Potential evaporation at eddy-covariance sites across the globe

Wouter H. Maes[1], Pierre Gentine[2], Niko E.C. Verhoest[1], Diego G. Miralles[1]
1 Laboratory of Hydrology and Water Management, Ghent University, Coupure Links 653, 9000 Gent, Belgium
2 Department of Earth and Environmental Engineering, Columbia University, New York, 10027, USA
Correspondence to: Wouter H. Maes (wh.maes@ugent.be)

Abstract. Potential evaporation (E_p) is a crucial variable for hydrological forecasting and drought monitoring. However, multiple interpretations of E_p exist, which reflect a diverse range of methods to calculate it. A comparison of the performance of these methods against field observations in different global ecosystems is urgently needed. In this study, potential evaporation was defined as the rate of land evaporation (or *evapotranspiration*) that the actual ecosystem would attain if it evaporates at maximal rate for the given atmospheric conditions. We use eddy-covariance measurements from the FLUXNET2015 database, covering eleven different biomes, to parameterize and inter-compare the most widely used E_p methods and to uncover their relative performance. For each site, we isolate days for which ecosystems can be considered as 'unstressed', based on both an energy balance and a soil water content approach. Evaporation measurements during these days are used as reference to calibrate and validate the different methods to estimate E_p. Our results indicate that a simple radiation-driven method, calibrated per biome, consistently performs best against *in situ* measurements (mean correlation of 0.93, unbiased RMSE of 0.56 mm day^{-1}, and bias of -0.02 mm day^{-1}). A Priestley and Taylor method, calibrated per biome, performed just slightly worse, yet substantially and consistently better than more complex Penman-, Penman-Monteith-based or temperature-driven approaches. We show that the poor performance of Penman-Monteith-based approaches largely relates to the fact that the unstressed stomatal conductance cannot be assumed to be constant in time at the ecosystem scale. On the contrary, the biome-specific parameters required by simpler radiation-driven methods are relatively constant in time and per biome type. This makes these methods a robust way to estimate E_p and a suitable tool to investigate the impact of water use and demand, drought severity and biome productivity.

(insert abbreviation list)

1 Introduction

Since its introduction 70 years ago by C. W. Thornthwaite (1948), the concept of potential evaporation (E_p), defined as the amount of water which would evaporate from a surface unconstrained by water availability, has been widely used in multiple fields. It has been incorporated in hydrological models dedicated to estimate runoff (e.g. Schellekens et al., 2017) or actual evaporation (Wang and Dickinson, 2012), as well as in drought severity indices (Sheffield et al., 2012; Vicente-Serrano et al., 2013). Long-term changes in E_p have been regarded as a driver of ecosystem distribution and aridity (Scheff and Frierson, 2013) and used to diagnose the influence of climate change on ecosystems based on climate model projections (e.g. Milly and Dunne, 2016). However, many different definitions of E_p exist, and consequently many different methods are available to calculate it. In recent years, there has been an increasing awareness of the impact of the underlying assumptions and caveats in traditional E_p formulations (Weiß and Menzel, 2008;Kingston et al., 2009;Sheffield et al., 2012;Seiller and Anctil, 2016;Bai et al., 2016;Milly and Dunne, 2016;Guo et al., 2017). As such, a global appraisal of the most appropriate method for assessing E_p is urgently needed. Yet, current formulations reflect a disagreement on the mere meaning of this variable, which requires the definition of some form of reference system (Lhomme, 1997). E_p has been typically defined as the evaporation which would occur in given meteorological conditions if water was not limited, either (*i*) over open water (Shuttleworth, 1993); (*ii*) over a reference crop, usually a wet (Penman, 1963) or irrigated (Allen et al., 1998) short green grass completely shading the ground; or (*iii*) over the actual ecosystem transpiring under unstressed conditions (Brutsaert, 1982;Granger, 1989).

A second source of disagreement on the definition of E_p relates to the spatial extent of the reference system and the consideration (or not) of feedbacks from the reference system on the atmospheric conditions. Several authors found it convenient to define E_p taking an extensive area as reference system, because this reduces the influence of advection and entrainment flows (Penman, 1963;Priestley and Taylor, 1972;Brutsaert, 1982;Shuttleworth, 1993). Such an idealized extensive and well-watered ecosystem evaporating at 'maximal rate' (for the given atmospheric conditions) can be expected to raise air humidity until the vapour pressure deficit tends to zero. In this case, evaporation is only driven by radiative and no longer by aerodynamic forcing. Meanwhile, others have defended the use of reference systems that are infinitesimally small (Morton, 1983;Pettijohn and Salvucci, 2009;Gentine et al., 2011b), in order to avoid the feedback of the reference system on aerodynamic forcing. The effect of the choice of reference system is best exemplified by the complementary relationship framework (Bouchet, 1964), which uses both approaches to link potential and actual evaporation, through $(1+b)\,E_{\mathrm{p0}} = E_{\mathrm{pa}} + \,b\,E_{\mathrm{a}}$, with b an empirical constant (Kahler and Brutsaert, 2006;Aminzadeh et al., 2016), E_{p0} the evaporation from an extensive well-watered surface (i.e. in which the feedback from the ecosystem on the VPD and aerodynamic forcing is considered and where evaporation is only driven by a radiative forcing), E_{pa} the evaporation from a well-watered but infinitesimally small surface (i.e. where evaporation is driven by both radiative and aerodynamic forcing) and E_a the actual evaporation (Morton, 1983).

Upon all this controversy, the net radiation of the reference system remains another point of discussion: some scientists argue that the (well-watered) reference system should have the same net radiation as the actual (water-limited) system (e.g. Granger, 1989;Rind et al., 1990;Crago and Crowley, 2005). Yet, this is inherently inconsistent as the surface temperature reflects the surface energy partitioning, thus a well-watered system transpiring at potential rate is expected to have a lower surface temperature (Maes and Steppe, 2012), and correspondingly a higher net radiation (e.g. Lhomme, 1997;Lhomme and Guilioni, 2006). Meanwhile, to some extent, the albedo also depends on soil moisture (Eltahir, 1998;Roerink et al., 2000;Teuling and Seneviratne, 2008) and it can be argued that it should be adjusted to reflect well-watered conditions (Shuttleworth, 1993). Finally, extensive reference surfaces can be expected to exert a feedback, not only on the aerodynamic forcing, but also on the incoming radiation (via impacts on air temperature, humidity and cloud formation). Yet, these larger-scale feedbacks are not acknowledged when computing E_p, even when considering extensive reference systems.

As it can be concluded from the above discussion, a unique and universally accepted definition of E_p does not exist, and the most appropriate definition remains tied to the specific interest and application. Nonetheless, as different applications make use of different E_p formulations, a good knowledge of the implications of the choice for a specific method is required (Fisher et al., 2011). For terrestrial ecosystems, the use of an open water reference system is uninformative of the actual available energy and the aerodynamic properties of the actual ecosystem (Shuttleworth, 1993;Lhomme, 1997). The approach of considering an idealised well-watered crop system only takes climate forcing conditions into account and not the actual land cover; as such, it has become the standard to estimate aridity and trends in global drought (Dai, 2011). When the actual ecosystem transpiring at unstressed rates is considered as reference system, both climate forcing conditions and ecosystem properties are taken into account. This has been the preferred approach when calculating E_p as an intermediate step to estimate actual evaporation, often by applying a multiplicative stress factor (S) varying between 0 and 1, such that E_a=S E_p (e.g. Barton 1979;Mu et al., 2007;Fisher et al., 2008;Miralles et al., 2011;Martens et al., 2017). This S factor can be considered analogous to the β factor used in some land surface models to incorporate the effect of soil moisture in the estimation of gross primary production and surface turbulent fluxes (Powell et al., 2013).

Several studies have attempted to compare and evaluate different E_p methods. Some of these studies have compared the performance of different E_p formulations in hydrological (Xu and Singh, 2002;Oudin et al., 2005a;Kay and Davies, 2008;Seiller and Anctil, 2016) or climate models (Weiß and Menzel, 2008;Lofgren et al., 2011;Milly and Dunne, 2016). Others considered the Penman-Monteith method as the benchmark to test less input-demanding formulations (e.g. Chen et al., 2005;Sentelhas et al., 2010). All these studies have their own merits, yet an evaluation of E_p methods based on empirical data of actual evaporation measurements is to be preferred (Lhomme, 1997). To date, such approaches have been hampered by limited data availability (Weiß and Menzel, 2008). Lysimeters provide arguably the most precise evaporation measurements available (e.g. Abtew, 1996;Pereira and Pruitt, 2004;Yoder et al., 2005;Katerji and Rana, 2011), but are sparsely distributed and not always representative of larger ecosystems. Pan evaporation measurements are more easy to perform and are broadly available (Zhou et al., 2006;Donohue et al., 2010;McVicar et al., 2012) but provide a proxy of open-water evaporation, rather than actual ecosystem potential evaporation; they also exhibit biases related to the location, shape and composition of the instrument (Pettijohn and Salvucci, 2009). Eddy-covariance measurements are an attractive alternative, but, apart from an unpublished study by Palmer et al. (2012), have so far only been used in E_p studies focusing on local to regional scales (Jacobs et al., 2004;Sumner and Jacobs, 2005;Douglas et al., 2009;Li et al., 2016).

The overall purpose of the present work is to identify the most suitable method to estimate E_p at the ecosystem-scale across the globe. Because we are using an empirical dataset of actual evaporation at FLUXNET sites, the reference system considered in this study is the actual ecosystem, so E_p is defined as the evaporation of the actual ecosystem when it is completely unstressed. As mentioned above, this definition is the most suitable for hydrological studies, studies of ecosystem drought, and derivations of actual evaporation through constraining E_p calculations. Following this definition, E_p is similar to E_{p0} in the complementary relationship. We used the most recent and complete eddy-covariance database available, i.e. the FLUXNET2015 archive (http://fluxnet.fluxdata.org/). The most frequently-adopted E_p methods are applied based on standard parameterizations as well as calibrated parameters by biome, and inter-compared in order to gain insights into the most adequate means to estimate E_p from ecosystem to global scales.

2. Material and Methods

2.1. Selection of E_p methods

Methods to calculate E_p can be categorized based on the amount and type of input data required. In this overview, we will only discuss the ones evaluated in our study, which are arguably the most frequently used.

Methods based on radiation, temperature, wind speed and vapour pressure

The well-known Penman-Monteith equation (Monteith, 1965) expresses latent heat flux λE_a (W m^{-2}) as:

$$\lambda E_\mathrm{a} = \frac{s\,(\mathrm{R_n}-\mathrm{G}) + \frac{\rho_\mathrm{a}\,\mathrm{c_p}\mathrm{VPD}}{\mathrm{r_{aH}}}}{s+\gamma+\gamma\frac{\mathrm{r_c}}{\mathrm{r_{aH}}}} \tag{1}$$

with λ being the latent heat of vaporisation (J kg^{-1}), E_a the actual evaporation (kg m^{-2} s^{-1}), s the slope of the Clausius-Clapeyron curve relating air temperature with the saturation vapour pressure (Pa K^{-1}), R_n the net radiation (W m^{-2}), G the ground heat flux (W m^{-2}), ρ_a the air density (kg m^{-3}), γ the psychrometric constant (Pa K^{-1}), c_p the specific heat capacity of the air (J kg^{-1} K^{-1}), VPD the vapour pressure deficit (Pa), r_{aH} the resistance of heat transfer to air (s m^{-1}), r_c the surface resistance of water transfer (s m^{-1}). While λ, c_p, s and γ are air temperature-dependent, r_{aH} is a complex function of wind speed, vegetation characteristics and the stability of the lower atmosphere (see Section 2.3). In most methods to estimate E_a or E_p, r_{aH} is estimated from a simple function of wind speed.

The Penman-Monteith equation is frequently used to calculate E_p by adjusting r_c to its minimum value (the value under unstressed conditions). If the reference system is the actual system or a reference crop, r_c is usually considered a fixed, constant value larger than zero (even if the soil is well watered). In this study, both a universal, fixed value of r_c for reference crops and a biome-specific constant value are used (see Sect. 2.5). When instead of a well-watered soil, a wet canopy (i.e. a canopy covered by water) is considered, r_c=0 and Eq. (1) collapses to:

$$\lambda E_\mathrm{p} = \frac{s\,(\mathrm{R_n}-\mathrm{G}) + \frac{\rho_\mathrm{a}\,\mathrm{c_p}\mathrm{VPD}}{\mathrm{r_{aH}}}}{s+\gamma} \tag{2}$$

Eq. (2) is often referred to as the Penman (1948) formulation, and can be conveniently rearranged as $\lambda E_\mathrm{p} = \frac{s\,(R_n-G)}{s+\gamma} + \frac{\rho_a\,c_p VPD}{(s+\gamma)\,r_{aH}}$ to illustrate that E_p is driven by a radiative (left term) and an aerodynamic (right term) forcing (Brutsaert and Stricker, 1979).

Methods based on radiation and temperature

When the reference system is considered an idealized extensive area, or when radiative forcing is very dominant, the aerodynamic component of Eq. (2) may become negligible, thus the whole equation collapses to $\lambda E_\mathrm{p} = \frac{s\,(\mathrm{R_n}-\mathrm{G})}{s+\gamma}$, which is commonly referred to as 'equilibrium evaporation' (Slatyer and McIlroy, 1961). Priestley and Taylor (1972) analysed time series of open water and water-saturated crops and grasslands and found that the evaporation over these surfaces closely matched the equilibrium evaporation corrected by a multiplicative factor, commonly denoted as α_{PT}:

$$\lambda E_\mathrm{p} = \alpha_\mathrm{PT}\frac{s\,(\mathrm{R_n}-\mathrm{G})}{s+\gamma} \tag{3}$$

This formulation is known as the Priestley and Taylor equation. Because usually a constant value of α_{PT} is adopted, it assumes that the aerodynamic term in the Penman equation (eq. 2) is a constant fraction of the radiative term. Typically, α_{PT} =1.26 is considered, as estimated by Priestley and Taylor (1972) in their original experiments. In this study, we also include a biome-specific value to extend its applicability to all biomes (see Sect 2.5). Since this method does not require wind speed or VPD as input, it is widely applied in hydrological models (Norman et al., 1995;Castellvi et al., 2001;Agam et al., 2010), remote sensing evaporation models (Norman et al., 1995;Fisher et al., 2008;Agam et al., 2010;Miralles et al., 2011) and drought monitoring (Anderson et al., 1997;Vicente-Serrano et al., 2018).

Methods based on radiation

Other studies such as Lofgren et al. (2011), or the more recent Milly and Dunne (2016), further simplified Eq. (3) to make it a linear function of the available energy by defining a constant multiplier here referred to as α_{MD}:

$$\lambda E_p = \alpha_{MD}\,(R_n - G) \tag{4}$$

In the case of Milly and Dunne (2016) this equation was applied to climate model outputs based on a constant and universal value of α_{MD}=0.8. On a daily scale, $(R_n - G)$ expresses the total amount of energy available for evaporation, and the fraction of this energy that is actually used for evaporation is typically referred to 'evaporative fraction', or $EF = \frac{\lambda E_a}{(H+\lambda E_a)} = \frac{\lambda E_a}{(R_n - G)}$. From Eq. (4), it follows that the parameter α_{MD} can be interpreted as the EF of the unstressed ecosystem. In this study, we test both the general value of α_{MD}=0.8 and a biome-specific constant (see Sect. 2.5).

Methods based on temperature

Of the many empirical methods to estimate E_p, temperature-based methods have arguably been the most commonly used because of the availability of reliable air temperature data. For an overview of these methods, we refer to Oudin et al. (2005a). In this study, three methods are included. First, Pereira and Pruitt (2004) formulated a daily version of the well-known Thornthwaite (1948) equation:

$$T_{eff}<0 \qquad \lambda E_p = 0 \tag{5a}$$

$$0<T_{eff}<26 \qquad \lambda E_p = \alpha_{Th}\left(\frac{10\,T_{eff}}{I}\right)^{b}\left(\frac{N}{360}\right) \tag{5b}$$

$$26<T_{eff} \qquad \lambda E_p = -c + d\,T_{eff} - e\,T_{eff}^{\,2} \tag{5c}$$

with T_{eff} the effective temperature, based on maximum and minimal temperatures (see further, Section 2.5), α_{Th} an empirical parameter (see below), I the yearly sum of $(T_{a\_month}/5)^{1.514}$, with T_{a_month} the mean air temperature for each month, N the number of daylight hours, b a parameter depending on I and c, d and e empirical constants (see Section 2.5). The general value of α_{Th}=16 is often adopted; in this study, we will also calculate and apply a biome-specific value.

The second temperature-based formulation is that proposed by Oudin et al. (2005a), selected and developed after comparing 27 physically-based and empirical methods with runoff data from 308 catchments:

$$T_a<5 \qquad \lambda E_p = 0 \tag{6a}$$

$$T_a>5 \qquad \lambda E_p = \frac{R_e}{\rho_a}\,\frac{(T_a - 5)}{\alpha_{Ou}} \tag{6b}$$

with T_a the air temperature (°C) and R_e top-of-atmosphere radiation (MJ m^{-2} day^{-1}), depending on latitude and Julian day. Oudin et al. (2005a) suggested to use α_{Ou}=100. This value will be used, in addition to a biome-specific value.

Finally, the third temperature-based method is the Hargreaves and Samani (1985) formulation, which includes minimum (T_{min}) and maximum (T_{max}) daily temperature, next to T_a and R_e:

$$\lambda E_p = \alpha_{HS}\,R_e\,(T_a + 17.8)\sqrt{T_{max} - T_{min}} \tag{7}$$

with α_{HS} a constant, normally assumed to equal 0.0023. As for the other methods, we additionally apply a biome-specific value. A detailed description of the calibration of all E_p methods is given in Section 2.5.

2.2. FLUXNET2015 database

The Tier2 FLUXNET2015 database based on half-hourly or hourly measurements from eddy-covariance sensors is used to evaluate the different E_p formulations (http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/). Sites lacking at least one of the basic measurements required for our analysis (i.e. R_n, G, λE_a, H, precipitation, wind speed (u), friction velocity (u_*), T_a and relative humidity (RH) or VPD) were not further considered. For latent and sensible heat fluxes, we used the data corrected by the Bowen ratio method. In this approach, the Bowen ratio is assumed to be correct, and the measured λE_a and H are multiplied by a correction factor derived from a moving window method; see http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/data-processing/ for a detailed description of this standard procedure. Nonetheless, taking the uncorrected λE_a instead did not impact the main findings (not shown). For the main heat fluxes (G, H, λE_a), medium and poor gap-filled data were masked out according to the flags provided by FLUXNET. As no quality flag was available for R_n measurements, the flag of the shortwave incoming radiation was used instead. All negative values for H or λE_a were masked out, as these relate to periods of interception loss and condensation when accurate measurements are not guaranteed (Mizutani et al., 1997). Similarly, all negative values of R_n were masked out.

Finally, sub-daily measurements were aggregated to daytime composites by applying a minimum threshold of 5 W m^{-2} of top-of-atmosphere incoming shortwave radiation, and after excluding the first and last (half-) hours from these aggregates. Based on these daytime aggregates, the daytime means of s, γ and ρ_a were calculated using the parameterisation procedure described by Allen et al. (1998). We used T_a to calculate s. Only days in which more than 70% of the data were measured directly were retained, and days with rainfall (between midnight and sunset) were removed from the analyses to avoid the effects of rainfall interception. Furthermore, only sites with at least 80 retained days were used for the further analysis. The global distribution of the final selection of sites is shown in Fig. 1 and detailed information about these sites is provided in Table S1 of the Supporting Information. The IGBP classification was used to assign a biome to each site.

(Insert Figure 1)

2.3. Calculation of resistance parameters

Estimates of r_{aH} are required for the Penman and Penman-Monteith equations. The resistance of heat transfer to air, r_{aH}, was calculated as:

$$r_{aH} = \frac{u}{u_*^2} + \frac{1}{k\,u_*}\left[\ln\left(\frac{z_{0m}}{z_{0h}}\right) + \Psi_m\left(\frac{z-d}{\mathrm{L}}\right) - \Psi_m\left(\frac{z_{0h}}{\mathrm{L}}\right) - \Psi_h\left(\frac{z-d}{\mathrm{L}}\right) + \Psi_h\left(\frac{z_{0h}}{\mathrm{L}}\right)\right] \tag{8}$$

in which k=0.41 is the von Karman constant, z the (wind) sensor height (m), d the zero displacement height (m), z_{0m} and z_{0h} the roughness lengths for momentum and sensible heat transfer (m), respectively, L the Obukhov length (m), and $\Psi_m(X)$ and $\Psi_h(X)$ the Businger-Dyer stability functions for momentum and heat for the variable X, respectively. These were calculated based on the equations given by Garratt (1992) and Li et al. (2017) for stable, neutral and unstable conditions. Note that in neutral and stable conditions, $\Psi_m(X) = \Psi_h(X)$ and that $\Psi_m\left(\frac{z-d}{\mathrm{L}}\right) - \Psi_m\left(\frac{z_{0h}}{\mathrm{L}}\right) - \Psi_h\left(\frac{z-d}{\mathrm{L}}\right) + \Psi_h\left(\frac{z_{0h}}{\mathrm{L}}\right) = 0$. This is not the case for the unstable conditions that usually prevail during the daytime. Daytime averages of all variables were used as input in Eq. (8). The sensor height z was collected individually for each tower through online and literature research, or personal communication with the towers' P.I. The Obukhov length L was calculated as:

$$\mathrm{L} = \frac{-\,u_*^3\,\rho_a\,\mathrm{T_a}\,(1 + 0.61\,\mathrm{q_a})\,\mathrm{c_p}}{k\,g\,\mathrm{H}} \tag{9}$$

with $\mathrm{q_a}$ being the specific humidity (kg $\mathrm{kg^{-1}}$) and g=9.81 m $\mathrm{s^{-2}}$ the gravitational acceleration.

The displacement height d and the roughness length for momentum flux z_{0m} were estimated as a function of the canopy vegetation height (VH), as d=0.66 VH and z_{0m}=0.1 VH (Brutsaert, 1982). VH was estimated from the flux tower measurements using the approach of Pennypacker and Baldocchi (2016):

$$\mathrm{VH} = \frac{z}{0.66 + 0.1\,\exp\left(\frac{k\,u}{u_*}\right)} \tag{10}$$

This equation was applied to the full (half-)hourly database, and only when conditions were near-neutral ($|z/\mathrm{L}| < 0.01$) and friction velocities lower than one standard deviation below the mean u_* at each site. The daily VH was then aggregated by averaging out the half-hourly estimates to daily values, excluding the 20% outliers of the data, and then calculating a 30-day window moving average on the dataset. When not enough (half-)hourly vegetation height observations (<160) were available, the site was excluded from the analysis. This gave robust results for all remaining sites; an example of VH temporal development for a specific location is given in Fig. 2a. For homogeneous canopies, the VH calculated this way represents the true vegetation height. For savannah-like ecosystems, it corresponds to the vegetation height that the vegetation would have if it was represented by a single big leaf model.

The Stanton number (defined as $k\mathrm{B}^{-1}$=ln(z_{0M}/z_{0H})) was calculated by assuming that the surface aerodynamic temperature $\mathrm{T_0}$ (defined by $\mathrm{H} = \rho_\mathrm{a}\,\mathrm{c_p}\,\frac{(\mathrm{T_0} - \mathrm{T_a})}{\mathrm{r_{aH}}}$) is equal to the radiative surface temperature $\mathrm{T_s}$ derived from the longwave fluxes. Then, through an iterative approach, an optimal value of z_{0h} was obtained, using the following equations for $\mathrm{T_0}$ (Garratt, 1992) and $\mathrm{T_s}$ (Maes and Steppe, 2012):

$$\mathrm{T_0} = \mathrm{T_a} + \left(\frac{\mathrm{H}}{k\,u_*\,\rho_\mathrm{a}\,\mathrm{c_p}}\right)\left[\ln\left(\frac{z-d}{z_{0h}}\right) - \Psi_h\left(\frac{z-d}{\mathrm{L}}\right) + \Psi_h\left(\frac{z_{0h}}{\mathrm{L}}\right)\right] \tag{11}$$

$$\mathrm{T_s} = \sqrt[4]{\frac{\mathrm{LW_{out}} - (1-\varepsilon)\,\mathrm{LW_{in}}}{\sigma\,\varepsilon}} \tag{12}$$

with σ the Stefan-Boltzmann constant and ε the emissivity (see further). The (half-)hourly data were used for this calculation. Following the approach of Li et al. (2017), only summertime data were used, and only when H was larger than 20 W $\mathrm{m^{-2}}$ and u_* larger than 0.01 m $\mathrm{s^{-1}}$. Summertime was defined as those months in which the maximal daily value of H is at least 85% of the 98$^{\mathrm{th}}$ percentile of H based on the entire tower time series. In addition, (half-)hourly observations with counter-gradient heat fluxes were excluded from the analysis. For each observation, z_{0h} was optimized by minimizing the difference between $\mathrm{T_0}$ and $\mathrm{T_s}$. Then, $k\mathrm{B}^{-1}$ was calculated at each site based on its relation with the observed Reynolds number (Re) by fitting the following function, based on the work by Li et al. (2017):

$$k\,\mathrm{B}^{-1} = a_0 + a_1\,\mathrm{Re}^{a2} \tag{13}$$

Note that Eq. (12) requires a value for ε, which is often assumed to be equal to 0.98 for all sites (e.g. Li et al., 2017;Rigden and Salvucci, 2015). Under the assumption that $\mathrm{T_0}$=$\mathrm{T_s}$, ε can also be calculated separately per site. If H=0, it follows that $\mathrm{T_0}$=$\mathrm{T_a}$ and from Eq. (12),

$$\varepsilon = \frac{\mathrm{LW_{out}} - \mathrm{LW_{in}}}{\sigma\,\mathrm{T_a^4} - \mathrm{LW_{in}}} \tag{14}$$

Here, ε was calculated for each site using (half-)hourly data, selecting those measurements where H was close to 0 ($-2<H<2$ W m^{-2}) and excluding rainy days as well as measurements in which the albedo (calculated as $\alpha=SW_{out}/SW_{in}$ with SW_{in} the incoming and SW_{out} the outgoing shortwave radiation) was above 0.4, to avoid influences of snow or ice. Negative estimates of ε were masked out, and the ε of the site was calculated as the mean, after excluding the outlying 20% of the data. Equation 3 was applied both with a fixed ε of 0.98 and with the observed ε, and the ε value yielding the lowest RMSE in Eq. (12) was retained for each site. An example of such a function between kB^{-1} and Re is shown in Fig. 2b.

(Insert Figure 2)

Finally, the surface resistance r_c (s m^{-1}) was calculated by inverting the Penman-Monteith equation as:

$$\mathrm{r_c} = \frac{s\,(\mathrm{R_n}-\mathrm{G})\,\mathrm{r_{aH}} + \rho_\mathrm{a}\,\mathrm{c_p}\mathrm{VPD}}{\gamma\,\lambda E_\mathrm{a}} - \frac{(s+\gamma)\,\mathrm{r_{aH}}}{\gamma} \quad (15)$$

We converted the r_c estimates to surface conductance g_c (mm s^{-1}) using $g_c=1000\ r_c^{-1}$; we will continue using g_c (rather than r_c) for the remainder of the study. Note that the approach of calculating kB^{-1} directly requires a separate measurement of LW_{in} and LW_{out}, which was only available in 95 of the 107 selected sites. For the remaining sites, an alternative approach was used to calculate kB^{-1} (see Supporting Information).

2.4. Selection of unstressed days

We include two different approaches to identify a subset of measurements per eddy-covariance site in which the ecosystem was unstressed and provide the results for both methods. A first approach is based on soil moisture levels. For those sites where soil moisture measurements were available, the maximal soil moisture level for each site was determined as the 98th percentile of all soil moisture measurements. We then split up the dataset of each site in 5 equal-size classes based on evaporation percentiles. For each class, days having soil moisture levels belonging to the highest 5% of soil moisture levels within each class were selected as unstressed days, but only if the soil moisture level of these selected days was above 75% of the maximum soil moisture. This guaranties the sampling of unstressed evaporation during all seasons.

As soil moisture data were not available for a large number of sites, and because using soil moisture data does not exclude days in which evaporation may be constrained by other kinds of biotic or abiotic stress, a second approach for defining unstressed days was applied based on an energy balance criterion. We calculated the EF from the daytime λE_a and H values, and considered it as a direct proxy for evaporative stress, i.e. we assumed that under unstressed conditions, a larger fraction of the available energy is used to evaporate water (Gentine et al., 2007; 2011a; Maes et al., 2011). This approach is similar to those of other E_p studies using eddy-covariance or lysimeter data, in which the Bowen ratio (e.g. Douglas et al., 2009) or the ratio of $\lambda E_a/(SW_{in}+LW_{in})$ (Pereira and Pruitt, 2004) are used to define unstressed days. The unstressed record was comprised of all days with EF exceeding the 95th percentile EF threshold for each particular site; or, if fewer than 15 days fulfilled this criterion, the 15 days with the highest EF. Consequently, we assume that at each site during at least 5% of the days the conditions are such that evaporation is unstressed and E_a reflects E_p. The measured E_a from the identified unstressed days is further referred to as E_{unstr} (mm day^{-1}) and used as reference data to evaluate the different E_p methods.

To assess whether the atmospheric conditions of the unstressed datasets are representative for the FLUXNET sites as such, a random bootstrap sample having the same number of records as the unstressed dataset was taken from the entire dataset of daily records. The mean, standard deviation, 2nd and 98th percentile of SW_{in}, T_a, VPD and u were calculated. This procedure was repeated 1000 times per site. A t-test comparing the values of the unstressed subsample with those of the 1000 random samples was used to analyse whether the atmospheric conditions of unstressed subsample were representative for the overall site conditions. This analysis was carried out for both methods to select unstressed days: the soil moisture threshold and the energy balance criterion.

2.5. Calculation and calibration of the different E_p methods

An overview of the different methods to calculate E_p is given in Table 1. If possible, three versions of each method were calculated: (1) a reference crop version, (2) a standard version, and (3) a biome-specific version. The reference crop version calculates E_p for the reference short turf grass crop, and with R_n and other properties of this reference crop as well. The standard version uses the same non-biome-specific parameters of the reference crop but considers R_n and other properties of the actual ecosystem. The biome-specific version of each method applies a calibration of the key parameters per biome (Table 1) and considers $R_n$and other properties of the actual ecosystem. These calibrated values per biome are based on the mean value of this key parameter for the unstressed dataset for each site, averaged out per biome type.

To estimate the radiation and crop properties of the reference crop versions, the equations described by Allen et al. (1998) in the FAO-56 method (Food and Agricultural Organization) were used and G was considered to be 0. R_n was calculated as:

$$R_n=SW_{in}(1-\alpha_{ref}) + LW^* \tag{16}$$

with α_{ref}=0.23 (Allen et al., 1998) and LW^* the net longwave radiation, calculated after Allen et al. (1998; Eq. (39), Chapter 3).

In the case of the reference crop version of the Penman-Monteith equation (Eq. (1)), the FAO-56 method was used as described by Allen et al. (1998), with g_{c_ref} fixed as 14.49 mm s^{-1} (corresponding with r_{c_ref}=69 s m^{-1}) and using Eq. (16) to calculate R_n. The standard version of the Penman-Monteith equation used observed (R_n, G, VPD) and calculated (s, γ, ρ_a, r_{aH}) daytime values as described in Section 2.2 in Eq. (1), and also assumed g_{c_ref}=14.49 mm s^{-1}. The biome-specific version was calculated with the same data but used a biome-dependent value of g_c. First, for each individual site, the unstressed g_c was calculated as the mean of the g_c values of the unstressed record (see Section 2.4). The mean value per biome was then calculated from these unstressed g_c values. Regarding the Penman method (Eq. (2)), the reference crop and standard versions were calculated using the same input data as for the Penman-Monteith methods; given Penman's consideration of no surface resistance, no biome-specific version was calculated.

The reference crop version of the Priestley and Taylor method was calculated from Eq. (3) with R_n from Eq. (16), s and γ from the FAO-56 calculations, and with α_{PT}=1.26. The standard version used the same value for α_{PT} but the observed daytime values for R_n and G. The biome-specific version followed a calibration of α_{PT} similar to the g_{c_ref} calculation. For each site, the unstressed α_{PT} was calculated as the average α_{PT}, obtained by solving Eq. (3) for α_{PT} using the unstressed dataset. Finally, the mean per biome was calculated and used in the E_p estimation. Regarding the method by Milly and Dunne (2017) (Eq. (4)), the reference crop, standard and biome-specific calculation were calculated accordingly, with R_n from Eq. (16) for the reference crop version, α_{MD}=0.8 for the reference crop and standard version, and a calibrated α_{MD} by biome type for the biome-specific version.

For Thornthwaite's, Oudin's and Hargreaves-Samani's formulations (Eqs. (5), (6) and (7)), only the standard and biome-specific versions were calculated. The standard version of Thornthwaite's formulation used α_{Th}=16 . In the biome-specific version, this parameter was again calculated per site as the mean value of the unstressed records (e.g. Xu and

Singh, 2001;Bautista et al., 2009), and then averaged per biome type. The effective temperature T_{eff} was calculated as $T_{eff} = 0.36\,(3\,T_{max} - T_{min})$ (Camargo et al., 1999). The parameter b was calculated as $b = (6.75\,10^{-7} I^3) - (7.71\,10^{-7} I^2) + 0.0179\,I + 0.492$ and the parameters c, d and e in Eq. (5c) were 415.85, 32.24 and 0.43, respectively (Pereira and Pruitt, 2004). For Oudin's temperature-based formulation, α_{Ou}=100 was taken for the standard version (Eq. (6)). In the biome-specific version, this value was recalculated by calculating α_{Ou} for the unstressed records through Eq. (6), calculating the mean α_{Ou} per site and finally the biome-dependent α_{Ou}. Similarly, for the Hargreaves-Samani method, α_{HS}=0.0023 is used for the standard version (Eq. (7)), whereas in the biome-specific version, this value was calculated using the unstressed records. Altogether, this exercise yielded a total of seventeen different methods to estimate E_p whose specificities are documented in Table 1.

(Insert Table 1)

The influence of climatic forcing data on E_{unstr} and on g_{c_ref}, α_{PT} and α_{MD} was investigated. This was done by calculating for each individual site the correlations between the daily estimates of the atmospheric conditions and the daily values of the unstressed datasets. Analyses were then performed on these correlations of all sites.

3. Results

3.1. Representativeness of climate forcing data of unstressed datasets

Table S2 provides an overview of the analyses used to verify if the climatic forcing data of the unstressed subsets were representative for the atmospheric conditions of the sites as such. For the subsets of both unstressed criteria, atmospheric conditions were very representative for the site conditions, including for VPD. For the energy balance criterion, for instance, only at one site, the unstressed subset of the 98th percentile was significantly different from the random sampling-based simulations and in only 2 sites, the 2nd percentile was significantly different from the simulations.

3.2. Key parameters by biome

We focus here on the parameter estimates of the unstressed record based on the energy balance criterion (Section 2.4). Of the full dataset, 107 flux sites meet all the selection criteria (i.e. at least 80 days without rainfall, good quality measurements of radiation and main fluxes, and at least 160 vegetation height observations – see Sections 2.2 and 2.3). Despite considerable variation, g_{c_ref} does not differ statistically across biomes, in contrast with α_{PT} and α_{MD}. Overall, croplands (CRO) are characterised by a higher measured E_{unstr}, which translates in the highest g_{c_ref}, α_{PT} and α_{MD} of all biomes. Grasslands (GRA), deciduous broadleaf forest (DBF) and evergreen broadleaf forest (EBF) also have relatively high g_{c_ref}, α_{PT} and α_{MD}, whereas mixed forests and savannah ecosystems (closed shrubland (CSH), woody savannah (WSA), open shrublands (OSH) and savannah (SAV) ecosystems) are characterised by lower g_{c_ref}, α_{PT} and α_{MD}. Only five sites (DE-Kli and IT-BCi, croplands (CRO); CA-SF3, OSH; AU-Rig, grassland and AU-Wac, evergreen broadleaf forest) have mean values of α_{PT} higher than the typically assumed 1.26 (Table 2). In contrast, 27 sites, including 9 croplands, have a mean value of α_{MD} above 0.80, and 42 sites have mean g_{c_ref} above 14.49 mm s^{-1}. Finally, wetlands (WET) show a large standard deviation of α_{PT} and α_{MD} (Table 2) due to their location in tropical, temperate as well as in arctic regions. The parameters of the three temperature-based differed significantly across biomes, but trends were different for each key parameter and did not always match those for g_{c_ref}, α_{PT} and α_{MD} (Table 2).

Next, the effect of the atmospheric conditions on E_{unstr} and on the key parameters g_{c_ref}, α_{PT} and α_{MD} of the unstressed dataset is investigated. Fig. 3 gives the distribution of the correlations between the climatic variables (R_n, T_a, VPD, u and [CO2]) and E_{unstr}, g_{c_ref}, α_{PT} and α_{MD} of the unstressed records at each site. We did not include α_{Th}, α_{HS} or α_{Ou} because the temperature-based methods did not perform well (see Section 3.3). E_{unstr} was strongly positively correlated with R_n, T_a and VPD in most sites, but less with u (Fig. 3a, Table 3). Considering all sites, the correlation between g_{c_ref} and the climatic variables is not significantly different from zero for any climate variable, yet g_{c_ref} is significantly negatively correlated with T_{air} and with VPD in 40% and 45% of the flux tower sites, respectively (Table 3, Fig. 3b). The two other parameters, α_{PT} and α_{MD}, appear less correlated to any of the climatic variables across all sites (Table 3b). In the case of α_{MD}, in particular, the distributions of the correlations against all climate forcing variables peak around zero (Fig. 3d): α_{MD} is hardly influenced by R_n, and is overall not dependent on u, T_a, [CO_2], or VPD in most sites (Table 3).

(Insert Table 2)
(Insert Table 3)
(Insert Figure 3)

3.3. Evaluation of different E_p methods

We first list the results of the analysis using the energy balance criterion for selecting the unstressed records (Section 2.4). The scatterplots of measured E_{unstr} versus estimated E_p based on the 17 different methods are shown in Fig. 4 for three sites belonging to different biomes. Despite the overall skill shown by the different E_p methods, considerable differences can be appreciated. In general, the methods designed for reference crops (PM_r, Pe_r, PT_r, MD_r) overestimate E_{unstr} and only two methods, MD_B and PT_B, match the measured E_{unstr} closely.

(Insert Figure 4)

Table 4 gives the mean correlation per biome for each method. The results are very consistent and reveal that the highest correlations for nearly all biomes are obtained with the standard and biome-specific radiation-based method (MD_s and MD_b), closely followed by the standard and biome-dependent Priestley and Taylor method (PT_s and PT_b). Temperature-based methods have the lowest overall mean correlation as well as lower mean correlations per biome, with the Hargreaves-Samani method performing slightly better than the other two temperature-based methods. Note that the correlations are the same for the standard and biome-specific version in the case of Priestley and Taylor (PT_s and PT_b), radiation-based (MD_s and MD_b), Oudin (Ou_s and Ou_b) and Hargreaves-Samani (HS_s and HS_b) methods (Table 4) – this is to be expected, as the only difference between the standard and biome-specific version of these methods is the value of their key parameters (α_{PT}, α_{MD}, α_{Ou}, α_{HS}) which are multiplicative (see Eqs. (3), (4), (6) and (7)). Differences are however reflected in the unbiased Root Mean Square Error (unRMSE) and mean bias – see Tables 5 and 6. The biome-specific versions of the radiation-based method (MD_b) and of the Priestley and Taylor method (PT_b) have consistently the lowest unRMSE for all biomes. Though the difference between these two methods is small, MD_b is performing slightly better. The standard Penman method (Pe_s) has the highest unRMSE. All reference crop methods (PM_r, Pe_r, PT_r, MD_r) have mean unRMSE above one mm day^{-1}, and the temperature-based methods (Th_s, Ou_s, HS_s, Th_b, Ou_b and HS_b) also have relatively high unRMSE. Finally, bias estimates are given in Table 6. Again, MD_b is overall the best performing method (mean bias closest to zero mm day^{-1}), closely followed by the PT_b method. Both methods have consistently the bias closest to zero among all biomes, except for wetlands. Most reference crop methods (PM_r, Pe_r, PT_r, MD_r) as well as Pe_s overestimate E_p in all biomes.

(Insert Table 4)
(Insert Table 5)
(Insert Table 6)

The use of soil water content as criterion to select unstressed days (see Sect. 2.4) is explored. In total, 62 sites have soil water content data and meet the other selection criteria documented in Section 2.2. The results of this analysis are given in Tables S3–S5 of the supporting section. To allow for a fair comparison, the same statistics have also been computed for just the same 62 tower sites with the energy balance-criterion (Tables S6-S8). Using the soil moisture criterion, the correlations are overall lower and the results of the mean correlation, unRMSE and biases are less consistent. However, the overall performance ranking of the different models remains similar: PT_b is the best performing method with overall the highest mean correlation (R=0.84) and the lowest unRMSE (0.78 mm/day) and a bias close to zero (-0.04), closely followed by the MD_b method, with R=0.81, unRMSE=0.89 and a mean bias of -0.12. More complex Penman-based models, and especially the empirical temperature-based formulations, show again a lower performance.

So far, all flux sites were used to calibrate the key parameters (Table 2) and those same sites were also used for the evaluation of the different methods. This was done to maximise the sample size. However, to avoid possible overfitting, we also repeated the analysis after separating calibration and validation samples. For each biome, two-thirds of the sites were randomly selected as calibration sites, and one third as validation sites. The key parameters were then calculated from the calibration subset, and applied to estimate E_p of the biome-specific methods of the validation subset. This procedure was repeated 100 times and the mean correlation, unbiased RMSE and bias per biome were calculated. Results show no substantial differences in overall correlation, unRMSE and bias of each method, and are provided in Tables S9–S11 for completeness.

4 Discussion

4.1. Comparison of criteria to define unstressed days

We prioritised the energy balance over the soil moisture criterion to select unstressed days (see Sect. 2.4), because it can be applied to sites without soil moisture measurements and because it implicitly allows the exclusion of days in which the ecosystem is stressed for reasons other than soil moisture availability (e.g. insect plagues, phenological leaf-out, fires, heat and atmospheric dryness stress, nutrient limitations). In addition, soil moisture at specific depths can be a poor indicator of water stress, as rooting depth can vary and is not accurately measured (Powell et al., 2006;Douglas et al., 2009;Martínez-Vilalta et al., 2014). This is confirmed by our results: sampling unstressed days based on the energy balance-based criterion resulted in higher correlations between E_p and E_{unstr} for all methods (Table S6 vs Table S3) and in lower unRMSE, with the exception of the temperature-based methods (Table S7 vs Table S4).

Nonetheless, it could be argued that, because the MD method assumes a constant evaporative fraction, the use of the evaporative fraction as a criterion for selecting unstressed days may favour the MD and even the closely related PT formulation. However, the soil moisture criterion adopted here provides an independent check of the results, and confirms the robust and superior performance of the energy-driven PT_b and MD_b methods, independently of the framework chosen to select unstressed days. In the following discussion, the primary focus is on the results of selecting unstressed days based on the energy balance criterion.

4.2. Estimation of key ecosystem parameters

The resulting biome-specific values of the key parameters in Table 2 are within the range of values used in reference crop and standard applications of the models (Table 1), with the exception of α_{PT}, which is typically lower than the frequently adopted value of 1.26. Other studies also found α_{PT} values far below 1.26 but within the range of our study, mainly for forests (e.g. Shuttleworth and Calder, 1979;Viswanadham et al., 1991;Eaton et al., 2001;Komatsu, 2005), but also for tundra (Eaton et al., 2001) and grassland sites (Katerji and Rana, 2011) – see McMahon et al. (2013) for an overview. Our results and these studies demonstrate that the standard level of α_{PT}=1.26 is close to the upper bound and will lead to an overestimation of E_p at most flux sites (Table 5).

4.3. Lower performance of the Penman-Monteith method

The poor performance of the PM_r, PM_s and PM_b methods was relatively unexpected. Because the Penman-Monteith method incorporates the effects of T_a, VPD, R_n and u, it is often considered superior (e.g. Sheffield et al., 2012), and is even used as reference to evaluate other E_p formulations (e.g. Xu and Singh, 2002;Oudin et al., 2005b;Sentelhas et al., 2010). However, in studies dedicated to estimate E_a at eddy-covariance sites, in which g_c is adjusted so it reflects the actual stress conditions in the ecosystem, the Penman-Monteith method has already been shown to perform worse than other, simpler methods, such as the Priestley and Taylor equation (e.g. Ershadi et al., 2014;Michel et al., 2016). It is well known that its performance depends on the reliability of a wide range of input data, and on the methods used to derive r_{aH} and g_c (Singh and Xu, 1997;Dolman et al., 2014;Seiller and Anctil, 2016). In our case, the strict procedure followed to select the samples of 107 FLUXNET sites (see Sect. 2.2) ensured that all relevant variables were available, and the meteorological measurements were quality-checked. Hence, in our analysis, input quality is unlikely to be the cause of low performance.

We believe that the underlying assumption of a constant g_c, typically adopted by PM methods (PM_r, PM_s, PM_b) when estimating E_p, is a more likely explanation of the poor performance. PM was the only method of which the biome-specific calibration did not improve the performance. This is partially because of the large variation in g_{c_ref} among the different flux sites of the same biome type (Table 2). In addition, of all the key parameters, g_{c_ref} showed the largest mean relative standard deviation of the unstressed datasets at individual sites (*results not shown*). Surface conductance of the unstressed dataset was significantly negatively correlated with VPD in 45% of the sites (Fig. 3b, Table 3). The relationship between g_c and VPD for two of these sites is illustrated in Fig. 5. It is clear that g_c of unstressed days (red dots) is always high for a given VPD, confirming the validity of the energy balance method. However, for these sites, it is also shown that g_c of the unstressed days becomes very high when VPD becomes very low. As a consequence, the mean value of g_c of the unstressed records, used to ultimately calculate g_{c_ref} per biome type, is highly influenced by local VPD and is not necessarily a representative ecosystem property.

The dependence of g_c on VPD, even when soil moisture is not limiting, has been well studied (e.g. Jones, 1992;Granier et al., 2000;Sumner and Jacobs, 2005;Novick et al., 2016) and incorporated in most conventional stomatal or surface conductance models (e.g. Jarvis, 1976;Ball et al., 1987;Leuning, 1995). Yet, out of practical reasons, g_{c_ref} is usually considered constant in E_p-methods using the Penman-Monteith approach, with the PM_r as best illustration. Our data confirm that in unstressed conditions stomata open maximally only when VPD is very low. As such, the VPD-dependence of g_c smooths the impact of VPD in the Penman-Monteith equation - drops in VPD are compensated by increases in g_c, and vice versa, lowering the impact of VPD on E_a (Eq. 1). As such, assuming a constant g_{c_ref} value overestimates the influence of VPD (and wind speed) on E_p.

Moreover, assuming a constant g_{c_ref}-value in the Penman-Monteith method also ignores the influence of CO_2 levels on g_c. As a result, Milly and Dunne (2016) found that the Penman-Monteith methods with constant g_{c_ref} overpredicted E_p in models estimating future water use. Calibrating the sensitivity of g_{c_ref} to VPD and [CO_2] in the Penman-Monteith equation is outside the scope of this study, but could certainly improve E_p calculations – yet, it would further increase the complexity of the model. Finally, we note that the above discussion also applies to Penman's method: taking a wet canopy as reference in the Penman method (g_c=∞ or r_c=0) may not only overestimate E_p (Table 6) but also the influence of VPD and u on E_p.

(Insert Figure 5)

4.4. Considerations regarding simple energy-based methods

The simpler Priestley and Taylor and radiation-based methods came forward as the best methods for assessing E_p with both criteria to define unstressed days. These observations are in agreement with studies highlighting radiation as the dominant driver of evaporation of saturated or unstressed ecosystems (e.g. Priestley and Taylor, 1972;Abtew, 1996;Wang et al., 2007;Song et al., 2017;Chan et al., 2018). It also agrees with Douglas et al. (2009) who found that PT outperformed the PM method for estimating unstressed evaporation in 18 FLUXNET sites.

Both PT_b and MD_b are attractive from a modelling perspective, as they require a minimum of input data. However, this simplicity can also hold risks. The Priestley and Taylor method has been criticised for the implicit assumptions, which are also present in the MD_b-method. For instance, by not incorporating wind speed explicitly, it is assumed that the effect of u on E_p is somehow embedded within α_{PT} (or α_{MD}). Yet, several studies indicate that wind speeds are decreasing ('stilling') globally (McVicar et al., 2008;Vautard et al., 2010;McVicar et al., 2012). McVicar et al. (2012) also reported an associated decreasing trend in observed pan evaporation worldwide as well as in E_p calculated with the PM_r method. With PT (or MD) methods, this trend cannot be captured. A similar criticism can be drawn with regards to the effect of [CO_2] on stomatal conductance, water use efficiency, and thus potential evaporation (Field et al., 1995). A separate question is whether more complex E_p methods that incorporate the effects of u, [CO_2] or VPD explicitly do this correctly; the above-mentioned issues about the fixed parameterisation of the Penman-Monteith methods for estimating E_p indicate that this may typically not be the case.

Regarding the non-explicit consideration of u by simpler methods, our records show a limited effect of u on E_a and E_p, even when considering larger temporal scales. Of the 16 flux towers with at least 10 years of evaporation data, we calculated the yearly average E_a as well as the annual mean climatic forcing variables. Yearly averages were calculated from monthly averages, which in turn were calculated if at least three daytime measurements were available. Despite a relatively large mean standard deviation in yearly average u of 7.0%, yearly average u was not significantly correlated with E_a in any of these sites. In contrast, yearly average R_n was positively correlated with yearly average E_a in seven of the 16 sites, with comparable mean standard deviation in annual R_n (8.5%). Moreover, looking at all individual towers and using the daily estimates, neither α_{MD} nor α_{PT} were correlated with u (Fig. 3c, d, Table 3). In fact, since α_{MD} appears hardly affected by any climatic variable, and given the relatively small range in α_{MD} values within each biome (Table 2), it appears that α_{MD} is a robust biome property that can be used in the seamless application of these methods at global scale. The robustness of α_{MD} as biome property is furthermore confirmed by the analysis with independent calibration and validation sites, which hardly affected the unRMSE and bias (Tables S10–11).

4.5. Use of available energy under stress conditions

The best performing methods rely heavily on measurements of available energy (R_n–G) (Eqs. 2 and 3). In Section 3.2, all E_p calculations used available energy obtained during unstressed conditions. The question is whether E_p can also be calculated correctly using the actual R_n and G when the ecosystems are not unstressed. As mentioned in Section 1, there is discussion whether SW_{out} and LW_{out} should be considered as forcing variables or as ecosystem responses (e.g. Lhomme, 1997;Lhomme and Guilioni, 2006;Shuttleworth, 1993). Among other considerations, it is clear that T_s will be lower if vegetation is healthy and soils are well watered (Maes and Steppe, 2012), which results in lower LW_{out} and higher R_n. Therefore, while using the observations of SW_{in} and LW_{in} as forcing variables in the computation of E_p is defendable (despite the potential atmospheric feedbacks that may derive from the consideration of unstressed conditions), we agree that SW_{out}, LW_{out} and G should ideally reflect unstressed rather than actual conditions to estimate E_p. Note that in that case, E_p deviates from E_{p0} in the complementarity relationship, in which atmospheric feedbacks affecting incoming radiation or VPD are implicitly considered (Kahler and Brutsaert, 2006).

A method to derive the 'unstressed' estimates of SW_{in} (through the albedo, α), LW_{in} (through T_s) and G under stressed conditions is presented in the Section S2 of the Supplement and is based the MD_b method and on flux data of the unstressed datasets. We further refer to this method as the unstressed E_p . It requires a large amount of input data and is not practically applicable at global scale. Comparing the mean unstressed E_p with the 'actual' E_p (i.e., E_p calculated from the actual R_n and G) for all sites reveals that the actual E_p is 8.2±10.1% lower than the unstressed E_p. There are no significant differences between biomes, but the distribution of the underestimation is left-skewed and the actual E_p is more than 10% lower than the unstressed E_p in 22% of the sites (Fig. 6). The main reason, explaining about 65% of the difference between the actual and the unstressed E_p, is the difference in T_s. Assuming that the unstressed T_s can be estimated as the mean of T_a and the actual T_s results in a straightforward alternative to approximate the unstressed E_p with only data of T_a and radiation:

$$E_p = \alpha_{MD}\left((1-\alpha)\,SW_{in} + \varepsilon\,LW_{in} - 0.5\,\varepsilon\,LW_{out} - 0.5\,\varepsilon\,\sigma\,T_a^4 - G\right) \tag{17}$$

This approach was also tested and resulted in a mean underestimation of E_p of 2.6±5.8% compared to the 'unstressed' E_p, with a mean median value at -0.1% (Fig. 6). Given the low error and the straightforward calculation, we recommend this method to calculate E_p at global scales.

(Insert Figure 6)

Fig. 7 gives an example of the seasonal evolution of E_a and E_p and the *S* factor ($S=E_a\,E_p^{-1}$) in a grassland (Fig 7a) and a deciduous forest site (Fig. 7b). The short growing season in the grassland site, when E_a is close to E_p and values of *S* are close to 1, stands in clear contrast with the winter period, when grasses have died off and E_a and consequently also *S* are very low. In the relatively wet broadleaf forest, E_a and E_p follow a similar seasonal cycle. In winter, when total evaporation is limited to soil evaporation, *S* is very low; in spring, when leaves are still developing, E_a lags E_p. In summer, *S* remains high and close to one.

(Insert Figure 7)

Conclusion

Based on a large sample of eddy-covariance sites from the FLUXNET2015 database, we demonstrated a higher potential of radiation-driven methods calibrated by biome type to estimate E_p than of more complex Penman-Monteith approaches or empirical temperature-based formulations. This was consistent across all 11 biomes represented in the database, and for two different criteria to identify unstressed days, one based on soil moisture and the other on evaporative fractions. Our analyses also showed that the key parameters required to apply the higher-performance radiation-driven methods are relatively insensitive to climate forcing. This makes these methods robust for incorporation into global offline models, e.g. for hydrological applications. Finally, we conclude that, at the ecosystem scale, Penman-Monteith methods for estimating E_p should only be prioritised if the unstressed stomatal conductance is calculated dynamically and high accuracy observations from the wide palate of required forcing variables are available.

Data availability

The FLUXNET2015 dataset can be downloaded from http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/. The main script for calculating potential evaporation with the different method as well as the daily flux data of one site (AU-How), for which permission of distribution was granted, is available as supplement. For further questions, we ask readers to contact the corresponding author at wh.maes@ugent.be.

Supplement

The supplement contains a description of the calculation method of kB^{-1} for sites where radiance fluxes are not separately measured (S1), a description of the method to estimate 'unstressed' E_p (See Section 4.5) from flux data (S2) and several tables (S3), including an overview of the FLUXNET sites used in the study (Table S1), a table on the representativeness of the climate forcing conditions of the unstressed datasets for the full dataset (Table S2) as well as correlation, unbiased RMSE and bias tables for different selection criteria (Tables S3-S11).

Acknowledgements

This study was funded by the Belgian Science Policy Office (BELSPO) in the frame of the STEREO III program project STR3S (SR/02/329). Diego G. Miralles acknowledges support from the European Research Council (ERC) under grant agreement no. 715254 (DRY–2–DRY). This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, FLUXNET-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization was carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux and AsiaFlux offices. The Atqasuk and Ivotuk towers in Alaska were supported by the Office of Polar Programs of the National Science Foundation (NSF) awarded to DZ, WCO, and DAL (award number 1204263) with additional logistical support funded by the NSF Office of Polar Programs, and by the Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), an Earth Ventures (EV-1) investigation, under contract with the National Aeronautics and Space Administration, and by the ABoVE (NNX15AT74A; NNX16AF94A) Program. The OzFlux network is supported by the Australian Terrestrial Ecosystem Research Network (TERN, http://www.tern.org.au). We would also like to thank three anonymous reviewers for their useful suggestions on an earlier version of the manuscript.

Author contribution. WHM and DGM designed the research. PG and NECV assisted in developing the optimal method for analysing all flux tower data. WHM performed the calculations and analyses. WHM and DGM wrote the manuscript, with contributions from PG and NECV.

References

Abtew, W.: Evapotranspiration measurements and modelling for three wetland systems in South Florida, JAWRA Journal of the American Water Resources Association, 32, 465-473, 10.1111/j.1752-1688.1996.tb04044.x, 1996.

Agam, N., Kustas, W. P., Anderson, M. C., Norman, J. M., Colaizzi, P. D., Howell, T. A., Prueger, J. H., Meyers, T. P., and Wilson, T. B.: Application of the Priestley–Taylor Approach in a Two-Source Surface Energy Balance Model, J. Hydrometeorol., 11, 185-198, 10.1175/2009jhm1124.1, 2010.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop Evapotranspiration - guidelines for computing crop water requirements, FAO Irrigation and Drainage Paper No. 56, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy, 1998.

Aminzadeh, M., Roderick, M. L., and Or, D.: A generalized complementary relationship between actual and potential evaporation defined by a reference surface temperature, Water Resources Research, 52, 385-406, 10.1002/2015WR017969, 2016.

Anderson, M. C., Norman, J. M., Diak, G. R., Kustas, W. P., and Mecikalski, J. R.: A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing, Remote Sensing of Environment, 60, 195-216, 1997.

Bai, P., Liu, X. M., Yang, T. T., Li, F. D., Liang, K., Hu, S. S., and Liu, C. M.: Assessment of the influences of different potential evapotranspiration inputs on the performance of monthly hydrological models under different climatic conditions, J. Hydrometeorol., 17, 2259-2274, 10.1175/jhm-d-15-0202.1, 2016.

Ball, J. T., Woodrow, I. E., and Berry, J. A.: A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in: Progress in photosynthesis research: Volume 4 proceedings of the VIIth international congress on photosynthesis providence, Rhode Island, USA, August 10–15, 1986, edited by: Biggins, J., Springer Netherlands, Dordrecht, 221-224, 1987.

Barton, I. J.: A parameterization of the evaporation from nonsaturated surfaces. J. Appl. Metereol., 18, 43-47, 1979.

Bautista, F., Baustista, D., and Delgado-Carranza, C.: Calibration of the equations of Hargreaves and Thornthwaite to estimate the potential evapotranspiration in semi-arid and subhumid tropical climates for regional applications, Atmósfera, 22, 331-348, 2009.

Bouchet, R. J.: Evapotranspiration réelle, évapotranspiration potentielle et production agricole, L'eau et la production végétale. INRA, Paris, 1964.

Brutsaert, W., and Stricker, H.: An advection-aridity approach to estimate actual regional evapotranspiration, Water Resources Research, 15, 443-450, 10.1029/WR015i002p00443, 1979.

Brutsaert, W.: Evaporation into the atmosphere. Theory, history and applications, D. Reidel Publishing Company, Dordrecht, 299 pp., 1982.

Camargo, A. P., R., M. F., Sentelhas, P. C., and Picini, A. G.: Adjust of the Thornthwaite's method to estimate the potential evapotranspiration for arid and superhumid climates, based on daily temperature amplitude, Revista Brasileira de Agrometeorologia, 7, 251-257, 1999.

Castellvi, F., Stockle, C., Perez, P., and Ibanez, M.: Comparison of methods for applying the Priestley–Taylor equation at a regional scale, Hydrological Processes, 15, 1609-1620, 2001.

Chan, F. C., Arain, M. A., Khomik, M., Brodeur, J. J., Peichl, M., Restrepo-Coupe, N., Thorne, R., Beamesderfer, E., McKenzie, S., and Xu, B.: Carbon, water and energy exchange dynamics of a young pine plantation forest during the initial fourteen years of growth, Forest Ecology and Management, 410, 12-26, 2018.

Chen, D., Gao, G., Xu, C.-Y., Guo, J., and Ren, G.: Comparison of the Thornthwaite method and pan data with the standard Penman-Monteith estimates of reference evapotranspiration in China, Climate Research, 28, 123-132, 2005.

Crago, R., and Crowley, R.: Complementary relationships for near-instantaneous evaporation, Journal of Hydrology, 300, 199-211, http://dx.doi.org/10.1016/j.jhydrol.2004.06.002, 2005.
Dai, A.: Drought under global warming: a review, Wiley Interdisciplinary Reviews: Climate Change, 2, 45-65, 10.1002/wcc.81, 2011.
Dolman, A. J., Miralles, D. G., and de Jeu, R. A. M.: Fifty years since Monteith's 1965 seminal paper: the emergence of global ecohydrology, Ecohydrology, 7, 897-902, 10.1002/eco.1505, 2014.
Donohue, R. J., McVicar, T. R., and Roderick, M. L.: Assessing the ability of potential evaporation formulations to capture the dynamics in evaporative demand within a changing climate, Journal of Hydrology, 386, 186-197, http://dx.doi.org/10.1016/j.jhydrol.2010.03.020, 2010.
Douglas, E. M., Jacobs, J. M., Sumner, D. M., and Ray, R. L.: A comparison of models for estimating potential evapotranspiration for Florida land cover types, Journal of Hydrology, 373, 366-376, http://dx.doi.org/10.1016/j.jhydrol.2009.04.029, 2009.
Eaton, A. K., Rouse, W. R., Lafleur, P. M., Marsh, P., and Blanken, P. D.: Surface energy balance of the Western and Central Canadian Subarctic: variations in the energy balance among five major terrain types, Journal of Climate, 14, 3692-3703, 10.1175/1520-0442(2001)014, 2001.
Eltahir, E. A.: A soil moisture–rainfall feedback mechanism: 1. Theory and observations, Water Resources Research, 34, 765-776, 1998.
Ershadi, A., McCabe, M. F., Evans, J. P., Chaney, N. W., and Wood, E. F.: Multi-site evaluation of terrestrial evaporation models using FLUXNET data, Agricultural and Forest Meteorology, 187, 46-61, https://doi.org/10.1016/j.agrformet.2013.11.008, 2014.
Field, C. B., Jackson, R. B., and Mooney, H. A.: Stomatal responses to increased CO_2 - Implications from the plant to the global scale, Plant Cell and Environment, 18, 1214-1225, 1995.
Fisher, J. B., Tu, K. P., and Baldocchi, D. D.: Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites, Remote Sensing of Environment, 112, 901-919, http://dx.doi.org/10.1016/j.rse.2007.06.025, 2008.
Fisher, J. B., Whittaker, R. J., and Malhi, Y.: ET come home: potential evapotranspiration in geographical ecology, Global Ecology and Biogeography, 20, 1-18, 10.1111/j.1466-8238.2010.00578.x, 2011.
Garratt, J. R.: The atmospheric boundary layer, Cambridge ; New York : Cambridge University Press, 1992.
Gentine, P., Entekhabi, D., Chehbouni, A., Boulet, G., and Duchemin, B.: Analysis of evaporative fraction diurnal behaviour, Agricultural and Forest Meteorology, 143, 13-29, https://doi.org/10.1016/j.agrformet.2006.11.002, 2007.
Gentine, P., Entekhabi, D., and Polcher, J.: The diurnal behavior of evaporative fraction in the soil–vegetation–atmospheric boundary layer continuum, J. Hydrometeorol., 12, 1530-1546, 10.1175/2011JHM1261.1, 2011a.
Gentine, P., Polcher, J., and Entekhabi, D.: Harmonic propagation of variability in surface energy balance within a coupled soil-vegetation-atmosphere system, Water Resources Research, 47, n/a-n/a, 10.1029/2010WR009268, 2011b.
Granger, R. J.: A complementary relationship approach for evaporation from nonsaturated surfaces, Journal of Hydrology, 111, 31-38, http://dx.doi.org/10.1016/0022-1694(89)90250-3, 1989.
Granier, A., Biron, P., and Lemoine, D.: Water balance, transpiration and canopy conductance in two beech stands, Agricultural and Forest Meteorology, 100, 291-308, 2000.
Guo, D., Westra, S., and Maier, H. R.: Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones, Hydrology and Earth System Sciences, 21, 2107-2126, 10.5194/hess-21-2107-2017, 2017.

Hargreaves, G. H., and Samani, Z.: Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture, 1, 96-99, 10.13031/2013.26773, 1985

Jacobs, J. M., Anderson, M. C., Friess, L. C., and Diak, G. R.: Solar radiation, longwave radiation and emergent wetland evapotranspiration estimates from satellite data in Florida, USA Hydrological Sciences Journal, 49, null-476, 10.1623/hysj.49.3.461.54352, 2004.

Jarvis, P. G.: Interpretation of variations in leaf water potential and stomatal conductance found in canopies in field, Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 273, 593-610, 1976.

Jones, H. G.: Plants and microclimate. A quantitative approach to environmental plant physiology, Second Edition ed., Cambridge University Press, Cambridge, UK, 1992.

Kahler, D. M., and Brutsaert, W.: Complementary relationship between daily evaporation in the environment and pan evaporation, Water Resources Research, 42, n/a-n/a, 10.1029/2005WR004541, 2006.

Katerji, N., and Rana, G.: Crop reference evapotranspiration: a discussion of the concept, analysis of the process and validation, Water Resources Management, 25, 1581-1600, 10.1007/s11269-010-9762-1, 2011.

Kay, A. L., and Davies, H. N.: Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts, Journal of Hydrology, 358, 221-239, http://dx.doi.org/10.1016/j.jhydrol.2008.06.005, 2008.

Kingston, D. G., Todd, M. C., Taylor, R. G., Thompson, J. R., and Arnell, N. W.: Uncertainty in the estimation of potential evapotranspiration under climate change, Geophysical Research Letters, 36, 10.1029/2009GL040267, 2009.

Komatsu, H.: Forest categorization according to dry-canopy evaporation rates in the growing season: comparison of the Priestley–Taylor coefficient values from various observation sites, Hydrological Processes, 19, 3873-3896, 10.1002/hyp.5987, 2005.

Leuning, R.: A critical appraisal of a combined stomatal-photosynthesis model for C3 plants, Plant, Cell & Environment, 18, 339-355, 10.1111/j.1365-3040.1995.tb00370.x, 1995.

Lhomme, J. P.: Towards a rational definition of potential evaporation, Hydrology and Earth System Sciences, 1, 257-264, 1997.

Lhomme, J. P., and Guilioni, L.: Comments on some articles about the complementary relationship, Journal of Hydrology, 323, 1-3, 10.1016/j.jhydrol.2005.08.014, 2006.

Li, D., Rigden, A., Salvucci, G., and Liu, H.: Reconciling the Reynolds number dependence of scalar roughness length and laminar resistance, Geophysical Research Letters, 44, 3193-3200, 10.1002/2017GL072864, 2017.

Li, S., Kang, S., Zhang, L., Zhang, J., Du, T., Tong, L., and Ding, R.: Evaluation of six potential evapotranspiration models for estimating crop potential and actual evapotranspiration in arid regions, Journal of Hydrology, 543, Part B, 450-461, http://dx.doi.org/10.1016/j.jhydrol.2016.10.022, 2016.

Lofgren, B. M., Hunter, T. S., and Wilbarger, J.: Effects of using air temperature as a proxy for potential evapotranspiration in climate change scenarios of Great Lakes basin hydrology, Journal of Great Lakes Research, 37, 744-752, http://dx.doi.org/10.1016/j.jglr.2011.09.006, 2011.

Maes, W. H., Pashuysen, T., Trabucco, A., Veroustraete, F., and Muys, B.: Does energy dissipation increase with ecosystem succession? Testing the ecosystem exergy theory combining theoretical simulations and thermal remote sensing observations, Ecological Modelling, 23-24, 3917-3941, 2011.

Maes, W. H., and Steppe, K.: Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review, Journal of Experimental Botany, 63, 4671-4712, 2012.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development, 10, 1903-1925, 10.5194/gmd-10-1903-2017, 2017.
Martínez-Vilalta, J., Poyatos, R., Aguadé, D., Retana, J., and Mencuccini, M.: A new look at water transport regulation in plants, New Phytologist, 204, 105-115, 10.1111/nph.12912, 2014.
McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., and McVicar, T. R.: Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, Hydrology and Earth System Sciences, 17, 1331-1363, 10.5194/hess-17-1331-2013, 2013.
McVicar, T. R., Van Niel, T. G., Li, L. T., Roderick, M. L., Rayner, D. P., Ricciardulli, L., and Donohue, R. J.: Wind speed climatology and trends for Australia, 1975-2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output, Geophysical Research Letters, 35, 6, 10.1029/2008gl035627, 2008.
McVicar, T. R., Roderick, M. L., Donohue, R. J., Li, L. T., Van Niel, T. G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N. M., Mescherskaya, A. V., Kruger, A. C., Rehman, S., and Dinpashoh, Y.: Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation, Journal of Hydrology, 416, 182-205, http://dx.doi.org/10.1016/j.jhydrol.2011.10.024, 2012.
Michel, D., Jiménez, C., Miralles, D., Jung, M., Hirschi, M., Ershadi, A., Martens, B., McCabe, M., Fisher, J., and Mu, Q.: TheWACMOS-ET project–Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms, Hydrology and Earth System Sciences, 20, 803-822, 2016.
Milly, P. C. D., and Dunne, K. A.: Potential evapotranspiration and continental drying, Nature Clim. Change, 6, 946-949, 10.1038/nclimate3046, 2016.
Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrology and Earth System Sciences, 15, 453, 2011.
Mizutani, K., Yamanoi, K., Ikeda, T., and Watanabe, T.: Applicability of the eddy correlation method to measure sensible heat transfer to forest under rainfall conditions, Agricultural and Forest Meteorology, 86, 193-203, https://doi.org/10.1016/S0168-1923(97)00012-9, 1997.
Monteith, J. L.: Evaporation and environment, Symposia of the Society for Experimental Biology, 19, 205-234, 1965.
Morton, F. I.: Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology, Journal of Hydrology, 66, 1-76, http://dx.doi.org/10.1016/0022-1694(83)90177-4, 1983.
Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W.: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sensing of Environment, 111, 519-536, https://doi.org/10.1016/j.rse.2007.04.015, 2007.
Norman, J. M., Kustas, W. P., and Humes, K. S.: Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature, Agricultural and Forest Meteorology, 77, 263-293, 1995.
Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., Papuga, S. A., Blanken, P. D., Noormets, A., Sulman, B. N., Scott, R. L., Wang, L., and Phillips, R. P.: The increasing importance of atmospheric demand for ecosystem water and carbon fluxes, Nature Clim. Change, 6, 1023-1027, 10.1038/nclimate3114, 2016.
Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling, Journal of Hydrology, 303, 290-306, http://dx.doi.org/10.1016/j.jhydrol.2004.08.026, 2005a.

Oudin, L., Michel, C., and Anctil, F.: Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 1—Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs?, Journal of Hydrology, 303, 275-289, http://dx.doi.org/10.1016/j.jhydrol.2004.08.025, 2005b.

Padrón, R. S., Gudmundsson, L., Greve, P., and Seneviratne, S. I.: Large-Scale Controls of the Surface Water Balance Over Land: Insights From a Systematic Review and Meta-Analysis, Water Resources Research, 53, 9659-9678, doi:10.1002/2017WR021215, 2017.

Palmer, C., Fisher, J. B., Mallick, K., and Lee, J.: The Potential of Potential Evapotranspiration, American Geophysical Union, Fall Meeting 2012, 2012,

Penman, H. L.: Natural evaporation from open water, bare soil and grass, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 193, 120-145, 10.1098/rspa.1948.0037, 1948.

Penman, H. L.: Vegetation and Hydrology, Soil Science, 96, 357, 1963.

Pennypacker, S., and Baldocchi, D.: Seeing the Fields and Forests: Application of Surface-Layer Theory and Flux-Tower Data to Calculating Vegetation Canopy Height, Boundary-Layer Meteorology, 158, 165-182, 10.1007/s10546-015-0090-0, 2016.

Pereira, A. R., and Pruitt, W. O.: Adaptation of the Thornthwaite scheme for estimating daily reference evapotranspiration, Agricultural Water Management, 66, 251-257, https://doi.org/10.1016/j.agwat.2003.11.003, 2004.

Pettijohn, J. C., and Salvucci, G. D.: A new two-dimensional physical basis for the complementary relation between terrestrial and pan evaporation, J. Hydrometeorol., 10, 565-574, 10.1175/2008jhm1026.1, 2009.

Powell, T. L., Bracho, R., Li, J., Dore, S., Hinkle, C. R., and Drake, B. G.: Environmental controls over net ecosystem carbon exchange of scrub oak in central Florida, Agricultural and Forest Meteorology, 141, 19-34, http://dx.doi.org/10.1016/j.agrformet.2006.09.002, 2006.

Powell, T. L., Galbraith, D. R., Christoffersen, B. O., Harper, A., Imbuzeiro, H. M. A., Rowland, L., Almeida, S., Brando, P. M., da Costa, A. C. L., Costa, M. H., Levine, N. M., Malhi, Y., Saleska, S. R., Sotta, E., Williams, M., Meir, P., and Moorcroft, P. R.: Confronting model predictions of carbon fluxes with measurements of Amazon forests subjected to experimental drought, New Phytologist, 200, 350-365, 10.1111/nph.12390, 2013.

Priestley, C. H. B., and Taylor, R. J.: On the assessment of surface heat flux and evaporation using large-scale parameters, Monthly Weather Review, 100, 81-92, 10.1175/1520-0493(1972)100<0081:otaosh>2.3.co;2, 1972.

Rigden, A. J., and Salvucci, G. D.: Evapotranspiration based on equilibrated relative humidity (ETRHEQ): Evaluation over the continental US, Water Resources Research, 51, 2951-2973, 10.1002/2014wr016072, 2015.

Rind, D., Goldberg, R., Hansen, J., Rosenzweig, C., and Ruedy, R.: Potential evapotranspiration and the likelihood of future drought, Journal of Geophysical Research: Atmospheres, 95, 9983-10004, 10.1029/JD095iD07p09983, 1990.

Roerink, G. J., Su, Z., and Menenti, M.: S-SEBI: a simple remote sensing algorithm to estimate the surface energy balance, Physics and Chemistry of the Earth Part B-Hydrology Oceans and Atmosphere, 25, 147-157, 2000.

Scheff, J., and Frierson, D. M. W.: Scaling potential evapotranspiration with greenhouse warming, Journal of Climate, 27, 1539-1558, 10.1175/JCLI-D-13-00233.1, 2013.

Schellekens, J., Dutra, E., Martinez-de la Torre, A., Balsamo, G., van Dijk, A., Weiland, F. S., Minvielle, M., Calvet, J. C., Decharme, B., Eisner, S., Fink, G., Florke, M., Pessenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W., and Weedon, G. P.: A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, Earth Syst. Sci. Data, 9, 389-413, 10.5194/essd-9-389-2017, 2017.

Seiller, G., and Anctil, F.: How do potential evapotranspiration formulas influence hydrological projections?, Hydrol. Sci. J.-J. Sci. Hydrol., 61, 2249-2266, 10.1080/02626667.2015.1100302, 2016.

Sentelhas, P. C., Gillespie, T. J., and Santos, E. A.: Evaluation of FAO Penman–Monteith and alternative methods for estimating reference evapotranspiration with missing data in Southern Ontario, Canada, Agricultural Water Management, 97, 635-644, https://doi.org/10.1016/j.agwat.2009.12.001, 2010.

Sheffield, J., Wood, E. F., and Roderick, M. L.: Little change in global drought over the past 60 years, Nature, 491, 435-438, 10.1038/nature11575, 2012.

Shuttleworth, W. J., and Calder, I. R.: Has the Priestley-Taylor equation any relevance to forest evaporation?, Journal of Applied Meteorology, 18, 639-646, 10.1175/1520-0450(1979)018<0639:htptea>2.0.co;2, 1979.

Singh, V. P., and Xu, C. Y.: Sensitivity of mass transfer-based evaporation equations to errors in daily and monthly input data, Hydrological Processes, 11, 1465-1473, 10.1002/(SICI)1099-1085(199709)11:11<1465::AID-HYP452>3.0.CO;2-X, 1997.

Slatyer, R., and McIlroy, I.: Practical micrometeorology, 1961.

Song, Q. H., Braeckevelt, E., Zhang, Y. P., Sha, L. Q., Zhou, W. J., Liu, Y. T., Wu, C. S., Lu, Z. Y., and Klemm, O.: Evapotranspiration from a primary subtropical evergreen forest in Southwest China, Ecohydrology, 10, 2017.

Sumner, D. M., and Jacobs, J. M.: Utility of Penman–Monteith, Priestley–Taylor, reference evapotranspiration, and pan evaporation methods to estimate pasture evapotranspiration, Journal of Hydrology, 308, 81-104, http://dx.doi.org/10.1016/j.jhydrol.2004.10.023, 2005.

Teuling, A. J., and Seneviratne, S. I.: Contrasting spectral changes limit albedo impact on land-atmosphere coupling during the 2003 European heat wave, Geophysical Research Letters, 35, doi:10.1029/2007GL032778, 2008.

Thornthwaite, C. W.: An approach toward a rational classification of climate, Geographical Review, 38, 55-94, 10.2307/210739, 1948.

Vautard, R., Cattiaux, J., Yiou, P., Thepaut, J. N., and Ciais, P.: Northern Hemisphere atmospheric stilling partly attributed to an increase in surface roughness, Nat. Geosci., 3, 756-761, 10.1038/ngeo979, 2010.

Vicente-Serrano, S. M., Gouveia, C., Camarero, J. J., Beguería, S., Trigo, R., López-Moreno, J. I., Azorín-Molina, C., Pasho, E., Lorenzo-Lacruz, J., Revuelto, J., Morán-Tejeda, E., and Sanchez-Lorenzo, A.: Response of vegetation to drought time-scales across global land biomes, Proceedings of the National Academy of Sciences, 110, 52-57, 10.1073/pnas.1207068110, 2013.

Vicente-Serrano, S. M., Miralles, Domínguez-Castro F., Azorín-Molina, C., El Kenawy, A., McVicar, T. R., Tomás-Burguera, M., Beguería, S., Maneta, M., Peña-Gallardo, M.: Global assessment of the Standardized Evapotranspiration Deficit Index (SEDI) for drought analysis and monitoring, Journal of Climate, 31, 5371-5393, 10.1175/jcli-d-17-0775.1.Viswanadham, Y., Silva Filho, V. P., and André, R. G. B.: The Priestley-Taylor parameter α for the Amazon forest, Forest Ecology and Management, 38, 211-225, https://doi.org/10.1016/0378-1127(91)90143-J, 1991.

Wang, K., Wang, P., Li, Z., Cribb, M., and Sparrow, M.: A simple method to estimate actual evapotranspiration from a combination of net radiation, vegetation index, and temperature, Journal of Geophysical Research: Atmospheres, 112, doi:10.1029/2006JD008351, 2007.

Wang, K. C., and Dickinson, R. E.: A review of global terrestrial evapotranspiration: observation, modeling, climatology, and climatic variability Reviews of Geophysics, 50, 10.1029/2011rg000373, 2012.

Weiß, M., and Menzel, L.: A global comparison of four potential evapotranspiration equations and their relevance to stream flow modelling in semi-arid environments, Advances in Geosciences, 18, 15-23, 10.5194/adgeo-18-15-2008, 2008.

Xu, C. Y., and Singh, V. P.: Evaluation and generalization of temperature-based methods for calculating evaporation, Hydrological Processes, 15, 305-319, 10.1002/hyp.119, 2001.

Xu, C. Y., and Singh, V. P.: Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland, Water Resources Management, 16, 197-219, 10.1023/A:1020282515975, 2002.

Yoder, R. E., Odhiambo, L. O., and Wright, W. C.: Evaluation of methods for estimating daily reference crop evapotranspiration at a site in the humid Southeast United States, Applied engineering in Agriculture, 21, 197-202, 10.13031/2013.18153, 2005.

Zhang, L., Potter, N., Hickel, K., Zhang, Y., and Shao, Q.: Water balance modeling over variable time scales based on the Budyko framework – Model development and testing, Journal of Hydrology, 360, 117-131, https://doi.org/10.1016/j.jhydrol.2008.07.021, 2008.

Zhou, M. C., Ishidaira, H., Hapuarachchi, H. P., Magome, J., Kiem, A. S., and Takeuchi, K.: 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, 151-173, http://dx.doi.org/10.1016/j.jhydrol.2005.11.013, 2006.

Tables

Abbreviation list

Symbol	Description	Unit
α	Albedo	[-]
α_{HS}	Parameter of Hargreaves-Samani equation (Eq. (7))	[-]
α_{MD}	Parameter of Milly & Dunne equation (Eq. (4))	[-]
α_{Ou}	Parameter of Oudin equation (Eq. (6))	[-]
α_{PT}	Parameter of Priestley & Taylor equation (Eq. (3))	[-]
α_{ref}	Albedo of reference crop (0.23)	[-]
α_{Th}	Parameter of Thornthwaite method (Eq. (5b))	[-]
c_p	Specific heat capacity of the air	J kg^{-1} K^{-1}
d	Zero displacement height	m
ε	Emissivity	[-]
E	Ecosystem evaporation, (or evapotranspiration, the sum of soil evaporation, transpiration, interception evaporation and snow sublimation)	kg m^{-2} s^{-1} or mm day^{-1}
E_a	Actual evaporation	kg m^{-2} s^{-1} or mm day^{-1}
E_p	Potential evaporation	kg m^{-2} s^{-1} or mm day^{-1}
E_{p0}	Evaporation from an extensive well-watered surface (Complementary Relationship)	mm day^{-1}
E_{pa}	Evaporation from an infinitesimally small well-watered surface (Complementary Relationship)	mm day^{-1}
E_{unstr}	Evaporation from an unstressed ecosystem	mm day^{-1}
EF	Evaporative fraction	[-]
γ	Psychrometric constant	Pa K^{-1}
G	Ground heat flux	W m^{-2}
g	Gravitational acceleration	m s^{-2}
g_c	Surface conductance to water transfer	m s^{-1} or mm s^{-1}
g_{c_ref}	g_c of reference crop	mm s^{-2}
H	Sensible heat flux	W m^{-2}
HS_b	Biome-specific version of the Hargreaves-Samani method (Table 1)	
HS_s	Standard version of the Hargreaves-Samani method (Table 1)	
I	parameter in Thornthwaite equation (Eq. (5))	
k	von Karman constant (0.41)	[-]
kB^{-1}	Stanton number	[-]
λ	Latent heat of vaporisation	J kg^{-1}
L	Obukhov length	m
LW_{in}	Incoming longwave radiation	W m^{-2}
LW_{out}	Outgoing longwave radiation	W m^{-2}
MD_b	Biome-specific version of the Milly-Dunne method (Table 1)	
MD_r	Reference crop version of the Milly-Dunne method (Table 1)	
MD_s	Standard version of the Milly-Dunne method (Table 1)	
Ou_b	Biome-specific version of the Oudin method (Table 1)	
Ou_s	Standard version of the Oudin method (Table 1)	

Pe_r	Reference crop version of the Penman method (Table 1)	
Pe_s	Standard version of the Penman method (Table 1)	
PM_b	Biome-specific version of the Penman-Monteith method (Table 1)	
PM_r	Reference crop version of the Penman-Monteith method (Table 1)	
PM_s	Standard version of the Penman-Monteith method (Table 1)	
PT_b	Biome-specific version of the Priestley and Taylor method (Table 1)	
PT_r	Reference crop version of the Priestley and Taylor method (Table 1)	
PT_s	Standard version of the Priestley and Taylor method (Table 1)	
q_a	Specific humidity	kg kg^{-1}
ρ_a	Air density	kg m^{-3}
$\Psi_h(X)$	Businger-Dyer stability function for heat exchange of variable X	[-]
$\Psi_m(X)$	Businger-Dyer stability function for momentum of variable X	[-]
r_{aH}	Resistance of heat transfer to air	s m^{-1}
r_c	Canopy resistance of water transfer	s m^{-1}
R_e	Top-of-atmosphere radiation	MJ m^{-2} day^{-1}
Re	Reynolds number	[-]
RH	Relative humidity	%
RH_{min}	Minimum daily RH	%
RH_{max}	Maximum daily RH	%
R_n	Net radiation	W m^{-2}
σ	Stefan-Boltzmann constant (5.67 10^{-8})	W m^{-2} K^{-4}
s	slope of the Clausius-Clapeyron curve	Pa K^{-1}
S	Ratio of E_a and E_p, confined to [0-1]	[-]
SW_{in}	Incoming shortwave radiation	W m^{-2}
SW_{out}	Outgoing shortwave radiation	W m^{-2}
SW_{TOA}	SW_{in} at the top of atmosphere	W m^{-2}
T_0	Aerodynamic temperature	°C or K
T_a	Air temperature	°C
T_{a_mean}	Mean air temperature for each month	°C
T_{eff}	Effective temperature (Thornthwaite equation, Eq. (5))	°C
Th_b	Biome-specific version of the Thornthwaite method (Table 1)	
Th_s	Standard version of the Thornthwaite method (Table 1)	
T_{min}	Minimum daily T_a	°C
T_{max}	Maximum daily T_a	°C
T_s	Surface temperature	°C or K
VH	Vegetation height	m
VPD	Vapour pressure deficit	Pa or hPa
u	Wind speed	m s^{-1}
$u*$	Friction velocity	m s^{-1}
z	Wind sensor height	m
z_{0h}	Roughness length for heat exchange	m

z_{0m}	Roughness length for momentum	m

Table 1. Overview of the different E_p methods used in this study and their calculation.

		Key parameter	**R_n**		**r_{aH}**	**T_a**	**RH or VPD**	**N**
			SW*	LW*				
Penman-Monteith		**g_{c_ref} (mm s^{-1})**						
PM_r	Reference crop	14.49	FAO-56 (α_{ref}=0.23)	FAO-56 $f(T_{max}, T_{min}$, SW, SW_{TOA}, e_a)	208 u_{2m}^{-1}	From T_{max}, T_{min}	From RH_{Max}, RH_{Min}	FAO-56
PM_s	Standard	14.49	measured	measured	calculated	Daytime mean	Daytime mean	
PM_b	Biome-specific	Biome-specific	measured	measured	calculated	Daytime mean	Daytime mean	
Penman		**g_{c_ref} (mm s^{-1})**						
Pe_r	Reference crop	∞ (r_{c_ref}=0)	FAO-56 (α_{ref} =0.23)	FAO-56 $f(T_{max}, T_{min}$, SW, SW_{TOA}, e_a)	208 u_{2m}^{-1}	From T_{max}, T_{min}	From RH_{Max}, RH_{Min}	
Pe_s	Standard	∞ (r_{c_ref}=0)	measured	measured	calculated	Daytime mean	Daytime mean	
Priestley and Taylor		**α_{PT} (-)**						
PT_r	Reference crop	1.26	FAO-56 (α_{ref} =0.23)	FAO-56 $f(T_{max}, T_{min}$, SW, SW_{TOA}, e_a)				
PT_s	Standard	1.26	measured	measured		Daytime mean		
PT_b	Biome-specific	Biome-specific	measured	measured		Daytime mean		
Milly and Dunne		**α_{MD} (-)**						
MD_r	Reference crop	0.8	FAO-56 (α_{ref} =0.23)	FAO-56 $f(T_{max}, T_{min}$, SW, SW_{TOA}, e_a)				
MD_s	Standard	0.8	measured	measured				
MD_b	Biome-specific	Biome-specific	measured	measured				
Thornthwaite		**α_{Th} (-)**						
Th_s	Standard	16				From T_{max}, T_{min}		Measured
Th_b	Biome-specific	Biome-specific				From T_{max}, T_{min}		Measured
Oudin		**α_{ou} (-)**						
Ou_s	Standard	100				Daily mean		
Ou_b	Biome-specific	Biome-specific				Daily mean		
Hargreaves-Samani		**α_{HS} (-)**						
HS_s	Standard	0.0023				Daily mean, and from T_{max}, T_{min}		
HS_b	Biome-specific	Biome-specific				Daily mean, and from T_{max}, T_{min}		

N=number of daylight hours; T_{max}, T_{min}, RH_{max} and RH_{min} the maximum and minimum daily air temperature or RH; SW* and LW* are net shortwave and net longwave radiation; SW_{TOA} is the shortwave incoming radiation at the top the atmosphere; FAO-56 refers to the methodology described by Allen et al. (1998).

Table 2. Overview of the difference of the key parameters (g_{c_ref}, α_{PT}, α_{MD}, α_{Th}, α_{Ou} and α_{HS}) during unstressed conditions per biome. The energy balance method was used for defining unstressed days (See section 2.4, see Table 1 for definition of key parameters). The p value of the ANOVA test is given, as well as the mean ± 1 standard deviation for each biome. Different alphabetic superscripts indicate significantly differing means (Tukey post-hoc test; p<0.05). The number of sites per biome is given between brackets. Different colours are used to group biomes into broader ecosystem types (in descending order: croplands, grasslands, forests, savannah ecosystems, wetlands).

	g_{c_ref} (mm s^{-1})	α_{PT} (-)	α_{MD} (-)	α_{Th} (-)	α_{Ou} (-)	α_{HS} (x10^{-3})
p (ANOVA)	0.47	0.004	<0.001	<0.001	<0.001	<0.001
CRO (10)	38.3 ± 23.0	1.15 ± 0.14^{a}	0.86 ± 0.09^{a}	38.7 ± 14.5^{ab}	77.0 ± 27.8^{b}	2.96 ± 0.69^{ab}
GRA (20)	30.5 ± 40.2	1.02 ± 0.16^{ab}	0.74 ± 0.12^{ab}	30.4 ± 13.9^{b}	103.2 ± 38.9^{b}	2.32 ± 0.70^{bc}
DBF (15)	32.6 ± 27.4	1.09 ± 0.14^{ab}	0.80 ± 0.08^{ab}	33.3 ± 7.8^{b}	70.5 ± 18.0^{b}	3.39 ± 0.83^{a}
EBF (9)	42.0 ± 36.6	1.09 ± 0.15^{ab}	0.74 ± 0.05^{abc}	53.1 ± 16.8^{a}	95.5 ± 22.9^{b}	3.07 ± 0.57^{ab}
ENF (26)	28.4 ± 52.1	0.89 ± 0.26^{b}	0.62 ± 0.09^{c}	40.3 ± 16.7^{ab}	92.0 ± 21.8^{b}	2.78 ± 0.76^{ab}
MF (4)	10.0 ± 7.1	0.88 ± 0.23^{ab}	0.64 ± 0.13^{bc}	26.1 ± 3.6^{b}	138.2 ± 91.6^{ab}	2.21 ± 0.97^{abc}
CSH (2)	8.5 ± 3.9	0.90 ± 0.10^{ab}	0.64 ± 0.15^{abc}	41.4 ± 13.7^{ab}	130.3 ± 36.1^{ab}	2.03 ± 0.68^{abc}
WSA (5)	8.4 ± 3.4	0.95 ± 0.09^{ab}	0.70 ± 0.10^{abc}	33.8 ± 6.4^{ab}	104.6 ± 19.7^{b}	2.25 ± 0.51^{abc}
OSH (5)	7.8 ± 3.7	0.87 ± 0.14^{b}	0.68 ± 0.15^{c}	35.0 ± 4.1^{ab}	147.1 ± 63.9^{ab}	1.88 ± 0.61^{c}
SAV (6)	4.3 ± 2.0	0.79 ± 0.11^{b}	0.58 ± 0.09^{bc}	31.3 ± 11.2^{ab}	147.7 ± 61.8^{ab}	1.59 ± 0.38^{bc}
WET (5)	20.0 ± 14.1	1.03 ± 0.47^{ab}	0.75 ± 0.11^{b}	17.8 ± 13.3^{b}	638.6 ± 1230.1^{a}	2.00 ± 0.54^{bc}

CRO=cropland; DBF=Deciduous Broadleaf Forest; EBF=Evergreen Broadleaf Forest; ENF=Evergreen Needleleaf Forest; MF=Mixed Forest; CSH=Closed Shrubland; WSA=Woody Savanna; SAV=Savanna; OSH=Open Shrubland; GRA=Grasslands; WET=Wetlands.

Table 3. Influence of atmospheric conditions on E_{unstr} and on selected key parameters (g_{c_ref}, α_{PT}, α_{MD}). (left) Mean ± 1 standard deviation of the correlations of E_{unstr}, g_c, α_{PT} and α_{MD} against the atmospheric conditions, and (right) number of sites (out of total of 107) with significant negative/positive correlations between E_{unstr}, α_{PT}, g_{c_ref} and α_{MD} and the climate forcing variables. Based on unstressed days only defined using the energy balance criterion.

	Mean ± 1 standard deviation of the correlations				Number of sites with significant negative/positive correlations			
	E_{unstr}	g_{c_ref}	α_{PT}	α_{MD}	E_{unstr}	g_{c_ref}	α_{PT}	α_{MD}
Wind	0.13 ± 0.26	0.03 ± 0.25	0.12 ± 0.31	0.01 ± 0.22	6/26	4/13	11/30	5/6
T_{air}	0.60 ± 0.24*	-0.22 ± 0.29	-0.21 ± 0.34	-0.02 ± 0.28	0/93	43/0	43/5	16/13
VPD	0.64 ± 0.20*	-0.27 ± 0.27	-0.11 ± 0.31	-0.01 ± 0.28	0/93	48/0	31/10	15/11
R_n	0.90 ± 0.08*	-0.05 ± 0.25	-0.13 ± 0.30	-0.10 ± 0.31	0/106	17/3	33/5	30/14
$[CO_2]$	-0.16 ± 0.30	-0.01 ± 0.23	-0.03 ± 0.22	-0.03 ± 0.25	34/5	7/5	9/4	12/4

*significantly different from 0

Table 4. Mean correlations per biome between the measured E_{unstr} and the different E_p methods. The methods with the highest correlation per biome are highlighted in bold and underlined. Based on unstressed days only defined using the energy balance criterion. Different colours are used to group biomes into broader ecosystem types (in descending order: croplands, grasslands, forests, savannah ecosystems, wetlands).

	Radiation, Temperature, VPD					*Radiation, Temperature*			*Radiation*			*Temperature*					
	Penman-Monteith			**Penman**		**Priestley and Taylor**			**Milly and Dunne**			**Thornthwaite**		**Oudin**		**Hargreaves-Samani**	
	Ref. crop	**Standard**	**Biome**	**Ref. crop**	**Standard**	**Ref. crop**	**Standard**	**Biome**	**Ref. crop**	**Standard**	**Biome**	**Standard**	**Biome**	**Standard**	**Biome**	**Standard**	**Biome**
CRO (10)	0.84	0.91	0.90	0.76	0.81	0.86	0.96	0.96	0.82	**0.96**	**0.96**	0.77	0.77	0.74	0.74	0.76	0.76
GRA (20)	0.79	0.87	0.87	0.77	0.84	0.82	0.93	0.93	0.80	**0.94**	**0.94**	0.55	0.54	0.55	0.55	0.66	0.66
DBF (15)	0.78	0.87	0.88	0.79	0.85	0.78	0.91	0.91	0.75	**0.91**	**0.91**	0.57	0.56	0.57	0.57	0.62	0.62
EBF (9)	0.88	0.89	0.88	0.86	0.85	0.87	**0.91**	**0.91**	0.83	0.90	0.90	0.71	0.79	0.57	0.57	0.75	0.75
ENF (26)	0.89	0.90	0.91	0.88	0.86	0.90	0.95	0.95	0.88	**0.95**	**0.95**	0.77	0.79	0.76	0.76	0.82	0.82
MF (4)	0.90	0.93	0.93	0.90	0.93	0.90	**0.94**	**0.94**	0.88	0.93	0.93	0.79	0.75	0.74	0.74	0.79	0.79
CSH (2)	0.90	0.94	0.93	0.89	0.90	0.90	0.95	0.95	0.89	**0.95**	**0.95**	0.80	0.78	0.75	0.75	0.79	0.79
WSA (5)	0.76	0.78	0.78	0.73	0.73	0.80	0.89	0.89	0.79	**0.90**	**0.90**	0.41	0.41	0.46	0.46	0.51	0.51
SAV (6)	0.79	0.82	0.81	0.77	0.79	0.83	**0.91**	**0.91**	0.81	0.91	0.91	0.52	0.52	0.56	0.56	0.71	0.71
OSH (5)	0.72	0.80	0.78	0.64	0.78	0.79	0.90	0.90	0.77	**0.90**	**0.90**	0.54	0.53	0.56	0.56	0.64	0.64
WET (5)	0.87	0.81	0.76	0.87	0.66	0.79	0.83	0.83	0.68	**0.85**	**0.85**	0.50	0.45	0.61	0.61	0.65	0.65
Overall (107)	**0.83**	**0.87**	**0.87**	**0.81**	**0.83**	**0.84**	**0.92**	**0.92**	**0.81**	**0.93**	**0.93**	**0.62**	**0.63**	**0.63**	**0.63**	**0.71**	**0.71**

CRO=cropland; DBF=Deciduous Broadleaf Forest; EBF=Evergreen Broadleaf Forest; ENF=Evergreen Needleleaf Forest; MF=Mixed Forest; CSH=Closed Shrubland; WSA=Woody Savanna; SAV=Savanna; OSH=Open Shrubland; GRA=Grasslands; WET=Wetlands.

Table 5. Unbiased Root Mean Square Error (UnRMSE) (in mm day^{-1}) for the E_p methods per biome. The methods with the lowest UnRMSE per biome are indicated in bold and are underlined. Based on unstressed days only defined using the energy balance criterion. Different colours are used to group biomes into broader ecosystem types (in descending order: croplands, grasslands, forests, savannah ecosystems, wetlands).

	Radiation, Temperature, VPD					*Radiation, Temperature*			*Radiation*			*Temperature*					
	Penman-Monteith			**Penman**		**Priestley and Taylor**			**Milly and Dunne**			**Thornthwaite**		**Oudin**		**Hargreaves-Samani**	
	Ref. crop	**Standard**	**Biome**	**Ref. crop**	**Standard**	**Ref. crop**	**Standard**	**Biome**	**Ref. crop**	**Standard**	**Biome**	**Standard**	**Biome**	**Standard**	**Biome**	**Standard**	**Biome**
CRO (10)	1.16	0.79	1.04	1.60	2.88	1.27	0.62	0.58	1.21	0.57	**<u>0.55</u>**	1.24	1.24	1.29	1.27	1.24	1.28
GRA (20)	1.22	0.70	0.81	1.75	1.04	1.40	0.58	0.47	1.13	0.44	**<u>0.44</u>**	1.07	1.03	1.05	1.04	0.97	0.97
DBF (15)	1.14	0.88	0.89	1.21	1.36	1.29	0.75	**<u>0.72</u>**	1.20	0.72	0.72	1.41	1.42	1.37	1.32	1.32	1.33
EBF (9)	0.84	0.62	0.93	1.07	1.33	1.09	0.75	0.59	0.96	0.55	**<u>0.54</u>**	1.04	0.98	1.15	1.14	1.00	1.03
ENF (26)	0.98	0.78	0.99	1.20	14.89	1.26	0.84	0.52	1.09	0.59	**<u>0.50</u>**	0.94	0.91	0.96	0.97	0.88	0.92
MF (4)	1.23	0.69	0.69	1.58	1.11	1.64	0.86	**<u>0.58</u>**	1.26	0.64	0.59	1.11	1.03	1.03	0.99	1.04	1.03
CSH (2)	0.82	0.59	0.59	0.98	0.92	1.12	0.75	**<u>0.48</u>**	0.91	0.55	0.49	0.90	0.96	0.83	0.81	0.91	0.87
WSA (5)	1.15	0.93	0.80	1.41	1.68	1.27	0.67	0.52	1.00	0.53	**<u>0.51</u>**	1.10	1.10	0.99	0.99	1.03	1.02
SAV (6)	1.22	1.02	0.83	1.53	1.88	1.39	0.76	0.52	1.07	0.58	**<u>0.52</u>**	1.22	1.21	1.10	0.97	1.02	0.94
OSH (5)	1.37	0.73	0.63	1.94	0.92	1.63	0.67	**<u>0.43</u>**	1.28	0.48	0.44	1.12	1.03	0.90	0.80	0.90	0.75
WET (5)	1.27	1.25	1.38	1.38	4.14	1.72	1.28	1.13	1.91	1.14	**<u>1.10</u>**	2.20	2.29	1.65	2.01	1.53	1.55
Overall (107)	**1.11**	**0.80**	**0.91**	**1.41**	**4.86**	**1.34**	**0.75**	**0.57**	**1.16**	**0.60**	**<u>0.56</u>**	**1.16**	**1.14**	**1.12**	**1.11**	**1.05**	**1.06**

CRO=cropland; DBF=Deciduous Broadleaf Forest; EBF=Evergreen Broadleaf Forest; ENF=Evergreen Needleleaf Forest; MF=Mixed Forest; CSH=Closed Shrubland; WSA=Woody Savanna; SAV=Savanna; OSH=Open Shrubland; GRA=Grasslands; WET=Wetlands.

Table 6. Mean bias (in mm day^{-1}) for the E_p methods per biome. The best performing method per biome is indicated in bold and is underlined. Based on unstressed days only defined using the energy balance criterion. Different colours are used to group biomes into broader ecosystem types (in descending order: croplands, grasslands, forests, savannah ecosystems, wetlands).

	Radiation, Temperature, VPD					*Radiation, Temperature*			*Radiation*			*Temperature*					
	Penman-Monteith			**Penman**		**Priestley and Taylor**			**Milly and Dunne**			**Thornthwaite**		**Oudin**		**Hargreaves-Samani**	
	Ref. crop	**Standard**	**Biome**	**Ref. crop**	**Standard**	**Ref. crop**	**Standard**	**Biome**	**Ref. crop**	**Standard**	**Biome**	**Standard**	**Biome**	**Standard**	**Biome**	**Standard**	**Biome**
CRO (10)	1.14	-0.49	0.84	2.83	2.64	2.20	0.47	**<u>-0.01</u>**	1.43	-0.24	0.12	-0.65	-0.59	-1.62	-0.62	-1.01	0.11
GRA (20)	2.65	0.53	1.16	4.37	1.90	3.69	1.11	**<u>0.02</u>**	2.57	0.22	-0.10	-0.14	-0.44	-0.61	-0.73	0.07	0.11
DBF (15)	0.30	-0.48	0.89	1.30	2.63	1.81	0.94	**<u>-0.06</u>**	0.74	-0.13	-0.15	-1.94	-2.03	-2.44	-0.71	-1.99	0.19
EBF (9)	0.70	**<u>0.04</u>**	0.95	1.39	1.74	1.39	0.79	0.16	0.79	0.17	-0.13	-0.83	-0.27	-0.53	-0.36	-0.84	0.20
ENF (26)	1.28	0.45	1.23	2.03	1.04	2.06	1.17	-0.05	1.88	0.90	**<u>0.02</u>**	-0.15	-0.05	-0.73	-0.54	-0.30	0.26
MF (4)	2.22	0.65	0.30	3.31	2.04	3.26	1.46	-0.07	2.53	0.87	-0.04	0.73	0.19	-0.01	-0.99	0.31	0.16
CSH (2)	1.01	0.49	**<u>0.00</u>**	1.61	1.79	1.46	1.10	-0.04	0.92	0.51	-0.14	0.18	0.39	0.14	-0.56	0.18	-0.18
WSA (5)	2.67	1.16	0.17	3.68	3.88	3.63	1.42	-0.03	2.33	0.40	-0.22	-0.14	-0.23	-0.20	-0.39	0.08	**<u>-0.02</u>**
SAV (6)	2.56	1.30	0.31	3.57	3.78	3.34	1.47	-0.15	2.21	0.54	-0.13	0.03	**<u>0.00</u>**	0.40	-0.93	0.53	-0.25
OSH (5)	4.32	1.68	0.37	6.20	2.73	5.08	2.00	0.10	3.89	1.15	**<u>0.00</u>**	1.13	0.84	0.81	-0.44	1.45	0.06
WET (5)	2.34	1.28	1.74	4.17	4.51	3.45	2.00	1.04	3.29	1.43	1.12	1.42	0.29	-0.52	-2.79	0.36	**<u>-0.11</u>**
Overall (107)	**1.69**	**0.40**	**0.93**	**2.88**	**2.21**	**2.67**	**1.14**	**0.04**	**1.92**	**0.45**	**<u>0.00</u>**	**-0.38**	**-0.45**	**-0.80**	**-0.72**	**-0.37**	**0.12**

CRO=cropland; DBF=Deciduous Broadleaf Forest; EBF=Evergreen Broadleaf Forest; ENF=Evergreen Needleleaf Forest; MF=Mixed Forest; CSH=Closed Shrubland; WSA=Woody Savanna; SAV=Savanna; OSH=Open Shrubland; GRA=Grasslands; WET=Wetlands.

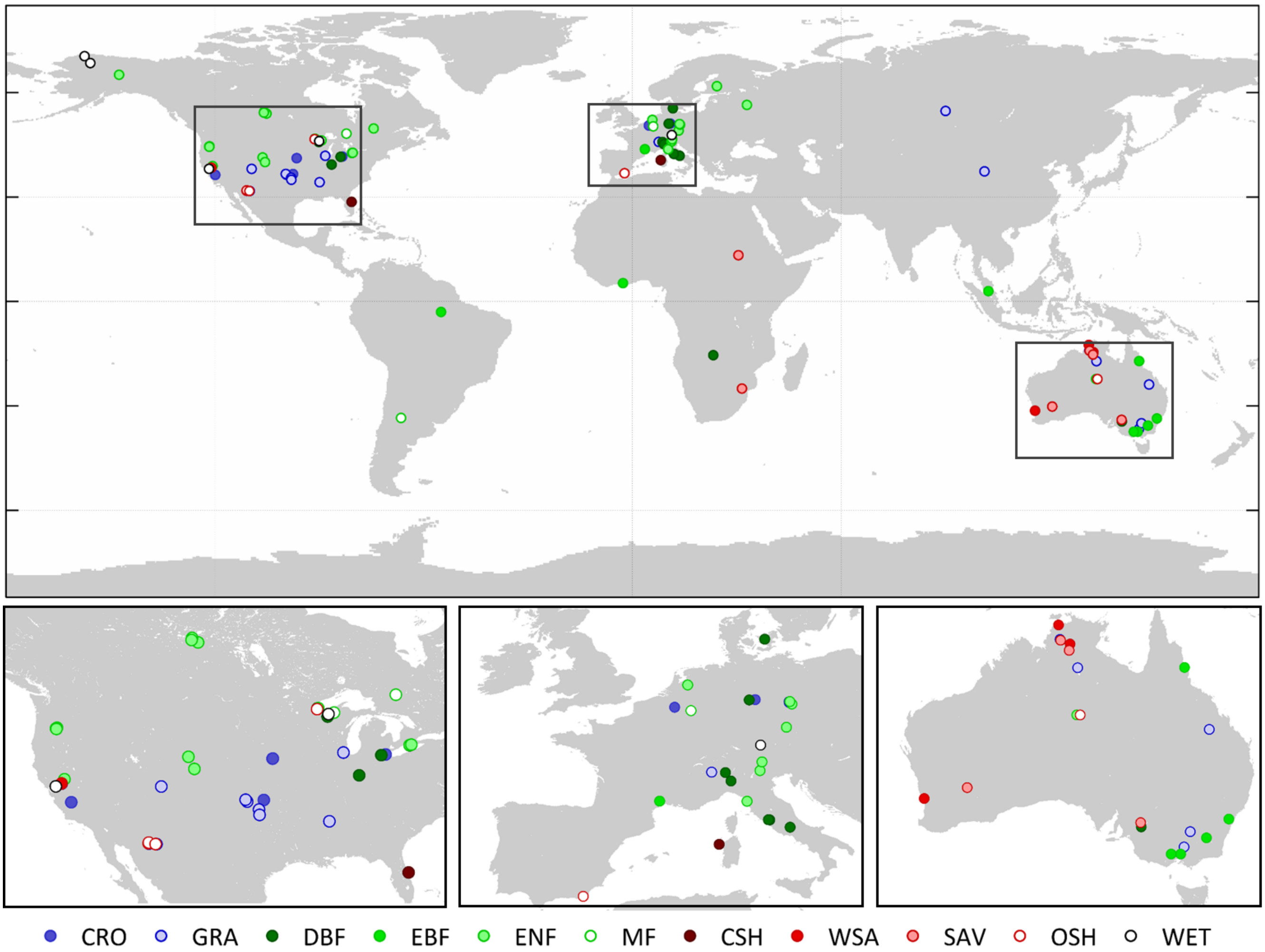

Figure 1. Location of the flux sites used in this study per biome. CRO=cropland; DBF=Deciduous Broadleaf Forest; EBF=Evergreen Broadleaf Forest; ENF=Evergreen Needleleaf Forest; MF=Mixed Forest; CSH=Closed Shrubland; WSA=Woody Savannah; SAV=Savannah; OSH=Open Shrubland; GRA=Grasslands; WET=Wetlands.

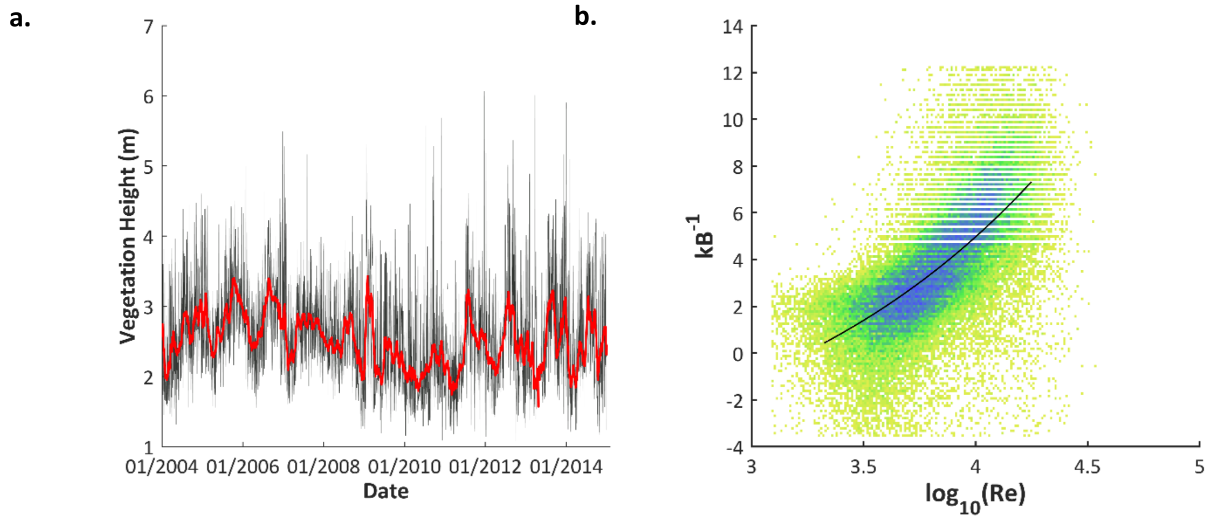

Figure 2. (a) Vegetation height dynamics in time (grey dots: half-hourly measurements; dark grey lines: daily mean vegetation height; red line: 30-day moving average, i.e. the final vegetation height dataset). (b) Relation between Stanton number (kB^{-1}) and Reynolds number (Re). Both plots correspond to the woody savannah site of Santa Rita Mesquite, US-SRM (Arizona, USA).

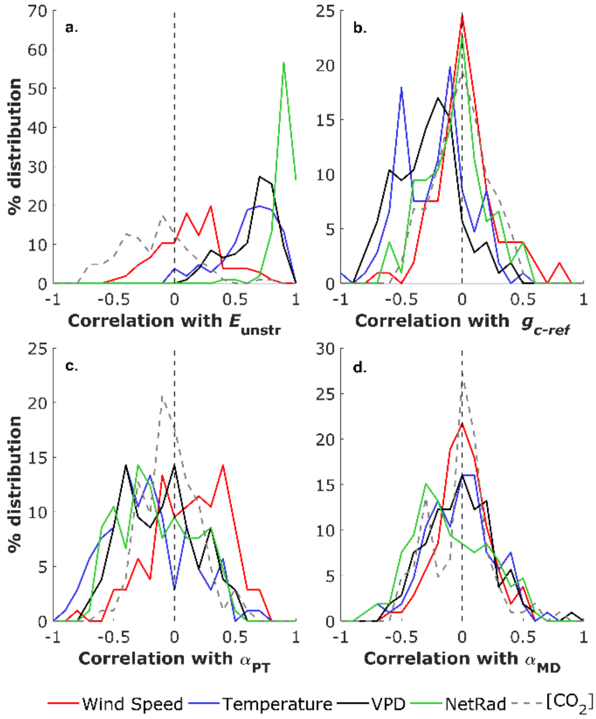

Figure 3. Histograms of correlations between the climate forcing variables and selected key parameters (a) E_{unstr} (b) g_{c_ref}, (c) α_{PT} and (d) α_{MD} measured in all flux tower sites. Based on unstressed days only defined using the energy balance criterion.

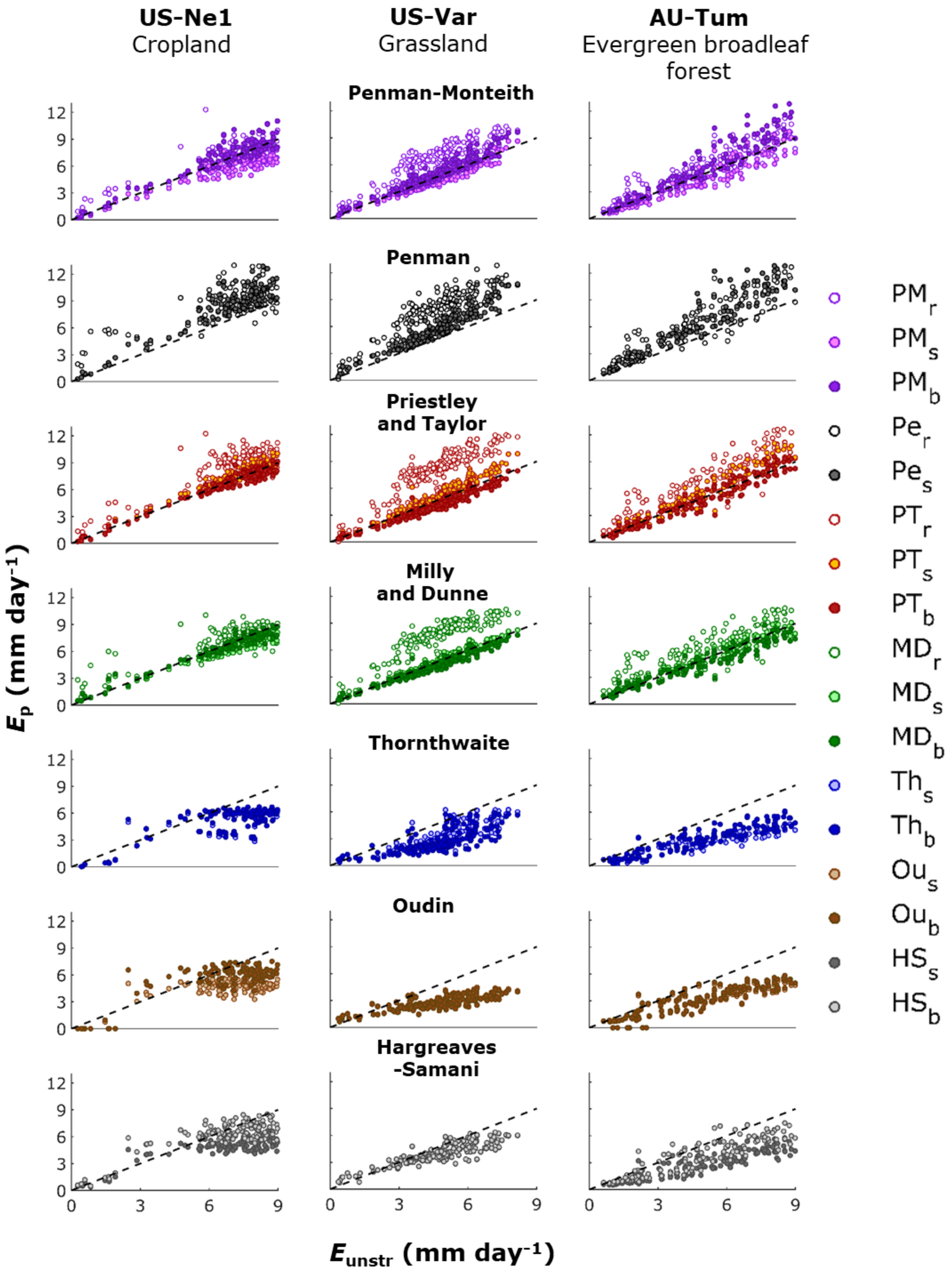

Figure 4. Scatterplot of the measured E_{unstr} versus E_p calculated with the different methods. The discontinuous line is the 1:1 line. Based on unstressed days only defined using the energy balance criterion.

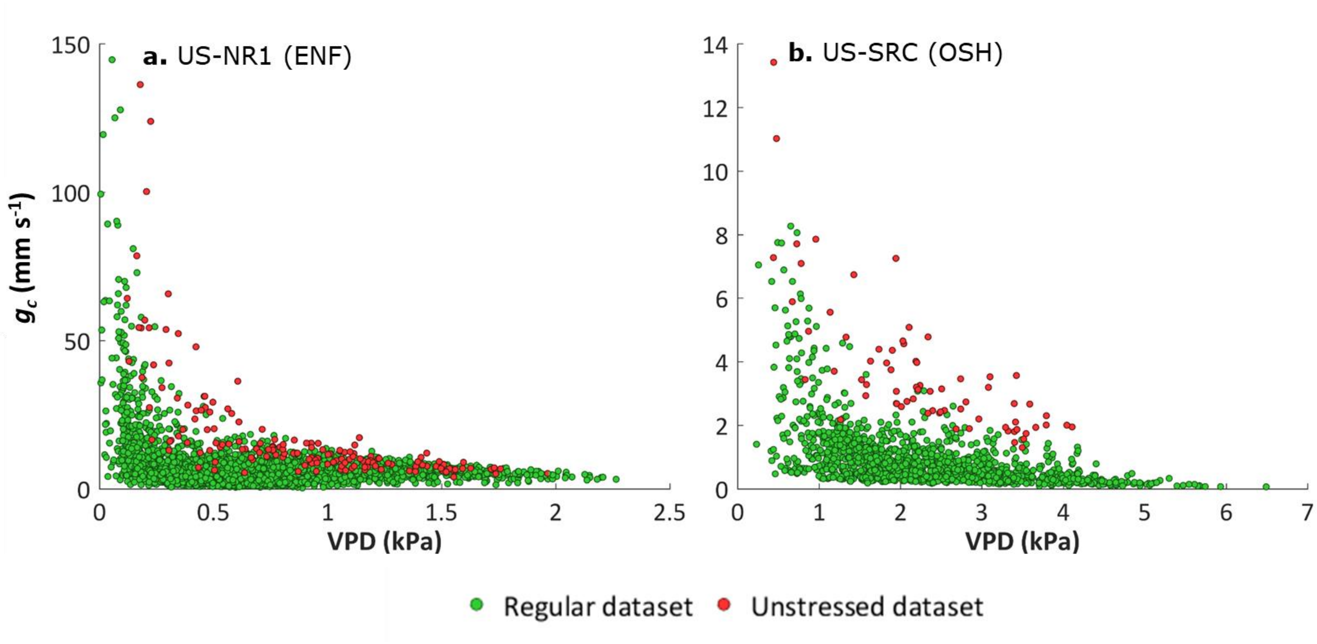

Figure 5. Surface conductance g_c as a function of vapour pressure deficit (VPD) of the regular and the unstressed datasets of two flux sites, (a) the evergreen needle forest Niwot Ridge Forest and (b) the open savannah woodland site Santa Rita Creosote.

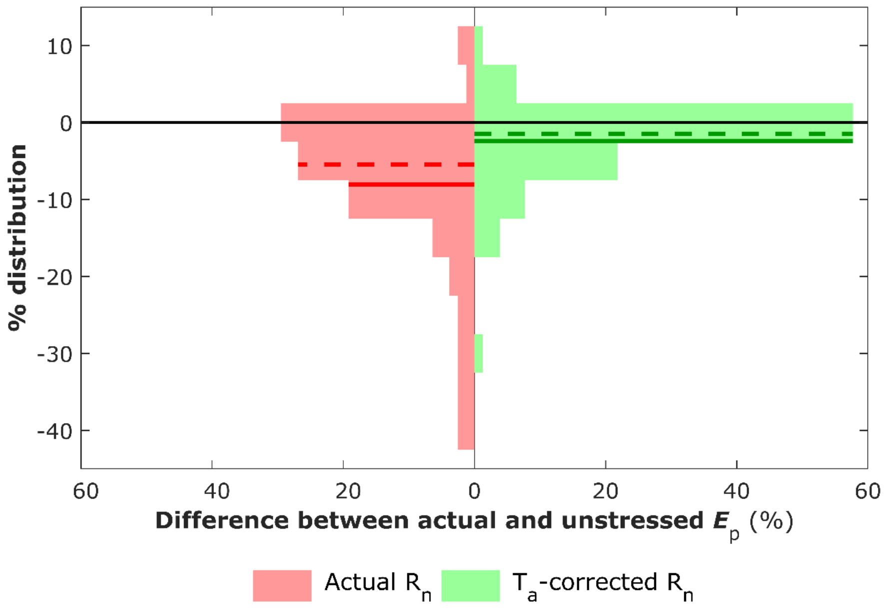

Figure 6. Comparison of the E_p calculated with a modelled method calculating (R_n-G) for unstressed conditions (Section S2) using flux tower data, and the actually observed (R_n-G) (red) or a simplified correction of (R_n-G) using T_a (Eq. (17)) (green). Negative Y-values indicate a lower estimation of E_p compared to the modelled method. For each distribution, the mean and median are indicated with a full and dashed line, respectively.

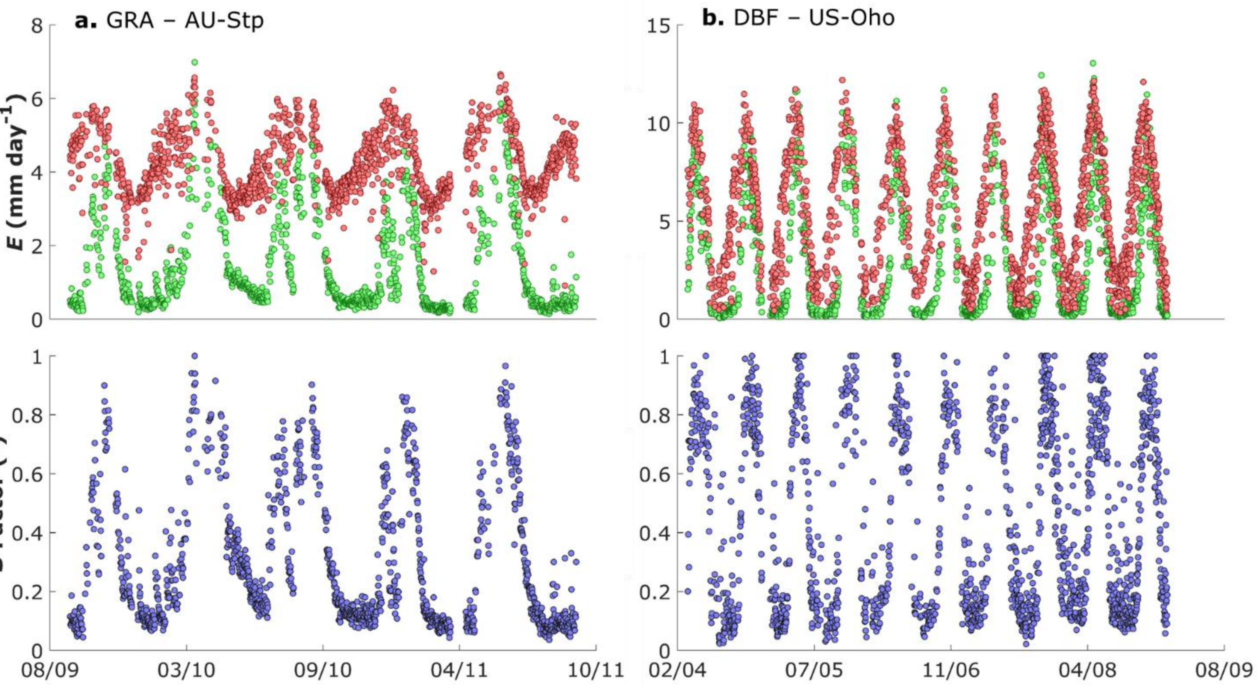

Figure 7. Seasonal evolution of E_a (top, green), E_p (top, red) and the *S* factor ($S=E_a\,E_p^{-1}$, confined between 0 and 1) for two ecosystems, (a) a grassland crop, Sturt Plains in Australia, and (b) a deciduous broadleaf forest, Ohio Oak forest in the USA. E_p was calculated with the MD_b-method and using the tower-based correction of (R_n-G) as presented in S2 of the supporting information.