Diagnosis of environmental controls on daily actual evapotranspiration across a global flux tower network: the roles of water and energy

Qiong Han¹, Qin Liu¹, Tiejun Wang¹,², Lichun Wang¹,², Chongli Di¹,², Xi Chen¹,², Keith Smettem³ and Shailesh Kumar Singh¹

¹ Institute of Surface-Earth System Science, School of Earth System Science, Tianjin University, Weijin Road 92, Tianjin 300072, People’s Republic of China
² Tianjin Key Laboratory of Earth Critical Zone Science and Sustainable Development in Bohai Rim, Tianjin University, Weijin Road 92, Tianjin 300072, People’s Republic of China
³ Institute of Agriculture, The University of Western Australia, Crawley, WA 6009, Australia
⁴ National Institute of Water and Atmospheric Research, Christchurch, New Zealand

E-mail: tiejun.wang@tju.edu.cn

Keywords: actual evapotranspiration, FLUXNET, soil water, net radiation, boosted regression tree, daily time scale

Abstract

Relative contributions from environmental factors to daily actual evapotranspiration (ETₐ) across a variety of climate zones is a widely open research question, especially regarding the roles played by soil water content (SWC); water supply) and net radiation (Rn; energy supply) in controlling ETₐ. Here, the boosted regression tree method scheme was employed to quantify environmental controls on daily ETₐ using the global FLUXNET dataset. Similar to the general trend suggested by the Budyko theory at annual scales, the results showed that the relative control of SWC on daily ETₐ increased with increasing aridity index (Φ); however, Rn played a major role at most FLUXNET sites (roughly Φ < 4), indicating that Rn could be a leading control on daily ETₐ even at water-limited sites. The variability in the relative controls of SWC and Rn also partly depended on factors affecting water availability for daily ETₐ (e.g. vegetation characteristics and groundwater depth). Our study showed that other than SWC and Rn, the net effect of environmental controls (particularly leaf area index) on daily ETₐ was more important at drier sites than at relatively humid sites. This suggests that near-surface hydrological processes are more sensitive to vegetation variations due to their ability to extract deep soil water and enhance ETₐ, especially under arid and semi-arid climatic conditions. Our findings illustrate how environmental controls on daily ETₐ change as the climate dries, which has important implications for many scientific disciplines including hydrological, climatic, and agricultural studies.

1. Introduction

Knowledge of actual evapotranspiration (ETₐ) is fundamental in understanding terrestrial water cycles and ecosystem functioning (Oki and Kanae 2006, Maxwell and Condon 2016), and also has important implications for hydrological, climatic, and agricultural studies (Wang and Dickinson 2012, Fisher et al 2017). Across different spatiotemporal scales, ETₐ intricately interacts with a multitude of land surface and meteorological processes through complex feedback mechanisms, e.g. the significant links between regional precipitation (P) and evaporation (Zveryaev and Allan 2010); changes in gross primary production coupling with surface energy partitioning (Gaib et al 2005). As a result, ETₐ tends to display significant variability over space and time (Vivoni et al 2008, Trambauer et al 2014). Therefore, it remains challenging to elucidate the impacts of different environmental factors (e.g. energy and water) on ETₐ under varying land surface and climatic conditions (Wang and Dickinson 2012).

At annual and mean annual time scales, ETₐ is primarily modulated by energy and water inputs
as illustrated by the well-known Budyko framework (Budyko 1974). This framework has been widely used to characterize water and energy exchanges across the land–atmosphere interface by linking ET$_a$ with potential evapotranspiration (ET$_p$; energy input) and P (water input), and also applied at monthly or even shorter time scales (Zhang et al 2008, Chen et al 2013). To this end, hydroclimatic regimes can be classified into (a) water-limited and (b) energy-limited environments, depending on relative supplies of energy and water (Haghighi et al 2018). Given that ET$_a$ responses to the changes in energy and water inputs differ considerably between the two environments, their differentiation is of importance for understanding a range of hydrological processes, such as the impacts of climate change on streamflow (Arora 2002) and regional water balance (Renner and Bernhofer 2012). In addition, the relationships of ET$_a$ with soil water content ((SWC); an indicator of water supply) and ET$_p$ vary significantly across sites (Wang and Dickinson 2012), which exhibit multidimensional functions modified by a suite of atmospheric, soil, eco-physiological, and biogeochemical factors (Zhang et al 2001, Vivoni et al 2008, Chang et al 2018, Haghighi et al 2018). For instance, Gu et al (2006) found that vapor pressure deficit (VPD) interfered with how SWC and net radiation (Rn) affected surface energy partitioning, which was also demonstrated by Novick et al (2016). Vivoni et al (2008) showed that vegetation greening affected the relationship between ET$_a$ and SWC along a latitudinal gradient of land covers. Apparently, complex environmental conditions provide a challenge for exploring the net effect of water and energy supplies on ET$_a$.

Previous studies have also suggested that the impacts of different controlling factors on ET$_a$ vary across different time scales (Cheng et al 2011, Ding et al 2013). At diurnal scales, the variability in ET$_a$ is mainly affected by atmospheric water demands, which are primarily determined by air temperature ($T_a$), Rn, and VPD (Wilson et al 2003, Shi et al 2008). At seasonal and interannual scales, the site-specific variability in ET$_a$ could be explained by changes in either annual P or Rn, depending on the water and energy availability (Brümmer et al 2012, Liu and Feng 2012). As a further demonstration, Scott and Biederman (2019) showed that the dominant controls on ET$_a$ changed from sub-daily (weather) to seasonal (soil moisture) scales in a semi-arid savanna ecosystem.

From the aforementioned studies, the question arises as to whether the conceptualization of the Budyko framework can be extended to daily time scales (e.g. whether the water input is the primary control on daily ET$_a$ in water-limited environments, and vice versa), as applied previously (e.g. Zhang et al 2008). This question is particularly challenging due to the complex interactions of daily ET$_a$ with various environmental factors, which has yielded mixed findings from previous studies. For instance, Ford et al (2014) found that daily ET$_a$ was strongly influenced by SWC when Rn was above normal in a semi-arid region, while Liu et al (2019) showed that Rn was the leading factor even at arid sites. Rim (2008) showed that the combined effect of Rn, SWC, and wind speed (WS) accounted for >70% of daily ET$_a$ variations in semi-arid watersheds with Rn being the dominant factor. In addition, Williams and Torn (2015) reported that leaf area index (LAI) had a significant control on daily surface heat flux partitioning instead of SWC in the U.S. Southern Great Plains. Gong et al (2007) reported a similar finding at a semi-arid site in northwest China. Note that those studies were based on data from a limited number of sites with certain climatic conditions. Due to the complex interplays between ET$_a$ and surrounding environments, the controls on daily ET$_a$ variability across different climate zones and ecosystems at the global scale are still not well understood, which calls for further research.

To this end, long-term observed data from a global network of eddy covariance (EC) towers (i.e. FLUXNET) were analyzed to systematically assess the relative controls of environmental variables on daily ET$_a$ within a quantitative framework, using the boosted regression tree (BRT) technique. Based on the BRT results, we evaluated the roles played by different factors across the globally distributed network, particularly by water and energy as stressed by the Budyko theory, in affecting ET$_a$ at the daily time scale under different land surface and climatic conditions.

2. Methods and materials

2.1. Data acquisition

The FLUXNET is a global network equipped with EC towers to measure site-level water, carbon, and energy exchanges between the atmosphere and biosphere (Baldocchi et al 2001). In situ data on meteorological and soil variables are collected at the FLUXNET sites. In this study, the FLUXNET 2015 Tier 1 dataset with daily values (e.g. latent heat (LE), sensible heat (H), soil heat flux (G), P, SWC, $T_a$, Rn, relative humidity (RH), VPD, and WS) was obtained. This dataset has been treated through rigorous procedures of data processing and quality control for research purposes (Pastorello et al 2014, Vuichard and Papale 2015). Full details regarding the descriptions of the FLUXNET and the data processing can be accessed at https://fluxnet.fluxdata.org/. Based on data availability, EC sites with at least 2 years of observations were initially chosen to ensure the inclusion of sufficient data points in the statistical analysis.

At the selected FLUXNET sites, daily ET$_p$ was calculated from measured weather data using the Penman–Monteith equation (Allen et al 1998). The aridity index $\Phi$ ($\Phi = \overline{ET_p}/\overline{P}$, where $\overline{ET_p}$ and $\overline{P}$ are mean annual ET$_p$ and P) was used to define the
climate dryness at the sites. Based on the classification from the Food and Agriculture Organization (FAO; Spinoni et al. 2015), the FLUXNET sites were divided into four categories: humid ($\Phi \leq 1.3$), sub-humid ($1.3 < \Phi \leq 1.5$), semi-arid ($1.5 < \Phi \leq 5$), and arid ($5 < \Phi < 20$) sites.

The land cover at the FLUXNET sites varied considerably, including cropland, forest, grassland, savanna, shrubland, and wetland according to the classification system given by the International Geosphere–Biosphere Programme (Loveland et al. 2001). LAI data that are routinely used to quantify vegetation coverage were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite land products. Specifically, based on the geographical coordinates of the FLUXNET sites, LAI data with a temporal resolution of 4 d and a spatial resolution of 500 m were retrieved from the MCD15A3H dataset (after 2002; version 6; https://e4ftl01.cr.usgs.gov/MOTA/)) to assess the impact of vegetation on ET$_a$.

The global soil database of Shangguan et al. (2014) with a spatial resolution of $30'' \times 30''$ and the global dataset of the depth to groundwater table (DTGT) of Fan et al. (2013) with a spatial resolution of ~1 km were used to obtain the soil texture and DTGT, respectively, at each FLUXNET site, according to their geographical coordinates.

### 2.2. Data pre-processing

All sites were initially screened to remove low-quality data. First, we only used measurements after 2002 as this is the starting date of the LAI dataset. Then, based on data availability, sites with no SWC or energy flux data were omitted. Secondly, the time periods with missing energy flux data (i.e. LE, $H$, $R_n$, and $G$, with a maximum gap of 51.6% of the raw data) were removed. Thirdly, to minimize the effect of freezing conditions on SWC measurements, the time periods with $T_s$ below 0°C were not included in the analysis. Fourthly, for the energy balance closure problem, data with $|ECR - 0.83| \geq 0.22$ (forest sites) and $|ECR - 0.87| \geq 0.15$ (other sites) were removed (Stoy et al. 2013), where the ECR (energy closure ratio; defined as $(LE + H)/(R_n - G)$ based on the daily data) is a standard measure to evaluate the performance of EC measurements. Finally, data with evaporative fraction $(EF)$; defined as $LE/(LE + H)$) values greater than 1 were also removed (Liu et al. 2019), where EF represents the percentage of incoming energy used for ET$_a$.

LAI data were then further processed to obtain a high-quality LAI dataset. First, on the basis of the quality control information, data with unfavorable conditions (e.g. the presence of cloud and/or snow cover) were first filtered out. Secondly, a cubic smoothing spline method was used for interpolating the 4 day interval data into daily values (Horn and Schulz 2010). Finally, a robust method that is based on the Savitzky–Golay filter for smoothing out noises in the LAI time series was used (Chen et al. 2004).

In addition, the multicollinearity in predictor variables (e.g. SWC, $R_n$, LAI, RH, VPD, WS, and $T_s$) may affect the performance of the BRT model for predicting response variables (i.e. ET$_a$). Before the BRT analysis, Pearson’s correlation coefficient ($r_p$) and variance inflation factor (VIF) analysis (i.e. $|r_p| < 0.7$ and VIF < 5) were used to test the multicollinearity among predictor variables (Dormann et al. 2013), which led to the use of five factors—SWC, $R_n$, LAI, RH (or VPD), and WS as predictor variables in this study. Note that due to the strong multicollinearity between RH and VPD at most sites, either RH or VPD was used based on two criteria: (a) no multicollinearity between the remaining RH or VPD and other variables; (b) the selected predictor variables at the sites for each aridity class remained the same as much as possible. Here, the sites existing high multicollinearity between any two of the five variables were also removed.

After the above procedures, a total of 78 FLUXNET sites were used in the following analysis with an average of 852 data points (i.e. 852 d) (see figure 1(a) and table S1 available online at http://stacks.iop.org/ERL/15/124070/mmedia, for site-specific details). Data retained after the filtering processes were used to assess the impact of different variables on daily ET$_a$.

### 2.3. Boosted regression tree method

The BRT method was adopted to quantify the relative controls of predictor variables (i.e. SWC, $R_n$, LAI, RH/VPD, and WS) on response variables (i.e. ET$_a$) at the selected sites (Breiman et al. 1984, Elith et al. 2008). The BRT method couples the regression tree method with boosting algorithms to predict the relationship between predictor and response variables, which does not require any specific data distributions. This method is particularly suitable for analyzing data with issues of outliers, interactions among predictor variables, and missing data (Elith et al. 2008). Given its advantages, the BRT method has been increasingly used in a variety of research fields, such as public health (Bhatt et al. 2013), oxynitride distributions (Sayegh et al. 2016), and carbon stocks (Lin et al. 2016).

In this study, the open-source BRT package developed by Elith et al. (2008) in R software (version 3.5.1, R Core Team, 2018) was employed to derive the BRT model at each site. The detailed introduction of the BRT method, and the procedures for constructing the BRT model and for computing the relative control of each predictor variable on ET$_a$ are provided in the supplemental text S1. The performance of the BRT model was evaluated using the Nash–Sutcliffe Efficiency (NSE), root mean square error (RMSE), and coefficient of determination ($R^2$) by comparing observed ET$_a$ with modeled ET$_a$ for each site (table S2).
3. Results and discussion

3.1. Two typical examples from the FLUXNET dataset

The dataset covered a wide range of hydroclimatic regimes as illustrated in figures 1(b)–(d), which show the distributions of $P$, $\overline{ET}_p$, and mean annual $ET_a$ ($ET_a$; calculated from the non-filtering daily latent heat flux data) across different climatic zones. The diversity of hydroclimatic conditions at the FLUXNET sites led to considerable variations in $ET_a$ from 122.0 to 1459.4 mm yr$^{-1}$, providing an ideal dataset for examining the impacts of various influencing variables on daily $ET_a$ under different environmental conditions.

To illustrate the roles played by water and energy inputs in controlling daily $ET_a$, figure 2 provides the time series of daily $P$, SWC, $Rn$, and $ET_a$ at two typical sites with contrasting conditions (e.g. an arid climate at AU-ASM with $\Phi = 5.43$ and a humid climate at US-KS2 with $\Phi = 1.17$). The temporal variations in daily $ET_a$ were largely controlled by the $P$ inputs at AU-ASM, while corresponded well with the changes in $Rn$ at US-KS2, clearly demonstrating the different roles of water and energy in controlling $ET_a$ under water-limited and energy-limited conditions according to the Budyko framework. In addition, the BRT results revealed that the respective contributions of SWC and $Rn$ to $ET_a$ were 59.9% and 8.9% at AU-ASM, and 21.0% and 64.5% at US-KS2, respectively, illustrating the viability of using the BRT method for diagnosing the controls on daily $ET_a$. More intriguingly, at AU-ASM, similar SWC levels during the same season in different years led to different magnitudes of daily $ET_a$ (as indicated by the open squared areas in figures 2(a)–(c)). This observation suggested that at this arid site, factors other than SWC were also non-negligible in controlling daily $ET_a$, which has decisive implications for hydrological cycles in arid regions (e.g. episodic groundwater recharge (GR); Crosbie et al. 2012).

3.2. Environmental controls on daily $ET_a$ across the FLUXNET sites

Figure 3 displays the relative contributions of different environmental factors in affecting daily $ET_a$
for four aridity classes from the BRT analysis (see table S2 for site-specific contributions). The BRT models exhibited good performance, with high NSE and $R^2$, and low RMSE values at most sites. Overall, $Rn$ was the dominant factor with an average contribution of 49.2% across all sites, followed by LAI (24.8%), SWC (14.8%), RH/VPD (7.0%), and WS (4.2%). As expected, the $Rn$ contribution tended to decrease with an increase in climate dryness, while the SWC contribution showed an opposite trend. Interestingly, the impact of LAI appeared to be greater under drier conditions; whereas, no clear patterns could be observed for RH/VPD and WS.

3.2.1. Environmental controls on daily $ET_a$ at humid and sub-humid sites
At the humid and sub-humid sites (i.e. $\Phi \leq 1.5$), daily $ET_a$ fluxes were predominantly affected by $Rn$ with an average contribution of 59.8%, while the average SWC contribution was comparatively less.

Figure 2. Example daily curves from two typical sites (an arid climate at AU-ASM and a humid climate at US-KS2) used in this study including precipitation ($P$), soil water content (SWC), net radiation ($Rn$) and actual evapotranspiration ($ET_a$).
Figure 3. Relative contributions of (a) net radiation ($R_n$), (b) soil water content (SWC), (c) leaf area index (LAI), (d) relative humidity or vapor pressure deficit (RH/VPD), and (e) wind speed (WS) to daily actual evapotranspiration for the four aridity classes from the BRT analysis. The 'violin' is a mirror-imaged kernel density plot. The width of the shapes showed the frequency of the data distributions. Boxes span the first to third quartiles; the horizontal lines inside the boxes represent the median values; black dots represent outliers in each group.
(9.4%), reinforcing the vital role of energy in determining daily \( \theta_\text{ET}_a \) at sites with relatively abundant water supplies.

Note that exceptions mainly existed in the cold region of the northern USA with the land cover of forests (see also figure S1), where LAI was the primary contributing factor; this is clearly demonstrated by the fact that the rapid increases in daily \( \theta_\text{ET}_a \) during early summers mainly coincided with the leaf emergence (figure S2). In this region, the climate is characterized by a continental type with short growing seasons and cold winters (Cook et al. 2004). The seasonal variations in LAI were obvious and \( \theta_\text{ET}_a \) occurred mainly in the growing seasons. One of the likely reasons for the dominant control of LAI on daily \( \theta_\text{ET}_a \) might be attributed to the dense coverage of forests, which could block the incoming solar radiation and thus suppress soil evaporation (Bagley et al. 2017). Therefore, \( \theta_\text{ET}_a \) primarily consisted of plant transpiration, which in turn was controlled by LAI.

3.2.2. Environmental controls on daily \( \theta_\text{ET}_a \) at semi-arid and arid sites

At the sites with relatively abundant energy inputs (i.e. \( \Phi > 1.5 \)), the SWC constraint on \( \theta_\text{ET}_a \) increased with increasing climate dryness, while the impact of Rn showed an opposite trend (figure 3). This trend was aligned with the general understanding of the roles that are played by water and energy in controlling \( \theta_\text{ET}_a \). However, interestingly, the BRT results indicated that the average contribution of Rn (32.1%) outweighed that of SWC (23.4%), which was consistent with the recent findings at a few water-limited sites in China (Liu et al. 2019), but was still unexpected at the global scale. This underscores the importance of energy in influencing \( \theta_\text{ET}_a \) and consequently other hydrological processes (e.g. GR), even under water-limited conditions across the globally distributed FLUXNET sites.

Several reasons might be attributed to the seemingly different finding that Rn was the leading factor at some sites with \( \Phi > 1.5 \). First, soil texture varied noticeably across the FLUXNET sites. It appeared that daily \( \theta_\text{ET}_a \) at some sites with high clay content tended to be more affected by Rn (e.g. the RU-Ha1 site with an average clay content of 38% in the topsoil (0–10 cm)). Wang et al. (2009) found that under similar climatic conditions, clayey soils had higher water holding capacities, resulting in higher opportunities of soil water for being evaporated. Secondly, the presence of any shallow groundwater could alleviate the constraint of water stress on \( \theta_\text{ET}_a \) in water-limited regions, primarily due to capillary rise into the root zone and/or direct root water uptake from groundwater for \( \theta_\text{ET}_a \) (Yue et al. 2016). For instance, at the CN-Cng site from northeast China where the land cover was grass with \( \Phi = 3.63 \) and the respective contributions of Rn and SWC were 63.6% and 5.5%, the long-term average DTGT was 0.83 m. As shallow groundwater can provide an additional water source for \( \theta_\text{ET}_a \), energy subsequently becomes a limiting factor for \( \theta_\text{ET}_a \) under water-limited conditions that are essentially only determined by climatic conditions (Yue et al. 2016). Therefore, the interplay between DTGT and rooting depth is another important factor to be considered in water-limited regions for quantifying \( \theta_\text{ET}_a \) (Maxwell and Condon 2016).

It should be emphasized that based on the BRT results, LAI was shown to have an important influence on daily \( \theta_\text{ET}_a \) with an average contribution of 34.6% across the semi-arid and arid sites (most notably in Australia and the USA). Note that the sites with LAI as the leading factor were mostly covered by grass or of savanna ecosystems (see figure S1). These sites exhibited similar seasonal patterns of vegetation growth accompanied by soil water and solar radiation status. In the growing season, transpiration made a major contribution to \( \theta_\text{ET}_a \), daily \( \theta_\text{ET}_a \) rate remained continuously high, and its amount accounted for the most of annual \( \theta_\text{ET}_a \). As an indicator of biomass, higher LAI generally leads to greater requirements for plant transpiration, which was strongly coupled to photosynthesis under arid conditions (Scott et al. 2015, Williams and Torn 2015). Moreover, plants in arid regions tend to access deeper soil water, where soil water is a temporally more stable (or less variable) source for plant transpiration (e.g. Schenk and Jackson 2002, Wang et al. 2015). The factor LAI seemed to be an improvement over available soil water as a measure of the land surface state relevant to surface heat flux partitioning (Williams and Torn 2015); thus, the \( \theta_\text{ET}_a \) responses to soil water status are mostly reflected by the changes in LAI. Therefore, our results suggested that the change in vegetation conditions would have a greater impact on \( \theta_\text{ET}_a \) and subsequently on hydrological cycles in drier regions than in humid regions. In short, the net effect of environmental controls, other than SWC and Rn, on daily \( \theta_\text{ET}_a \) appeared to be more critical at drier sites (\( \Phi > 1.5 \)) than at humid sites (\( \Phi \leq 1.5 \)).

3.3. Roles of energy and water in controlling daily \( \theta_\text{ET}_a \)

To compare the impacts of Rn and SWC on daily \( \theta_\text{ET}_a \), figure 4(a) shows the changes in the Rn and SWC contributions with \( \Phi \) across the FLUXNET sites. Although the impact of SWC on daily \( \theta_\text{ET}_a \) expectedly increased as the climate became progressively drier, Rn was the leading factor at most sites. Based on the fitted curves in figure 4(a), the overall trend suggested that the impact of SWC exceeded that for Rn when \( \Phi \) was roughly greater than 4 within the selected FLUXNET sites. Although the point \( \Phi > 4 \) would likely vary depending on the choice of the sites, our study provides strong evidence that Rn can play a major role in controlling daily \( \theta_\text{ET}_a \) even in water-limited environments (e.g. \( \Phi > 1.5 \) as defined...
by the FAO for semi-arid and arid climates) from the globally-distributed FLUXNET network. As an effective tool for diagnosing the coupling mechanism between water and energy exchanges across the land–atmosphere interface, the Budyko hypothesis states that the limiting factor for controlling ETa gradually shifts from energy in humid environments to water in arid environments (Budyko 1974, Haghighi et al 2018). Following Williams et al (2012) who also used the FLUXNET dataset, the relationship between Φ and ETa/P at mean annual time scales is demonstrated in figure 4(b) for the globally-distributed FLUXNET sites. Those sites were divided into two groups based on the relative contributions of Rn and SWC. Note that ETa/P ratios were greater than 1 at some sites, particularly at water-limited sites, owing to additional water supplies from other sources (e.g. shallow groundwater). In addition, it should be emphasized that long-term data were used to quantify the controls of Rn and SWC on daily ETa in this study (e.g. an average of 852 days for the selected FLUXNET sites), which was in line with the annual time scale required by the Budyko framework.

Nevertheless, our results further demonstrated that the contribution of Rn could exceed that of SWC in controlling ETa at water-limited sites. We proposed that the differences between our findings at the daily scale and the Budyko theory at the annual scale are also highly dependent on site-specific conditions (as indicated by the scatter of the data points). Part of the reason was that at daily or shorter time scales (unlike annual time scales for the Budyko theory; Istanbulluoglu et al 2012), concentrated rainfall events could suppress daily ETa due to lower Rn and higher RH during rainfall events at water-limited sites (e.g. the AU-ASM example in figure 2(a)); therefore, rainfall water might have better opportunities to pass the root zones or become surface runoff. Therefore, it is reasonable to argue that the temporally averaged SWC state does not necessarily reflect the amount of the available soil water for daily ETa, especially from a modeling perspective.

4. Conclusions

The global FLUXNET dataset was analyzed in this study using the BRT method to assess the contributions of various influencing factors in controlling daily ETa across different climate zones and ecosystems. The average contributions of Rn, LAI, SWC, RH/VPD, and WS were 49.2%, 24.8%, 14.8%, 7.0%, and 4.2%, respectively, across the selected FLUXNET sites. Overall, the BRT results showed that SWC became more important in controlling daily ETa with increasing Φ; meanwhile, the daily FLUXNET data revealed that Rn still played a pivotal role at water-deficient sites, which was tightly related to other environmental factors, such as vegetation, soil texture, and groundwater depth. Moreover, LAI exerted a stronger influence on daily ETa at drier sites than at humid sites, suggesting that hydrological processes in drier regions are more sensitive to the variations in vegetation conditions. As a result, the net effect of environmental controls other than SWC and Rn on ETa was more important at drier sites.

Data availability statements

LAI data are available at https://e4ftl01.cr.usgs.gov/MOTA/, the global soil database is available at http://globalchange.bnu.edu.cn/, the depth to groundwater table data should be obtained from the authors of the reference provided in this paper.

The data that support the findings of this study are openly available at the following URL/DOI: https://fluxnet.org/data/fluxnet2015-dataset/.
Acknowledgments

This work used EC data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, ICOS, OzFlux–TERN, Swiss Fluxnet, TCOS-Siberia, and USCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET EC data processing and harmonization was carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux and AsiaFlux offices. This study was supported by the National Key R&D Program of China (Grant Nos. 2019YFC1804400 and 2017YFC0406101) and the National Natural Scientific Foundation of China (Grant No. U1612441). T Wang acknowledges the financial support from the Thousand Talent Program for Young Outstanding Scientists and from Tianjin University, China.

ORCID iD

Shailesh Kumar Singh © https://orcid.org/0000-0002-3521-3917

References

Allen R G, Pereira L S, Raes D and Smith M 1998 FAO irrigation and drainage paper 56 (Rome)
Arora V K 2002 The use of the aridity index to assess climate change effect on annual runoff J. Hydrol. 265 164–77
Bagley J E, Kueppers L M, Billesbach D P, Williams I N, Biraud S C and Torn M S 2017 The influence of land cover on surface energy partitioning and evaporative fraction regimes in the U.S. Southern Great Plains J. Geophys. Res. 122 5793–807
Baldocchi D D et al 2001 FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities Bull. Am. Meteorol. Soc. 82 2415–34
Bhatt S et al 2013 The global distribution and burden of dengue Nature 496 504–7
Breiman L, Friedman J H, Olshen R A and Stone C J 1984 Classification and regression tree (Belmont, CA, USA: Wadsworth International Group)
Brümer C et al 2012 How climate and vegetation type influence evapotranspiration and water use efficiency in Canadian forest, peatland and grassland ecosystems Agric. Forest Meteorol. 153 14–30
Budyko M I 1974 Climate and Life (New York: Academic)
Chang K-Y, Paw U K T and Chen S-H 2018 The importance of carbon-nitrogen biogeochemistry on water vapor and carbon fluxes as elucidated by a multiple canopy layer higher order closure land surface model Agric. Forest Meteorol. 259 60–74
Chen J, Jonsson P, Tamura M, Gu Z, Matsushita B and Eklundh L 2004 A simple method for reconstructing a highquality NDVI time-series data set based on the Savitzky–Golay filter Remote Sens. Environ. 91 332–44
Chen X, Alimohammadi N and Wang D 2013 Modeling interannual variability of seasonal evaporation and storage change based on the extended Budyko framework Water Resour. Res. 49 6607–78
Cheng L, Xu Z, Wang D and Cai X 2011 Assessing interannual variability of evapotranspiration at the catchment scale using satellite-based evapotranspiration data sets Water Resour. Res. 47 W09509
Ciais P et al 2005 Europewide reduction in primary productivity caused by the heat and drought in 2003 Nature 437 529–33
Cook B D et al 2004 Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA Agric. Forest Meteorol. 126 271–95
Crosbie R S, McCallum J L, Walker G R and Chiew F H S 2012 Episodic recharge and climate change in the Murray–Darling Basin, Australia Hydrogeol. J. 20 245–61
Ding R, Kang S, Vargas R, Zhang Y and Hao X 2013 Multiscale spectral analysis of temporal variability in evapotranspiration over irrigated cropland in an arid region Agric. Water Manage. 130 79–89
Dormann C F et al 2013 Collinearity: a review of methods to deal with it and a simulation study evaluating their performance Ecoscary 36 27–46
Elith J, Leathwick J R and Hastie T 2008 A working guide to boosted regression trees J. Animal Ecol. 77 802–13
Fan Y, Li H and Miguez-Macho G 2013 Global patterns of groundwater table depth Science 339 940–3
Fisher J B et al 2017 The future of evapotranspiration: global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources Water Resour. Res. 53 2618–26
Ford T W, Wulf D O and Quiring S M 2014 Assessment of observed and model-derived soil moisture-evaporative fraction relationships over the United States Southern Great Plains J. Geophys. Res. 119 6278–91
Gong D, Kang S, Yao L and Zhang L 2007 Estimation of evapotranspiration and its components from an apple orchard in northwest China using sap flow and water balance methods Hydrol. Process. 21 931–8
Gu L et al 2006 Direct and indirect effects of atmospheric conditions and soil moisture on surface energy partitioning revealed by a prolonged drought at a temperate forest site J. Geophys. Res. 111 D16102
Haghighi E, Short Gianotti D J, Akbar R, Salvucci G D and Entekhabi D 2018 Soil and atmospheric controls on the land surface energy balance: a generalized framework for distinguishing moisture-limited and energy-limited evaporation regimes Water Resour. Res. 54 1831–51
Horn J and Schulz K 2010 Post-processing analysis of MODIS leaf area index subsets J. Appl. Remote Sens. 4 043557
Istanbulluolu E, Wang T, Wright O M and Lenters J D 2012 Interpretation of hydrologic trends from a water balance perspective: the role of groundwater storage in the Budyko hypothesis Water Resour. Res. 48 W00F16
Lin D, Anderson-Teixeira K J, Lai J, Mi X, Ren H and Ma K 2016 Traits of dominant tree species predict local scale variation in forest aboveground and topsoil carbon stocks Plant Soil 409 435–46
Liu H and Feng J 2012 Seasonal and interannual variations of evapotranspiration and energy exchange over different land surfaces in a semi-arid area of China J. Appl. Meteorol. Climatol. 51 1875–88
Liu Q, Wang T, Han Q, Sun S, Liu C-Q and Chen X 2019 Diagnosing environmental controls on actual evapotranspiration and evaporative fraction in a water-limited region from northwest China J. Hydrol. 578 124045
Loveland T, Reed B, Brown J, Ohlen D, Zhu Z, Yang I. and Merchant J 2001 Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data Int. J. Remote Sens. 21 1303–30
Maxwell R M and Condon L E 2016 Connections between groundwater flow and transpiration partitioning Science 353 377
Novick K A et al 2016 The increasing importance of atmospheric demand for ecosystem water and carbon fluxes Nat. Clim. Change 6 1–5
Oki T and Kanae S 2006 Global hydrological cycles and world water resources Science 313 1068
Pastorello G et al 2014 Observational data patterns for time series data quality assessment 2014 IEEE 10th Int. Conf. on e-Science 1 271–8
Renner M and Bernhofer C 2012 Applying simple water-energy balance frameworks to predict the climate sensitivity of streamflow over the continental United States Hydrol. Earth Syst. Sci. 16 2531–46
Rim C-S 2008 Estimating evapotranspiration from small watersheds using a water and energy balance approach Hydrol. Process. 22 703–14
Sayegh A, Tate J E and Ropkins K 2016 Understanding how roadside concentrations of NOx are influenced by the background levels, traffic density, and meteorological conditions using boosted regression trees Atmos. Environ. 127 163–75
Schenk H J and Jackson R B 2002 Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems J. Ecol. 90 480–94
Scott R L and Biederman J A 2019 Critical zone water balance over 13 years in a semiarid savanna Water Resour. Res. 55 574–88
Scott R L, Biederman J A, Hamerlynck E P and Barron-Gafford G A 2015 The carbon balance pivot point of southwestern U.S. semiarid ecosystems: insights from the 21st century drought J. Geophys. Res. 120 2612–24
Shangguan W, Dai Y, Duan Q, Liu B and Yuan H 2014 A global soil data set for earth system modeling J. Adv. Model. Earth Syst. 6 249–63
Shi T, Guan D, Wang A, Wu J, Jin C and Han S 2008 Comparison of three models to estimate evapotranspiration for a temperate mixed forest Hydrol. Process. 22 3431–43
Spinoni J, Vogt J, Naumann G, Carrao H and Barbosa P 2015 Towards identifying areas at climatological risk of desertification using the Köppen–Geiger classification and FAO aridity index Int. J. Climatol. 35 2210–22
Stoy P C et al 2013 A data driven analysis of energy balance closure across FLUXNET research sites: the role of landscape scale heterogeneity Agric. Forest Meteorol. 171–172 137–52
Tramhauer P et al 2014 Comparison of different evaporation estimates over the African continent Hydrol. Earth Syst. Sci. 18 193–212
Vivoni E R et al 2008 Observed relation between evapotranspiration and soil moisture in the North American monsoon region Geophys. Res. Lett. 35 L22403
Vivichard N and Papale D 2015 Filling the gaps in meteorological continuous data measured at FLUXNET sites with ERA-Interim reanalysis Earth Syst. Sci. Data 7 157–71
Wang K and Dickinson R E 2012 A review of global terrestrial evapotranspiration: observation, modeling, climatology, and climatic variability Rev. Geophys. 50 RG2005
Wang T, Istanbulluoglu E, Lenters J and Scott D 2009 On the role of groundwater and soil texture in the regional water balance: an investigation of the Nebraska Sand Hills, USA Water Resour. Res. 45 82–90
Wang T, Wedin D A, Franz T E and Hiller J 2015 Effect of vegetation on the temporal stability of soil moisture in grass-stabilized semi-arid sand dunes J. Hydrol. 521 447–59
Williams C A et al 2012 Climate and vegetation controls on the surface water balance: synthesis of evapotranspiration measured across a global network of flux towers Water Resour. Res. 48 W06523
Williams I N and Torn M S 2015 Vegetation controls on surface heat flux partitioning, and land-atmosphere coupling Geophys. Res. Lett. 42 9416–24
Wilson K B et al 2003 Diurnal centroid of ecosystem energy and carbon fluxes at FLUXNET sites J. Geophys. Res. 108 4664
Yue W, Wang T, Franz T E and Chen X 2016 Spatiotemporal patterns of water table fluctuations and evapotranspiration induced by riparian vegetation in a semiarid area Water Resour. Res. 52 1948–60
Zhang L, Dawes W R and Walker G R 2001 Response of mean annual evapotranspiration to vegetation changes at catchment scale Water Resour. Res. 37 701–8
Zhang L, Potter N, Hickel K, Zhang Y and Shao Q 2008 Water balance modeling over variable time scales based on the Budyko framework-Model development and testing J. Hydrol. 360 117–31
Zveryaev I I and Allan R P 2010 Summertime precipitation variability over Europe and its links to atmospheric dynamics and evaporation J. Geophys. Res. 115 D12102