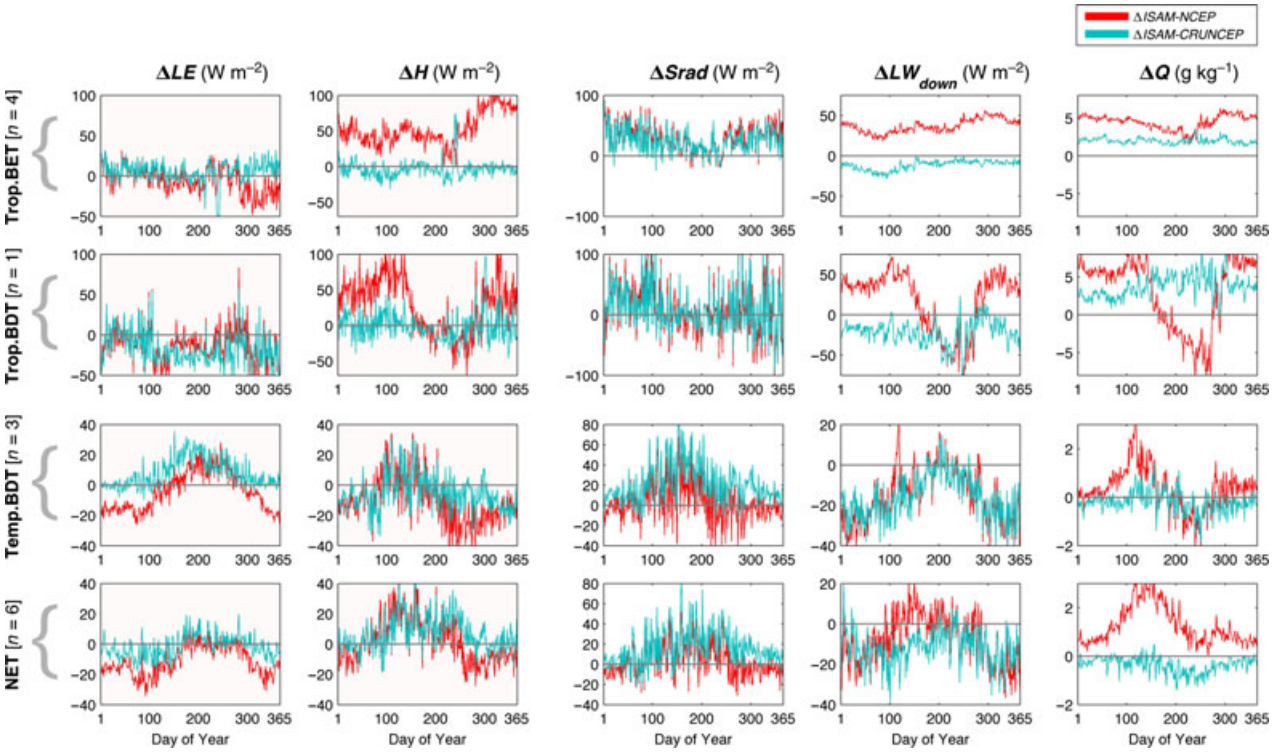

Fig. 3 Analysis for tree PFTs (Trop.BET, Trop.BDT, Temp.BDT, NET): daily climatology of reanalysis-driven ΔLE and ΔH, along with $\Delta Srad$, ΔLW_{down} and ΔQ. All variables were averaged over the available number of sites (n) for each PFT. All the biases (Δ) were calculated with respect to the *ISAM-FLUXNET* counterpart. Each row corresponds to a PFT group (name on left corner). Each column shows a variable (name on top). For each subplot, the *x*-axis is the 'Day of year' and the *y*-axis is the respective variable.

biases (e.g., seasonal amplitude of ca. −20 to 20 W m^{-2}). As a result, the ± seasonal flux biases may partly be mitigated in the annual timescale. Because of this, an important caveat is that a low annual energy flux bias may or may not imply correctness of the reanalysis data. (iii) For the Temp.BDT and NET, the systematically negative annual Δ*LE* using the NCEP/NCAR data was caused mostly by the negative wintertime *LE* biases, coincident with wintertime $-\Delta LW_{down}$ in the data. Because there is no GPP in these PFTs during the winter (and hence negligible LE_{veg}), the entire wintertime −Δ*LE* could be attributed to the corresponding $-\Delta LE_{ground}$. Finally, (iv) the mean annual −Δ*H* in the Temp.BDT and NET could also be mostly attributed to negative wintertime Δ*H* (driven by the input energy deficit during the winter).

Global uncertainties in modeled R_n, LE and H

Accurately quantifying the global extent of reanalysis, climate-driven biases in *LE* and *H* fluxes (as performed at the site-level) are not possible in a LSM framework, due to the lack of observed subdaily meteorological data at each model grid cell (e.g., Henderson-Sellers *et al.*, 2003). However, it may still be possible to investigate the modeling uncertainties in comparison with other observationally derived global estimates. For this purpose, we used data from two globally gridded sources: (i) FLUXNET-MTE (Jung *et al.*, 2011) – providing *LE* and *H* and (ii) MODIS (Mu *et al.*, 2011) – providing *LE*. Using these datasets and the two model simulations (*ISAM-NCEP*, *ISAM-CRUNCEP*) (see Figs 4 and 5 for spatial comparison), we constructed the corresponding zonally averaged mean annual R_n (=*LE* + *H*), *LE*, and *H* estimates (Fig. 6a–c) for the vegetated land surface (Fig. S5). Based on this plot, several features consistent with our site-level analysis are evident, which we describe below.

In comparison with FLUXNET-MTE, the mean annual R_n *anomaly* using the NCEP/NCAR meteorology was highly positive in the tropics and subtropics, but gradually transitioned to negative for the mid- and northern high-latitudes. In both the Northern and Southern Hemispheres, this transition happened at approximately 40°N and 40°S, respectively (Fig. 6a). In terms of the driving meteorological variables of *Srad* and LW_{down}, the relative differences between the CRUNCEP and the NCEP/NCAR were also very consistent with our site-level analysis. For example, in comparison with the CRUNCEP data, the annual

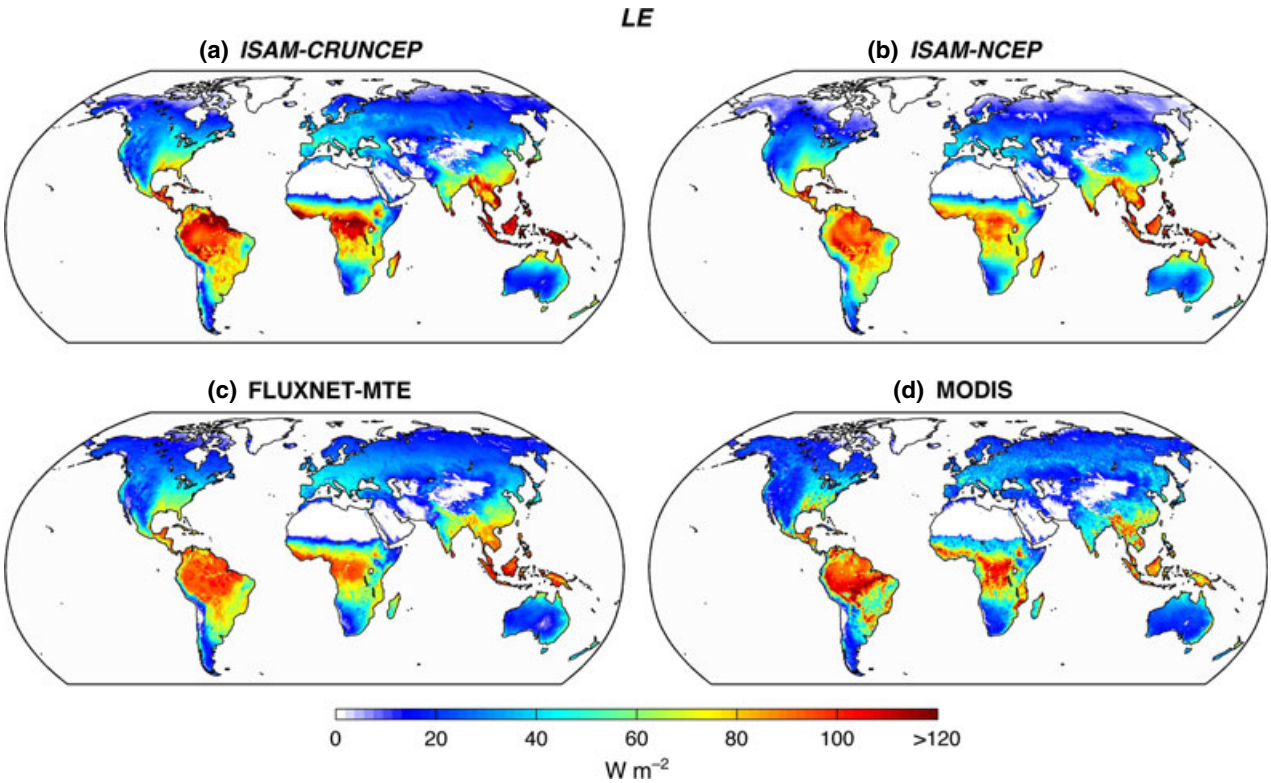

Fig. 4 Maps of mean annual estimates of *LE*, for two reanalysis-driven model simulations [*ISAM-CRUNCEP* (a), *ISAM-NCEP* (b)] and two observationally derived datasets [FLUXNET-MTE (c), MODIS (d)]. All the results are based on averaged output for 2000–2004, and are only for vegetated land surfaces (Fig. S5).

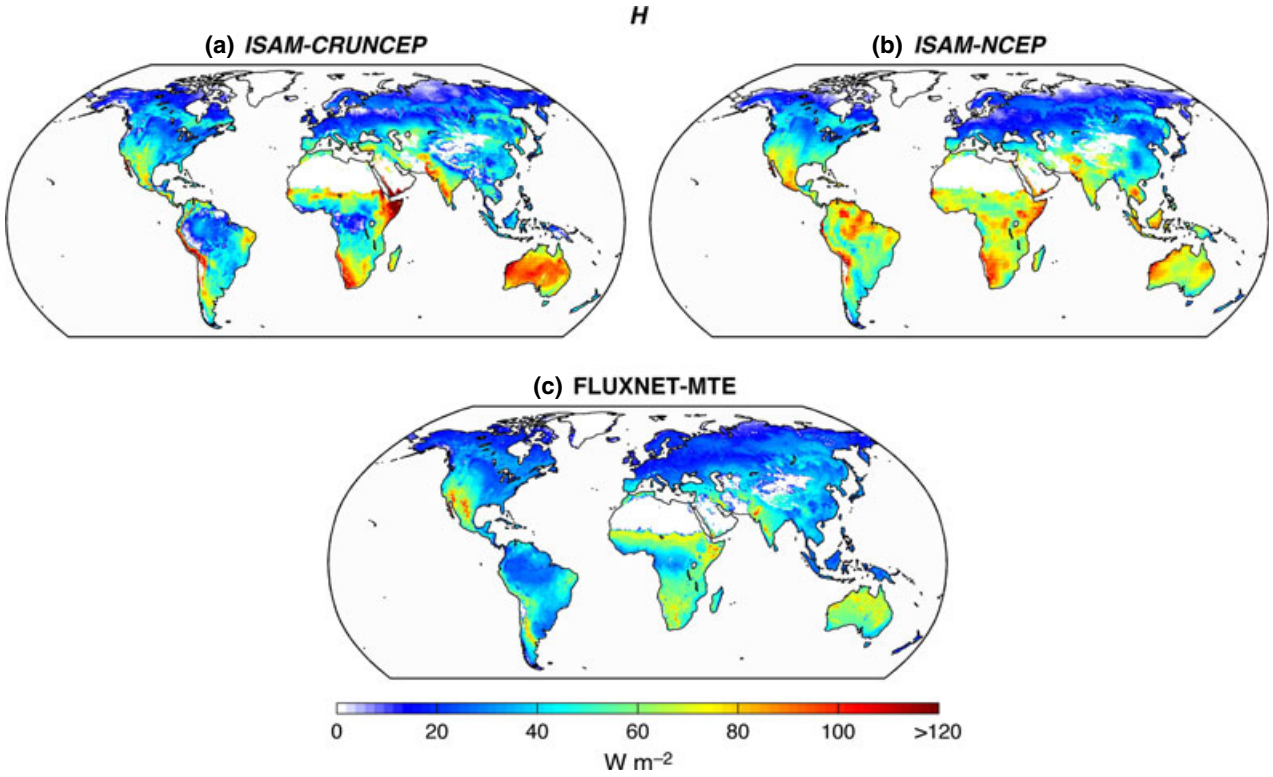

Fig. 5 Maps of mean annual estimates of *H*, for two reanalysis-driven model simulations [*ISAM-CRUNCEP* (a), *ISAM-NCEP* (b)] and an observationally derived dataset [FLUXNET-MTE (c)]. All the results are based on averaged output for 2000–2004, and are only for vegetated land surfaces (Fig. S5).

LW_{down} in the NCEP/NCAR showed very high positive differences in the tropics/subtropics; this difference gradually diminished for the midlatitudes, and eventually becoming negative at the upper high-latitudes (Fig. S1d). Next, the annual *Srad* in the NCEP/NCAR was also consistently lower in the mid/high-latitude vegetated surface (Fig. S1c). As discussed previously, the *Srad* + LW_{down} in the CRUNCEP dataset was generally in closer agreement with the site observed meteorology; correspondingly, the zonal mean R_n from *ISAM-CRUNCEP* was also generally in good agreement with FLUXNET-MTE across various latitudes (Fig. 6a). Nonetheless, between 20°S and 40°S, the R_n from *ISAM-CRUNCEP* was high in comparison with both FLUXNET-MTE and *ISAM-NCEP*. A limitation in our current model calibration is that the southernmost site used in our analysis was located at ca. 21.75°S (LBA-Pdg) and all other sites were above ca. 11°S; hence more Southern Hemispheric flux tower sites need to be integrated into ISAM, to better quantify the modeled energy/water flux biases in this region.

As in the site-level analysis, the annual *LE* estimates in the deep tropics from *ISAM-CRUNCEP* were higher than that of *ISAM-NCEP*, with a maximum relative difference of ca. 10 W m^{-2} at the equator (Fig. 6b), similar to the relative differences at the Trop.BET sites. In comparison with the observationally derived datasets, the zonally averaged *LE* from *ISAM-CRUNCEP* was also notably higher between 10°S and 10°N with a maximum positive anomaly of 10–15 W m^{-2} at the equator, while the corresponding tropical *LE* from *ISAM-NCEP* was in good agreement. Consistent with our site-level analysis, these relative *LE* biases between *ISAM-NCEP* and *ISAM-CRUNCEP* could be primarily attributed to the persistently higher *Q* in the *NCEP/NCAR* (Fig. S1e), which suppressed the *ET* to lower the *LE*. Also, we could rule out Precip differences between the two reanalyses as a cause of the larger modeled *LE* using the CRUNCEP, because the total annual Precip in the tropics was actually slightly higher in the NCEP/NCAR (data not shown).

Above approximately 40°N, the model simulation using the NCEP/NCAR data showed consistently low *LE* in comparison with both FLUXNET-MTE and MODIS (anomaly of −10 to −15 W m^{-2} in 52–70°N latitude) (Fig. 6b). Such zonal-scale negative *LE* anomalies were very similar to those previously simulated using the NCEP/NCAR data at most of the upper midlatitude and high-latitude sites (e.g., NET, Temp.BDT, tundra). As in the site simulations, the low zonal *LE* using

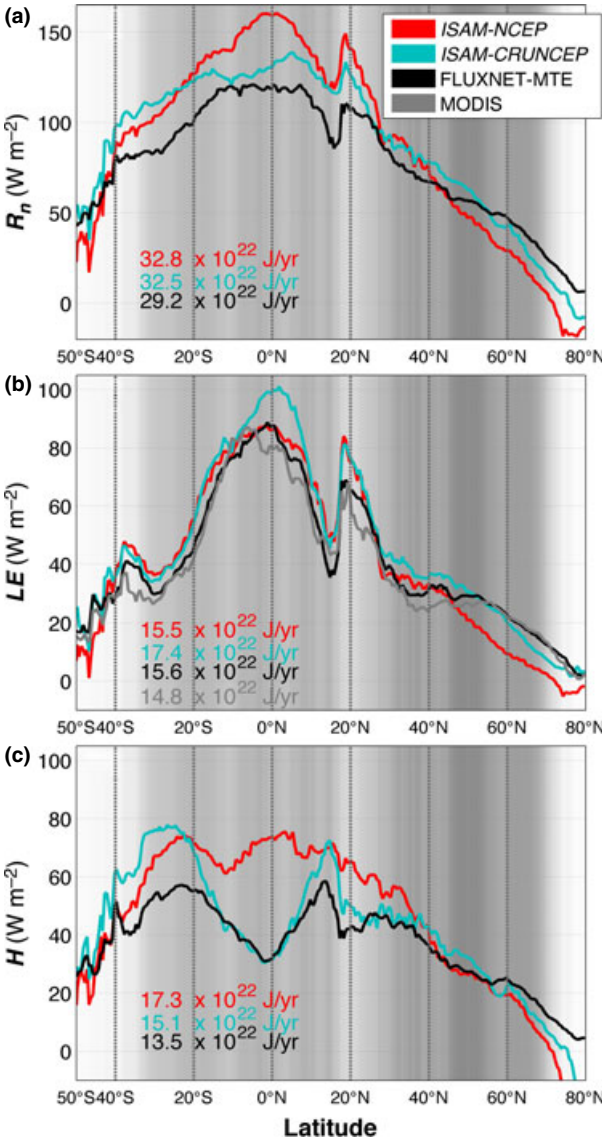

Fig. 6 Zonally averaged mean annual estimates for (a) R_n (=*LE* + *H*), (b) *LE*, and (c) *H*. Estimates are shown for two reanalysis-driven model simulations (*ISAM-NCEP* and *ISAM-CRUNCEP*) and observationally derived data (FLUXNET-MTE, and/or MODIS-DAO). Numbers in colors are respective global estimates. All the results are based on averaged output for 2000–2004, and are only for vegetated land surfaces (Fig. S5). The fractional land area (vegetated) at each latitude is shown as a gray scale, where darker shades represent more vegetated areas.

the NCEP/NCAR could primarily be attributed to the corresponding low R_n in *ISAM-NCAP* at above 40°N (Fig. 6a). These biases were much smaller in the CRUNCEP simulations (due to a better heat budget), showing that the accuracy of meteorological forcings is very important to determine the *LE* (especially in the northern high-latitudes).

Based on the strongly positive tropical R_n anomaly in *ISAM-NCEP* (relative to FLUXNET-MTE, Fig. 6a), the corresponding *H* biases were also particularly strong in the model using the NCEP/NCAR (Fig. 6c). Notably, the *ISAM-NCEP* consistently overestimated the mean annual *H* with respect to FLUXNET-MTE, with a maximum difference of 40 W m^{-2} the equator. Such biases were absent using the CRUNCEP data at most latitudes, except between 20°S and 40°S where the R_n anomaly in *ISAM-CRUNCEP* was largely positive (as discussed above). Additionally, due to a better northern high-latitude energy budget using the CRUNCEP, the *H* from *ISAM-CRUNCEP* also appeared to be in better agreement with FLUXNET-MTE than the *ISAM-NCEP* counterpart.

Finally, we also investigated the impact of the *LE* and *H* biases on the partitioning of the energy fluxes, using the metric of evaporative fraction: LE/R_n (Fig. 7). In our model simulation using the NCEP/NCAR, the LE/R_n between 15°S and 10°N was largely underestimated in comparison with FLUXNET-MTE (up to 20% lower at the equator, due to the anomalously low tropical *H*). The NCEP/NCAR simulation also strongly underestimated the LE/R_n in the northern high-latitudes over ca. 50°N (up to 22% lower at 60°N). At these latitudes, the underestimation in LE/R_n was due to the negative *LE* anomaly in *ISAM-NCEP*.

Discussion

Here, we used a LSM framework to show that potentially large uncertainties in terrestrial energy/water fluxes can arise from direct biases in reanalysis climate. To consistently quantify the modeled flux biases, we first analyzed the site-level biases in input meteorology

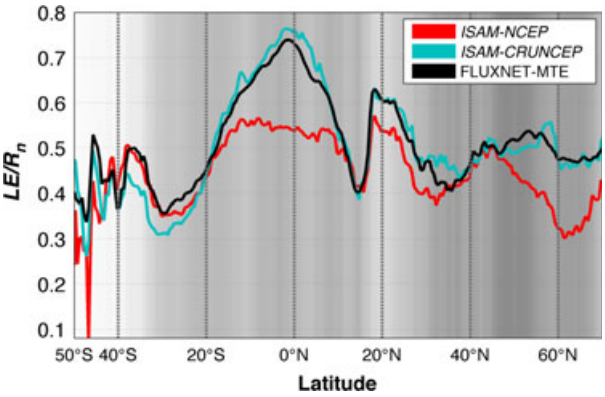

Fig. 7 Zonally averaged mean annual evaporative fraction: LE/R_n, computed using data from Fig. 5. The fractional land area (vegetated) at each latitude is shown as a gray scale, where darker shades represent more vegetated areas. For numerical stability, only 50°S to 70°N is shown.

from two reanalysis datasets: the NCEP/NCAR and the CRUNCEP. Using these datasets, several consistent patterns in the mean annual ΔR_n, ΔLE and ΔH were evident at the site-level as well as in the global simulations. Besides the mean annual biases in meteorology and fluxes, we also highlight the importance of investigating the seasonality of the biases, which have important consequences for the seasonal coupling among terrestrial carbon, energy and water fluxes.

With respect to the observed site meteorology, notable biases in the reanalysis variables were as follows: (i) high $+\Delta Srad$ in the tropics (both in NCEP/NCAR and CRUNCEP); (ii) high $+\Delta LW_{down}$ in the tropics but negative ΔLW_{down} in the mid/high-latitudes in NCEP/NCAR; and (iii) high $+\Delta Q$ in the NCEP/NCAR (Fig. 1). There were also varying degrees of biases in *Tavg* and *Precip*. Mostly, the mean annual ΔLW_{down} and ΔQ were smaller in the CRUNCEP. Specifically, based on the site-level modeling results of this study, the CRUNCEP driven *LE* and *H* fluxes were generally in better agreement (than the NCEP/NCAR counterparts) with the respective FLUXNET estimates. Similar patterns in *LE* and *H* response/biases were also evident in the zonal mean fluxes from global simulations (Figs 6 and 7).

We also analyzed the driving factors and mechanisms of the modeled biases in the R_n, *LE* and *H* fluxes. In the model, these flux biases could be primarily attributed to: (i) biases in total energy inputs to the surface (*Srad*, LW_{down}); (ii) biases in *Q* and *Tavg,* which modulate the atmospheric dryness and hence influence *LE*/*H* partitioning; (iii) *Precip,* which may be especially important for nontree/herbaceous ecosystems. Our model response is typically consistent with several existing observational analysis from literature that suggest: (i) R_n controls the seasonal variation of *LE* over the rain forest in Amazonia (Hasler & Avissar, 2007; Hutyra *et al*., 2007; Fisher *et al*., 2009; Costa *et al*., 2010), and they may not be primarily water stressed (Juárez *et al*., 2007; Barman *et al*., 2013); (ii) available energy is the most important parameter in determining *LE* in the high-latitude boreal forests (Admiral *et al*., 2006), which are not predominantly water stressed because of their slow transpiration rates (Baldocchi *et al*., 2000; Admiral *et al*., 2006); and (iii) in arid and semiarid ecosystems (e.g., nontree PFTs), *Precip* and factors controlling atmospheric dryness (*Tavg, Q*) are the dominant factors in determining *LE* (e.g., Chang *et al*., 2006; Ferguson & Veizer, 2007; Hasler & Avissar, 2007). However, our results show that the environmental control(s) determining ΔLE can be different from those controlling the response in absolute *LE*. For example, even though R_n may be the dominant factor determining the *LE* response for the tropical forests, the ΔLE was often controlled by ΔQ (unless the positive anomaly in R_n was very high). Here we also note that, as opposed to the usually positive mean annual ΔGPP using the NCEP/NCAR and CRUNCEP datasets (Barman *et al*., 2013), the corresponding ΔR_n, ΔLE, and ΔH were not uniformly positive or negative. This suggests that the ΔGPP may not be correlated with ΔLE even though the GPP and *LE* are largely coupled through the stomata. Additionally, because biotic factors such as canopy physiology/morphology, and environmental factors such as soil thermal and hydrological processes also determine the *LE* and *H* response (and biases) in the model, any *single* factor individually should *not* be expected to fully explain the energy/water flux biases and the variations in model response.

Land surface models are ultimately designed to study complex land–atmosphere interaction processes, and for application into future climate/environmental change scenarios, at regional to global scales. Hence, given the magnitudes of biases in the *LE* and *H* fluxes using reanalysis data, we feel that further study is warranted to quantify the associated impacts on various land–atmosphere exchange parameterizations dependent on partitioning of energy fluxes. Also, several important questions do arise on the philosophy of LSM calibration using site-scale FLUXNET data. For example, what are the net impacts of site-level calibration on global estimates of various fluxes and reservoirs from LSMs, and should we calibrate to optimize at the sites or should we rather optimize the model based on global datasets to counteract such biases in the first place? In this context, a comparable study by Zhao *et al.* (2006) investigating the reanalysis-driven biases in MODIS estimated GPP suggested two approaches to reduce the GPP biases: (i) by adjusting the reanalysis variables at each grid cell based on weather station data and/or (ii) by modifying the model parameters to optimize the output fluxes using the biased meteorology as the forcing data. In their study, the authors partly adopted the latter option for model calibration (though they acknowledged the associated caveats). But using a LSM framework such as in this study, this is most likely to result in propagating errors in various model parameterizations, due to the interactions among the full suite of carbon, energy, water, and momentum fluxes. This is also most likely to result in tunable parameter values that are inconsistent with the respective theoretical optimums, hence compromising the physical basis of model formulations. Subsequently, the model may also become susceptible to producing unreliable trends in fluxes, for future climate change simulations. Hence, in the long run, the only reliable alternative may be to improve the global reanalysis products to consistently force the model simulations. As for the current reanalysis products, it may be useful to develop potential

strategies to indirectly account for the LSM output biases postmodel simulations – such as the scaling of fluxes to account for the established model biases. For such purposes, estimation of site-level flux biases as demonstrated in our study (and in Barman *et al.* 2013) is a useful first step to formulate the respective scaling factors.

While continued model evaluation is necessary to improve the representation of carbon, energy and water cycles in the ISAM, here (along with Barman *et al.* 2013) we demonstrate the need to systematically investigate the flux uncertainties from forcing datasets itself, such as from meteorology. Better quantification of uncertainties should lead to better attribution of uncertainty sources, which can ultimately help to reduce the errors in future modeling efforts. Because LSMs usually use many similar schemes across models (due to shared model development or through infusion of sophisticated schemes from other models when available), the magnitude and range of flux uncertainties presented in this study is expected to be of interest to other LSM modelers, and to the ESM community in general.

Acknowledgements

This work was partly supported by the National Aeronautics and Space Administration (NASA) Land Cover and Land Use Change Program (No. NNX08AK75G), and U.S. Department of Energy (DOE) Office of Science (DOE-DE-SC0006706). RB was funded by the NASA Earth and Space Science Fellowship (NNX11AP85H). We thank Dr. Bassil El-Masri and Yang Song for contributing in various stages our model development. We specially thank the FLUXNET community and the P.I.s of the sites used in this study, for their continuous efforts of data collection.

References

Admiral SW, Lafleur PM, Roulet NT (2006) Controls on latent heat flux and energy partitioning at a peat bog in eastern canada. *Agricultural and Forest Meteorology*, **140**, 308–321.

Arora VK, Boer GJ (2003) A representation of variable root distribution in dynamic vegetation models. *Earth Interactions*, **7**, 1–19.

Asner GP (1998) Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, **64**, 234–253.

Asner GP, Wessman CA, Schimel DS, Archer S (1998) Variability in leaf and litter optical properties: implications for BRDF model inversions using AVHRR, MODIS, and MISR. *Remote Sensing of Environment*, **63**, 243–257.

Baldocchi D, Kelliher FM, Black TA, Jarvis P (2000) Climate and vegetation controls on boreal zone energy exchange. *Global Change Biology*, **6**, 69–83.

Baldocchi D, Falge E, Gu L, Olson R, Hollinger D, Running S, Evans R (2001) FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, **82**, 2415–2434.

Ban-Weiss GA, Bala G, Cao L, Pongratz J, Caldeira K (2011) Climate forcing and response to idealized changes in surface latent and sensible heat. *Environmental Research Letters*, **6**, 034032.

Barman R, Jain AK, Liang M (2013) Climate-driven uncertainties in modeling terrestrial gross primary production: a site-level to global scale analysis *Global Change Biology*, doi: 10.1111/gcb.12474.

Blyth E, Gash J, Lloyd A, Pryor M, Weedon GP, Shuttleworth J (2010) Evaluating the JULES land surface model energy fluxes using FLUXNET data. *Journal of Hydrometeorology*, **11**, 509–519.

Blyth E, Clark D, Ellis R, Huntingford C, Los S, Pryor M, Sitch S (2011) A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale. **4**, 255–269.

Bonan GB, Lawrence PJ, Oleson KW, Levis S, Jung M, Reichstein M, Swenson SC (2011) Improving canopy processes in the community land model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *Journal of Geophysical Research*, **116**, G02014.

Chang X, Zhao W, Zhang Z, Su Y (2006) Sap flow and tree conductance of shelter-belt in arid region of china. *Agricultural and Forest Meteorology*, **138**, 132–141.

Chen H, Dickinson RE, Dai Y, Zhou L (2011) Sensitivity of simulated terrestrial carbon assimilation and canopy transpiration to different stomatal conductance and carbon assimilation schemes. *Climate Dynamics*, **36**, 1037–1054.

Collatz GJ, Ball JT, Grivet C, Berry JA (1991) Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. *Agricultural and Forest Meteorology*, **54**, 107–136.

Costa MH, Biajoli MC, Sanches L, Malhado A, Hutyra LR, da Rocha HR, de Araújo AC (2010) Atmospheric versus vegetation controls of amazonian tropical rain forest evapotranspiration: are the wet and seasonally dry rain forests any different? *Journal of Geophysical Research: Biogeosciences (2005–2012)*, **115**, G04021. doi: 10.1029/2009JG001179.

Da Rocha HR, Manzi AO, Cabral OM, Miller SD, Goulden ML, Saleska SR, Artaxo P (2009) Patterns of water and heat flux across a biome gradient from tropical forest to savanna in brazil. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, **114**, G00B12. doi: 10.1029/2007JG000640.

Dai Y, Dickinson RE, Wang Y (2004) A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *Journal of Climate*, **17**, 2281–2299.

Dickinson RE (2011) Coupling atmospheric circulation models to bio-physical, biochemical, and biological processes at the land surface. In: *The Development of Atmospheric General Circulation Models: Complexity, Synthesis and Computation* (eds Donner L, Schubert W, Somerville R), pp. 177–197. Cambridge Univ. Press, Cambridge, UK.

Dorman J, Sellers PJ (1989) A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB). *Journal of Applied Meteorology*, **28**, 833–855.

El-Masri B, Barman R, Meiyappan P, Song Y, Liang M, Jain AK (2013) Carbon dynamics in the amazonian basin: integration of eddy covariance and ecophysiological data with a land surface model. *Agricultural and Forest Meteorology*, **19**, 1759–1779.

Falge E, Baldocchi D, Olson R, Anthoni P, Aubinet M, Bernhofer C, Dolman H (2001) Gap filling strategies for long term energy flux data sets. *Agricultural and Forest Meteorology*, **107**, 71–77.

FAO/IIASA/ISRIC/ISSCAS/JRC (2012) Harmonized World Soil Database (version 1.10), FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Fekete BM, Vörösmarty CJ, Roads JO, Willmott CJ (2004) Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Climate*, **17**, 294–304.

Ferguson PR, Veizer J (2007) Coupling of water and carbon fluxes via the terrestrial biosphere and its significance to the earth's climate system. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **112**, D24S06. doi: 10.1029/2007JD008431.

Fisher JB, Malhi Y, Bonal D, Da Rocha HR, De Araújo AC, Gamo M, Kondo H (2009) The land–atmosphere water flux in the tropics. *Global Change Biology*, **15**, 2694–2714.

Hasler N, Avissar R (2007) What controls evapotranspiration in the amazon basin? *Journal of Hydrometeorology*, **8**, 380–395.

Henderson-Sellers A, Irannejad P, McGuffie K, Pitman A (2003) Predicting land-surface climates-better skill or moving targets? *Geophysical Research Letters*, **30**, 1777. doi: 10.1029/2003GL017387.

Hollinger D, Richardson A (2005) Uncertainty in eddy covariance measurements and its application to physiological models. *Tree Physiology*, **25**, 873–885.

Hutyra LR, Munger JW, Saleska SR, Gottlieb E, Daube BC, Dunn AL, Wofsy SC (2007) Seasonal controls on the exchange of carbon and water in an amazonian rain forest. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, **112**, G03008. doi: 10.1029/2006JG000365.

Jain A, Yang X, Kheshgi H, McGuire AD, Post W, Kicklighter D (2009) Nitrogen attenuation of terrestrial carbon cycle response to global environmental factors. *Global Biogeochemical Cycles*, **23**, GB4028. doi: 10.1029/2009GB003519.

Janowiak JE, Gruber A, Kondragunta C, Livezey RE, Huffman GJ (1998) A comparison of the NCEP-NCAR reanalysis precipitation and the GPCP rain gauge-satellite combined dataset with observational error considerations. *Journal of Climate*, **11**, 2960–2979.

Jiménez C, Prigent C, Mueller B, Seneviratne S, McCabe M, Wood E, Dirmeyer P (2011) Global intercomparison of 12 land surface heat flux estimates. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **116**, D02102. doi: 10.1029/2010JD014545.

Juárez RIN, Hodnett MG, Fu R, Goulden ML, von Randow C (2007) Control of dry season evapotranspiration over the amazonian forest as inferred from observations at a southern amazon forest site. *Journal of Climate*, **20**, 2827–2839.

Jung M, Vetter M, Herold M, Churkina G, Reichstein M, Zaehle S, Chen Y (2007) Uncertainties of modeling gross primary productivity over Europe: a systematic study on the effects of using different drivers and terrestrial biosphere models. *Global Biogeochemical Cycles*, **21**, GB4021. doi: 10.1029/2006GB002915.

Jung M, Reichstein M, Margolis HA, Cescatti A, Richardson AD, Arain MA, Chen J (2011) Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, **116**, doi: 10.1029/2010JG001566.

Law B, Falge E, Gu LV, Baldocchi D, Bakwin P, Berbigier P, Fuentes J (2002) Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation. *Agricultural and Forest Meteorology*, **113**, 97–120.

Lawrence PJ, Chase TN (2009) Climate impacts of making evapotranspiration in the community land model (CLM3) consistent with the simple biosphere model (SiB). *Journal of Hydrometeorology*, **10**, 374–394.

Lawrence DM, Slater AG (2008) Incorporating organic soil into a global climate model. *Climate Dynamics*, **30**, 145–160.

Lawrence DM, Thornton PE, Oleson KW, Bonan GB (2007) The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy evaporation in a GCM: impacts on land-atmosphere interaction. *Journal of Hydrometeorology*, **8**, 862–880.

Lawrence DM, Slater AG, Romanovsky VE, Nicolsky DJ (2008) Sensitivity of a model projection of near-surface permafrost degradation to soil column depth and representation of soil organic matter. *Journal of Geophysical Research*, **113**, F02011.

Lawrence DM, Oleson KW, Flanner MG, Thornton PE, Swenson SC, Lawrence PJ, Sakaguchi K (2011) Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, **3**, doi: 10.1029/2011MS000045.

Mu Q, Zhao M, Running SW (2011) Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, **115**, 1781–1800.

Mu Q, Zhao M, Running SW (2012) Remote sensing and modeling of global evapotranspiration. In: *Multi-scale Hydrologic Remote Sensing: Perspectives and Applications* (eds Chang NB, Hong Y), pp. 443–479. The CRC Press, Boca Raton, FL.

Mueller B, Seneviratne S, Jimenez C *et al.* (2011) Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations. *Geophysical Research Letters*, **38**, L06402. doi: 10.1029/2010GL046230.

Oleson K, Niu G, Yang Z, Lawrence D, Thornton P, Lawrence P, Levis S (2008) Improvements to the community land model and their impact on the hydrological cycle. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, **113**, G01021. doi: 10.1029/2007JG000563.

Overgaard J, Rosbjerg D, Butts M (2006) Land-surface modelling in hydrological perspective? a review. *Biogeosciences*, **3**, 229–241.

Qian T, Dai A, Trenberth KE, Oleson KW (2006) Simulation of global land surface conditions from 1948 to 2004. part I: forcing data and evaluations*Journal of Hydrometeorology*, **7**, 953–975.

Randow C, Manzi A, Kruijt B, de Oliveira P, Zanchi F, Silva R, Waterloo M (2004) Comparative measurements and seasonal variations in energy and carbon exchange over forest and pasture in south west amazonia. *Theoretical and Applied Climatology*, **78**, 5–26.

Richardson AD, Hollinger DY, Burba GG, Davis KJ, Flanagan LB, Katul GG, Suyker AE (2006) A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. *Agricultural and Forest Meteorology*, **136**, 1–18.

Sakaguchi K, Zeng X (2009) Effects of soil wetness, plant litter, and under-canopy atmospheric stability on ground evaporation in the Community Land Model (CLM3.5). *Journal of Geophysical Research*, **114**, D01107. doi: 10.1029/2008JD010834.

Santanello JA, Peters-Lidard CD, Kumar SV, Alonge C, Tao W-K (2009) A modeling and observational framework for diagnosing local land–atmosphere coupling on diurnal time scales. *Journal of Hydrometeorology*, **10**, 577–599.

Schaefer K, Zhang T, Slater AG, Lu L, Etringer A, Baker I (2009) Improving simulated soil temperatures and soil freeze/thaw at high-latitude regions in the simple Biosphere/Carnegie-Ames-Stanford approach model. *Journal of Geophysical Research: Earth Surface (2003–2012)*, **114**, F02021. doi: 10.1029/2008JF001125.

Schenk HJ, Jackson RB (2002) The global biogeography of roots. *Ecological Monographs*, **72**, 311–328.

Sellers P, Randall D, Collatz G, Berry J, Field C, Dazlich D, Bounoua L (1996) A revised land surface parameterization (SiB2) for atmospheric GCMs. part I: model formulation. *Journal of Climate*, **9**, 676–705.

Sellers P, Dickinson R, Randall D, Betts A, Hall F, Berry J, Nobre C (1997) Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science*, **275**, 502–509.

Simmons A, Uppala S, Dee D, Kobayashi S (2007) ERA-interim: new ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter*, **110**, 25–35.

Song Y, Jain AK, McIsaac GF (2013) Implementation of dynamic crop growth processes into a land surface model: evaluation of energy, water and carbon fluxes under corn and soybean rotation. *Biogeosciences*, **10**, 8039–8066.

Stöckli R, Lawrence D, Niu G, Oleson K, Thornton P, Yang Z, Running S (2008) Use of FLUXNET in the community land model development. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, **113**, G01025. doi: 10.1029/2007JG000562.

Todd-Brown KEO, Randerson JT, Post WM, Hoffman FM, Tarnocai C, Schuur EAG, Allison SD (2013) Causes of variation in soil carbon simulations from CMIP5 earth system models and comparison with observations. *Biogeosciences*, **10**, 1717–1736.

Trenberth KE, Smith L, Qian T, Dai A, Fasullo J (2007) Estimates of the global water budget and its annual cycle using observational and model data. *Journal of Hydrometeorology*, **8**, 758–769.

Trenberth KE, Fasullo JT, Kiehl J (2009) Earth's global energy budget. *Bulletin of the American Meteorological Society*, **90**, 311–323. doi:10.1175/2008BAMS2634.1.

Twine TE, Kustas W, Norman J, Cook D, Houser P, Meyers T, Wesely M (2000) Correcting eddy-covariance flux underestimates over a grassland. *Agricultural and Forest Meteorology*, **103**, 279–300.

Viovy N, Ciais P (2009) A combined dataset for ecosystem modelling, Available at: http://dods.extra.cea.fr/data/p529viov/cruncep/readme.htm (accessed March 2011).

Wang K, Dickinson RE (2012) A review of global terrestrial evapotranspiration: observation, modeling, climatology, and climatic variability. *Reviews of Geophysics*, **50**, RG2005. doi: 10.1029/2011RG000373.

Wei Y, Liu S, Post WM *et al.* (2013) 'The North American Carbon Program (NACP) Multi-Scale Synthesis and Terrestrial Model Intercomparison (MsTMIP) Project: Environmental Driver Data.' *Journal of Geoscientific Model Development, Discussion*, **6**, 5375–5422.

Wilson K, Goldstein A, Falge E *et al.* (2002) Energy balance closure at FLUXNET sites. *Agricultural and Forest Meteorology*, **113**, 223–243.

Yang X, Wittig V, Jain AK, Post W (2009) Integration of nitrogen cycle dynamics into the integrated science assessment model for the study of terrestrial ecosystem responses to global change. *Global Biogeochemical Cycles*, **23**, GB4029. doi: 10.1029/2009GB003474.

Zhao M, Running SW, Nemani RR (2006) Sensitivity of moderate resolution imaging spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, **111**, G01002. doi: 10.1029/2004JG000004.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Latent heat (*LE*), Sensible heat (*H*), and net radiation (R_n) data from FLUXNET.
Data S2. Flux correction of *LE* and *H*.
Table S1. Percentage (%) of available 3-hourly FLUXNET data at NACP sites.
Table S2. Calculation of *LE* and *H* for the LBA sites used in this study, based on flux tower data from published studies.
Table S3. Annual mean climate variables from site/station data for each site used in this study.
Figure S1. Relative differences in mean annual meteorology variables between two global reanalyses datasets used in this study. Computations are based on annually averaged data during 2000–2004.
Figure S2. Monthly *ET* vs. monthly R_n: best-fit quadratic polynomials for individual PFTs, using model output from the *ISAM-FLUXNET* simulations. Only PFTs with more than one site-year of data were shown.
Figure S3. Analysis for individual nontree C3 and C4 sites: Daily climatology of Δ*LE* in reanalyses-driven simulations for the (a) nontree C3 and (b) C4 sites. For each subplot, the *x*-axis is the 'day of year' and the *y*-axis is Δ*LE*. All the biases (Δ) were calculated with respect to the *ISAM-FLUXNET* counterpart.
Figure S4. Analysis for individual nontree C3 and C4 sites: Daily climatology of Δ*H* in reanalyses-driven simulations for the (a) nontree C3 and (b) C4 sites. For each subplot, the *x*-axis is the 'day of year' and the *y*-axis is Δ*H*. All the biases (Δ) were calculated with respect to the *ISAM-FLUXNET* counterpart.
Figure S5. Map of vegetated land area mask used for comparison of R_n, *LE*, and *H* estimates.