Model representation of the coupling between evapotranspiration and soil water content at different depths

Jianxiu Qiu1,2, Wade T. Crow3, Jianzhi Dong3, and Grey S. Nearing4

1Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510275, China
2Southern Laboratory of Ocean Science and Engineering (Guangdong, Zhuhai), Zhuhai, 519000, China
3USDA ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD 20705, USA
4Department of Geological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA

Correspondence: Jianxiu Qiu (qiujianxiu@mail.sysu.edu.cn)

Received: 5 June 2019 – Discussion started: 15 July 2019
Revised: 25 November 2019 – Accepted: 6 January 2020 – Published: 7 February 2020

Abstract. Soil water content (θ) influences the climate system by controlling the fraction of incoming solar and long-wave energy that is converted into evapotranspiration (ET). Therefore, investigating the coupling strength between θ and ET is important for the study of land surface–atmosphere interactions. Physical models are commonly tasked with representing the coupling between θ and ET; however, few studies have evaluated the accuracy of model-based estimates of θ/ET coupling (especially at multiple soil depths). To address this issue, we use in situ AmeriFlux observations to evaluate θ/ET coupling strength estimates acquired from multiple land surface models (LSMs) and an ET retrieval algorithm – the Global Land Evaporation Amsterdam Model (GLEAM). For maximum robustness, coupling strength is represented using the sampled normalized mutual information (NMI) between θ estimates acquired at various vertical depths and surface evaporation flux expressed as a fraction of potential evapotranspiration (fPET, the ratio of ET to potential ET). Results indicate that LSMs and GLEAM are generally in agreement with AmeriFlux measurements in that surface soil water content (θs) contains slightly more NMI with fPET than vertically integrated soil water content (θv). Overall, LSMs and GLEAM adequately capture variations in NMI between fPET and θ estimates acquired at various vertical depths. However, GLEAM significantly overestimates the NMI between θ and ET, and the relative contribution of θ to total ET. This bias appears attributable to differences in GLEAM’s ET estimation scheme relative to the other two LSMs considered here (i.e., the Noah model with multi-parameterization options and the Catchment Land Surface Model, CLSM). These results provide insight into improved LSM model structure and parameter optimization for land surface–atmosphere coupling analyses.

1 Introduction

Soil water content (θ) modulates water and energy feedbacks between the land surface and the lower atmosphere by determining the fraction of incoming solar energy that is converted into evapotranspiration (ET; Seneviratne et al., 2010, 2013). In water-limited regimes, θ exhibits a dominant control on ET, and therefore exerts significant terrestrial control on the Earth’s water and energy cycles. Accurately representing θ/ET coupling in land surface models (LSMs) is therefore expected to improve our ability to project the future frequency of extreme climates (Seneviratne et al., 2013).

A key question is how the constraint of θ on ET and sensible heat (H) varies as θ is vertically integrated over deeper vertical soil depths. Given the tendency for the timescales of θ dynamics to vary strongly with depth, the degree to which the ET is coupled with vertical variations in θ determines the temporal scale at which θ variations are propagated into the lower atmosphere. Therefore, in order to represent θ/ET coupling, and thus land–atmosphere interactions in general, LSMs must accurately capture the relationship between vertically varying θ values and ET. Unfortunately, their ability to do so remains an open question.
Recently, land surface–atmosphere coupling strength has been investigated by sampling mutual information proxies (e.g., correlation coefficient or other coupling indices) between time series of $\theta$ and ET observations (or air temperature proxies for ET). Results suggest that, even when confined to very limited vertical support (e.g., within the top 5 cm of the soil column), surface $\theta$ estimates retain significant information for describing overall $\theta$ control on local climate (Ford and Quiring, 2014; Qiu et al., 2014; Dong and Crow, 2018, 2019). These findings are in contrast with the common perception that ET is constrained only by $\theta$ values within deeper soil layers (Hirschi et al., 2014). Hence, it is necessary to examine whether LSMs can realistically reflect observed variations of $\theta$/ET coupling strength within the vertical soil profile.

Previous studies examining the $\theta$/ET relationship have generally been based on Pearson product–moment correlation (Basara and Crawford, 2002; Ford et al., 2014), which captures only the strength of a linear relationship between two variables. However, the coupling between $\theta$ and ET is generally nonlinear. Therefore, non-parametric mutual information measures are generally more appropriate. Nearing et al. (2018) used information theory metrics (transfer entropy, in particular) to measure the strength of direct couplings between different surface variables, including soil water content, and surface energy fluxes at short timescales in several LSMs. They found that the LSMs are generally biased as compared with the strengths of couplings in observation data, and that these biases differ across different study sites. However, they did not look specifically at the effect of vertical water content profiles or of subsurface soil water content on partitioning surface energy fluxes.

Here we apply the information theory-based methodology of Qiu et al. (2016) to examine the relationship between the vertical support of $\theta$ estimates and their mutual information (MI) with respect to ET. Our approach is based on analyzing the MI content between ET and $\theta$ time series acquired from both LSMs, and ET retrieval algorithm – the Global Land Evaporation Amsterdam Model (GLEAM) – and AmeriFlux in situ observations. MI values are then normalized by entropy in the corresponding ET time series to remove the effect of inter-site variations to generate estimates of normalized mutual information (NMI) between $\theta$ and ET. Both surface (roughly 0–10 cm) soil water content ($\theta_s$) and vertically integrated (0–40 cm) soil water content ($\theta_v$) are considered to capture the impact of depth on NMI results. AmeriFlux-based NMI results are then compared with analogous NMI results obtained from LSM-based and GLEAM-based $\theta$ and ET time series.

2 Data and methods

The AmeriFlux network provides temporally continuous measurements of $\theta$, surface energy fluxes and related environmental variables for sites located in a variety of North American ecosystem types, e.g., forests, grasslands, croplands, shrublands and savannas (Boden et al., 2013). To minimize sampling errors, AmeriFlux sites lacking a complete 3-year summer months (June, July and August) daily time series between the years of 2003 and 2015 (i.e., $3 \times 92 = 276$ daily observations in total) of $\theta_s$, $\theta_v$ and latent heat flux (LE) are excluded here, resulting in the 34 remaining eligible AmeriFlux sites listed in Table 1. These sites cover a variety of climate zones within the contiguous United States (CONUS). Table 1 gives background information on these 34 sites including local land-cover information. Hydro-climatic conditions in each site are characterized using the aridity index (AI), calculated using CRU (Climate Research Unit, v4.02) monthly precipitation and potential evaporation (PET) datasets.

As described above, $\theta$/ET coupling assessments made using AmeriFlux observations are compared with those using state-of-the-art LSMs including the Noah model with multi-parameterization options (NOAHMP) and Catchment Land Surface Model (CLSM). In addition, $\theta$ and ET retrievals provided by the Global Land Evaporation Amsterdam Model (GLEAM) are also considered. See below for details on all three approaches. To avoid any spurious correlations between $\theta$ and ET due to seasonality, all NMI analyses are performed on $\theta$ and ET time series anomalies acquired during the period 2003–2015. The $\theta$ and ET anomalies are calculated by removing the seasonal cycle – defined as 31 d window averages centered on each day of the year sampled across all years of the 2003–2015 historical data record – from the raw $\theta$ and ET time series data. The analysis is limited to the CONUS during summer months (June, July and August) when $\theta$/ET coupling is expected to be maximized.

2.1 Ground-based AmeriFlux measurements

The Level 2 (L2) AmeriFlux LE and H flux observations are based on high-frequency (typically > 10 Hz) eddy covariance measurements processed into half-hourly averages by individual AmeriFlux investigators. LE and $\theta$ observations at a half-hour time step and without gap-filling procedures are collected from the AmeriFlux Site and Data Exploration System (see http://amerifluxornl.gov/, last access: November 2018). The LE and $\theta$ observations are further aggregated into daily (00:00 to 24:00 UTC) values, and daily LE is converted into daily ET using the latent heat of vaporization. Daily ET values based on less than 30 % half-hourly coverage (i.e., <15 half-hourly observations per day) are considered not representative at a daily timescale and are therefore excluded.

Soil water content measurements are generally available at two discrete depths that vary between the AmeriFlux sites (Table 1). Here, the top (i.e., closest to the surface) soil water content observation is always used to represent surface soil water content ($\theta_s$). Since the depth of this top-layer measure-
Table 1. Attributes of selected AmeriFlux sites.

<table>
<thead>
<tr>
<th>AmeriFlux sites</th>
<th>Land cover</th>
<th>Elevation (m)</th>
<th>Top-layer depth (cm)</th>
<th>Bottom-layer depth (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARM SGP Main</td>
<td>Cropland</td>
<td>314</td>
<td>10(^a)</td>
<td>20(^b)</td>
</tr>
<tr>
<td>ARM USDA UNL OSU Woodward Switchgrass 1</td>
<td>Grassland</td>
<td>611</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>Audubon Research Ranch</td>
<td>Grassland</td>
<td>1469</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Bondville</td>
<td>Cropland</td>
<td>219</td>
<td>10(^c)</td>
<td>20</td>
</tr>
<tr>
<td>Brookings</td>
<td>Grassland</td>
<td>510</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Chimney Park</td>
<td>Evergreen needleleaf forest</td>
<td>2750</td>
<td>0–15</td>
<td>15–45</td>
</tr>
<tr>
<td>Duke Forest Hardwoods</td>
<td>Deciduous broadleaf forest</td>
<td>168</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>Duke Forest Open Field</td>
<td>Grassland</td>
<td>168</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>Fermi Agricultural</td>
<td>Cropland</td>
<td>225</td>
<td>2.5</td>
<td>10</td>
</tr>
<tr>
<td>Fermi Prairie</td>
<td>Grassland</td>
<td>226</td>
<td>2.5</td>
<td>10</td>
</tr>
<tr>
<td>Flagstaff Managed Forest</td>
<td>Evergreen needleleaf forest</td>
<td>2160</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Flagstaff Unmanaged Forest</td>
<td>Woody savannas</td>
<td>2180</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Flagstaff Wildfire</td>
<td>Grassland</td>
<td>2270</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Fort Peck</td>
<td>Grassland</td>
<td>634</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Freeman Ranch Woodland</td>
<td>Woody savannas</td>
<td>232</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Glacier Lakes Ecosystem Experiments Site</td>
<td>Evergreen needleleaf forest</td>
<td>3190</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Howland Forest Main</td>
<td>Mixed forest</td>
<td>60</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Lucky Hills Shrubland</td>
<td>Open shrubland</td>
<td>1372</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Marys River Fir Site</td>
<td>Evergreen needleleaf forest</td>
<td>263</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Metolius Intermediate Pine</td>
<td>Evergreen needleleaf forest</td>
<td>1253</td>
<td>0–30</td>
<td>NA</td>
</tr>
<tr>
<td>Missouri Ozark</td>
<td>Deciduous broadleaf forest</td>
<td>219</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Nebraska SandHills Dry Valley</td>
<td>Grassland</td>
<td>1081</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>Quebec Boreal Cutover Site</td>
<td>Evergreen needleleaf forest</td>
<td>400</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Quebec Mature Boreal Forest Site</td>
<td>Evergreen needleleaf forest</td>
<td>400</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Santa Rita Creosote</td>
<td>Open shrubland</td>
<td>991</td>
<td>2.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Santa Rita Mesquite</td>
<td>Woody savannas</td>
<td>1116</td>
<td>2.5–5</td>
<td>5–10</td>
</tr>
<tr>
<td>Sherman Island</td>
<td>Grassland</td>
<td>–5</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Sylvania Wilderness</td>
<td>Mixed forest</td>
<td>540</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Tonzi Ranch</td>
<td>Woody savannas</td>
<td>169</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>University of Michigan Biological Station</td>
<td>Deciduous broadleaf forest</td>
<td>234</td>
<td>0–30</td>
<td>NA</td>
</tr>
<tr>
<td>Vaira Ranch</td>
<td>Grassland</td>
<td>129</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Walker Branch</td>
<td>Deciduous broadleaf forest</td>
<td>343</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Willow Creek</td>
<td>Deciduous broadleaf forest</td>
<td>515</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Wind River Field Station</td>
<td>Evergreen needleleaf forest</td>
<td>371</td>
<td>30(^d)</td>
<td>50(^e)</td>
</tr>
</tbody>
</table>

\(^a\) Was 5 cm prior to 13 April 2005. \(^b\) Was 25 cm prior to 13 April 2005. \(^c\) Was 5 cm prior to 1 January 2006. \(^d\) Was 0–30 cm prior to 2007. \(^e\) Unavailable prior to 2007. NA = not available.

ment varies between 0 and 15 cm (see Table 1), we consider the surface-layer measurement \(\theta_s\) to be roughly representative of 0–10 cm (vertically integrated) \(\theta\). For more details on AmeriFlux sites utilized here, see Raz-Yaseef et al. (2015).

Given variations in the depth of the lower AmeriFlux \(\theta\) observations (see Table 1), we applied a variety of approaches for estimating vertically integrated soil water content (\(\theta_v\)). Our first approach, hereinafter referred to as Case I, is based on the application of an exponential filter (Wagner et al., 1999; Albergel et al., 2008) to extrapolate \(\theta_s\) to a consistent 40 cm bottom-layer depth. Therefore, only \(\theta_s\) is used to derive \(\theta_v\) and the bottom-layer (or second-layer) AmeriFlux \(\theta\) measurement is neglected in this case. The application of the exponential filter requires a single timescale parameter \(T\). Since \(\theta\) measurements from the United States Department of Agriculture’s Soil Climate Analysis Network (SCAN) are taken at fixed soil depth, we utilized this dataset to determine the most appropriate parameter \(T\) at AmeriFlux sites. Following Qiu et al. (2014), first, we estimated the optimal parameter \(T (T_{opt})\) for the extrapolation of \(\theta\) measurements from 10 to 40 cm depth and established a global relationship between \(T_{opt}\) and site-based NDVI (MOD13Q1 v006, 250 m, 16 d; \(T_{opt} = 2.098 \times \exp(-1.895 \times (\text{NDVI}+0.6271)) + 2.766\)). Then, this global relationship (goodness of fit \(R^2\): 0.85) is applied to AmeriFlux sites to extrapolate 0–10 cm \(\theta_s\) time series into 0–40 cm \(\theta_v\).

Previous research has suggested that such a filtering approach does not significantly squander ET information present in actual measurements of \(\theta_v\) (Qiu et al., 2014, 2016). Nevertheless, since the quality of \(\theta_v\) estimates is important in...
our analysis, we also calculated two additional cases where 0–40 cm $\theta_i$ is estimated using (1) the bottom-layer soil water content measurement at each AmeriFlux site (hereinafter, Case II) and (2) linear interpolation of $\theta_i$, and the bottom-layer AmeriFlux soil water content measurement (hereinafter, Case III). The sensitivity of key results to these various cases is discussed below.

2.2 LSM- and GLEAM-based simulations

Simulations are acquired from the NOAHMP (Niu et al., 2011) and CLSM (Koster et al., 2000) LSMs embedded within the NASA Land Information System (LIS, Kumar et al., 2006) and the GLEAM ET retrieval algorithm (Miralles et al., 2011). Both NOAHMP and CLSM are set up to simulate 0.125° $\theta$ profiles at a 15 min time step using North America Land Data Assimilation System, Phase 2 (NLDAS-2) forcing data. A 10-year model spin-up period (1992 to 2002) is applied for NOAHMP and CLSM.

NOAHMP numerically solves the one-dimensional Richards equation within four soil layers of thicknesses of 0–10, 11–30, 31–60 and 61–100 cm. Major parameterization options relevant to $\theta$ simulation include options for canopy stomatal resistance parameterization and schemes controlling the effect of $\theta$ on the vegetation stress factor $\beta$. Here we employed the Ball–Berry-type stomatal resistance scheme and Noah-type soil water content factor controlling the $\beta$ factor. The specific expressions are as follows:

$$\beta = \frac{\sum N_{\text{root}} \Delta Z_i}{Z_{\text{root}}} \min \left(1.0, \frac{\theta_i - \theta_{\text{wilt}}}{\theta_{\text{ref}} - \theta_{\text{wilt}}} \right),$$

(1)

where $\theta_{\text{wilt}}$ and $\theta_{\text{ref}}$ are, respectively, soil water content at the wilting point ($\text{m}^3\text{m}^{-3}$) and reference soil water content ($\text{m}^3\text{m}^{-3}$), which is set as field capacity during parameterization. $\theta_i$ and $\Delta Z_i$ are soil water content ($\text{m}^3\text{m}^{-3}$) and soil depth (cm) at $i$th layer, $N_{\text{root}}$ and $Z_{\text{root}}$ are total number of soil layers with roots and total depth (cm) of root zone, respectively.

Following the Ball–Berry stomatal resistance scheme, the $\theta$-controlled $\beta$ factor and other multiplicative factors including temperature and foliage nitrogen simultaneously determine the maximum carboxylation rate $V_{\text{max}}$ as follows:

$$V_{\text{max}} = V_{\text{max25}} \alpha_{\text{vmax}} f(N) f(T_v) \beta,$$

(2)

where $V_{\text{max25}}$ is maximum carboxylation rate at 25°C (mmol CO$_2$ m$^{-2}$ s$^{-1}$), $\alpha_{\text{vmax}}$ is a parameter sensitive to vegetation canopy surface temperature $T_v$, $f(N)$ is a factor representing foliage nitrogen and $f(T_v)$ is a function that mimics thermal breakdown of metabolic processes. Based on $V_{\text{max}}$, photosynthesis rates per unit leaf area index (LAI) including carboxylase-limited (Rubisco limited, denoted by $AC$) type and export-limited (for C$_3$ plants, denoted by $AS$) type are calculated, respectively. The minimum of $AC$, $AS$ and the light-limited photosynthesis rate determines stomatal resistance $r_s$, and consequently affects ET over vegetated areas.

For the complete NOAHMP configuration, please see Table S1 in the Supplement.

CLSM simulates the 0–2 and 0–100 cm soil water content and evaporative stress as a function of simulated $\theta$ and environmental variables. ET is then estimated based on the estimated evaporative stress and land–atmosphere humidity gradients. Energy and water flux estimates are iterated with soil state estimates (e.g., $\theta$ and soil temperature) to ensure closure of surface energy and water balances. For a detailed explanation of CLSM physics, please refer to Koster et al. (2000).

GLEAM is a set of algorithms dedicated to the estimation of terrestrial ET and root-zone $\theta$ from satellite data. In this study, the latest version of this model (v3.2a) is employed. In GLEAM, the configuration of soil layers varies as a function of the land-cover type. Soil stratification is based on three soil layers for tall vegetation (0–10, 10–100 and 100–250 cm), two layers for low vegetation (0–10 and 10–100 cm) and only one layer for bare soil (0–10 cm; Martens et al., 2017).

The cover-dependent PET (mm d$^{-1}$) of GLEAM is calculated using the Priestley and Taylor (1972) equation based on observed air temperature and net radiation. Following this, estimates of PET are converted into actual transpiration or bare soil evaporation (depending on the land-cover type, ET (mm d$^{-1}$)), using a cover-dependent, multiplicative stress factor $S$ (–), which is calculated as a function of microwave vegetation optical depth (VOD) and root-zone $\theta$ (Miralles et al., 2011). The related expressions are as follows:

$$\text{ET} = \text{PET} \times S + E_i,$$

(3)

$$S = \sqrt{\frac{\text{VOD}}{\text{VOD}_{\text{max}}}} \left(1 - \left(\frac{\theta_c - \theta_{\text{wilt}}}{\theta_c - \theta_{\text{wilt}}}\right)^2\right),$$

(4)

where $E_i$ is rainfall interception (mm), $S$ essentially represents the fPET (see Sect. 2.3) estimated by GLEAM, $\theta_c$ ($\text{m}^3\text{m}^{-3}$) is the critical soil water content and $\theta_{\text{wilt}}$ ($\text{m}^3\text{m}^{-3}$) is the soil water content of the wettest layer, assuming that plants withdraw water from the layer that is most accessible. Based on Eq. (4), GLEAM $S$ (or fPET) tends to become more sensitive to $\theta$ in areas of low VOD seasonality (i.e., low differences between VOD and VOD$_{\text{max}}$). As for bare soil conditions, $S$ is linearly related to surface soil water content ($\theta_1$):

$$S = 1 - \frac{\theta_c - \theta_1}{\theta_c - \theta_{\text{wilt}}}.$$

(5)

To resolve variations in the vertical discretization of $\theta$ applied by each model, we linearly interpolated NOAHMP, CLSM and GLEAM outputs into daily 0–10 and 0–40 cm soil water content values using depth-weighted averaging.

2.3 Variable indicating soil water content and surface flux coupling

Soil water content–ET coupling can be diagnosed using a variety of different variables derived from ET, e.g., the fraction
of PET (fPET, the ratio of ET and PET) or the evaporative fraction (EF, the ratio of LE and the sum of LE and sensible heat). Since ET is strongly tied to net radiation ($R_n$; Koster et al., 2009), both fPET and EF are advantageous in that they normalize ET by removing the impact of non-soil water content influences on ET (e.g., net radiation, wind speed and soil heat flux ($G$)). However, since sensible heat flux is not provided in the GLEAM dataset, we are restricted here to using fPET.

It should be noted that the applied meteorological forcing data for NOAHMP and CLSM are somewhat different from those used for GLEAM. Therefore, to minimize the impact of this difference, NOAHMP and CLSM fPET are computed from North American Regional Reanalysis (NARR) using the modified Penman scheme of Mahrt and Ek (1984), while GLEAM fPET is calculated using its own internal PET estimates. To examine the impact of the PET source on the results, AmeriFlux fPET calculations are duplicated using both GLEAM- and NARR-based PET values.

### 2.4 Information measures

Mutual information (MI; Cover and Thomas, 1991) is a non-parametric measure of correlation between two random variables. MI and the related Shannon-type entropy (SE; Shannon, 1948) are calculated as follows. Entropy about a random variable $\zeta$ is a measure of uncertainty according to its distribution $p_{\zeta}$ and is estimated as the expected amount of information from $p_{\zeta}$ sample:

$$SE(p_{\zeta}) = E\{ - \ln(p_{\zeta}(\zeta)) \}. \quad (6)$$

Likewise, MI between $\zeta$ and another variable $\psi$ can be thought of as the expected amount of information about variable $\zeta$ contained in a realization of $\psi$ and is measured by the expected Kullback–Leibler (KL) divergence (Kullback and Leibler, 1951) between the conditional and marginal distributions over $\zeta$:

$$MI(\zeta; \psi) = E_{\psi}[D(p_{\zeta} \mid O \psi \mid |p_{\zeta})]. \quad (7)$$

In this context, the generic random variables $\zeta$ and $\psi$ represent fPET and $\theta$ (soil water content), respectively. The observation space of the target random variable fPET is discretized using a fixed bin width. As bin width decreases, entropy increases but mutual information asymptotes to a constant value. On the other hand, increased bin width requires a greater sample size, which cannot always be satisfied. The trick is choosing a bin width where the NMI values stabilize with sample size. After a careful sensitivity analysis, we choose a fixed bin width of 0.25 [–] for fPET and make sure that each AmeriFlux site had enough samples to accurately estimate the NMI, and change of this constant bin width from 0.1 to 0.5 [–] will not significantly alter our conclusions. Following Nearing et al. (2016), a bin width of 0.01 m$^3$ m$^{-3}$ (1 % volumetric water content) for $\theta$ is applied.

Integrations required for MI calculation in Eq. (7) are then approximated as summations over the empirical probability distribution function bins (Paninski, 2003).

By definition, the MI between two variables represents the amount of entropy (uncertainty) in either of the two variables that can be reduced by knowing the other. Therefore, the MI normalized by the entropy of the AmeriFlux-based fPET measurements represents the fraction of uncertainty in fPET that is resolvable given knowledge of the soil water content state (Nearing et al., 2013). Unlike Pearson’s correlation coefficient, MI is insensitive to the impact of nonlinear variable transformations. Therefore, it is well suited to describe the strength of the (potentially non-linear) relationship between $\theta$ and fPET.

Here, we applied this approach to calculate the MI content between soil water content representing different vertical depths (as reflected by $\theta_\varepsilon$ and $\theta_v$) and fPET at each AmeriFlux site. All estimated site-specific MI are normalized by the entropy of the corresponding AmeriFlux-based fPET measurements to remove the effect of inter-site entropy variations on the magnitude of NMI differences. The resulting normalized MI calculations between both $\theta_\varepsilon$ and $\theta_v$ and fPET are denoted as NMI($\theta_\varepsilon$, fPET) and NMI($\theta_v$, fPET), respectively.

The underestimation of observed $\theta$ / ET coupling via the impact of mutually independent $\theta$ and ET errors in AmeriFlux observations (Crow et al., 2015) is minimized by focusing on the ratio between NMI($\theta_\varepsilon$, fPET) and NMI($\theta_v$, fPET). Therefore, relative comparisons between NMI($\theta_\varepsilon$, fPET) and NMI($\theta_v$, fPET) are based on examining the size of their mutual ratio NMI($\theta_\varepsilon$, fPET) / NMI($\theta_v$, fPET). To quantify the standard error of NMI differences between various soil water content products, we applied a nonparametric, 500-member bootstrapping approach and calculated the pooled average of sampling errors across all sites assuming spatially independent sampling error.

Finally, we also examined the impact of potential nonlinearity in the $\theta$ / ET relationship by comparing non-parametric NMI results with comparable inferences based on a conventional Pearson’s correlation calculation. The correlation-based coupling strength between $\theta_\varepsilon$ and fPET is denoted as $R(\theta_\varepsilon, fPET)$ and between $\theta_v$ and fPET as $R(\theta_v, fPET)$.

### 3 Results

#### 3.1 Comparison of NMI($\theta_\varepsilon$, fPET) and NMI($\theta_v$, fPET)

Figure 1 contains boxplots of modeled and observed NMI($\theta_\varepsilon$, fPET) and NMI($\theta_v$, fPET), i.e., the relative magnitude of fPET information contained in surface soil water content and vertically integrated (0–40 cm) soil water content estimated from Case I, sampled across all the AmeriFlux locations listed in Table 1. According to the AmeriFlux ground mea-
measurements, median values of NMI(θv, fPET) and NMI(θs, fPET) individually and (b) NMI(θv, fPET) normalized by NMI(θv, fPET).

As mentioned in Sect. 2.1, an important concern is the impact of interpolation errors used to estimate 0–40 cm θv from AmeriFlux θv observations acquired at non-uniform depths. To ensure that different methods for calculating AmeriFlux θv values do not affect the main conclusion of this study, we configured three cases for NMI(θs, fPET) and NMI(θv, fPET) estimated by the NOAHMP and CLSM models are similar to these (observation-based) AmeriFlux values. With the single exception that the CLSM predicts much larger site-to-site variation in NMI(θs, fPET).

In contrast, NMI(θs, fPET) and NMI(θv, fPET) values sampled from GLEAM θ and fPET estimates show positive biases (with median θ of about 0.5 and 0.4 [–] for NMI(θs, fPET) and NMI(θv, fPET), respectively) with respect to all other estimates.

Using the 34 AmeriFlux site-collocated samples pixels for a paired t test, both LSMs and GLEAM overall exhibit significantly (at the 0.05 level) higher NMI(θv, fPET) compared to NMI(θs, fPET), implying the surface soil water content observations contain more fPET information than vertically integrated soil water content observations. However, the observed difference between NMI(θs, fPET) and NMI(θv, fPET) is less discernible in AmeriFlux measurements (Fig. 1a).

Here, AmeriFlux observations are used as a baseline for LSM and GLEAM evaluation. However, it should be stressed that random observation errors in θ and fPET will introduce a low bias into AmeriFlux-based estimates of both NMI(θs, fPET) and NMI(θv, fPET; Crow et al., 2015) and thus their difference as well. To address this concern, Fig. 1b plots the ratio of NMI(θs, fPET) and NMI(θv, fPET), which effectively normalizes (and therefore minimizes) the impact of random observation errors. As discussed above, these ratios results illustrate the general tendency for NMI(θs, fPET) to exceed NMI(θv, fPET). They also highlight the tendency for GLEAM to overvalue θv (relative to θs) when estimating fPET. A second approach for reducing the random error of θ and fPET measurement errors is the correction based on triple collocation (TC) applied in Crow et al. (2015). However, this approach is currently restricted to linear correlations and cannot be applied to estimate NMI. Future work will examine extending the information-based TC approach of Nearing et al. (2017) to the examination of NMI.

3.2 Sensitivity of AmeriFlux-based NMI(θs, fPET) / NMI(θv, fPET)

Figure 3 plots the spatial distribution of NMI(θs, fPET) and NMI(θv, fPET) results for each of the individual 34 AmeriFlux sites listed in Table 1. The climatic regime is represented by AI (aridity index) values plotted as the background color in Fig. 3. It can be seen in Fig. 3 that NMI(θs, fPET)
estimates from LSMs and GLEAM are spatially related to hydro-climatic conditions, as NOAHMP and CLSM predict that $\theta_s$ is moderately coupled with fPET (i.e., NMI($\theta_s$, fPET) of 0.3–0.5 [–]) in the arid southwestern USA (AI<0.2) and only loosely coupled with fPET in the relatively humid eastern USA. A similar decreasing trend of NMI($\theta_s$, fPET) from the southwestern to eastern USA is also captured by GLEAM. However, as noted above, GLEAM generally overestimates NMI($\theta_s$, fPET) and NMI($\theta_v$, fPET) compared to NOAHMP, CLSM and AmeriFlux. In contrast, a relatively weaker spatial pattern emerges in AmeriFlux-based NMI($\theta$, fPET) results. In addition, spatial patterns for NMI($\theta_s$, fPET) are less defined than for NMI($\theta_v$, fPET) in all four datasets.

Scatterplots in Fig. 4 summarize the spatial relationship between LSM- and GLEAM-based NMI($\theta_s$, fPET) and NMI($\theta_v$, fPET) results versus AmeriFlux observations across different land use types. While observed levels of correlation in Fig. 4 are relatively modest, there is a significant level ($p < 0.05$) of spatial correspondence between modeled and observed NMI results only over forest sites, motivating the need to better understand processes responsible for spatial variations in NMI results. In addition, stratifying NMI($\theta_s$, fPET) / NMI($\theta_v$, fPET) ratio results according to vegetation type (Fig. A1 in the Appendix) confirms that NMI($\theta_s$, fPET) slightly exceeds NMI($\theta_v$, fPET) across all vegetation types (and thus all rooting depths characterizing each vegetation type). This suggests that our analysis is not severely affected by variations in the depth of $\theta$ measurements. For further discussion on the impact of land cover on NMI results, please see Appendix A.

3.4 Sensitivity of NMI($\theta_s$, fPET) / NMI($\theta_v$, fPET) ratio to climatic conditions

Figure 5 further summarizes the NMI($\theta_s$, fPET) / NMI($\theta_v$, fPET) ratio as a function of AI for all four products (NOAHMP, CLSM, GLEAM and AmeriFlux). Error bars represent the standard deviation of sampling errors calculated from a 500-member bootstrapping analysis. With increasing AI, there is a significant decreasing trend in both NMI($\theta_s$, fPET) and NMI($\theta_v$, fPET) for all three simulations, with a goodness of fit above 0.5 (figure not shown). For all cases, the NMI($\theta_s$, fPET) / NMI($\theta_v$, fPET) ratios are consistently greater than unity under all climatic conditions. However, the estimated NMI($\theta_s$, fPET) / NMI($\theta_v$, fPET) ratios from all three simulations (NOAHMP, CLSM and GLEAM) exhibit quite different trends with respect to AI. The NMI($\theta_s$, fPET) / NMI($\theta_v$, fPET) ratio for CLSM decreases with increasing AI, with a moderate goodness-of-fit value of 0.28, while GLEAM estimates of NMI($\theta_s$, fPET) / NMI($\theta_v$, fPET) shows an opposite increasing trend with increasing AI. Conversely, there is relatively lower sensitivity of the NMI($\theta_s$, fPET) / NMI($\theta_v$, fPET) ratio to AI captured in the AmeriFlux measurements.

Connecting these findings to the spatial distribution of NMI($\theta_s$, fPET) and NMI($\theta_v$, fPET; Fig. 3) confirms that the relative magnitudes of NMI($\theta_s$, fPET) and NMI($\theta_v$, fPET) for both LSMs and GLEAM are spatially related to hydro-climatic regimes. In contrast, this link is weaker in the AmeriFlux measurements which, except for a small fraction of very low AI sites, do not appear to vary as a function of AI. These conclusions are not qualitatively impacted by looking at NMI($\theta_s$, fPET) and NMI($\theta_v$, fPET) differences, as opposed to their ratio as in Fig. 5, or by looking at $R(\theta_s$, fPET) and $R(\theta_v$, fPET) instead of NMI.

4 Discussion and conclusion

Since transpiration dominates the global ET (Jasechko et al., 2013), deep-layer soil water content ($\theta_s$) is generally considered to contain more ET information than that of surface soil water content ($\theta_v$), given that plant transpiration is balanced by root water uptake from deeper soils (Seneviratne et al., 2010). However, this assumption is rarely tested using models and/or observations. Here, we apply normalized mutual
Figure 3. NMI($\theta_s$, fPET) (left panels) and NMI($\theta_v$, fPET) (right panels) estimates at AmeriFlux sites for (a, e) NOAHMP, (b, f) CLSM, (c, g) GLEAM and (d, h) AmeriFlux. Marker color reflects NMI magnitudes and symbol type reflects local land-cover type at each site. Background color shading reflects aridity index (AI) values.
Figure 4. Scatterplot of LSM- and GLEAM-based (a, c, e) NMI($\theta_s$, fPET) and (b, d, f, h) NMI($\theta_v$, fPET) results versus AmeriFlux observations. Red symbols represent simulations from NOAHMP36; blue symbols represent simulations from CLSM2 and green symbols represent GLEAM retrievals.
information (NMI) to examine how the vertical support of a soil water content product affects its relationship with concurrent surface ET.

Specifically, using AmeriFlux ground observations, we examine whether (NMI-based) estimates of LSMs and GLEAM $\theta_S$ versus ET and $\theta_V$ versus ET coupling strength accurately reflect observations acquired at a range of AmeriFlux sites. In general, compared to the baseline case of exponential filter extrapolated 40 cm bottom-layer $\theta_V$, LSMs and GLEAM agree with AmeriFlux observations in that the overall fPET information contained in $\theta_S$ is slightly higher than that of $\theta_V$ (Fig. 1). However, the sensitivity analysis showed this difference between NMI($\theta_S$, fPET) and NMI($\theta_V$, fPET) diminishes when using different methods for calculating $\theta_V$ using AmeriFlux observations (Fig. 2). As a result, this result should be viewed with caution.

While NOAHMP- and CLSM-derived NMI($\theta_S$, fPET) and NMI($\theta_V$, fPET) results are generally consistent with the AmeriFlux observations, GLEAM overestimates NMI($\theta_S$, fPET), NMI($\theta_V$, fPET) and the ratio NMI($\theta_S$, fPET) / NMI($\theta_V$, fPET) relative to observations. Although both LSMs and GLEAM are based on the same classical two-section (soil water content-limited and energy-limited) ET regimes framework (Sect. 2.2), they differ in two fundamental aspects. First, the evaporative stress factor $S$ is represented as a more direct and strong function of soil water content in GLEAM – see Eqs. (4) and (5) – which leads to the overestimation of the $\theta$ / ET coupling strength. This is consistent with our results that GLEAM generally overestimates NMI($\theta_S$, fPET) and NMI($\theta_V$, fPET) consistently across all land covers, compared to AmeriFlux-based estimates. On the other hand, NOAHMP and CLSM approximate ET in the manner of biophysical models, and expresses biophysical control on ET through the stomatal resistance $r_S$, which is a function of multiple limiting factors including $\theta$. Therefore, the more complex ET scheme employed by NOAHMP and CLSM would seem to mitigate the overestimation of NMI($\theta_S$, fPET) and NMI($\theta_V$, fPET), as other relevant factors besides $\theta$ (such as temperature, foliage nitrogen) are also considered in determining maximum carboxylation rate $V_{\text{max}}$ and stomatal resistance $r_S$, and consequently more realistic actual ET.

Second, the stress factor $\beta$ in both LSMs considers the cumulative effects of $\theta$ conditions along different layers (Eq. 1), while the corresponding factor $S$ in GLEAM only uses the wettest soil layer condition, which is top layer at most sites. This likely explains the overestimation of the NMI($\theta_S$, fPET) / NMI($\theta_V$, fPET) ratio by GLEAM.

Nevertheless, we would like to stress that all approaches considered in our paper contain (at their core) a parameterized relationship between $\theta$ and ET. While the implications of mis-parameterizing this relationship are arguably more severe for a land surface model, we argue that the issue remains relevant for any approach (such as GLEAM) that utilizes

Figure 5. For (a) NOAHMP, (b) CLSM, (c) GLEAM and (d) AmeriFlux estimates, the ratio of NMI($\theta_S$, fPET) and NMI($\theta_V$, fPET) as a function of AI across all AmeriFlux sites.
J. Qiu et al.: Model representation of the coupling

Figure 6. Daily ET error in GLEAM as a function of GLEAM bias in NMI(θ_s, fPET) / NMI(θ_v, fPET) ratio across 34 AmeriFlux sites.

Although the median values of NMI(θ_s, fPET) and NMI(θ_v, fPET) predicted by NOAHMP and CLSM are generally in line with AmeriFlux observations, they are more spatially related to hydro-climatic conditions (as summarized by AI) than their counterparts acquired from AmeriFlux measurements. Seen from the plot of NMI(θ_s, fPET) / NMI(θ_v, fPET) ratio as a function of AI (Fig. 5), the modeled and observed median of NMI(θ_s, fPET) / NMI(θ_v, fPET) ratio decreases with increasing AI, and the decreasing trend is particularly clear when AI is lower than 1.0 [-]. In contrast, there is relatively lower sensitivity to aridity exhibited in the AmeriFlux measurements.

These results provide several key insights into future land–atmosphere coupling analysis and LSMs, as well as ET algorithm development. First, all the datasets – both model-based and ground-observed – indicate that θ_s contains at least as much ET information as θ_v. Hence, remote-sensing land surface soil water content datasets are suitable, and should be considered, for analyzing the general interaction between land and atmosphere, e.g., soil water content–air temperature coupling (Dong and Crow, 2019) and the interplay of soil water content and precipitation (Yin et al., 2014). Additionally, future generations of GLEAM may consider more sophisticated evaporation stress functions, which may improve its accuracy in representing the soil’s control on local ET. This may, in turn, improve the accuracy of the GLEAM ET product. Finally, our results demonstrate that modeled θ / ET is more sensitive to hydro-climates than the observed relationship. Modifying the model structures to reduce such sensitivity might be necessary for accurately representing the interaction of land surface and atmosphere across different climate zones. This may lead to more realistic projections of future drought-induced heat waves, when coupled with general circulation models.
Appendix A

We performed an additional sensitivity analysis to explicitly demonstrate the effect of different vegetation land-cover types and consequently different rooting depths (or $\theta_v$ measurement depths) on the $\text{NMI}(\theta_S, \text{fPET}) / \text{NMI}(\theta_v, \text{fPET})$ ratio, and plotted these results in Fig. A1. The figure confirms that, consistent with AmeriFlux, both LSMs and GLEAM predict that $\text{NMI}(\theta_v, \text{fPET})$ is slightly higher than $\text{NMI}(\theta_S, \text{fPET})$ over most vegetation types, and GLEAM overestimates $\text{NMI}(\theta_S, \text{fPET}) / \text{NMI}(\theta_v, \text{fPET})$ for most vegetation types.

![Figure A1](image)

**Figure A1.** For (a) NOAHMP, (b) CLSM, (c) GLEAM and (d) AmeriFlux estimates, the ratio of $\text{NMI}(\theta_S, \text{fPET})$ and $\text{NMI}(\theta_v, \text{fPET})$ as a function of vegetation types across all AmeriFlux sites. “ENF”, “DBF”, “MF”, “OS” and “WS” represent evergreen needleleaf forests, deciduous broadleaf forests, mixed forests, open shrubland, and woody savannas, respectively.
Ground-based soil water content and surface flux data are available from https://ameriflux.lbl.gov/login/?redirect_to=/data/download-data/ (last access: February 2020) (Raz-Yaseef et al., 2015). GLEAM dataset is available from https://www.gleam.eu/ (last access: February 2020). LSMs simulations of NOAHMP and CLSM used in this study are available by contacting the authors.

The supplement related to this article is available online at: https://doi.org/10.5194/hess-24-581-2020-supplement.

JQ and WTC conceptualized the study. JD helped preparing the LSMs simulation. GSN assisted in the mutual information analysis. JQ carried out the analysis and wrote the first draft manuscript and WTC refined the work. All authors contributed to the analysis, interpretation and writing.

The authors declare that they have no conflict of interest.

This research has been supported by the National Natural Science Foundation of China (grant nos. 41971031, 41501450 and 51779278), the Natural Science Foundation of Guangdong Province (grant no. 2016A030310154).

This paper was edited by Dominic Mazvimavi and reviewed by two anonymous referees.

References


