An Enhanced SPEI Drought Monitoring Method Integrating Land Surface Characteristics

Liqing Peng[1,2,3*], Justin Sheffield[4], Zhongwang Wei[5], Michael Ek[6], Eric F. Wood[7†]

[1] Food Program, World Resources Institute, Washington DC, United States.
[2] Department of Geography, The University of Hong Kong, Hong Kong, China.
[3] Institute for Climate and Carbon Neutrality, The University of Hong Kong, Hong Kong, China.
[4] School of Geography and Environmental Science, University of Southampton, Southampton, United Kingdom.
[5] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Guangdong Province Key Laboratory

10 for Climate Change and Natural Disaster Studies, School of Atmospheric Sciences, Sun Yat-Sen University, Guangzhou, China.
[6] Joint Numerical Testbed, Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado, United States.
[7] Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, United States.

15 (†Deceased)

Correspondence to: Liqing Peng (pengliqing51@gmail.com)

20 **Abstract.** Atmospheric evaporative demand is a key metric for monitoring agricultural drought. The existing ways of estimating evaporative demand in drought indices do not faithfully represent the constraints of land surface characteristics and become less accurate over non-uniform land surfaces. This study proposes incorporating surface vegetation characteristics, such as vegetation dynamics data, aerodynamic and physiological parameters, into existing potential evapotranspiration (PET) methods. This approach is

25 implemented over the Continental United States (CONUS) for the period of 1981-2017 and is tested in a recently developed drought index the Standardized Precipitation Evapotranspiration Index (SPEI). We show that activating realistic maximum surface and aerodynamic conductance could improve prediction of soil moisture dynamics and drought impacts by 29-41% on average compared to the widely used simple methods, especially effective in the forests and humid regions, by 86-89%. Our approach only requires the minimum

30 amount of ancillary data, while permitting both historical reconstruction and real-time forecast of drought. This offers a physically meaningful, yet easy-to-implement way to account for the vegetation control in drought indices.

 1 Introduction

Drought is one of the most costly hydrological hazards (Wilhite, 2000; Ross & Lott, 2003; Piao et al., 2019), with devastating impacts on croplands and pastures (Kogan, 1995), forests ecosystems (Clark et al., 2016; Xu et al., 2022), electricity production, water quality, and soil fertility (Loon, 2015). Monitoring the changes in water availability is critical for providing early warnings of drought and for risk management (Wilhite, 40 Sivakumar, & Pulwarty, 2014). Many physical or probabilistic measures have been developed (Heim, 2002) to quantify drought, such as Palmer Drought Severity Index (PDSI, Palmer, 1965), Standardized Precipitation Index (SPI, McKee, Doesken, Kleist, & others, 1993), Vegetation Condition Index (VCI, Kogan, 1995), and multiple remote sensing drought indices (Zhang, Jiao, Zhang, Huang, & Tong, 2017; Yang et al., 2023).

Atmospheric evaporative demand (AED) is a key input to drought indices because it is a measure of water 45 demand, namely, how thirsty the atmosphere is (Peng, Li, & Sheffield, 2018). AED typically reflects the effect of temperature and humidity, and is considered a major driver of drought stress on vegetation and tree mortality (Williams et al., 2012; McDowell et al., 2018). Among the drought indices, the recently developed Standardized Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano, Beguería, & López-Moreno, 2010) factors in water demand (AED) in addition to water supply (precipitation). Compared to the SPI that 50 only considers precipitation, the SPEI is more suitable for quantifying the drought impacts on agriculture (Potop, 2011; Potop, Možný, & Soukup, 2012), and ecosystems (Vicente-Serrano et al., 2012; Vicente-Serrano et al., 2013; Barbeta, Ogaya, & Peñuelas, 2013). In addition, the SPEI is more flexible than the PDSI because it is not sensitive to soil water field capacity and can be implemented on various time scales (Vicente-Serrano, der Schrier, Beguería, Azorin-Molina, & Lopez-Moreno, 2015; Zhao et al., 2017). It has been widely 55 used for both drought reconstruction and monitoring (Paulo, Rosa, & Pereira, 2012; Beguería, Vicente-Serrano, Reig, & Latorre, 2013).

The way of estimating AED in drought indices has a significant impact on drought quantification (Sheffield, Wood, & Roderick, 2012; Trenberth et al., 2013; Yang, Roderick, Zhang, McVicar, & Donohue, 2018; Dewes et al., 2017). AED is approximated by potential evapotranspiration (PET), the maximum rate of 60 evapotranspiration when surface water supply is unlimited. Previous work has used various PET formulations for AED in the SPEI since it was first proposed in 2010 (Vicente-Serrano, Beguería, & López-Moreno, 2010; Beguería, Vicente-Serrano, Reig, & Latorre, 2013). These conventional PET methods do not factor in the effects of surface characteristics, which often assume no or simple universal vegetation control on transpiration (e.g., the Thornthwaite, Hargreaves-Samani, and Penman methods). Without vegetation control, 65 the maximum surface conductance is overestimated and the PET rate during the onset and retreat of the growing season is unrealistically high. Furthermore, by assuming an smooth reference surface, some methods

do not account for surface roughness, hence downplay aerodynamic conductance and suppress the PET estimate (Peng et al., 2019). Even though the reference evapotranspiration (ET_0) method (Allen, Pereira, Raes, & Smith, 1998) considers the biophysical limitation of transpiration by assigning a surface resistance under well-watered condition, it does not account for vegetation phenology (Lorenz, Davin, Lawrence, Stöckli, & Seneviratne, 2013) and assumes a fixed reference height and a constant surface resistance for all vegetation types. This approach is not physically meaningful for forests, where canopy height is relatively large and vegetation cover varies significantly. A recent study by Sun et al. (2023) highlighted the importance of incorporating surface properties especially vegetation control in PET and used a two source Shuttleworth-Wallace (SW) model designed and validated for sparse and fragmented vegetation surfaces. However, without further calibration and parameterization, the SW model's broader applicability beyond sparse vegetation is uncertain, and additionally it may increase data requirements and associated uncertainties (Gao et al., 2021; Abeysiriwardana et al., 2022).

We hypothesize that adding the surface vegetation characteristics to an existing drought quantification approach will improve the spatial and temporal accuracy of drought prediction. The goals of this study are to explore which surface features are the most useful for enhancing drought prediction, and which vegetation types benefit most from incorporating these features. We propose incorporating realistic vegetation restrictions into existing PET methods, while not increasing much cost and uncertainty caused by additional data sources and complex formulations. Then we use independent soil moisture observations (Dai, Trenberth, & Qian, 2004) from satellite to evaluate the drought depictions by various forms of PET approaches across different temporal scales. The evaluation against observed soil moisture allows the direct diagnosis of the most sensitive surface characteristics and the most effective approach for drought quantification (Vicente-Serrano et al., 2012).

In this study, we focus on the continental U.S. (CONUS) primarily because the drought events hitting this region have raised interest in variability, trends, and future risks of drought (Andreadis & Lettenmaier, 2006; Hobbins et al., 2012; Dewes et al., 2017). Several most severe droughts hit the western U.S. in the recent decade, including the 2012 Great Plains drought (Hoerling et al., 2014) and the 2012-2016 California drought (Dong et al., 2019). The western U.S. has been experiencing the most severe drought period after the 1930s and 1950s (Andreadis, Clark, Wood, Hamlet, & Lettenmaier, 2005), and its vulnerability to drought continued to grow (Andreadis & Lettenmaier, 2006). Besides, high-quality meteorological datasets are available over the CONUS (Daly et al., 2008; Xia et al., 2012) and can help reduce the uncertainty of drought prediction originating from input forcings.

2 Data

2.1 Meteorology

100 To calculate the SPEI, PET is estimated on daily scale over the period of 1981-2017 using high-quality daily meteorology data from PRISM (Parameter-elevation Regressions on Independent Slopes Model) that employs weather stations and digital elevation model (Daly, Neilson, & Phillips, 1994; Daly et al., 2008). We acquire daily precipitation, daily mean, maximum, minimum, and dew point temperature on a 4 km grid for the period of 1981-2017. Surface downward shortwave and net longwave radiation, pressure, and wind

105 speed are taken from the NLDAS-2 (North American Land Data Assimilation System phase 2 (Xia et al., 2012). All data are spatially restricted to the continental United States (25–50ºN, 67–125ºW) and regridded to the 0.125º NLDAS-2 grid using the first-order conservative remapping tool provided by Climate Data Operators (https://code.zmaw.de/projects/cdo).

2.2 Soil moisture

110 The European Space Agency Climate Change Initiative (ESA CCI) v4.3 surface soil moisture (SMsurf) is used to evaluate the drought severity quantified by the SPEI time series (https://www.esa-soilmoisture-cci.org/). This dataset combines several active and passive microwave soil moisture products into a harmonized surface layer soil moisture (2-5 cm) in $m^3 m^{-3}$ (Liu et al., 2012; Gruber et al., 2017). The dataset is chosen for its enhanced data reliability by integrating multiple single-sensor active and passive microwave

115 soil moisture products to minimize uncertainty (Gruber et al., 2019). The version 4.3 provides soil moisture on a 0.25º grid at daily time step for the 1979-2017 period and has been widely used in ET and drought studies (Dorigo et al., 2017; Martens et al., 2017).

2.3 Land surface ancillary data

The land surface data used for deriving biophysical parameters include gridded land cover type, leaf area

120 index, and surface albedo. The land cover type is provided by the 0.5 km MODIS-based Global Land Cover Climatology during the 2001-2010 period (Broxton, Zeng, Sulla-Menashe, & Troch, 2014, https://archive.usgs.gov/archive/sites/landcover.usgs.gov/global_climatology.html). This dataset has 17 land cover classes based on the International Geosphere‐Biosphere Program (IGBP) classification. This land cover climatology dataset is displayed in Fig. 1.

125 The monthly climatology of leaf area index is obtained from the 15-day, 1 km AVHRR GIMMS LAI3g product that covers the period of 1982-2016 (Zhu et al., 2013).

The monthly climatology of surface albedo is derived from the 8-day, 0.05° GLASS (Global Land Surface Satellite) albedo product. This GLASS02A05/06 product combines MODIS and AVHRR (Advanced Very-High-Resolution Radiometer) to provide a gap-filled land surface shortwave black-sky and white-sky albedo (Qu et al., 2014; Liu et al., 2013) that covers the period of 1982-2012. We resample the 8-day albedo to a daily resolution and obtain daily albedo by averaging the black- and white-sky albedos. Missing data are gap-filled using the average of adjacent years.

This study uses the newly developed 10-m global canopy height dataset that merges the Global Ecosystem Dynamics Investigation (GEDI) space-borne LiDAR height data with Sentinel-2 satellite data (Lang et al., 2023). The original 10-m resolution was remapped to 0.125° using the average. Additionally, this study uses a global tree height dataset at 1-km for 2005 using spaceborne lidar (Simard et al., 2011) for complementary analysis in the forests (Appendix B).

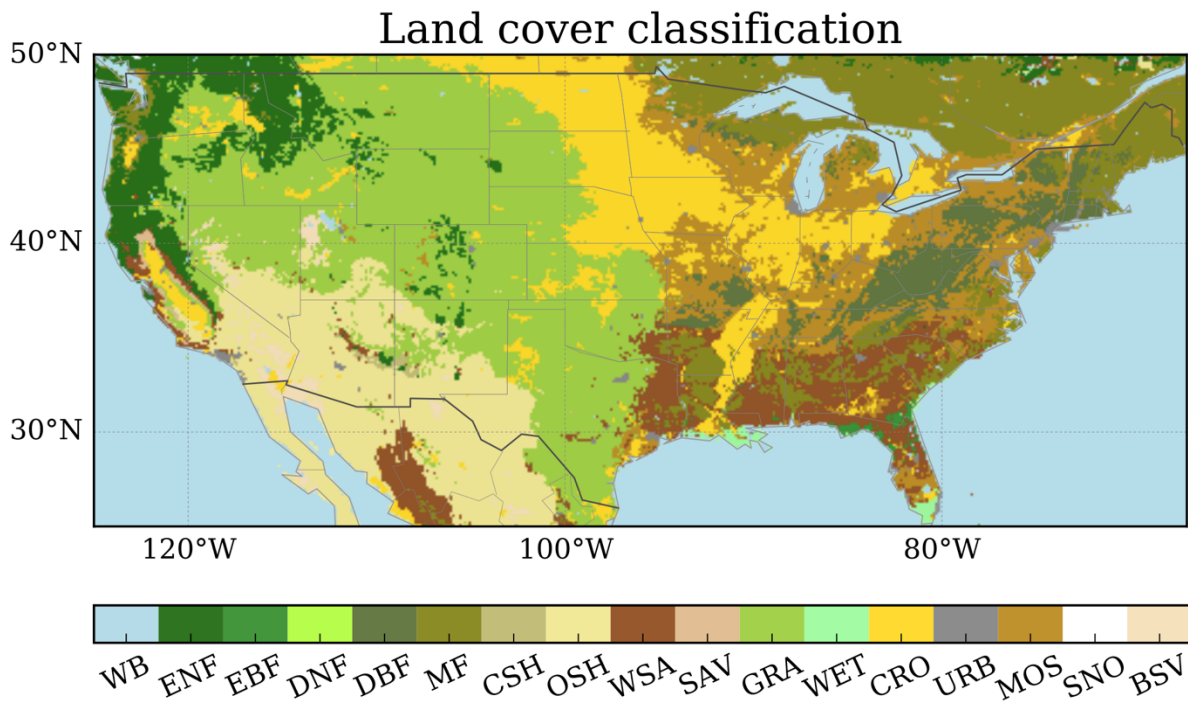



Figure 1. The land cover classification over the Continental United States used for surface vegetation parameter inference. The classification is based on the satellite retrieval of land cover climatology during 2001-2010 (see Table 1 for a list of land cover full names).

3 PET methods

3.1 Current PET methods

PET can be estimated from univariate empirical models such as temperature-based methods (Thornthwaite, 1948) and physically-based models. Empirically based methods can induce large uncertainty in the drought projection (Sheffield et al., 2012; Feng, Trnka, Hayes, & Zhang, 2017) and are therefore not considered in the study. Physically-based methods can account for multiple input variables such as surface net radiation, near-surface temperature, wind speed, or specific humidity. The Penman equation (Penman, 1948) is the most comprehensive physically-based method to estimate PET by combining the radiative and aerodynamic components:

$$PET_{Penman} = \frac{\Delta(R_n - G) + \rho_a C_p DGa}{\lambda(\Delta + \gamma)} \tag{1}$$

where PET is expressed as water mass fluxes (kg m^{-2} s^{-1}), R_n is the surface net radiation (W m^{-2}), G is the surface ground heat flux (W m^{-2}), Δ is the slope of the saturation vapor pressure curve at the temperature of interest (Pa K^{-1}), γ is the psychrometric constant (Pa K^{-1}), λ is the latent heat of vaporization (J kg^{-1}), ρ_a is the air density (kg m^{-3}), C_p is the specific heat of air (J kg^{-1} K^{-1}), D is the vapor pressure deficit (VPD, Pa), and Ga is the aerodynamic conductance (m s^{-1}). The variants of the Penman equation have been widely used to estimate PET in hydrological and land surface modeling (Sellers et al., 1996; Liang et al., 1994; Ek et al., 2003; Peng, Li, & Sheffield, 2018; Peng et al., 2019; Yang et al., 2019).

The open-water Penman (OW) equation is a simplified Penman equation to calculate PET over an open water surface, re-parameterized by Shuttleworth (1993):

$$PET_{OW} = \frac{\Delta}{(\Delta + \gamma)} \frac{(R_n - G)}{\lambda} + \frac{\gamma}{\Delta + \gamma} \frac{6.43(1 + 0.536u_2)D}{\lambda} \tag{2}$$

where PET_{OW} is typically in mm d^{-1} (kg m^{-2} s^{-1} = 86400 mm d^{-1}), $(R_n - G)$ is daily available energy (J m^{-2} d^{-1}), u_2 is the wind speed at 2-m height (m s^{-1}), λ is J kg^{-1}, and D is in kPa. Note that the OW equation provides daily estimates, and therefore some of the variables have different units compared to those in Equation 1.

The Priestley-Taylor (PT) equation is also a simplified form of the Penman equation, which describes evaporation from a well-watered surface based on the equilibrium evaporation under conditions of minimal advection (Priestley & Taylor, 1972):

$$PET_{PT} = 1.26 \frac{\Delta(R_n - G)}{\lambda(\Delta + \gamma)} \tag{3}$$

where PET_{PT} is in mm d^{-1} and $(R_n - G)$ is in J m^{-2} d^{-1}.

The Penman-Monteith (PM) equation (Monteith, 1965) is an extended version of the Penman equation to estimate actual ET (kg m^{-2} s^{-1}), which introduces the surface conductance (Gs, m s^{-1}):

$$PET_{PM} = \frac{\Delta(R_n - G) + \rho_a C_p D Ga}{\lambda(\Delta + \gamma\left(1 + \frac{Ga}{Gs}\right))}$$ (4)

The reference crop evapotranspiration (PET_{RC}) recommended by the UN Food and Agricultural Organization (FAO) is a specific application of the Penman-Monteith equation (Allen, Pereira, Raes, & Smith, 1998). It is designed for calculating the maximum ET of reference crop under well-watered condition. The general formula is given by Allen et al. (2005):

$$PET_{RC} = \frac{0.408\Delta(R_n - G) + \frac{C_n u_2}{T_a + 273}\gamma D}{\Delta + \gamma(1 + C_d u_2)}$$ (5)

where PET_{RC} is also in mm d^{-1}, $(R_n - G)$ is daily available energy (MJ m^{-2} d^{-1}), Δ and γ are in kPa ^oC $^{-1}$, T_a is the air temperature at 2-m height (^oC), D is in kPa, C_n (K mm s^3 Mg^{-1} d^{-1}) is a constant describing the effect of aerodynamic conductance (Ga) that increases with canopy height. The denominator $\Delta + \gamma(1 + C_d u_2)$ is a special form of the denominator of the Penman-Monteith equation $\Delta + \gamma(1 + Rs/Ra)$. C_d ($\frac{Rs}{Ra\,u_2}$, s m^{-1}) is a constant that increases with the ratio of surface resistance ($Rs = 1/Gs$) to aerodynamic resistance ($Ra = 1/Ga$). There are two sets of C_n and C_d, tall crop (C_n=1600, C_d=0.38) and short crop (C_n=900, C_d=0.34). The FAO short crop equation is used in the recent version of the SPEI calculation (Beguería, Vicente-Serrano, Reig, & Latorre, 2013).

The above-mentioned equations treat the surface vegetation as a "big leaf" by considering the canopy resistance and soil resistance together as the bulk surface resistance, and therefore require fewer parameters and less computational costs. One challenge of the big-leaf assumption is to infer bulk surface resistance from canopy resistance when the surface is not fully covered by vegetation (Leuning et al., 2008). Additionally, we compare the big leaf models with the Shuttle-Wallace (SW) two source model (Shuttleworth and Wallace, 1985; Sun et al., 2023), incorporating vegetation cover and separating ET into the sum of transpiration and soil evaporation:

$$PET_{SW} = C_c PET_{PMc} + C_s PET_{PMs}$$ (6)

where the formulas and parameterizations of PET_{PMc}, PET_{PMs}, C_c, and C_s are given in the Appendix A.

3.2 Surface characteristics formulas

Classical PET definitions rely on surface meteorology and do not faithfully represent the vegetation conditions and biophysical constraints and become less accurate over non-uniform land surfaces (Moran et al., 1996). This section introduces the major options of formulas for aerodynamic conductance and surface conductance.

3.2.1 Aerodynamic conductance

Aerodynamic conductance Ga in the OW and PT methods (Equations 2 and 5) are implicitly derived from a smooth surface with low roughness length, which can underestimate the Ga and PET values in the forests (Peng et al., 2019). Open water aerodynamic conductance Ga_{OW} can be obtained by inverting the open water Penman equation (Equation 2) to match the Penman equation (Equation 1), given by Peng et al. (2019):

$$Ga_{OW} = \frac{6.43(1 + 0.536u_2) \cdot P_s}{86.4\epsilon\lambda\rho_a} \tag{6}$$

where u_2 is converted from wind speed at 10-m to 2-m height following the wind profile relationship in Allen, Pereira, Raes, & Smith (1998). P_s is near-surface atmospheric pressure (Pa), ϵ is the ratio of molecular weight of water to dry air (= 0.622).

Short and tall reference crop aerodynamic conductance $Ga_{RC-short}$ and $Ga_{RC-tall}$ are given by

$$Ga_{RC-short} = \frac{u_2}{208} \tag{7}$$

$$Ga_{RC-tall} = \frac{u_2}{110} \tag{8}$$

where u_2 is converted from wind speed at 10-m to 2-m height (m s^{-1}).

Instead of the low Ga in OW and the fixed Ga in RC, it is better to generate more realistic surface roughness varying by land cover type, hereafter called Ga_{LC} (Brutsaert & Stricker, 1979; Allen, Pereira, Raes, & Smith, 1998; Shuttleworth, 1993):

$$Ga_{LC} = \frac{k^2 u_z}{\ln\left(\frac{z_m - d_0}{z_{0m}}\right)\ln\left(\frac{z_h - d_0}{z_{0h}}\right)} \tag{9}$$

where z_m is the measurement height (m) for wind speed, z_h is the measurement height (m) for temperature and humidity, u_z is the wind speed at measurement height (m s^{-1}), k is the von Karman constant, d_0 is the zero-plane displacement height (m), z_{0m} and z_{0h} are the roughness lengths for momentum and heat (m). d_0 and z_{0m} can be estimated from canopy height (h) following $d_0 = 2h/3$ and $z_{0m} = h/8$ (Brutsaert, 1982). h is

based on the typical value for each land cover type. When estimating z_{0h}, instead of assuming $z_{0h} = 0.1z_{0m}$ as in Allen, Pereira, Raes, & Smith (1998), it is common to introduce a concept of excess resistance (Verma, 1989) and characterize the relationship between z_{0h} and z_{0m}:

$$z_{0h} = \frac{z_{0m}}{\exp(kB^{-1})} \tag{10}$$

The $\ln(z_{0m}/z_{0h})$ term, also known as kB^{-1}, depends on the roughness Reynold's number $Re*$ or frictional velocity ($u*$), LAI (Yang & Friedl, 2003), and land cover type (Rigden, Li, & Salvucci, 2018).

On top of the above land cover based roughness (Equation 9), it is possible to further incorporate realistic canopy height (h) to account for its effect on wind speed, hereafter called Ga_{CH}:

$$Ga_{CH} = \frac{k^2 u_r}{\ln\left(\frac{z_r - d_0}{z_{0m}}\right)\ln\left(\frac{z_r - d_0}{z_{0h}}\right)} \tag{11}$$

Similar to Equation 9, d_0 and z_{0m} are estimated from canopy height (h) following $d_0 = 2h/3$ and $z_{0m} = h/8$ (Brutsaert, 1982), and z_{0h} is estimated using Equation 10; the only difference is that this CH approach use actual canopy height data instead of look up table. Different from Equation 9, this approach assumes a reference level, where z_r is the reference height (2m above canopy height) for wind speed, temperature, and humidity (Zhou et al., 2006). The reference height wind speed u_r (m s^{-1}) is converted from the measured wind speed u_z following the wind profile relationship. The internal boundary layer (z_b) on top of the measurement height ($z = 10$m) and canopy reference height are matched (Zhou et al., 2006; Brutsaert, 1982; Federer et al., 1996):

$$u_r = u_z \frac{\ln\left(\frac{z_b}{z_{0g}}\right)}{\ln\left(\frac{z_b}{z_{0m}}\right)} \frac{\ln\left(\frac{z_r - d_0}{z_{0m}}\right)}{\ln\left(\frac{z}{z_{0g}}\right)} \tag{12}$$

where ground roughness length z_{0g} is 0.005 m, z_r is the reference height at ($h+2$) m, z is the measurement height at 10 m. The internal boundary layer height z_b is estimated at about 4.4 m:

$$z_b = 0.334F^{0.875}z_{0g}^{0.125} \tag{13}$$

where F is the fetch at 5000 m, the effective distance over which the wind blows do not change in direction.

For the SW method, the two aerodynamic resistances are given by Eq. A11-17 (Appendix A).

3.2.2 Surface conductance

In previous PET methods, surface conductance is either not considered or assumed to be constant across vegetation types and over time. LAI plays a dominant role in determining the canopy-atmosphere coupling and ET partitioning (Peng et al, 2019; Wei et al., 2017; Forzieri et al., 2020). The OW and PT approach does not consider the role of LAI. The FAO approach uses a constant LAI throughout the growing season. Here we adopt a widely used method in estimating actual ET and assume a well-watered condition. The maximum surface conductance Gs_{max} can be obtained by scaling the maximum stomatal conductance (Gst_{max}) with LAI (Yan et al., 2012):

$$Gs_{max} = Gst_{max} \cdot LAI \tag{14}$$

An alternative formula for Gs_{max} is from Zhou et al. (2006):

$$Gs_{max} = \frac{LAI_e}{Rst_{min}} \tag{15}$$

where LAI_e is the effective LAI, which is equal to LAI/2 when LAI is greater than 4. We introduce two options to incorporate an average LAI or the seasonal cycle of LAI into the surface conductance.

3.3 Parameterizations of surface characteristics

For Eq. 9, given that NLDAS-2 provides wind speed at a 10 m level, we used a measurement height = 10 m for both wind speed and temperature because the variation in the vertical temperature profile (2-10 m) is negligible compared to that of wind speed. For z_{0m}, we apply the typical values based on median canopy height for different land cover types, and estimated d_0 from z_{0m} ($d_0 \approx 16z_{0m}/3$). We use a simple look-up table approach to provide parameters based on land cover type (Fig. 1), summarized in Table 1.

For kB^{-1}, we adopt estimates from a collection of literature as below. The forests generally have lower kB^{-1} values ($kB^{-1} = 1$ for needleleaf or mixed forest, $kB^{-1} = 0.5$ for broadleaf) than shrublands ($kB^{-1} = 3.75$) and croplands ($kB^{-1} = 1.75$), based on the values of Rigden et al. (2018) for the medium emissivity case ($\epsilon = 0.96$). For grasslands, $kB^{-1} = 2.25$ is computed as the average of short grass ($kB^{-1} = 2.0$) and medium-length grass ($kB^{-1} = 2.5$), based on Brutsaert (1982). For barren or bare soil, we estimate $kB^{-1} = 3$ by taking the average of all observed kB^{-1} in Yang et al. (2008). Nadeau et al. (2009) suggested $kB^{-1} = 6$ for the urban area. For water body, wetlands, and snow, we adopt the widely-used $kB^{-1} = 2$, as Zilitinkevich et al. (2001) showed that kB^{-1} over the water surface is within the 0-4 range. There are large variations in the observed kB^{-1} for savannas. Troufleau et al. (1997) reported $kB^{-1} = 7.9$ for fallow savanna; Kustas et al.

(1989) provided a range of 1 to 11; Stewart et al. (1994) found an average value of $kB^{-1} = 5.8$, similar to the study by Lhomme et al. (1997) that reported $kB^{-1} = 5.9$ for Sahelian vegetation; Verhoef et al. (1997) suggested a high value of $kB^{-1} = 12.4$. We choose $kB^{-1} = 7$ as most of these observed values fall into the range of 6-8. z_{0h} is then estimated based on land cover specific z_{0m} and kB^{-1} (Eq. 10).

Table 1. Ga and Gs parameters by IGBP land cover*.

ID	Code	Name	z_{0m} (m)	d_0 (m)	kB^{-1}	Gst_{max} [j] (mm s^{-1})	Rst_{min} [k] (s m^{-1})
0	WB	Water body	0.0004 [a]	0.002	2.0 [e]	NA	NA
1	ENF	Evergreen needleleaf	1.1 [b]	5.9	1.0 [f]	9.3	150
2	EBF	Evergreen broadleaf	1.1 [b]	5.9	0.5 [f]	9.3	150
3	DNF	Deciduous needleleaf	0.9 [b]	4.8	1.0 [f]	9.3	150
4	DBF	Deciduous broadleaf	0.9 [b]	4.8	0.5 [f]	9.3	150
5	MF	Mixed forest	0.9 [b]	4.8	1.0 [f]	9.3	150
6	CSH	Closed shrublands	0.2 [a]	1.1	3.75 [f]	9.3	150
7	OSH	Open shrublands	0.2 [a]	1.1	3.75 [f]	9.3	100
8	WSA	Woody savannas	0.4 [a]	2.1	7.0 [g]	9.3	180
9	SAV	Savannas	0.4 [a]	2.1	7.0 [g]	9.3	120
10	GRA	Grasslands	0.05 [a]	0.27	2.25 [a]	12	115
11	WET	Permanent wetlands	0.04 [c]	0.21	2.0 [e]	12	65
12	CRO	Croplands	0.12 [d]	0.64	1.75 [f]	12.2	90
13	URB	Urban and built up	1.1 [b]	5.9	6.0 [h]	NA	NA
14	MOS	Cropland/vegetation	0.12 [d]	0.64	1.75 [f]	12.2	120
15	SNO	Snow/ice	0.00001 [a]	5.3E-05	2.0 [e]	NA	NA
16	BSV	Barren	0.01 [d]	0.053	3.0 [i]	NA	NA

*The above estimates are collected from [a]Brutsaert (1982), [b]Campbell and Norman (1998), [c]Acreman et al. (2003), [d]Monteith and Unsworth (2013), [e]Zilitinkevich et al. (2001), [f]Rigden et al. (2018), [g]Kustas et al. (1989), Stewart et al. (1994), Troufleau et al. (1997), Lhomme et al. (1997), and Verhoef et al. (1997), [h]Nadeau et al. (2009), [i]Yang et al. (2008), [j]Kelliher et al. (1995), [k]Zhou et al. (2006).

Canopy height (h) is a key parameter in determining aerodynamic conductance and is eventually used to obtain d_0 and z_{0m} for Eq.9. The OW and FAO methods generally assume it to be constant across vegetation types and temporal scales. To address this limitation, we evaluate two methods for canopy height parameterization.

The first method uses literature values and is adopted in the Land Cover approach (LC, Eq.9). For most of the land cover types (ID 6-16), we applied the values from the look up table except for the forests, where we

determined canopy height by calculating the median height within each land cover from the tree height lidar data (Simard et al., 2011).

The second more comprehensive method is adopted in the Canopy Height approach (CH, Eq. 11) and the SW two source model (Appendix A, Eq. A9-10). It takes into account three factors: land cover type, measured canopy height, and dynamic LAI. We overlaid the land cover map (Fig. 1) and the canopy/tree height data (Lang et al., 2023; Simard et al., 2011) to obtain the distribution in each land cover type (Appendix B). Based on the distribution of the two datasets, land cover definition, and literature ranges, we estimated the minimum canopy height (h_{min}) and maximum canopy height (h_{max}) by land cover type (Table 2). As for quality control, we set the outlier (smaller than h_{min} or greater than h_{max}) to a typical value of canopy height given land cover type (h_{typ}, obtained through the mode of the distribution). The actual canopy height is then determined by assuming a linear relationship with dynamic LAI following Zhou et al. (2006).

$$h = h_{min} + \frac{(h_{max} - h_{min})LAI}{LAI_{max}} \tag{13}$$

where LAI_{max} represents the annual maximum value at the grid cell level, obtained from the satellite data. Note that h is set to zero if LAI_{max} is zero.

Table 2. Canopy height parameters by IGBP land cover*.

ID	Code	Name	h_{min} (m)	h_{max} (m)	h_{typ} (m)
0	WB	Water body	0.001	0.02	0.01
1	ENF	Evergreen needleleaf	2 [a]	48 [b]	13 [b]
2	EBF	Evergreen broadleaf	2 [a]	45 [b]	17 [b]
3	DNF	Deciduous needleleaf	7 [b]	23 [b]	17 [b]
4	DBF	Deciduous broadleaf	6 [b]	37 [c]	9.5 [c]
5	MF	Mixed forest	2 [a]	32 [b]	25 [b]
6	CSH	Closed shrublands	1 [b]	39 [b]	14.9 [b]
7	OSH	Open shrublands	2 [b]	17 [b]	6 [b]
8	WSA	Woody savannas	1 [b]	23 [b]	1 [b]
9	SAV	Savannas	1 [b]	26 [b]	17.7 [b]
10	GRA	Grasslands	0.1	3	1.5
11	WET	Permanent wetlands	0.1	5	0.5
12	CRO	Croplands	0.1	5	1
13	URB	Urban and built up	2	50	13
14	MOS	Cropland/vegetation	0.1	21	12
15	SNO	Snow/ice	0.001	0.02	0.01
16	BSV	Barren	0.01	0.1	0.05

*The above estimates are collected from [a]IGBP classification, [b]Lang et al. (2023), [c]Simard et al. (2023), otherwise based on authors' best estimates.

To calculate surface conductance in Eq. 11-12, we provide two set of parameterizations based on land cover type. The first set is derived from the findings of Kelliher, Leuning, Raupach, & Schulze (1995). For Gst_{max}, the measured values are ranging from 9 mm/s for natural vegetation to 12 mm/s for crops, as detailed in Table 1. They also found that Gs_{max} estimates are at most three times of the Gst_{max} estimates, therefore we set a maximum limit for $LAI = 4$. The second set uses the minimum stomatal resistance Rst_{min}, following Zhou et al. (2006), also listed in Table 1.

Current PET methods generally apply a uniform grass albedo value of 0.23 regardless of the underlying land cover type (Allen, Pereira, Raes, & Smith, 1998). To improve upon this assumption, we also introduce an option of introducing seasonal albedo cycle from satellite observations to both align albedo with specific land cover type and reflect temporal variations accurately.

4 Evaluation of the PET methods and parameterizations

4.1 Drought quantification: SPEI vs. soil moisture

Given the substantial divergence in the PET magnitudes among different models (Peng et al., 2019), a direct comparison of the absolute values among methods is not meaningful. However, the performance in representing drought between PET methods should be comparable. We hypothesize that incorporating the parameters or model structures in Sect. 3 into the existing methods will increase the accuracy of drought quantification.

We integrate the PET methods into the SPEI drought index across 1-, 3-, 6-, and 12-month time scales over the CONUS for the period of 1981-2017. The SPEI is based on the climatological water balance (water supply – atmospheric evaporative demand) cumulated over multiple time scales (e.g., 1, 3, 6, 12 months) following a similar procedure as in the SPI computation (Vicente-Serrano, Beguería, & López-Moreno, 2010). The accumulated water balances are fit using the log-logistic distribution and the probability distribution function is normalized to a standardized variable with mean = 0 and standard deviation = 1, termed as 1-, 3-, 6-, 12-month SPEI, respectively. We calculate the monthly SPEI with the SPEI R package (https://cran.r-project.org/web/packages/SPEI/) using daily meteorological data. We choose the SPEI driven by zero PET as a control scenario to showcase the net effect of introducing existing PET methods into traditional SPI drought index. We choose the SPEI driven by the Open Water (OW) method as the reference method, because the OW approach is the simplest scenario with minimal surface characteristics.

Soil moisture is a direct measure of drought severity. Therefore, we use the correlation between SPEI and soil moisture observations to quantify the skill of PET methods. We aggregated the daily ESA CCI surface soil moisture (SMsurf, m^3 m^{-3}) to monthly averages between 1981-2017 over the CONUS. To match the SPEI on multiple time scales, we calculated the moving average of SMsurf for 1, 3, 6, 12-month periods,
325 respectively. Our analysis focuses on the growing season (April-September), because PET is close to zero during the cold season (not shown). Given the monthly SPEI and SMsurf series during the growing season, Pearson correlation coefficient (R) is calculated for each pair of the SPEI and SMsurf monthly series in each grid cell of the CONUS on the time scale of 1, 3, 6, and 12 months. Then we calculate the change of correlation for each method from the control scenario or the reference. This change can identify whether a
330 PET method causes an improvement in drought quantification relative to the reference approach.

4.2 Initial examination of surface characteristics

We conducted a pilot analysis to identify the relative importance of different surface characteristics. To test the hypotheses, we use the PM algorithm for big leaf methods, so that we can easily control a specific set of parameters that represents a process option. Each of the above processes are regarded as different options:
335 (i) using active surface roughness or open water surface, (ii) seasonally varying or fixed surface conductance, and (iii) seasonally varying or fixed surface albedo. Table 3 provides the PET methods and the parameters in the preliminary analysis. We selected four existing PET methods and seven testing methods. The first set of methods (a, e, i, j) are the existing physically-based PET approaches: the open-water Penman equation (OW), the Priestley-Taylor equation (PT), the FAO reference crop evapotranspiration for short and tall crop.

340 First, in methods (b-d), the aerodynamic conductance module is not active as we set Ga to the open water Ga_{OW}, indicating a smooth surface with low roughness (Eq. 6). In methods (f-h), we activate the aerodynamic conductance using land cover based surface roughness as Ga_{LC} (Eq. 9). In methods (k-l), we activate the aerodynamic conductance using the formula of reference short crop (Eq.7). Second, in methods (b), (g), (k), the surface conductance parameter is unconstrained as we set Gst_{max} to infinity (Gs_{OW}). In methods (c), (d),
345 (g), (h), (l), we activate the surface conductance using seasonal LAI dynamics and Gst_{max} from Kelliher et al. (1995). Lastly, in methods (c), (g), (k), (l), the albedo parameter is not active as we set α to a constant (RC: 0.23 for grass, OW: 0.08 for water). In methods (b), (d), (f), (h), we activate the albedo parameter using seasonal albedo dynamics (α_{CLM}).

Table 3. Summary of the PET methods for initial assessment with their ID, name and abbreviation
350 **code, and details about surface characteristics.**

ID	Method (Code)	Ga		Gs_{max}			Albedo (α)	
		Open Water	Rough surface	Infinite	Constant	Seasonal	Constant	Seasonal
a	Open Water (OW)	X		X			X	
e	Priestley-Taylor (PT)						X	
i	FAO Short reference crop (RC-short)		X		X		X	
j	FAO Tall reference crop (RC-tall)		X		X		X	
b	Ga OW|Gs OW|α CLM	X		X				X
c	Ga OW|Gs LAI|α OW	X				X	X	
d	Ga OW|Gs LAI|α CLM	X				X		X
f	Ga LC|Gs OW|α CLM		X	X				X
g	Ga LC|Gs LAI|α OW		X			X	X	
h	Ga LC|Gs LAI|α CLM		X			X		X
k	Ga RC|Gs OW|α RC		X	X			X	
l	Ga RC|Gs LAI|α RC		X			X	X	

*Note that many methods in these experiments are unrealistic due to the inconsistencies of the surface conditions. Our attention is to include as many combinations as possible for a preliminary analysis.

In Sect. 5.1, we compare the CONUS averaged R values between the pairs of PET methods that share the same surface characteristics except for one of the features (see Fig. 3). The first feature surface roughness is determined by the way Ga is estimated. We compare the parameter set between rough and the open water surface by calculating the differences (Rough - Open Water) for the following pairs of experiments including (f) Ga LC|Gs OW|α CLM – (b) Ga OW|Gs OW|α CLM, (g) Ga LC|Gs LAI|α OW – (c) Ga OW|Gs LAI|α OW, and (h) Ga LC|Gs LAI|α CLM – (d) Ga OW|Gs LAI|α CLM. In terms of the surface conductance, we calculate the differences between seasonal and infinite Gs_{max} (Seasonal - Infinite) for the following pairs of experiments: (c) Ga OW|Gs LAI|α OW – (a) OW, (d) Ga OW|Gs LAI|α CLM – (b) Ga OW|Gs OW|α CLM, and (h) Ga LC|Gs LAI|α CLM – (f) Ga LC|Gs OW|α CLM. We also compare the differences between all consistent and inconsistent surfaces and between seasonal and constant albedo.

4.3 Comparison of PET parameterizations

Based on the results of Sect. 5.1, we further examine different parameterizations for Ga and Gs in order to identify optimal PET algorithms (Table 4, results in Sect. 5.2). We establish a control scenario where PET is not considered at all in the SPEI, equivalent to the traditional SPI. The PET methods under consideration have two categories, the big leaf model and the two source model. The big leaf model include three traditional methods (Open Water [OW], Reference Crop for short [RC-short], and tall [RC-tall] crops), land cover dependent (LC), and canopy height dependent (CH) methods. The LC method uses the same aerodynamic conductance method (Eq. 9) but differ in their surface conductance parameterizations: LC-K, which adopts

Gst_{max} from Kelliher, Leuning, Raupach, & Schulze (1995), and LC-Z, which uses Rst_{min} from Zhou (2006). The CH method also has the two parameterizations: CH-K and CH-Z. We then calculated ΔR between each PET method and the control scenario (set PET to zero).

 Table 4. Summary of PET methods with their formula and parameterization

type	Model code (equation)	Ga		Gs	
		formula	parameter	formula	parameter
	OW (Eq. 2)	Eq. 6	Peng et al. 2019	-	-
	PT (Eq. 3)	-	-	-	-
	RC-short (Eq. 5)	Eq. 7	Allen et al. 2005	Eq. 5	Allen et al. 2005
Big leaf	RC-tall (Eq. 5)	Eq. 8	Allen et al. 2005	Eq. 5	Allen et al. 2005
	LC-K (Eq. 4)	Eq. 9-10	z_{0m} (m), d_0 (m), kB^{-1} (Table 1); Tree height (Simard et al. 2011)	Eq. 14	Kelliher et al. 1995
	LC-Z (Eq. 4)			Eq. 15	Zhou et al. 2006
	CH-K (Eq. 4)	Eq. 11-13	Canopy height data (Lang et al. 2023)	Eq. 14	Kelliher et al. 1995
	CH-Z (Eq. 4)			Eq. 15	Zhou et al. 2006
Two source	SW (Eq. A1-5)	Eq. A6	Zhou et al. 2006; Canopy height data (Lang et al. 2023)	Eq. A7-8	Zhou et al. 2006

5 Results

5.1 Initial assessment of surface characteristics

We conducted a preliminary analysis to identify the relative importance of different surface characteristics. We examine eight algorithms to isolate the effects of surface characteristics on PET (Table 2). Fig. 2 displays the spatial patterns of growing season averages of these methods. For the classical Penman/Penman-Monteith methods (Fig. 2a, i, j), the highest mean growing season AED values are found in southern California, Arizona, and Texas, while the PT method (Fig. 2e) predicts the largest AED values in Texas and Florida. The spatial patterns of PET based on the rough surface (Fig. 2f-h, Rough Ga) are very different from those methods that assume a universal reference height (Fig. 2i-j, reference crop) or open water surface (Fig. 2a, b-d). Specifically, the regions which exhibit large PET estimates (> 250 mm/mon, Fig. 2h-k) are forests, such as ENF in the Pacific Northwest, DBF in the Northeast, and MF in the southeastern U.S.. Interestingly, although the methods using constant albedo (α=0.08) have generally larger AED values than those using seasonal albedo, the differences in the spatial pattern between the two are almost negligible (Fig. 2c vs. d, g vs. h). The combination of the rough aerodynamic and unconstrained surface conductance, represented by (f), produces extremely high monthly PET values with means at 330 to 340 mm/mon. The remaining methods also predict a wide range of mean monthly totals.

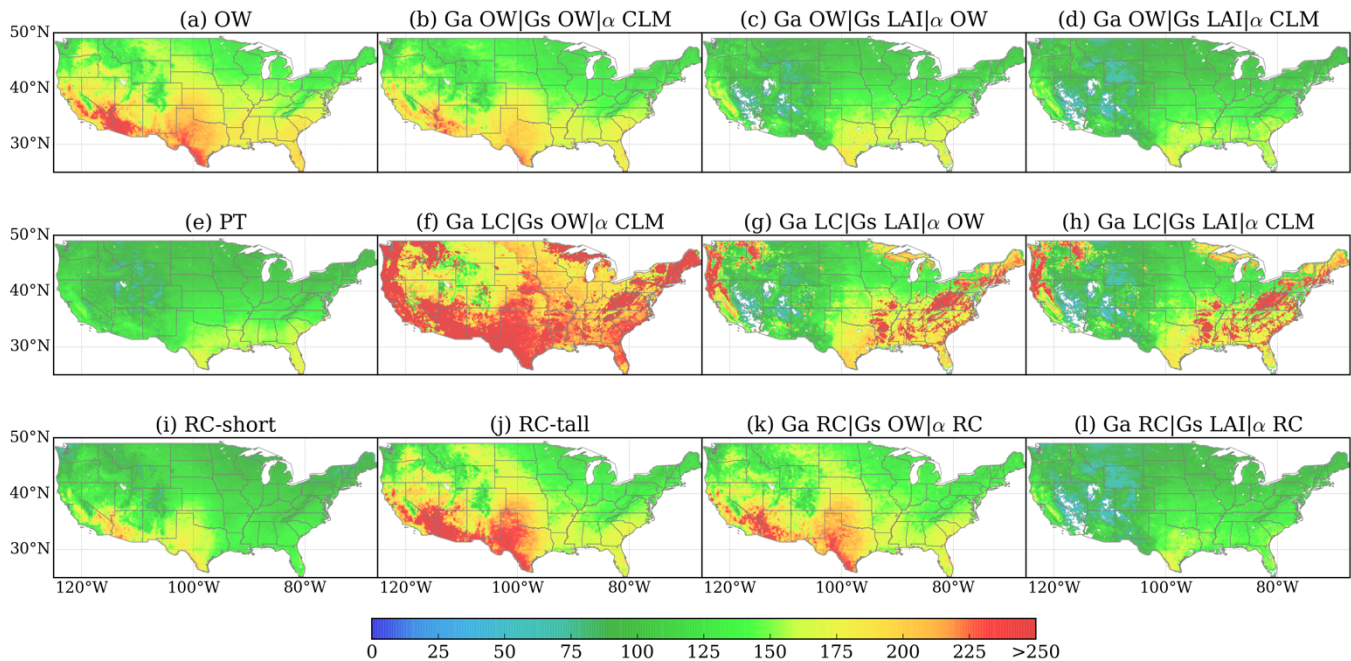



Figure 2. Growing season averages of AED derived from four PET methods and eight testing algorithms over the CONUS. Details and ID for each method are listed in Table 2.

400   Assessing the change between pairs of the above methods can identify whether adding/removing a surface feature eventually causes an improvement in drought quantification (Fig. 3). For 1-month time scale, surface roughness stands out to be the most important feature for enhancing the skill of drought index (ΔR = 0.01-0.025). For 6-month time scale, interestingly, activating realistic surface roughness does not necessarily increase the correlation with SMsurf, while activating dynamic surface conductance improves the correlation,

405   meaning that adding the plant phenology driven by LAI can improve the seasonal variations of drought index over longer time scale. We compare ΔR of five pairs with an inconsistent surface (e.g., a combination of open water Ga and seasonal Gc) subtracting from a consistent surface and find that methods with consistent surface features have persistently higher correlations with SMsurf (ΔR = 0 - 0.02). Given the steadily better performance, we only focus on the consistent surface approaches in subsequent sections. Surprisingly,

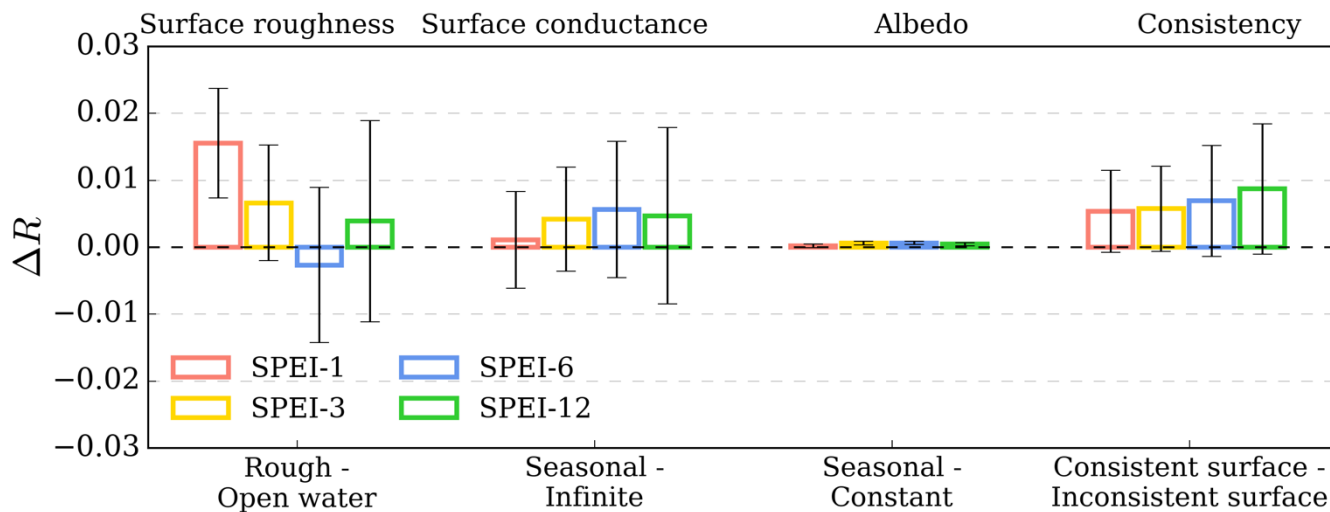
410   seasonal and constant albedo showed no significant difference on the correlations, possibly because of the little variation of albedo during the growing season. The differences in the spatial pattern between constant and seasonal albedo are almost negligible (Fig. 2). In subsequent sections, we default to using the seasonal albedo in our PET methods to fully represent the surface characteristics.

415

Figure 3. Differences in spatially averaged correlation (ΔR) of pairs of PET methods that share the same surface characteristics except for one of the surface features: surface roughness, surface conductance, albedo, and overall consistency among the above features.

5.2 Performance of PET parameterizations

Fig. 4 shows ΔR between each PET method in Table 4 and the control scenario (set PET to zero) for all grids, forested grids, and nonforested grids using 1-month SPEI. Incorporating the benchmark OW method into the SPEI increases R by 0.042, shown by the top horizontal bars. Among the conventional PET methods, the tall reference crop (RC-tall) method stands out. Over the CONUS, it improved ΔR relative to control scenario by 29% more than the OW method (0.054 versus 0.042). The short reference crop (RC-short) method has an identical averaged R with the OW method. Although the RC-tall algorithm (Allen et al., 2005) is less known than the widely used RC-short algorithm (Allen, Pereira, Raes, & Smith, 1998), our results suggest that the SPEI driven by RC-tall correlates better with the SMsurf dynamics.

430

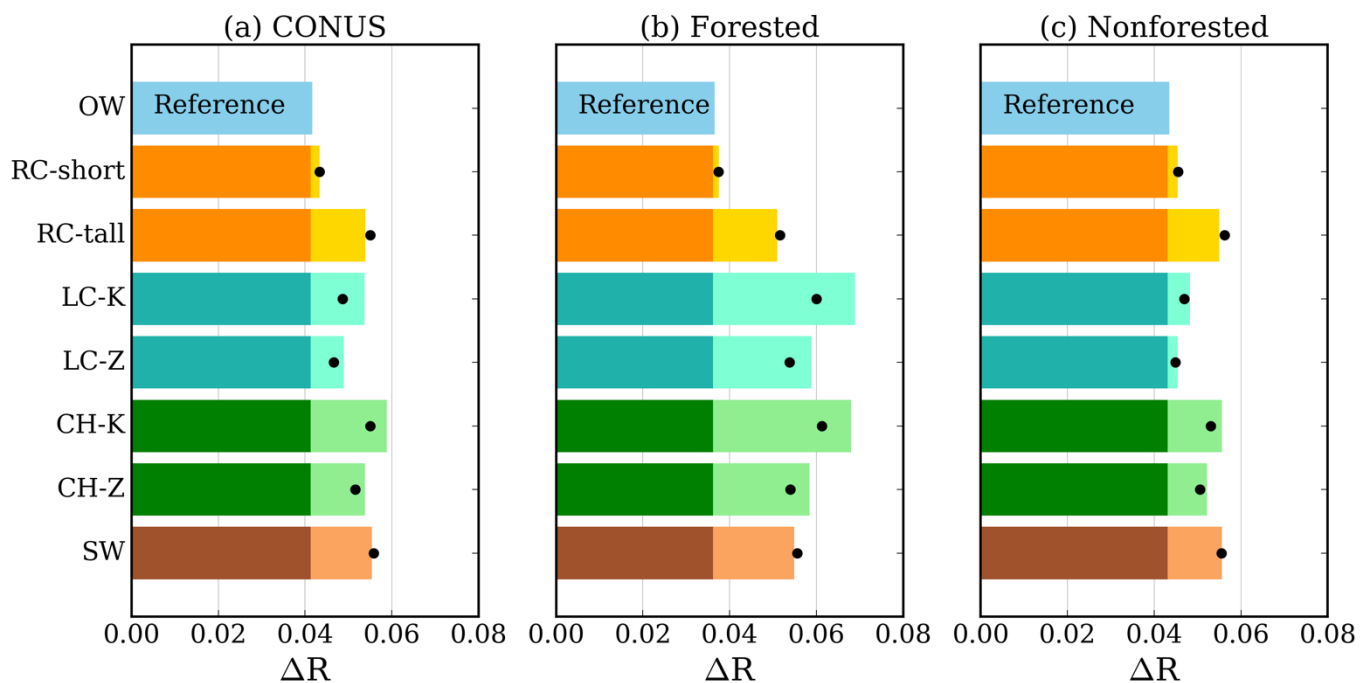



Figure 4. Differences in correlations (ΔR) for selected PET methods versus the control scenario (PET = 0). Correlations were computed between the 1-month SPEI and SMsurf series across: (a) CONUS, (b) forested grids, and (c) nonforested grids. The bars represent the mean ΔR and the black dots represent the median ΔR. The top blue bars show ΔR in the OW approach versus PET = 0 as a reference. For each bar, the darker shade indicates the reference ΔR and the lighter shade represents any improvement (or decline) relative to the reference.

One encouraging outcome is the performance improvement seen in the two big leaf LC and CH algorithms incorporating realistic surface conductance. Activating both surface roughness and seasonal Gs produces high correlations of SPEI with SMsurf. These algorithms improve the OW method ($\Delta R = 0.042$) by 29-41% (ΔR is 0.053-0.059). Methods where Ga is determined by canopy height (the green bars in Fig. 3a) especially improve the correlations with SMsurf. Methods where Gs is determined by Eq. 14 and parameterized using Kelliher et al. produce higher correlations in both LC and CH algorithms. It confirms our hypothesis that incorporating realistic vegetation information in atmospheric evaporative demand can enhance drought characterization. Finally, the two source Shuttleworth-Wallace (SW) method outperforms the OW method as expected. However, the SW method produces a smaller R than the CH-K method. This suggests that the simple big leaf model in combination with the land cover details can achieve the same efficacy of the more complicated two source model.

Over the CONUS, RC-tall, LC-K, CH-K, and SW are the top methods with similar average R. However, when we evaluate the performance in the forested areas (Fig. 3b), LC-K and CH-K exhibit the most significant enhancement in ΔR to control scenario, with increases of 86-89% over OW's improvement (0.068 relative to 0.036). RC-tall and SW improve ΔR to control scenario by 39% (0.05) and 50% (0.054), respectively. In nonforested areas (Fig. 3c), RC-tall has the best performance, followed by SW and CH-K. The SW method, designed for sparse vegetation, naturally demonstrates strong performance in these regions. Similarly, the CH-K method uses the dataset by Lang et al., which includes better quality canopy height measurements in short vegetated areas. Conversely, LC-K only exhibits a moderate improvement in ΔR. This suggests that the performance of the land cover based approach in the sparse vegetation is strongly influenced by the uncertainty of the roughness parameters. On the other hand, it is surprising to see that the simple parametrized RC-tall can outperform SW. This suggests that, particularly in the sparsely vegetated areas, RC-tall can serve as a strong yet simple approach for PET estimates and drought characterization.

5.3 Spatial patterns analysis

In the subsequent sections, we compare the LC-K (referred to as LC) and CH-K (referred to as CH) methods, RC-tall, and SW with the three widely used methods: OW, PT, RC-short. The time series of these PET methods as well as the SMsurf time series are shown in the Fig. 5. The spatial patterns of the mean PET monthly values are shown in Fig. C1 (Appendix C). The OW approach serves as the reference. The highest R is observed for long-term drought (12-month, average $R = 0.73$) and the lowest is found in medium-term

drought (3- and 6-month, average R = 0.48). This suggests that the meteorology-driven SPEI can generally reproduce soil moisture dynamics, especially on an annual time scale.

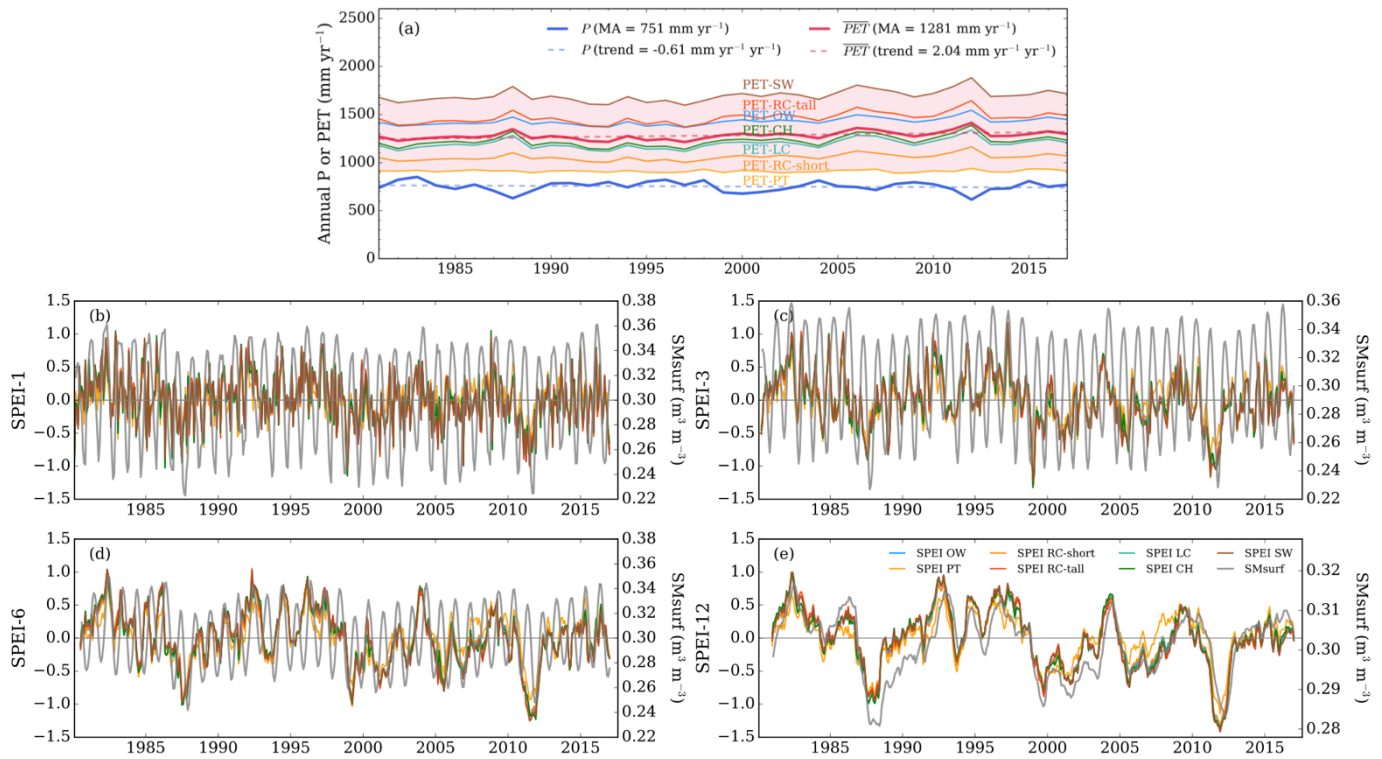



Figure 5. Temporal evolution of PET methods, SPEIs, and SMsurf. a) The annual precipitation and PET (mm yr⁻¹) from PET methods between 1981-2017. b)-e) SPEI series driven by the PET methods, aligning with the SMsurf time series for four time scales: 1, 3, 6, 12-month.

Fig. 6 displays the spatial distribution of correlations between SPEI driven by OW and SMsurf, along with the differences in correlations of PT, RC-short, RC-tall, LC, CH, and SW compared to OW. PT consistently exhibits lower correlations than OW over most regions, with an average decrease of 0.04, and has especially weak correlations in the southwest U.S. (lower by 0.15). Interestingly, the widely used RC-short method for SPEI presents little improvement over OW with minimal increases in correlation, while RC-tall method has overall better performance across CONUS and time scales. Both CH and LC show substantial improvements in some areas, with ΔR exceeding 0.16, notably in the eastern and pacific western U.S.. The enhancements of LC and CH are prominent but can be diluted when averaged across CONUS, with ΔR relative to the control scenario 0.012 higher than OW (Fig. 4a). This is especially true when considering their less favorable performance in the wouthwest and midwest U.S.. SW also exhibits notable improvements in the eastern and

pacific western U.S., with a magnitude of improvement falling between CH and RC-short. It is encouraging to see that LC and CH outperform SW in many eastern US grid cells ($\Delta R = 0.15$ versus $\Delta R = 0.05$), given their much simpler parameterization. Though it is worth noting that all LC, CH, and SW experience performance declines in the Southwest, with LC and CH slightly worse than SW. On the other hand, RC-tall robustly displays improvements in this particular area.

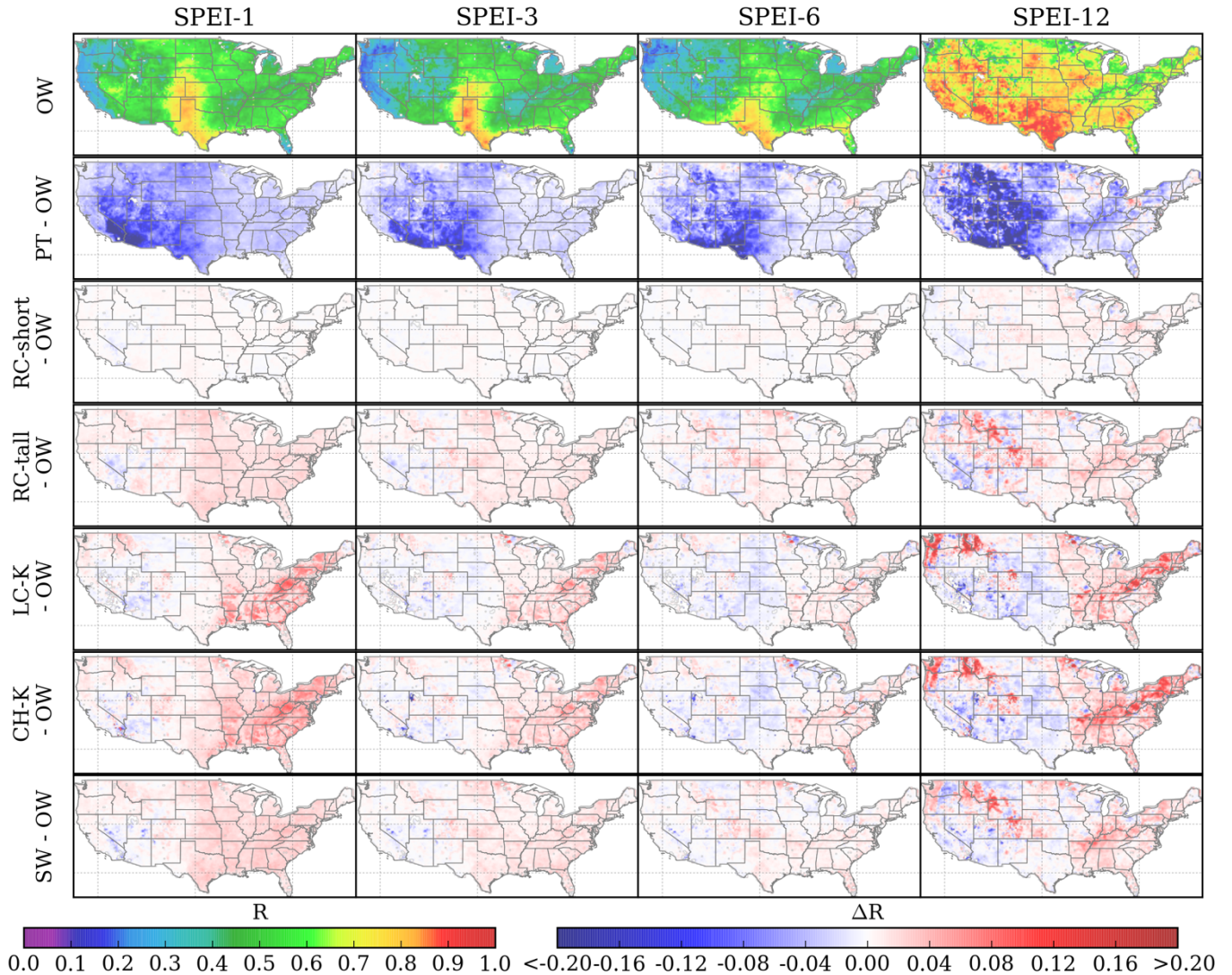



Figure 6. The first row displays the correlations between SPEI driven by OW and SMsurf. The rows below show the differences in correlations (ΔR) of PET methods relative to OW.

We further delve into the relative performance of the top four methods summarized by major vegetation types and by aridity (Fig. 7). Both LC and CH increase R significantly in forests, especially in evergreen broadleaf, deciduous broadleaf, and mixed forests, where the largest ΔR exceeds 0.1, and the average ΔR hovers around

or above 0.05 for 1-month scale (Fig. 7a). Notable improvements in evergreen needleleaf forest, woody

500   savanna, croplands, and mosaic lands compared to the OW are also observed. LC performs slightly better

than CH in forests, while CH performs slightly better in shorter vegetation.

For the time scale of 12-month (Fig. 7c), OW has an already high the average R of 0.73 across the CONUS.

LC and CH's performance are outstanding in forests, with an average ΔR of about 0.05 and the largest ΔR

even exceeding 0.25. In evergreen needleleaf forests, CH and LC's performance are significantly higher than

505   that of the 1-month time scale. In humid regions, CH and LC's improvements over SW becomes even more

apparent compared to the 1-month time scale (Fig. 7d).

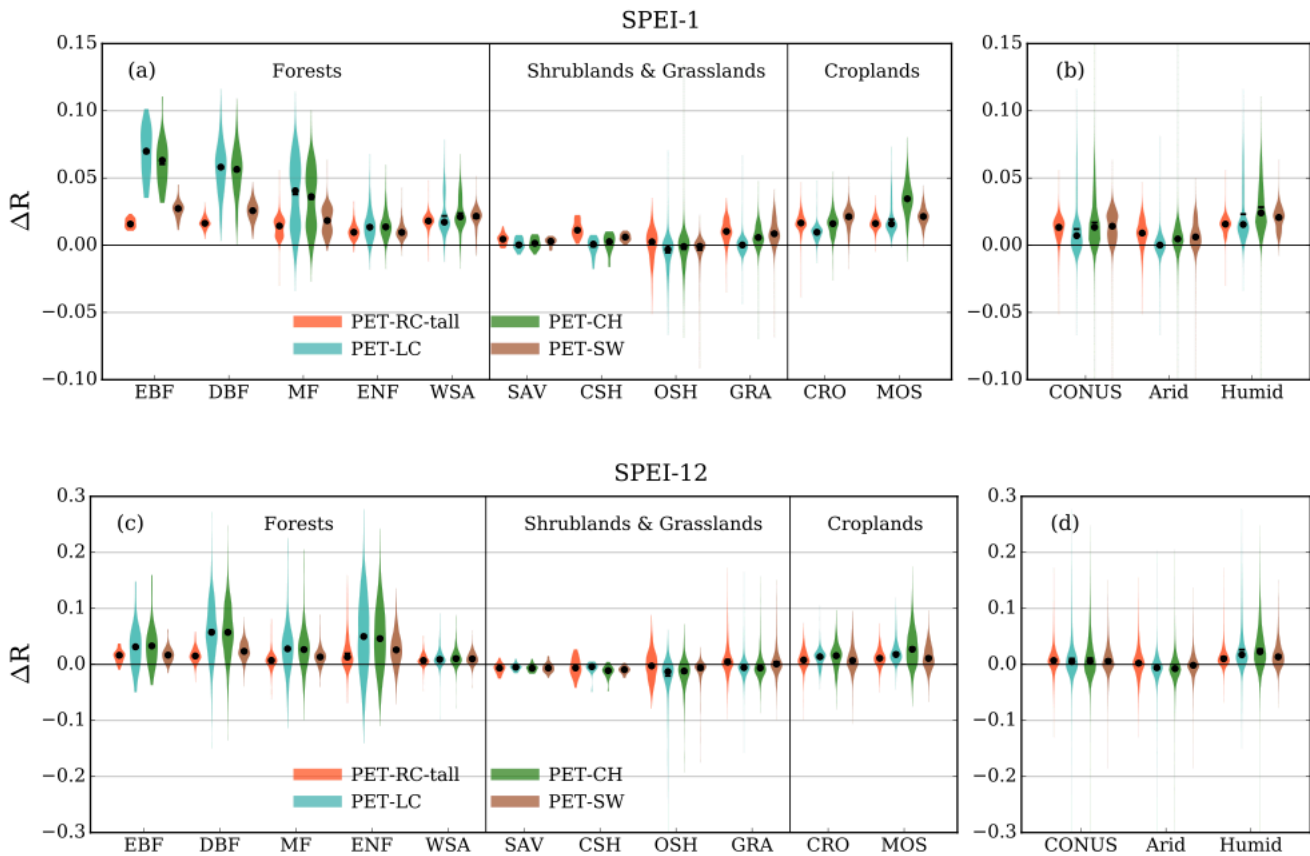



Figure 7. Violin plots of differences in correlations of three PET methods relative to OW,
grouped by vegetation types and aridity. In each violin plot, the black dot represents the

510   **median and the black line represents the mean.**

In contrast, the average performance of LC and CH in grasslands, shrublands, and savannas (Fig. 7a and c),

which are the dominant vegetation types in the western CONUS, are equivalent to or slightly lower than OW.

The magnitude of averaged ΔR of LC and CH are slightly smaller than RC-tall and SW, mainly due to their weaker performance in the arid shrublands and grasslands, which cover large portions of the CONUS. The more complicated LC, CH, and SW methods show less advantage or even worse performance than RC-tall and OW in nonforested and arid grid cells (Fig. 7c-d).

6 Discussion

6.1 Interaction between surface features

Fig. 3 provides important insights into the SPEI sensitivity to different surface features. Introducing Gs with seasonal vegetation dynamics accounts for most of the total improvement of PET algorithm. This confirms that the FAO approaches are more favored than the OW approach due to its constraints on Gs. This highlights the importance of leaf area index (LAI) as a vegetation feature for drought depiction. LAI is a scaling factor to upscale Gst_{max} to maximum canopy conductance. This is different from the drought index based on the normalized difference vegetation index (NDVI) or LAI, which requires the real-time dynamics of satellite data. This approach only requires the climatology of LAI, which can be easily implemented for drought forecasting where real-time or near-future data are not available.

Using realistic surface roughness does not necessarily improve the overall performance of the SPEI. In fact, the consistency between aerodynamic conductance and surface conductance is more critical for the skill of PET method. Previous study by Peng et al. (2019) explains the linkage between the ratio of actual ET to PET and the ratio of Ga to Gs. When Ga/Gs is large, the ratio of actual ET to PET becomes smaller. Although our study focused on the maximum evapotranspiration given the realistic vegetation condition, such a relationship remains valid. Thus, a large Ga/Gst_{max} ratio should better limit PET with realistic surface constraints. In fact, the LC approach activates surface roughness and increases Ga, while constraining Gst_{max} and reducing Gs; the CH approach further incorporates canopy height and LAI in the representation of surface roughness. Altogether these factors increase the Ga/Gs ratio and result in significant improvement in capturing the temporal evolution of SMsurf.

6.2 Surface characteristics matter in the forests

Our analysis concludes that incorporating surface features can largely improve the accuracy of drought monitoring in the forests. There are two vegetation groups with significantly improved correlation after incorporating the realistic surface characteristics. Forests over the eastern and pacific western U.S., such as evergreen broadleaf and deciduous broadleaf forests, the LC and CH methods exhibit large ΔR compared to OW (up to 0.12 for 1-month and up to 0.25 for 12-month, Fig. 7a, c). While OW has a ΔR at about 0.036 compared to the zero PET control scenario (Fig. 4b), LC has an average ΔR of 0.032 relative to OW in these forests. This means the improvement of LC over control scenario is almost doubling of OW. CH and LC also display a significant increase in R at about 0.025 in woody savanna. The enhancements in the forests or woody savannas are the most predominant since LAI in forests is relatively variable, and surface roughness is also the strongest. Although the southeastern U.S. has a humid subtropic climate, this region also suffered

from periodic droughts in 1986–1988, 1998–2002 and 2006–2009 (Seager, Tzanova, & Nakamura, 2009; Pederson et al., 2012), which is consistent with the increased forest drought severity from 1987-2013 (Peters, Iverson, & Matthews, 2014; Clark et al., 2016). Drought monitoring in these regions is also critical and can benefit from our approach that significantly improve the spatial and temporal accuracy in the forests.

In contrast, the short-grass regions (grasslands, shrublands, and savannas) located in the western U.S. exhibit minimal improvements for LC. The CH method, incorporating the newly available Lang et al. (2023) canopy height dataset, improves the correlations in grasslands, croplands, and mosaic lands. Given that the RC-tall method—a similar big leaf model—performs better than LC and CH in shrublands (Fig. 4, 7), it suggests that uncertainties in LC and CH's Gst_{max} could result in these outcomes. Additionally, a comparison between Gst_{max} and Rst_{min} (used in SW) highlights uncertainties in this parameter. For instance, Rst_{min} in shrublands, grasslands, and savannas ranges from 100-180 s m^{-1} (equivalent to Gst_{max} of 5-10 mm s^{-1}), which is generally lower than 9-12 mm s^{-1} reported by Kelliher et al. (1995). These findings highlight the need for in-situ measurements of surface conductance in these areas.

Furthermore, these areas have sparse vegetation cover, and thus LAI plays a less effective role in determining the seasonal dynamics of PET. In the meantime, these areas are located in the arid regions (Fig. 7), the improvements of PET do not have significant effects on modeling the soil moisture, and precipitation dynamics may dominate the soil moisture variations.

6.3 Strategies for PET method selection

Both the CH and LC methods not only provide modest absolute PET values (Fig. 5a, C2) but also display better performance across many areas (Fig. 6). Specifically, LC and CH estimate an annual PET of roughly 1200 mm, which is within the range of the higher OW value (1424 mm) and the lower values around 1100 mm from RC-short as well as Sun et al. PET dataset (Fig. C2).

As Ershadi et al. (2015) pointed out, no single model consistently outperformed any other when considered across all land cover types. the selection of PET for model simulation varies depending on the region (Pimentel et al., 2023). We recommend the use of both LC and CH parameterization for drought monitoring in the forests, in which the roughness and surface conductance parameters vary with realistic vegetation conditions. Both are superior than OW or RC-short because of better performance, and compared to SW, they are both better performing and a simpler approach in the forested areas. Between CH and LC, we recommend CH because it factors in the dynamic change of vegetation structure and provides slightly better performance in woody savanna.

For shrublands and grasslands, we recommend the use of RC-tall to replace the more widely used RC-short for drought monitoring. We found that the RC-tall approach has a higher skill than the RC-short approach that is more widely used. The main difference between these two methods is the C_n constant that describes the effect of aerodynamic conductance (Allen et al, 2005). The implementation of tall reference ($C_n = 1600$) seems to work better than the short reference ($C_n = 900$) over the CONUS. It is worth noting, however, that the FAO approaches assume a universal C_n regardless of actual vegetation type. The better skill of RC-tall will not always hold, which may overestimate PET in semi-arid non-vegetated regions.

For sparse vegetation, since the responses of the components of evapotranspiration to the environmental drivers are different (Katul, Oren, Manzoni, Higgins, & Parlange, 2012; Or & Lehmann, 2019), the partitioning between canopy and soil can also play a role in determining AED. The SW model significantly improves the SPEI skill driven by the OW approach. It outperforms LC and CH in the croplands and grasslands. Despite its complexity, it is a good choice for drought monitoring in these vegetation types (Sun et al., 2023).

For croplands, we recommend choosing between RC-tall versus RC-short based on the actual crop canopy height. The more realistic approach is to use RC-tall for higher crops. Lastly, the PT method has the poorest correlation with soil moisture and is unlikely to capture drought dynamics.

6.4 Bridging gaps in drought prediction

Motivated by the question of whether incorporating surface characteristics can improve drought prediction, we overcome several limitations of previous drought quantification methods. Firstly, our study presents a different approach whereby we focus on the maximum possible evapotranspiration for a given vegetation condition. This concept allows a physically meaningful definition of evaporative demand for the non-uniform land surfaces.

Secondly, the ultimate goal of PET calculation is to simulate ET and to quantify drought. Despite the simplicity of calculating PET using the existing Penman-type methods, the biggest challenge for assessing these methods is validation. Since the real evaporative demand rate is unattainable from observations, it is challenging to validate which PET method is superior directly. Even using ET observations for PET validation can be problematic because biased PET estimates and wrong surface biophysical parameters can still produce accurate ET estimates for locations with ET measurements (Peng et al., 2019). Our study evaluates the PET methods by comparing drought index with independently observed soil moisture (Vicente-Serrano et al., 2012). This approach helps diagnose the most appropriate PET approach for drought quantification directly while avoiding the complexity and divergence caused by various PET definitions.

While the absolute improvements in correlation with soil moisture appear modest, they represent significant percentage changes of average 25-30% and notable local improvements of 86-89% in forests. We acknowledge the need for evaluation of the effectiveness in addition to the temporal correlations. Specifically, future studies should evaluate the capability of the land cover specific approaches to accurately capture extreme events.

Finally, our approach bridges the gap between the two methodologies for quantifying soil-moisture drought, which is of most relevance to agriculture (Seneviratne, 2012). Since soil moisture observations are limited by inadequate measurement networks, drought indices such as the SPEI are often used to quantify drought. In hydrology, a drought index is a simple water balance model driven by surface meteorology without the use of any surface characteristics. Its shortcomings are the neglect of seasonally varying vegetation cover and the incapability to capture the vegetation control on transpiration. An alternative is to use land surface models to estimate large-scale soil moisture (Sheffield, & Wood, 2007). This approach often builds in vegetation dynamics and can provide temporally consistent soil moisture simulations, but it also requires substantial efforts to prepare meteorological forcings at high temporal resolution, set up the domain, spin up, and calibrate. Our approach is a compromise between the above two types of models, which is more realistic and process-based than the commonly used drought index while being easy-to-implement and less data-intensive than a land surface model.

7 Conclusions

630

To understand whether incorporating surface characteristics can improve drought prediction, we revise current PET methods in a newly developed drought index (SPEI), using the concept of maximum ET for any given vegetation condition. We use a simple look-up table approach combining in situ measurements and large-scale data fusion products for the key surface and aerodynamic parameters,. This study also presents a

635

novel application of independent soil moisture observations to diagnose the most appropriate PET approach for drought quantification. Our approach is proved to be more effective than widely used big leaf methods and two source model in accurately predicting soil moisture spatiotemporal dynamics in the forests and humid regions. This new yet simple approach strikes a balance between a meteorology-driven water balance model and a complex land surface model for drought prediction. It could improve the accuracy of the drought

640

reconstruction in forests and displays great potential to improve real-time drought forecast.

Appendix A. Shuttleworth-Wallace Model

The Shuttleworth-Wallace (SW) two source model was developed to more accurately represent evapotranspiration from the sparse vegetation. Different from the big leaf models, SW treats the surface as a two-component structure: sparse vegetation (e.g., row crops) and soil. The following formulas are adapted from Equations 11-18 in Shuttleworth and Wallace (1985).

$$PET_{SW} = C_c PET_{PM}^c + C_s PET_{PM}^s \tag{A1}$$

where PET_{PM}^c and PET_{PM}^s are Penman-Monteith like combined equations (Eq. 4) for a closed canopy and bare soil. Each term is given by the following formulas

$$PET_{PM}^c = \frac{\Delta(R_n - G) + (\rho_a C_p D - \Delta r_a^c (R_n^s - G))/(r_a^a + r_a^c)}{\lambda \left(\Delta + \gamma \left(1 + \frac{r_s^c}{r_a^a + r_a^c}\right)\right)} \tag{A2}$$

$$PET_{PM}^s = \frac{\Delta(R_n - G) + (\rho_a C_p D - \Delta r_a^s (R_n - R_n^s))/(r_a^a + r_a^s)}{\lambda \left(\Delta + \gamma \left(1 + \frac{r_s^s}{r_a^a + r_a^s}\right)\right)} \tag{A3}$$

$$C_c = \frac{1}{1 + \frac{R_c R_a}{R_s(R_c + R_a)}} \tag{A4}$$

$$C_s = \frac{1}{1 + \frac{R_s R_a}{R_c(R_s + R_a)}} \tag{A5}$$

$$R_a = (\Delta + \gamma) r_a^a \tag{A6}$$

$$R_s = (\Delta + \gamma) r_a^s + \gamma r_s^s \tag{A7}$$

$$R_c = (\Delta + \gamma) r_a^c + \gamma r_s^c \tag{A8}$$

where many terms have been given by Eq.1-2, except

R_n^s = net radiation over soil surface = $R_n^s (1 - f_{veg}) = R_n^s \cdot \exp(-0.5 \cdot LAI)$

r_a^a = aerodynamic resistance between canopy height and reference level (s m⁻¹)

r_s^s = surface resistance of the substrate (s m⁻¹)

r_a^s = aerodynamic resistance between substrate and the canopy (s m⁻¹)

r_s^c = bulk stomatal resistance of the canopy (s m⁻¹)

r_a^c = bulk boundary layer resistance of the vegetative elements in the canopy (s m⁻¹).

In this study, the resistances are parameterized for the feasible minimal values based on the water-unlimited assumption for estimating PET. The substrate resistance r_s^s is set to zero s m⁻¹ as a saturated surface. The canopy resistances are dependent on LAI (Shuttleworth and Wallace, 1985, Equations 19-20).

$$r_s^c = Rst \cdot \frac{1}{LAI_e} \tag{A9}$$

$$r_a^c = r_b \cdot \frac{1}{2LAI} \tag{A10}$$

Stomatal resistance Rst is set to Rst_{min} obtained by the land cover types in Table 1. The effective leaf area index LAI_e is LAI/2 and is capped to 2 (even when LAI is greater than 4). Note that, r_s^c does not have valid values for non-vegetated grid cells (at a specific time of the year or location). The leaf boundary layer resistance r_b is set to a value of 50 s m^{-1} (Brisson et al., 1998).

The formulas of aerodynamic resistances are given as follows (Shuttleworth and Gurney, 1990; Zhou et al., 2006).

$$r_a^s = \frac{h \cdot \exp(n) \ln\left(\frac{z_m - d_0}{z_0}\right)}{nk^2(h - d_0)} (\exp\left(-\frac{n z_{0g}}{h}\right) - \exp\left(-\frac{n(z_{0m} + d_p)}{h}\right)) \tag{A11}$$

$$r_a^a = \frac{\ln\left(\frac{z_m - d_0}{z_0}\right) \ln\left(\frac{z_m - d_0}{h - d_0}\right)}{k^2 u_z} + \frac{\ln\left(\frac{z_m - d_0}{z_0}\right) h}{nk^2(h - d_0)} (\exp(n\left(1 - \frac{z_{0m} + d_p}{h}\right)) - 1) \tag{A12}$$

where h is canopy height (m), k is the von Karman constant, z_{0m} is the "preferred" roughness length (m), $z_{0m} = h/8$, d_p is the "preferred" zero plane displacement height (m), $d_p = 0.63h$, z_{0g} is the roughness length of ground (m), u_z is the wind speed from the measurement height (m s^{-1}), and z_m is the measurement height (m), assuming $z_m = h + 2$.

d_0 is the zero plane displacement of canopy (m), n is the eddy diffusivity decay constant of the vegetation, and z_0 is the canopy roughness length (m). These terms are parameterized as following (Equations 22-26, Zhou et al., 2006):

$$n = \begin{cases} 2.5, & h \leq 1 \\ 2.306 + 0.194h, & 1 < h < 10 \\ 4.25, & h \geq 10 \end{cases} \tag{A13}$$

$$d_0 = \begin{cases} h - z_{0c}/0.3, & LAI \geq 4 \\ 1.1h \cdot \ln(1 + (C_d LAI)^{0.25}), & LAI < 4 \end{cases} \tag{A14}$$

$$z_0 = \min(0.3(h - d_0), z_{0g} + 0.3h(C_d LAI)^{0.5}) \tag{A15}$$

$$C_d = \begin{cases} 1.4 \times 10^{-3}, & h = 0 \\ 0.25\left(-1 + \exp\left(0.909 - \frac{3.03 z_{0c}}{h}\right)\right)^4, & h > 0 \end{cases} \tag{A16}$$

$$z_{0c} = \begin{cases} 0.13h, & h \leq 1 \\ 0.139h - 0.009h^2, & 1 < h < 10 \\ 0.05h, & h \geq 10 \end{cases} \tag{A17}$$

where z_{0c} is the roughness length for a closed canopy (m), C_d is the mean drag coefficient for individual leaves.

Appendix B. Canopy height data

We evaluated the newly available global canopy height dataset (Lang et al., 2023) and the widely used
global tree height dataset (Simard et al., 2011). Although the datasets are highly consistent with each other
(Fig. B1), some discrepancies exist for low-vegetation cases, which is expected due to the fact that Simard
et al. focused on the tree height estimates. We further compared the histogram of canopy height in different
land cover types between Lang et al. (Fig. B2) and Simard et al. (Fig. B3). The two datasets are highly
consistent in the forests, while Lang et al. provides valuable information in the short vegetation types.

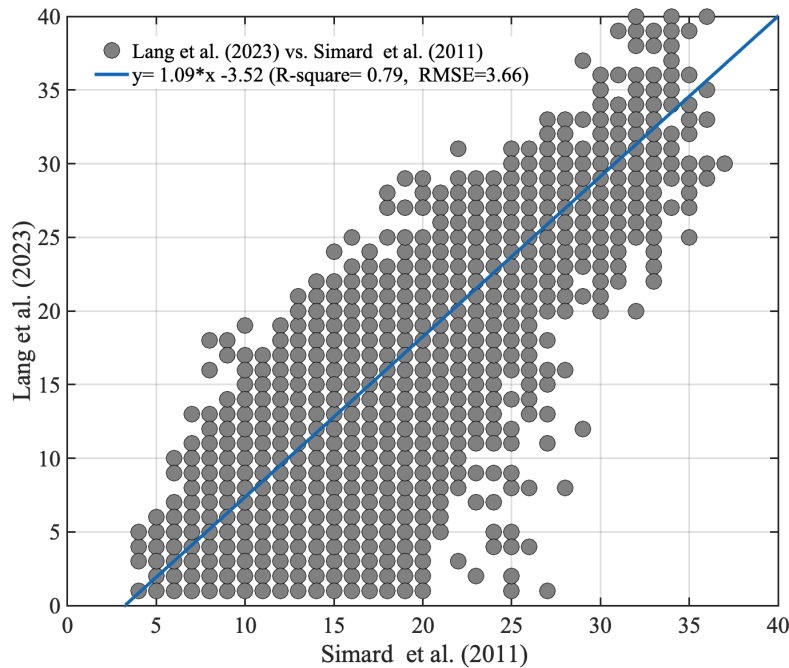



**Figure B1. The comparison in canopy height between Lang et al. (2023) and Simard et al.
(2011).**

We reconstruct the canopy height in each grid cell by comparing the value in Lang et al. with the ranges
given the land cover type, if it is out of the range (smaller than h_{min} or greater than h_{max}) then we give the
grid cell a typical value of canopy height (h_{typ}). For forests, we continue to follow the definitions from the
IGBP land cover classification that the forests are more than 2 m, which we supersede the range in Lang et
al. when it gave a range less than 2 m. Typical canopy height is taken from the value of the peak (mode)
instead of median for forests. For DBF, Lang et al. only has 3 data points, so we use the distribution of Simard
et al. instead, while keeping the lower limit of 6 m from Lang et al.. For grasslands, wetlands, and croplands,
the lidar estimates from Lang et al. or Simard et al. are typically more than 3-5 meters, possibly due to the
overestimation of the grid cell by the sampling of tall trees. Considering the difficulties in separating trees
from the grass and pastures, we did not adopt the high canopy height values in these land cover types. We
use conservative estimates from the literature, 1.5 m (mean value of 0-3m) for grasslands and 0.5 m for
wetland.

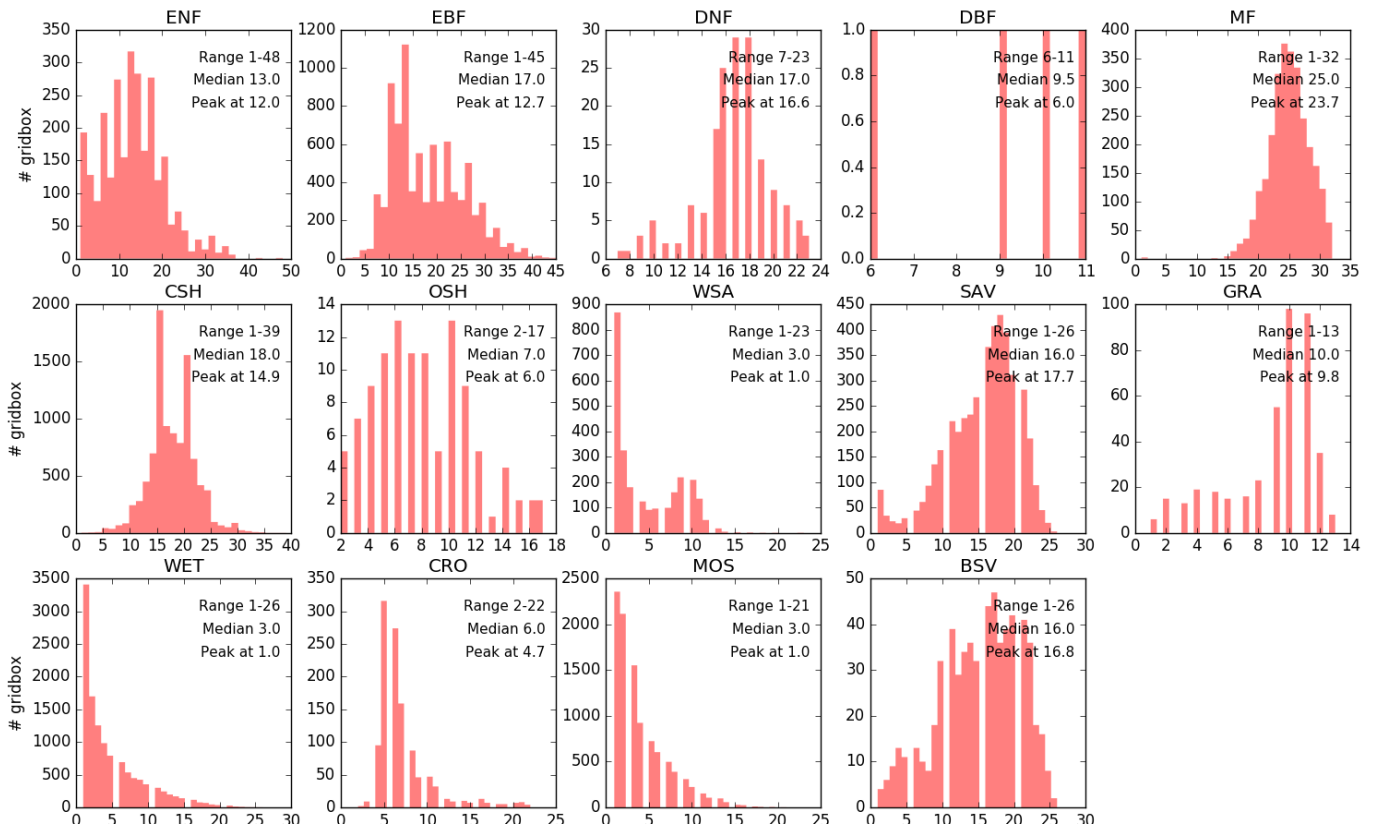
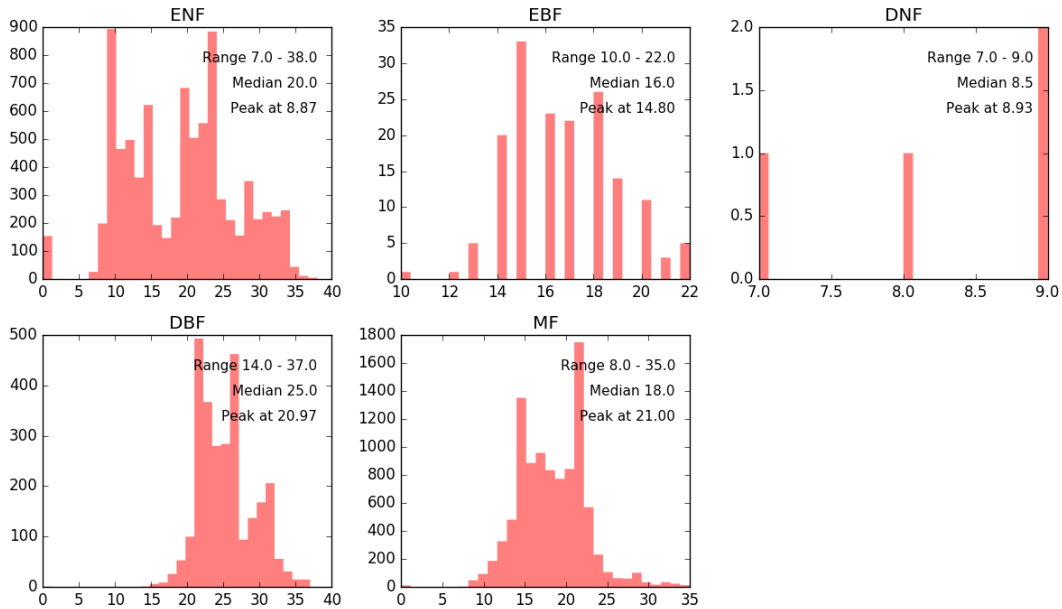



Figure B2. The histogram of canopy height (Lang et al., 2023) by land cover type over the CONUS (excluding non-vegetated land cover: URB, MOS, SNO).

 Figure B3. The histogram of tree height (Simard et al., 2023) by forest type over the CONUS.

Appendix C. Comparison of multiple PET datasets (mm mon⁻¹)

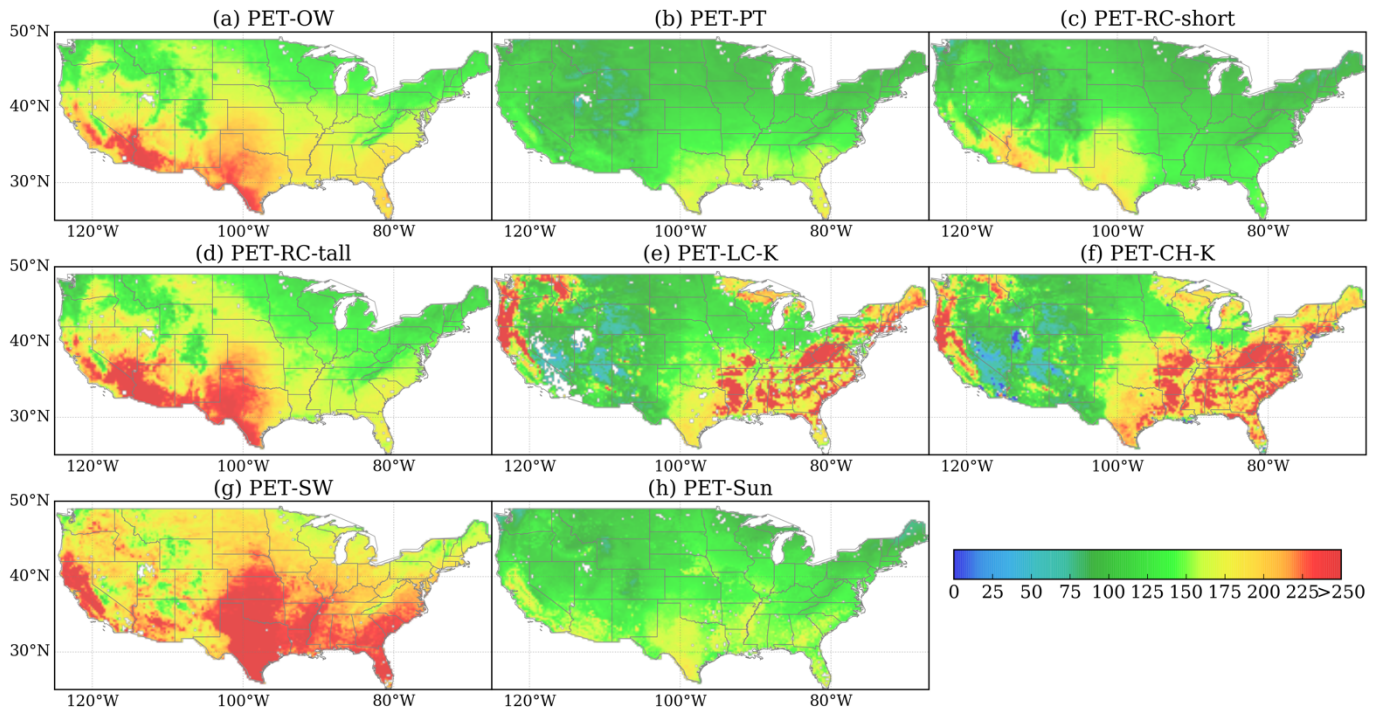



Figure C1. Growing season averages of the PET methods in this study (Table 4) and the PET dataset by Sun et al. (2023) over the CONUS.

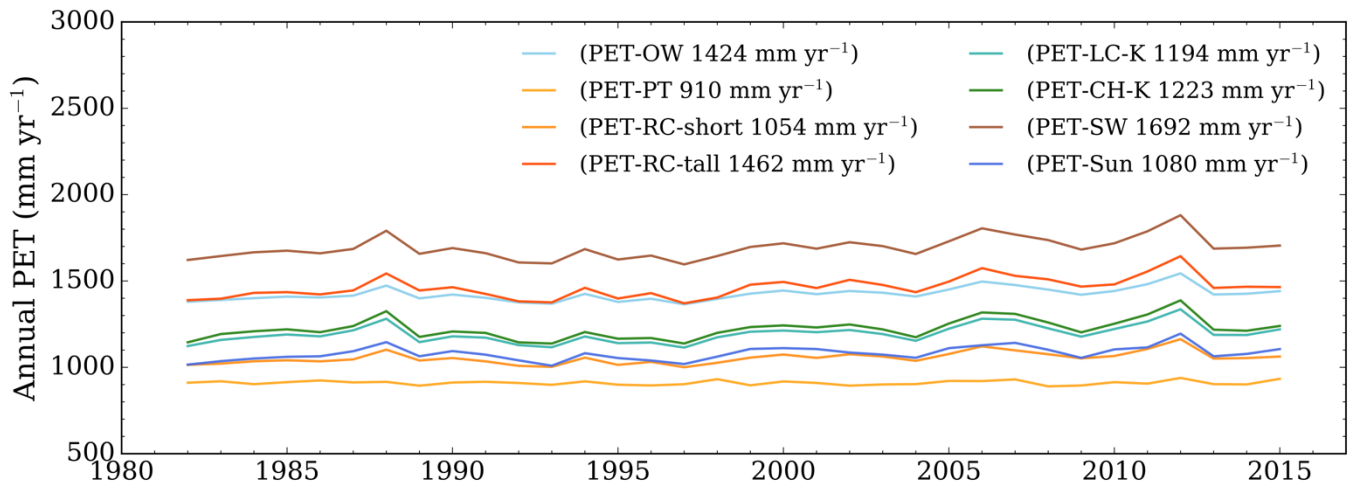



Figure C2. Annual times series of the PET methods in this study (Table 4) and the PET dataset by Sun et al. (2023) over the CONUS.

Code and data availability

The code used to process data and perform analysis for this study is available in the public repository at
https://github.com/pitcheverlasting/spei-pet-evaluation/

720 The data provided along with this study include the key surface parameters, PET annual data from the main methods, precipitation, and SPEI dataset, available in this public repository: https://doi.org/10.6084/m9.figshare.12132696.v1.

The primary data and tools can be downloaded from the PRISM Climate Group at Oregon State University (http://www.prism.oregonstate.edu), the ESA CCI soil moisture project team (https://www.esa-

725 soilmoisture-cci.org/node/145), the GIMMS LAI3g product team (https://drive.google.com/open?id=0BwL88nwumpqYaFJmR2poS0d1ZDQ), the Global Land Surface Satellite project (http://www.glass.umd.edu/Download.html), the SPEI R package released by Santiago Beguería and Sergio M. Vicente-Serrano at CSIC in Spain (https://cran.r-project.org/web/packages/SPEI/), the Global Land Cover Climatology project

730 (https://archive.usgs.gov/archive/sites/landcover.usgs.gov/global_climatology.html), and the CDO software (https://code.zmaw.de/projects/cdo).

Author contribution

LP and JS conceived the idea, LP designed and implemented the PET experiments and analyzed the data, ZW processed some input data, LP wrote the paper with contributions from ZW, ME, and EFW.

735 **Competing interests**

The authors declare that they have no conflict of interest.

Acknowledgements

This research has been supported by NASA under grant NNX14AB36A. We thank Dan Li at Boston University for the helpful discussions. We thank Philip Lu for proofreading.

740

References

Abeysiriwardana, H. D., Muttil, N., & Rathnayake, U. (2022). A comparative study of potential evapotranspiration estimation by three methods with FAO Penman–Monteith method across Sri Lanka. *Hydrology*, *9*(11), 206.

745   Acreman, M. C., Harding, R. J., Lloyd, C. R., & McNeil, D. D. (2003). Evaporation characteristics of wetlands: experience from a wetgrassland and a reedbed using eddy correlation measurements. *Hydrology and Earth System Sciences*, *7*(1), 11–21. http://doi.org/10.5194/hess-7-11-2003

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *FAO*, *300*(9), D05109.

750   Allen, R. G., Walter, I. A., Elliott, R., Howell, T. A., Itenfisu, D., & Jensen, M. (2005). *The ASCE standardized reference evapotranspiration equation* (p. 4). Reston, VA 20191: American Society of Civil Engineers.

Andreadis, K. M., Clark, E. A., Wood, A. W., Hamlet, A. F., & Lettenmaier, D. P. (2005). Twentieth-Century Drought in the Conterminous United States. *Journal of Hydrometeorology*, *6*(6), 985–1001.
755   https://doi.org/10.1175/jhm450.1

Andreadis, K. M., & Lettenmaier, D. P. (2006). Trends in 20th century drought over the continental United States. *Geophysical Research Letters*, *33*(10). https://doi.org/10.1029/2006gl025711

Barbeta, A., Ogaya, R., & Peñuelas, J. (2013). Dampening effects of long-term experimental drought on growth and mortality rates of a Holm oak forest. *Global Change Biology*, *19*(10), 3133–3144.
760   https://doi.org/10.1111/gcb.12269

Beguería, S., Vicente-Serrano, S. M., Reig, F., & Latorre, B. (2013). Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, *34*(10), 3001–3023. https://doi.org/10.1002/joc.3887

765   Brisson, N., Itier, B., L'Hotel, J. C., & Lorendeau, J. Y. (1998). Parameterisation of the Shuttleworth-Wallace model to estimate daily maximum transpiration for use in crop models. *Ecological Modelling*, *107*(2-3), 159-169.

Broxton, P. D., Zeng, X., Sulla-Menashe, D., & Troch, P. A. (2014). A Global Land Cover Climatology Using MODIS Data. *Journal of Applied Meteorology and Climatology*, *53*(6), 1593–1605.

770   https://doi.org/10.1175/jamc-d-13-0270.1

Brutsaert, W. (1982). The Surface Roughness Parameterization. In A. J. Davenport, B. B. Hicks, G. R. Hilst, R. E. Munn, & J. D. Smith (Eds.), *Evaporation into the Atmosphere* (pp. 113–127). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-1497-6_5

Brutsaert, W., & Stricker, H. (1979). An advection-aridity approach to estimate actual regional

775   evapotranspiration. *Water Resources Research*, *15*(2), 443–450. https://doi.org/10.1029/wr015i002p00443

Campbell, G. S., & Norman, J. M. (1998). Wind. In *An introduction to environmental biophysics* (pp. 68-70). New York, NY: Springer.

Clark, J. S., Iverson, L., Woodall, C. W., Allen, C. D., Bell, D. M., Bragg, D. C., … Zimmermann, N. E. (2016). The impacts of increasing drought on forest dynamics structure, and biodiversity in the United

780   States. *Global Change Biology*, *22*(7), 2329–2352. https://doi.org/10.1111/gcb.13160

Dai, A., Trenberth, K. E., & Qian, T. (2004). A Global Dataset of Palmer Drought Severity Index for 18702002: Relationship with Soil Moisture and Effects of Surface Warming. *Journal of Hydrometeorology*, *5*(6), 1117–1130. https://doi.org/10.1175/jhm-386.1

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., … Pasteris, P. P. (2008).

785   Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*(15), 2031–2064. https://doi.org/10.1002/joc.1688

Daly, C., Neilson, R. P., & Phillips, D. L. (1994). A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal of Applied Meteorology*, *33*(2), 140–158.

790   https://doi.org/10.1175/1520-0450(1994)033<0140:astmfm>2.0.co;2

Dewes, C. F., Rangwala, I., Barsugli, J. J., Hobbins, M. T., & Kumar, S. (2017). Drought risk assessment under climate change is sensitive to methodological choices for the estimation of evaporative demand. *Plos One*, *12*(3), e0174045. http://doi.org/10.1371/journal.pone.0174045

Dong, C., MacDonald, G. M., Willis, K., Gillespie, T. W., Okin, G. S., & Williams, A. P. (2019).

795   Vegetation Responses to 20122016 Drought in Northern and Southern California. *Geophysical Research Letters*, *46*(7), 3810–3821. https://doi.org/10.1029/2019gl082137

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., … Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

800   Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., et al. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research*, *108*(D22), 2002JD003296. http://doi.org/10.1029/2002JD003296

Fang, L., Zhan, X., Hain, C. R., Yin, J., & Liu, J. (2018). Impact of GVF Derivation Methods on Noah
805   Land Surface Model Simulations and WRF Model Forecasts. *Journal of Hydrometeorology*, *19*(12), 1917–1933. https://doi.org/10.1175/jhm-d-18-0075.1

Federer, C. A., Vörösmarty, C., & Fekete, B. (1996). Intercomparison of methods for calculating potential evaporation in regional and global water balance models. *Water Resources Research*, *32*(7), 2315-2321. https://doi.org/10.1029/96WR00801

810   Feng, S., Trnka, M., Hayes, M., & Zhang, Y. (2017). Why Do Different Drought Indices Show Distinct Future Drought Risk Outcomes in the U.S. Great Plains?. *Journal of Climate*, *30*(1), 265–278. https://doi.org/10.1175/jcli-d-15-0590.1

Forzieri, G., G Miralles, D., Ciais, P., Alkama, R., Ryu, Y., Duveiller, G., et al. (2020). Increased control of vegetation on global terrestrial energy fluxes. *Nature Climate Change*, *4*(2), 1–7.
815   http://doi.org/10.1038/s41558-020-0717-0

Gao, G., Feng, Q., Liu, X., & Zhao, Y. (2021). Measuring and modeling evapotranspiration of a Populus euphratica forest in northwestern China. *Journal of Forestry Research*, *32*(5), 1963-1977.

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple collocation-based merging of satellite soil moisture retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(12), 6780-6792.
820   https://doi.org/10.1109/TGRS.2017.2734070.

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., & Dorigo, W. (2019). Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. *Earth System Science Data*, *11*, 717-739. https://doi.org/10.5194/essd-11-717-2019

Heim, R. R. (2002). A Review of Twentieth-Century Drought Indices Used in the United States. *Bulletin of*
825   *the American Meteorological Society*, *83*(8), 1149–1166. https://doi.org/10.1175/1520-0477-83.8.1149

Hobbins, M. T., Wood, A., Streubel, D., & Werner, K. (2012). What Drives the Variability of Evaporative Demand across the Conterminous United States? *Journal of Hydrometeorology*, *13*(4), 1195–1214. http://doi.org/10.1175/JHM-D-11-0101.1

830   Hoerling, M., Eischeid, J., Kumar, A., Leung, R., Mariotti, A., Mo, K., … Seager, R. (2014). Causes and Predictability of the 2012 Great Plains Drought. *Bulletin of the American Meteorological Society*, *95*(2), 269–282. https://doi.org/10.1175/bams-d-13-00055.1

Katul, G. G., Oren, R., Manzoni, S., Higgins, C., & Parlange, M. B. (2012). Evapotranspiration: A process driving mass transport and energy exchange in the soil-plant-atmosphere-climate system. *Reviews of Geophysics*, *50*(3). https://doi.org/10.1029/2011rg000366

835   Kelliher, F. M., Leuning, R., Raupach, M. R., & Schulze, E.-D. (1995). Maximum conductances for evaporation from global vegetation types. *Agricultural and Forest Meteorology*, *73*(1-2), 1–16. https://doi.org/10.1016/0168-1923(94)02178-m

Kogan, F. N. (1995). Droughts of the Late 1980s in the United States as Derived from NOAA Polar-Orbiting Satellite Data. *Bulletin of the American Meteorological Society*, *76*(5), 655–668.
840   https://doi.org/10.1175/1520-0477(1995)076<0655:dotlit>2.0.co;2

Kustas, W. P., Choudhury, B. J., Moran, M. S., Reginato, R. J., Jackson, R. D., Gay, L. W., & Weaver, H. L. (1989). Determination of sensible heat flux over sparse canopy using thermal infrared data. *Agricultural and Forest Meteorology*, *44*(3-4), 197–216. https://doi.org/10.1016/0168-1923(89)90017-8

Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2023). A high-resolution canopy height model of the
845   Earth. *Nature Ecology & Evolution*, *7*(11), 1778-1789.

Leuning, R., Zhang, Y. Q., Rajaud, A., Cleugh, H., & Tu, K. (2008). A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman-Monteith equation. *Water Resources Research*, *44*(10), 1872. http://doi.org/10.1029/2007WR006562

Lhomme, J.-P., Troufleau, D., Monteny, B., Chehbouni, A., & Bauduin, S. (1997). Sensible heat flux and
850   radiometric surface temperature over sparse Sahelian vegetation II. A model for the kB-1 parameter. *Journal of Hydrology*, *188-189*, 839–854. https://doi.org/10.1016/s0022-1694(96)03173-3

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *99*(D7), 14415–14428. http://doi.org/10.1029/94JD00483

855   Liu, Q., Wang, L., Qu, Y., Liu, N., Liu, S., Tang, H., & Liang, S. (2013). Preliminary evaluation of the long-term GLASS albedo product. *International Journal of Digital Earth*, *6*(sup1), 69–95. https://doi.org/10.1080/17538947.2013.804601

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., … van Dijk, A. I. J. M. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote*

860   *Sensing of Environment*, *123*, 280–297. https://doi.org/10.1016/j.rse.2012.03.014

Loon, A. F. V. (2015). Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water*, *2*(4), 359–392. https://doi.org/10.1002/wat2.1085

Lorenz, R., Davin, E. L., Lawrence, D. M., Stö Ckli, R., & Seneviratne, S. I. (2013). How important is vegetation phenology for European climate and heat waves? *Journal of Climate*, *26*(24), 10077–10100.

865   http://doi.org/10.1175/JCLI-D-13-00040.1

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., … Verhoest, N. E. C. (2017). GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

McDowell, N. G., Allen, C. D., Teixeira, K. A., Brando, P., Brienen, R., Chambers, J., et al. (2018).

870   Drivers and mechanisms of tree mortality in moist tropical forests. *New Phytologist*, *219*(3), 851–869. http://doi.org/10.1111/nph.15027

McKee, T. B., Doesken, N. J., Kleist, J., & others. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology* (Vol. 17, pp. 179–183). Boston, MA: American Meteorological Society.

875   Monteith, J. L. (1965). Evaporation and environment. *Symposia of the Society for Experimental Biology*, *19*, 205–234.

Monteith, J., & Unsworth, M. (2013). Micrometeorology. In *Principles of environmental physics: plants, animals, and the atmosphere* (pp. 301-310). Academic Press.

Moran, M. S., Rahman, A. F., Washburne, J. C., Goodrich, D. C., Weltz, M. A., & Kustas, W. P. (1996).

880   Combining the Penman-Monteith equation with measurements of surface temperature and reflectance to estimate evaporation rates of semiarid grassland. *Agricultural and Forest Meteorology*, *80*(2-4), 87–109. http://doi.org/10.1016/0168-1923(95)02292-9

Nadeau, D. F., Brutsaert, W., Parlange, M. B., Bou-Zeid, E., Barrenetxea, G., Couach, O., … Vetterli, M. (2009). Estimation of urban sensible heat flux using a dense wireless network of observations. *Environmental Fluid Mechanics*, *9*(6), 635–653. https://doi.org/10.1007/s10652-009-9150-7

Norman, J. M., Kustas, W. P., & Humes, K. S. (1995). Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agricultural and Forest Meteorology*, *77*(3-4), 263–293. https://doi.org/10.1016/0168-1923(95)02265-y

Or, D., & Lehmann, P. (2019). Surface Evaporative Capacitance: How Soil Type and Rainfall Characteristics Affect Global-Scale Surface Evaporation. *Water Resources Research*, *55*(1), 519–539. https://doi.org/10.1029/2018wr024050

Palmer, W. C. (1965). *Meteorological drought* (Vol. Research paper No. 45, p. 58). Washington, DC: U.S. Weather Bureau.

Paulo, A. A., Rosa, R. D., & Pereira, L. S. (2012). Climate trends and behaviour of drought indices based on precipitation and evapotranspiration in Portugal. *Natural Hazards and Earth System Sciences*, *12*(5), 1481–1491. https://doi.org/10.5194/nhess-12-1481-2012

Pederson, N., Bell, A. R., Knight, T. A., Leland, C., Malcomb, N., Anchukaitis, K. J., … Riddle, J. (2012). A long-term perspective on a modern drought in the American Southeast. *Environmental Research Letters*, *7*(1), 014034. https://doi.org/10.1088/1748-9326/7/1/014034

Peng, L., Li, D., & Sheffield, J. (2018). Drivers of Variability in Atmospheric Evaporative Demand: Multiscale Spectral Analysis Based on Observations and Physically Based Modeling. *Water Resources Research*, *54*(5), 3510–3529. http://doi.org/10.1029/2017WR022104

Peng, L., Zeng, Z., Wei, Z., Chen, A., Wood, E. F., & Sheffield, J. (2019). Determinants of the ratio of actual to potential evapotranspiration. *Global change biology*, *25*(4), 1326-1343. https://doi.org/10.1111/gcb.14577

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London Series A*, *193*, 120–145.

Peters, M. P., Iverson, L. R., & Matthews, S. N. (2014). *Spatio-temporal trends of drought by forest type in the conterminous United States, 1960-2013*. U.S. Department of Agriculture Forest Service, Northern Research Station. https://doi.org/10.2737/nrs-rmap-7

Piao, S., Zhang, X., Chen, A., Liu, Q., Lian, X., Wang, X., ... & Wu, X. (2019). The impacts of climate extremes on the terrestrial carbon cycle: A review. *Science China Earth Sciences*, *62*, 1551-1563. https://link.springer.com/article/10.1007/s11430-018-9363-5

Pimentel, R., Arheimer, B., Crochemore, L., Andersson, J. C. M., Pechlivanidis, I. G., & Gustafsson, D. (2023). Which potential evapotranspiration formula to use in hydrological modeling world-wide? *Water Resources Research*, 59, e2022WR033447. https://doi.org/10.1029/2022WR033447

Potop, V. (2011). Evolution of drought severity and its impact on corn in the Republic of Moldova. *Theoretical and Applied Climatology*, *105*(3-4), 469–483. https://doi.org/10.1007/s00704-011-0403-2

Potop, V., Možný, M., & Soukup, J. (2012). Drought evolution at various time scales in the lowland regions and their impact on vegetable crops in the Czech Republic. *Agricultural and Forest Meteorology*, *156*, 121–133. https://doi.org/10.1016/j.agrformet.2012.01.002

Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, *100*(2), 81–92.

Qu, Y., Liu, Q., Liang, S., Wang, L., Liu, N., & Liu, S. (2014). Direct-Estimation Algorithm for Mapping Daily Land-Surface Broadband Albedo From MODIS Data. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(2), 907–919. https://doi.org/10.1109/tgrs.2013.2245670

Rigden, A., Li, D., & Salvucci, G. (2018). Dependence of thermal roughness length on friction velocity across land cover types: A synthesis analysis using AmeriFlux data. *Agricultural and Forest Meteorology*, *249*, 512–519. https://doi.org/10.1016/j.agrformet.2017.06.003

Ross, T., & Lott, N. (2003). A climatology of 1980-2003 extreme weather and climate events. *National Climatic Data Center Technical Report*, *1*, 14.

Seager, R., Tzanova, A., & Nakamura, J. (2009). Drought in the Southeastern United States: Causes Variability over the Last Millennium, and the Potential for Future Hydroclimate Change. *Journal of Climate*, *22*(19), 5021–5045. https://doi.org/10.1175/2009jcli2683.1

Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., et al. (1996). A revised land surface parameterization (SiB2) for atmospheric GCMs .1. Model formulation, *9*(4), 676–705.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3-4), 125–161. http://doi.org/10.1016/j.earscirev.2010.02.004

940   Seneviratne, S. I. (2012). Historical drought trends revisited. *Nature*, *491*(7424), 338–339. http://doi.org/10.1038/491338a

Sheffield, J., & Wood, E. F. (2007). Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *112*(D17), D17115. http://doi.org/10.1029/2006JD008288

945   Sheffield, J., Wood, E. F., & Roderick, M. L. (2012). Little change in global drought over the past 60 years. *Nature*, *491*(7424), 435–438. https://doi.org/10.1038/nature11575

Shuttleworth, W. J., & Wallace, J. S. (1985). Evaporation from sparse crops-an energy combination theory. *Quarterly Journal of the Royal Meteorological Society*, *111*(469), 839-855.

Shuttleworth, W. J., & Gurney, R. J. (1990). The theoretical relationship between foliage temperature and
950   canopy resistance in sparse crops. *Quarterly Journal of the Royal Meteorological Society*, *116*(492), 497-519.

Shuttleworth, W. J. (1993). Chapter 4 Evaporation. In D. R. Maidment (Ed.), *Handbook of hydrology* (Vol. 9780070). Sydney: McGraw-Hill.

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with
955   spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*, *116*(G4). https://doi.org/10.1029/2011JG001708

Stewart, J. B., Kustas, W. P., Humes, K. S., Nichols, W. D., Moran, M. S., & de Bruin, H. A. R. (1994). Sensible Heat Flux-Radiometric Surface Temperature Relationship for Eight Semiarid Areas. *Journal of Applied Meteorology*, *33*(9), 1110–1117. https://doi.org/10.1175/1520-
960   0450(1994)033<1110:shfrst>2.0.co;2

Sun, S., Bi, Z., Xiao, J., Liu, Y., Sun, G., Ju, W., ... & Chen, H. (2023). A global 5 km monthly potential evapotranspiration dataset (1982–2015) estimated by the Shuttleworth–Wallace model. *Earth System Science Data*, *15*(11), 4849-4876.

Thornthwaite, C. W. (1948). An Approach toward a Rational Classification of Climate. *Geographical*
965   *Review*, 38(1):55.

Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., & Sheffield, J. (2013). Global warming and changes in drought. *Nature Climate Change*, *4*(1), 17–22. https://doi.org/10.1038/nclimate2067

Troufleau, D., Lhomme, J. P., Monteny, B., & Vidal, A. (1997). Sensible heat flux and radiometric surface temperature over sparse Sahelian vegetation. I. An experimental analysis of the kB-1 parameter. *Journal of Hydrology*, *188-189*, 815–838. https://doi.org/10.1016/s0022-1694(96)03172-1

Verhoef, A., Bruin, H. A. R. D., & Hurk, B. J. J. M. V. D. (1997). Some Practical Notes on the Parameter kB-1 for Sparse Vegetation. *Journal of Applied Meteorology*, *36*(5), 560–572. https://doi.org/10.1175/1520-0450(1997)036<0560:spnotp>2.0.co;2

Verma, S. B. (1989). Aerodynamic resistances to transfers of heat, mass and momentum. In: Black, T. A., Spittlehouse, D. L., Novak, M. D., Price, D. T. (Eds.), *Estimation of Areal Evapotranspiration* (pp. 13-20). IAHS Press.

Vicente-Serrano, S. M., Beguería, S., Lorenzo-Lacruz, J., Camarero, J. J., López-Moreno, J. I., Azorin-Molina, C., … Sanchez-Lorenzo, A. (2012). Performance of Drought Indices for Ecological Agricultural, and Hydrological Applications. *Earth Interactions*, *16*(10), 1–27. https://doi.org/10.1175/2012ei000434.1

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, *23*(7), 1696–1718. https://doi.org/10.1175/2009jcli2909.1

Vicente-Serrano, S. M., Gouveia, C., Camarero, J. J., Begueria, S., Trigo, R., Lopez-Moreno, J. I., … Sanchez-Lorenzo, A. (2013). Response of vegetation to drought time-scales across global land biomes. *Proceedings of the National Academy of Sciences*, *110*(1), 52–57. https://doi.org/10.1073/pnas.1207068110

Vicente-Serrano, S. M., der Schrier, G. V., Beguería, S., Azorin-Molina, C., & Lopez-Moreno, J.-I. (2015). Contribution of precipitation and reference evapotranspiration to drought indices under different climates. *Journal of Hydrology*, *526*, 42–54. https://doi.org/10.1016/j.jhydrol.2014.11.025

Wei, Z., Yoshimura, K., Wang, L., G Miralles, D., Jasechko, S., & Lee, X. (2017). Revisiting the contribution of transpiration to global terrestrial evapotranspiration. *Geophysical Research Letters*, *44*(6), 2792–2801. http://doi.org/10.1002/2016GL072235

Wilhite, D. A. (2000), Drought as a natural hazard: Concepts and definitions, In *Drought: A Global Assessment* (pp. 3-18). London: Routledge.

Wilhite, D. A., Sivakumar, M. V. K., & Pulwarty, R. (2014). Managing drought risk in a changing climate: The role of national drought policy. *Weather and Climate Extremes*, *3*, 4–13. https://doi.org/10.1016/j.wace.2014.01.002

Williams, A. P., Allen, C. D., Macalady, A. K., Griffin, D., Woodhouse, C. A., Meko, D. M., et al. (2012). Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature Climate Change*, *3*(3), 292–297. http://doi.org/10.1038/nclimate1693

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., … Mocko, D. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, *117*(D3), n/a–n/a. https://doi.org/10.1029/2011jd016048

Xu, H., Lian, X., Slette, I. J., Yang, H., Zhang, Y., Chen, A., & Piao, S. (2022). Rising ecosystem water demand exacerbates the lengthening of tropical dry seasons. *Nature Communications*, *13*(1), 4093. https://www.nature.com/articles/s41467-022-31826-y

Yang, H., Munson, S. M., Huntingford, C., Carvalhais, N., Knapp, A. K., Li, X., ... & Chen, A. (2023). The detection and attribution of extreme reductions in vegetation growth across the global land surface. *Global Change Biology*, *29*(8), 2351-2362. https://onlinelibrary.wiley.com/doi/full/10.1111/gcb.16595

Yan, H., Wang, S. Q., Billesbach, D., Oechel, W., Zhang, J. H., Meyers, T., … Scott, R. (2012). Global estimation of evapotranspiration using a leaf area index-based surface energy and water balance model. *Remote Sensing of Environment*, *124*, 581–595. https://doi.org/10.1016/j.rse.2012.06.004

Yang, K., Koike, T., Ishikawa, H., Kim, J., Li, X., Liu, H., … Wang, J. (2008). Turbulent Flux Transfer over Bare-Soil Surfaces: Characteristics and Parameterization. *Journal of Applied Meteorology and Climatology*, *47*(1), 276–290. https://doi.org/10.1175/2007jamc1547.1

Yang, R., & Friedl, M. A. (2003). Determination of Roughness Lengths for Heat and Momentum Over Boreal Forests. *Boundary-Layer Meteorology*, *107*(3), 581–603. https://doi.org/10.1023/a:1022880530523

Yang, Y., Roderick, M. L., Zhang, S., McVicar, T. R., & Donohue, R. J. (2019). Hydrologic implications of vegetation response to elevated CO2 in climate projections. *Nature Climate Change*, *9*(1), 44–48. https://doi.org/10.1038/s41558-018-0361-0

Zhang, L., Jiao, W., Zhang, H., Huang, C., & Tong, Q. (2017). Studying drought phenomena in the Continental United States in 2011 and 2012 using various drought indices. *Remote Sensing of Environment*, *190*, 96–106. https://doi.org/10.1016/j.rse.2016.12.010

Zhao, H., Gao, G., An, W., Zou, X., Li, H., & Hou, M. (2017). Timescale differences between SC-PDSI and SPEI for drought monitoring in China. *Physics and Chemistry of the Earth Parts A/B/C*, *102*, 48–58. https://doi.org/10.1016/j.pce.2015.10.022

Zhou, M. C., Ishidaira, H., Hapuarachchi, H. P., Magome, J., Kiem, A. S., & Takeuchi, K. (2006). Estimating potential evapotranspiration using Shuttleworth–Wallace model and NOAA-AVHRR NDVI data to feed a distributed hydrological model over the Mekong River basin. *Journal of Hydrology*, 327(1-2), 151-173. https://doi.org/10.1016/j.jhydrol.2005.11.013

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., … Myneni, R. (2013). Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011. *Remote Sensing*, *5*(2), 927–948. https://doi.org/10.3390/rs5020927

Zilitinkevich, S. S., Grachev, A. A., & Fairall, C. W. (2001). Scaling Reasoning and Field Data on the Sea Surface Roughness Lengths for Scalars. *Journal of the Atmospheric Sciences*, *58*(3), 320–325. https://doi.org/10.1175/1520-0469(2001)058<0320:nacraf>2.0.co;2

1030

1035