Bridging scales: a temporal approach to evaluate global transpiration products using tree-scale sap flow data

Paulo R.L. Bittencourt[1], Lucy Rowland[2], Stephen Sitch[3], Rafael Poyatos[4], Diego G. Miralles[5], and Maurizio Mencuccini[4]

[1]College of Life and Environmental Sciences
[2]University of Exeterty
[3]University of Exeter
[4]CREAF
[5]Ghent University

December 7, 2022

Abstract

Transpiration is a key process driving energy, water and thus carbon dynamics. Global T products are fundamental for understanding and predicting vegetation processes. However, validation of these transpiration products is limited, mainly due to lack of suitable datasets. We propose a method to use SAPFLUXNET, the first quality-controlled global tree sap flow database, for evaluating transpiration products at global scale. Our method is based on evaluating temporal mismatches, rather than absolute values, by standardizing both transpiration and sap flow products. We evaluate how transpiration responses to hydro-meteorological variation from the Global Land Evaporation Amsterdam Model (GLEAM), a widely used global transpiration product, compare to in-situ responses from SAPFLUXNET field data. Our results show GLEAM and SAPFLUXNET temporal trends are in good agreement, but diverge under extreme conditions. Their temporal mismatches differ depending on the magnitude of transpiration and are not random, but linked to energy and water availability. Despite limitations, we show that the new global SAPFLUXNET dataset is a valuable tool to evaluate T products and identify problematic assumptions and processes embedded in models. The approach we propose can, therefore, be the foundation for a wider use of SAPFLUXNET, a new, independent, source of information, to understand the mechanisms controlling global transpiration fluxes.

Bridging scales: a temporal approach to evaluate global transpiration products using tree-scale sap flow data

Paulo Bittencourt[1], Lucy Rowland[1], Stephen Sitch[1], Rafael Poyatos[2], Diego G. Miralles[3], Maurizio Mencuccini[2,4]

[1] College of Life and Environmental Sciences, University of Exeter, Exeter, EX4 4RJ, UK

[2] CREAF, Bellaterra (Cerdanyola del Vallès), Catalonia, E08193, Spain

[3] Hydro-Climate Extremes Lab (H-CEL), Ghent University, Coupure links 653, 9000 Ghent, Belgium

[4] ICREA, Pg. Lluís Companys 23, Barcelona, 08010, Spain

Corresponding author: Paulo Bittencourt (paulo09d@gmail.com)

Key points

- Transpiration products are vital for understanding land-atmosphere processes, but their validation is limited by lack of suitable datasets.
- We propose a method to use SAPFLUXET - the first global database of tree sap flow data - to evaluate transpiration products at global scale.
- We show SAPFLUXNET to be a valuable tool to evaluate potential errors in the assumptions and processes embedded in transpiration models.

Abstract

Transpiration is a key process driving energy, water and thus carbon dynamics. Global T products are fundamental for understanding and predicting vegetation processes. However, validation of these transpiration products is limited, mainly due to lack of suitable datasets. We propose a method to use SAPFLUXNET, the first quality-controlled global tree sap flow database, for evaluating transpiration products at global scale. Our method is based on evaluating temporal mismatches, rather than absolute values, by standardizing both transpiration and sap flow products. We evaluate how transpiration responses to hydro-meteorological variation from the Global Land Evaporation Amsterdam Model (GLEAM), a widely used global transpiration product, compare to in-situ responses from SAPFLUXNET field data. Our results show GLEAM and SAPFLUXNET temporal trends are in good agreement, but diverge under extreme conditions. Their temporal mismatches differ depending on the magnitude of transpiration and are not random, but linked to energy and water availability. Despite limitations, we show that the new global SAPFLUXNET dataset is a valuable tool to evaluate T products and identify problematic assumptions and processes embedded in models. The approach we propose can, therefore, be the foundation for a wider use of SAPFLUXNET, a new, independent, source of information, to understand the mechanisms controlling global transpiration fluxes.

Plain language summary

Transpiration, the water evaporating from leaves, is a key element in the energy, water and carbon cycles of terrestrial ecosystems. Understanding patterns of transpiration at global scales is fundamental for prediction of future climates. Several models are used for estimating global transpiration, however identifying limitations and biases in these models is difficult, because we lack field data to compare them against. In this work, we propose a new method to enable tree-level sap flow data from SAPFLUXNET, the first global sap flow database, to be used to evaluate transpiration products and models. We evaluated how well GLEAM, a widely used transpiration product, matches

SAPFLUXNET field data. We found GLEAM and SAPFLUXNET data to be in reasonable agreement however, mismatches occur under extreme dry or wet meteorological conditions, conditions which are likely to become more common under future climates. The detection of mismatches between SAPFLUXNET and GLEAM data is valuable for the identification of model processes and assumptions which could be reasonable within current climate, but inadequate for future climate conditions. The method we propose allows the use of SAPFLUXNET to understand the true mechanisms controlling global transpiration providing a new, independent, source of information to evaluate transpiration products and models.

Index terms: 3322 Land/atmosphere interactions, 1840 Hydrometeorology, 1878 Water/energy interactions, 0426 Biosphere/atmosphere interactions

Keywords: transpiration, sap flow, SAPFLUXNET, GLEAM, transpiration scaling, product validation

1 Introduction

Transpiration (T), the evaporation of water from within plants, is a key process linking ecosystem energy, water and carbon dynamics, and accounts for ~60% of global terrestrial evaporation, or 'evapotranspiration' (ET) (Wei *et al.*, 2017; Stoy *et al.*, 2019). T is regulated by a complex combination of energy availability and soil and atmospheric water stresses (Dolman *et al.* 2014). The responses of T to drought stress, at leaf, plant, and ecosystem scales, remain a huge source of uncertainty in understanding biosphere-atmosphere feedbacks (Maes *et al.* 2020).Understanding T responses under climate change is an even more challenging task, as responses to combined environmental changes, for example changes in water, nitrogen and CO_2 availability, alongside land use changes additively and interactively modulate the way T is controlled by vegetation (Lemordant *et al.* 2018, Keenan *et al.* 2013). Additionally, ongoing global changes are causing plants to acclimate and communities to change, which might be shifting or modifying the way T is regulated by vegetation (Kumarathunge *et al.* 2019, Stephens *et al.* 2021). Recent studies indicate climate change is making global T fluxes more sensitive to vegetation responses (Forzieri *et al.* 2020). Global T products are therefore key to help us determine the mechanisms driving plant and ecosystem T at global scales and to monitor vegetation responses as climate changes. However, without quality-controlled T products, validated against empirical data, our capabilities to predict land surface interactions may be limited (Stoy *et al.*, 2019).

In the past decade, multiple models have been developed to derive global T and ET largely from remotely sensed (RS) data (Fisher et al. 2017). These RS-derived ET products, such as the Global Land Evaporation Amsterdam Model (GLEAM; Miralles *et al.*, 2011; Martens *et al.*, 2017) are used for a diversity of purposes, e.g., quantification of water resources (Immerzeel *et al.*, 2020), driving basin hydrological models (Dembélé *et al.*, 2020), studying global climate (Miralles *et al.*, 2014; Martens *et al.*, 2018) and benchmarking climate models, such as those from CMIP6 (Wang *et al.*, 2021). These RS models retrieve ET indirectly by applying process-based (Miralles *et al.*, 2016) or machine learning (Jung *et al.*, 2019) algorithms. This modelling induces errors, which are tightly related to the difficulties to properly capture the T component of ET, whose uncertainties can be two to three times larger than for the total ET (Miralles *et al.*, 2016; Talsma *et al.*, 2018; Feng *et al.*, 2020). Model improvement is limited by a lack of suitable datasets to directly validate T products, test the model`s embedded mechanisms and constrain its parameters (Stoy *et al.*, 2019). In fact, validation exercises are often insufficient (Bayat *et al.*, 2021), hindered by the sparseness of *in situ* data (Fisher *et al.*, 2017) and the limited availability of measurement techniques and datasets at the necessary spatial and temporal scales (Kool *et al.*, 2014; Talsma *et al.*, 2018; Bayat *et al.*, 2021).

Plant gas exchange measurements in the field provide accurate T data at leaf or branch level (e.g., Sabater *et al.*, 2020), but are difficult to scale and monitor continuously. Isotope-based methods can be used to unravel the T components of ET and provide information at ecosystem scale (Williams *et al.*, 2004), but are expensive and require additional information for end-member analysis. Most commonly, the validation of T products involves the use of latent heat flux measurements from eddy covariance, basin-level water balances, soil lysimeters or soil water balance approaches – yet all these methods involve explicit assumptions regarding the partitioning of ET. Carbonyl-sulphide flux (Whelan *et al.*, 2018) and solar-induced fluorescence (Maes *et al.*, 2020) measurements have also been used to independently evaluate T products, however both rely on physiological modelling assumptions to derive T.

On the other hand, sap flow (SF) measurements are a promising source of information to directly evaluate T products and model mechanisms (Wang & Dickinson, 2012; Stoy *et al.*, 2019; Poyatos *et al.*, 2021). At daily or longer time scales, average SF can be equated to T with minimal errors (Kumagai *et al.*, 2009; Kool *et al.*, 2014). To date, SF data have never been used to evaluate T products globally, due to limitations in data availability (Stoy *et al.*, 2019). However, a new coordinated network of SF data (SAPFLUXNET; (Poyatos *et al.*, 2016, 2021)) has recently generated the first quality-controlled SF dataset at a global scale. SAPFLUXNET opens new opportunities to validate T products directly (Bright *et al.* 2022). However, new generalised procedures need to be developed to enable the comparison between tree level T and T at larger spatial scales (Nelson *et al.*, 2020). SF is usually measured on a unit-sapwood-area basis, and scaling SF to tree level is a common procedure with known sources of uncertainty, requiring estimation of tree sapwood area and knowledge of wood thermal and anatomical traits (Forster, 2017; Flo *et al.*, 2019). However, scaling tree-level SF to stand-level poses a more difficult challenge, as it requires within and between species replication of SF measurements to account for individual, size and species variations, as well as forest inventory and structure information to weigh the importance of trees of different sizes and species to stand SF (Čermák *et al.*, 2004). Scaling from stand-level (hundreds of meters to a few kilometres) to global datasets spatial scales (10–50km), requires further consideration of landscape heterogeneity, which increases uncertainty (Ford *et al.* 2008; Mackay *et al.* 2010). Consequently, the use of sap flow data to evaluate T products has so far been limited to few sites (Nelson *et al.*, 2020).

In this study, we use the novel SAPFLUXNET dataset to evaluate the GLEAM T product under different climate conditions, and explore potential mismatches between the two estimates of T. We develop a new procedure which shortcuts the challenges of scaling site SF to grid cell T by focussing on temporal mismatches rather than absolute values. We use SF data from >80 sites across the globe and analyse temporal mismatches between GLEAM and SAPFLUXNET to demonstrate the capacity of our new approach to contribute to validating global T products and testing their assumptions. While comparisons between grid-scale and individual scale T at individual sites may be subject to large sources of systematic biases caused by lack of representativeness of the temporal trends in the sampled trees relative to the entire pixel, we propose here that, by analysing a sufficient large number of sites under different environmental conditions, these systematic site-specific biases will average out allowing to identify general differences between the behaviour of ground SF data and modelled T data. We assess, for days with low, median, and high transpiration values, (i) how GLEAM and SAPFLUXNET compare over time, (ii) whether GLEAM and SAPFLUXNET sensitivity to vapour pressure deficit and radiation match, and (iii) whether temporal mismatches between the products can be explained by site model parameters and meteorological conditions. Although our analysis is limited to GLEAM, the generic approach that we present could easily be applied to validate other remotely sensed T products, as well as T fields and models from land-surface, climate and hydrological models.

2 Material and Methods

2.1 Sap flow and transpiration datasets

We use the SAPFLUXNET global database of tree SF (SFN v0.1.5; Poyatos *et al.*, 2021). SAPFLUXNET contains half-hourly tree-level SF data and is accompanied by tree metadata (size, species, SF sensor type), site information (vegetation type, soil, elevation, etc) and local hydro-meteorological data. Normally, multiple trees of different species are sampled per site and SF data are given per unit xylem area, per unit leaf area or per tree. We use all SAPFLUXNET data available after filtering out sites which either (i) had non-native vegetation, (ii) were affected by experimental manipulations or recent fire, or (iii) had less than 6 months of data available, considering only months with at least 20 days of data. After this filtering, the total number of sites available was 83 and the total number of trees was 1195 (Table S1).

We use the outputs from the GLEAM model (Miralles *et al.*, 2011; Martens *et al.*, 2017). GLEAM uses remote sensing data to calculate potential ET based on the Priestley & Taylor (1972) model. Potential ET is converted into actual ET using models of water stress derived from vegetation optical depth and root-zone soil moisture; the latter is calculated based on retrievals of precipitation and surface soil moisture. This procedure is applied at a daily time step to each land fraction of a 0.25° (~25km at equator) grid cell (water, soil, short and tall vegetation); these fractions are derived based on the Moderate Resolution Imaging Spectroradiometer (MODIS) product MOD44B (DiMiceli, Charlene *et al.*, 2015). For each grid cell, the contribution per land fraction is then aggregated, and rainfall interception based on the (Gash, 1979) model is added to yield the total ET. Here, we use the GLEAM v3.5b tall vegetation T product. For each SAPFLUXNET site, we extracted the GLEAM time series from the corresponding 0.25° grid-cell.

2.2 Meteorological data

To describe the sensitivity of SAPFLUXNET and GLEAM to environmental drivers and site climate, we obtain time series of mean monthly incoming surface solar radiation ($S_{\downarrow}$), air temperature and vapour pressure deficit (VPD) from 2003 to 2018 for each site. For $S_{\downarrow}$ and air temperature we use the ERA5 reanalysis (Hersbach *et al.*, 2020) at the monthly time scale. We calculate VPD from the CRUJRA monthly dataset of air vapour pressure and air temperature (Harris *et al.*, 2020) after standardizing it to each site elevation.

2.3 Scaling sap flow temporal patterns from tree to site

To scale SF temporal variability from tree level to stand level, we first average hourly to daily SF for each tree after filtering out nighttime data. We define nighttime as any hour in which solar altitude –

the angle between the sun and the horizon – is lower than 0°. We calculate solar altitude for each hour using the site latitude, longitude and astronomical geometry (Michalsky, 1988) using the "sunAngle" method in the R package "oce" (Kelley & Richards 2020). We then standardize the daily average SF per tree by calculating its Z-score (i.e., subtracting the mean and dividing by the standard deviation of the entire time series; Fig. 1a, b). Z-scores remove differences in absolute values across sites while preserving information on temporal variability, facilitating comparisons among heterogeneous samples. Therefore, this standardization has the effect of removing size- and species-dependent effects on SF mean and variance, while retaining the full temporal variability of the data. We then scaled SF temporal variability to site level by averaging the standardized SF of all trees for each site (Fig. 1c). We performed analogous experiments using diameter-at-breast height weighted mean but found no differences in results and thus decided to report site-level scaling using mean only.

2.4 Extraction of low, median and high transpiration and sap flow days

To evaluate the agreement between GLEAM and SAPFLUXNET for days with contrasting conditions, we extract T and SF values representative of days with low, median and high T and SF conditions. We first quantify the monthly distribution, for each site, of SF and T using R`s base function quantile with default arguments (i.e. method 7 of Hyndman & Fan 1996, based on modal position). Then, from each distribution of SF and T, we extracted the 5th, 50th and 95th percentiles of T and SF (Fig. 1c-d to Fig. 1e-f). The resulting time series reflect the monthly dynamics of the days with low, median and high T and SF. Then, for each site-level time series of monthly percentiles, we standardize the values by calculating Z-scores so that T and SF temporal variability could be compared (Fig. 1e-f to Fig. 1g-h). This is the same process used to standardize tree-level SF values within a site (see previous section). Here, the Z-score standardization removes any information on absolute values from both SF and T, so that the variability in SF and T is now in the same scale (i.e., standard deviation units) and can be directly compared. Hereafter, we refer to these Z-score standardized values as GLEAM-T and SAPFLUXNET-SF consistently.

2.5 Site level GLEAM and SAPFLUXNET agreement indexes

For each site, we calculate two indices to evaluate how well GLEAM-T matches SAPFLUXNET-SF over time: 1) the root mean squared difference (RMSD) of T in relation to SF (Fig. S2c) and 2) the bivariate correlation between T and SF (r - the Pearson`s correlation):

1) $\text{RMSD} = \frac{\sum_i^j \sqrt{(T_m^2 - SF_m^2)}}{n}$

2) $r = \frac{\sum_i^j (T_m - \overline{T})(SF_m - \overline{SF})}{\sqrt{\sum_i^j (T_m - \overline{T})^2 \sum_i^j (T_m - \overline{T})^2}}$

Where "i" and "j" are the first and last month in the time series, "m" indicates a given month, "n" the total number of months and the overline symbol for T and SF indicates the mean of the time series. Both indices were calculated for each of the time series (i.e., low, median and high T and SF percentiles).

2.6 Sensitivity to vapour pressure deficit and solar radiation

For each site, we calculate the sensitivity of T and SF to VPD and $S_{\downarrow}$, by fitting the data using a linear mixed-effect model (Zuur *et al.*, 2009), with VPD and $S_{\downarrow}$ having both a fixed effect on T or SF (first two terms on right-hand side on equations 3 and 4, overall intercept and slope), as well as a random effect depending on site (two terms following the vertical bar, indicating that intercepts (the 1s) and slopes vary by site):

3) T or SF = a + b*VPD + (1 + VPD|site)

4) T or SF = a + b* $S_{\downarrow}$ + (1 + $S_{\downarrow}$|site)

Mixed-effects models produce both population-level estimates of the mean intercepts and slopes for all sites, as well as site-level estimates of these same quantities (best linear unbiased predictions). These site-dependent intercepts and slopes of the response functions against VPD or $S_{\downarrow}$, allow us to compare T versus SF sensitivities across sites. VPD and $S_{\downarrow}$ values were centred prior to use in the model. Procedures for fitting the linear mixed models are the same as those used in hypothesis testing and described in the next section. We calculate the VPD or $S_{\downarrow}$ sensitivity mismatch (VPD_{sm} and $S_{\downarrow sm}$), for each site, as GLEAM-T`s sensitivity to VPD or $S_{\downarrow}$ minus the site SAPFLUXNET-SF sensitivity to VPD or $S_{\downarrow}$.

2.7 Analysis

We evaluate whether GLEAM-T scales proportionally to SAPFLUXNET-SF (i.e., whether the scaling relationship is consistent with a 1:1 relationship) and whether the scaling is different among days with low, median and high transpiration (i.e., whether the scaling relationship changes with the percentile analysed) using standardized major axis regression (SMA; Smith, 2009). We then test whether site-level indices of mismatching between T and SF (RMSD and r) are different for different percentiles using a mixed-effect model, where the mismatching indices are the response variable, the percentile is the fixed effect and site is a random effect on the intercept, which allows pairing percentiles by site and controlling for site effects. We use the same approach to evaluate how VPD_{sm} and $S_{\downarrow sm}$ scale and whether the scaling is affected by percentiles. Moreover, we evaluate whether mismatches between GLEAM-T and SAPFLUXNET-SF were explained by site climatology (long-term site-averages of VPD, $S_{\downarrow}$, temperature and precipitation) and GLEAM input variables (S, potential and actual ET) using linear fixed effect models. We use principal component analysis (PCA) to collapse the variables into principal components as they were highly correlated. We evaluate the first and second PCA axis capacity to explain variability of the mismatch indices for the different percentiles.

We used the R programming environment (v3.6; R Core Team 2019) for all analysis and data processing; R base package for linear fixed-effects models (function "lm") and PCA (function "prcomp"); the SMATR3 package (Warton *et al.*, 2012) for SMA analysis; the NLME package (Pinheiro *et al.* 2020) for mixed-effect models. We followed the guidelines of (Zuur *et al.*, 2009) and Thomas *et al.* (2017) for assessing significance of model terms, validating model assumptions and verifying model sensitivity to outliers using Cook's distance. We tested for significance of fixed variables in mixed-effect models using likelihood ratio tests between the model with and without the fixed effect.

3 Results

3.1 GLEAM and SAPFLUXNET *scaling and occurrence of temporal mismatches*

Analysing the agreement between GLEAM-T and SAPFLUXNET-SF using standardized major-axis regression, we found their temporal variability scales with a slope of 1.06 ± 0.007 (mean ± confidence interval here and in following values) and with an intercept of 0.20 ± 0.008 ($p < 0.001$; Fig. 2. This indicates a good match in temporal patterns between GLEAM-T and SAPFLUXNET -SF, despite a high overall variability ($R^2 = 0.30$). The scaling for days with low, median and high transpiration (i.e., the 5^{th}, 50^{th} and 95^{th} percentiles – P05, P50 and P95) differed across percentiles ($p < 0.001$; Fig. 3). The percentiles had significantly different slopes (0.94 ± 0.03, 1.03 ± 0.04 and 1.01 ± 0.04 for P05, P50 and P95, respectively; $p < 0.001$) and the intercept of the relationship was close to zero for all percentiles (-0.04 ± 0.04, -0.004 ± 0.04 and -0.003 ± 0.03 for P05, P50 and P95). Their agreement explained 32% of the variability of P05, 39% of P50 and 34% of P95. These results indicate that GLEAM-T captures the overall SAPFLUXNET -SF temporal variability, but the match differs for different transpiration conditions as shown by the slope between SAPFLUXNET-SF and GLEAM-T being lower than one for low transpiration conditions. We also found this result to be robust when the analysis was repeated using tree diameter at breast height to calculate site SF using weighted mean, instead of simple mean (data not shown).

We tested whether site-level statistics of the match between the variability of GLEAM-T and SAPFLUXNET-SF (root mean squared deviation, RMSD and bivariate correlation, r) were different across percentiles (Fig. 4a-c). We found RMSD of the P50 to be 0.18 ± 0.01, which is 10.4% and 9.5% lower than the RMSD of P05 and P95 ($p <= 0.03$; Fig. 4a). Similarly, the bivariate correlation of SF and T (r) was greater for the P50 (0.62) and lower for the P05 and P95 (0.54 and 0.56; $p <= 0.01$; Fig. 4c), indicating GLEAM-T has a better temporal match to SAPFLUXNET-SF under median conditions.

3.2 Differences in sensitivity to VPD and $S_{\downarrow}$ *between GLEAM-T and* SAPFLUXNET *-SF*

We analysed how site-specific sensitivities of GLEAM-T and SAPFLUXNET-SF to VPD and $S_{\downarrow}$ relate to each other and whether this relationship was different across daily conditions with low, median and high transpiration, using standardized major axis regression. Our results show sensitivity to VPD scaled with a similar slope of 0.76 for all percentiles ($p = 0.15$ for slope differences across percentiles; Fig. 5a), but with different intercepts of -0.34, 0.14 and 0.07 for P05, P50 and P95 ($p < 0.001$), causing GLEAM-T sensitivity to VPD to approach SAPFLUXNET -SF sensitivity at lower VPD sensitivity sites. The scaling between GLEAM-T and SAPFLUXNET-SF sensitivity to VPD is significant for all percentiles ($p < 0.001$) and explained 39%, 49% and 49% of the variability in the relationship for P05, P50 and P95. The VPD sensitivity mismatch (VPD_{sm}) is higher for P05 than P50 and P95 ($p < 0.001$; Fig. 4d) but was always above 0, indicating a higher VPD sensitivity overall for GLEAM-T across all percentiles.

Regarding radiation responses, GLEAM-T and SAPFLUXNET -SF show again a good scaling to the 1:1 line, with a slope of 0.91 for all percentiles ($p = 0.87$; Fig. 5b). The intercepts were significantly different across the percentiles (-0.030, -0.008 and -0.008 for P05, P50 and P95; $p < 0.001$). The $S_{\downarrow}$ sensitivity mismatch ($S_{\downarrow sm}$) increases from P95 to P05 ($p < 0.01$; Fig. 4e).

3.3 Drivers of mismatches between GLEAM-T and SAPFLUXNET-SF

We evaluated whether mismatches between GLEAM-T and SAPFLUXNET -SF (RMSD and r), and their VPD_{sm} and $S_{\downarrow sm}$, were related to site-level climate data (VPD, $S_{\downarrow}$, air temperature and precipitation) or model variables (potential ET, actual ET and GLEAM`s stress factor S). To simplify the analysis, we collapsed the predictor variable space onto two principal component analysis (PCA) axes (Fig. 6). The first and second axis of the PCA (PC1 and PC2) explained most of the dataset variability (50% and 38%) and we restricted our analysis to these axes. PC1 inversely reflected variables which control a site's evaporative demand (VPD, $S_{\downarrow}$ and temperature) while the PC2 directly water limitation related variables (precipitation and actual ET; Table 1). GLEAM`s water stress factor and potential ET were distributed across both axes. We found the different predictors of mismatch between GLEAM-T and SAPFLUXNET -SF to be related to both the first and the second PCA axes (Table 2). The GLEAM-T to SAPFLUXNET-SF bivariate correlation for all percentiles and the VPD_{sm} for the P5 and P95 increase with PC1 (i.e., they decrease with increased evaporative demand). RMSD, VPD_{sm} and $S_{\downarrow sm}$ increased with PC2 (i.e. site actual ET and precipitation). Our results indicate GLEAM-T mismatches relative to SAPFLUXNET-SF are not random and are related to site level differences in evaporative demand and water availability, generally increasing with them. However, the way in which both site level evaporative demand and water availability influenced the GLEAM-T vs. SAPFLUXNET -SF mismatches varied depending on the percentile analysed (P5, P50, P95). This suggests the driver was often different for different transpiration conditions and, thus, the capacity of GLEAM to capture T is not the same for mean and extreme, low and high, T conditions.

4 Discussion

Evaluating T products has been a major challenge preventing improvements in our capabilities to understand and predict water and energy dynamics (Stoy *et al.*, 2019). While the use of sap flow has been proposed as a mean to evaluate T datasets, constraints in spatially scaling these fluxes have limited these evaluations to a handful of sites globally (Nelson *et al.*, 2020). Using the recently assembled and quality-controlled SAPFLUXNET database (Poyatos *et al.*, 2021), combined with a novel approach to allow stand-scale comparisons to global T products, we provided the first global evaluation of a widely used transpiration model – GLEAM (Martens *et al.*, 2017). Our new technique can be used to infer GLEAM-T and SAPFLUXNET-SF have a strong temporal agreement (Fig. 2 and 3) with a scaling close to 1:1 and an intercept close to 0. Interestingly, days with different transpiration levels scale differently, with low transpiration days scaling with a slope of 0.94, leading to higher mismatches at extreme values. Therefore, the mismatch will be greater for extreme low and high transpiration conditions within a site and between sites with different conditions, highlighting the limitations of T products to capture extreme patterns (Miralles *et al.*, 2016; Talsma *et al.*, 2018; Feng *et al.*, 2020).

Our work has shown that a quality controlled, standardized, SF product can be used for large-scale evaluation of the temporal trends in T products at monthly time scales. While the analysis of temporal patterns constitutes only a partial validation of a product, it provides valuable information on mechanisms which should be targeted for product improvement. Our results show, for example, days with low transpiration to be particularly problematic for GLEAM`s current model. GLEAM-T generally captures the VPD and $S_{\downarrow}$ sensitivities well, but overestimates them slightly but systematically relative to SAPFLUXNET-SF (Fig. 4d and e), especially for low transpiration conditions. Lower agreement between GLEAM and eddy-covariance data in arid conditions has been reported previously (Michel *et al.*, 2016), but to our knowledge, this is the first time T mismatches under low evaporative conditions have been identified generally. Ultimately, the fact that GLEAM is overly-responsive to radiation under low transpiration conditions relates to the use of the Priestley and Taylor formulation, which has difficulties to properly capture ET at low radiation conditions (Fisher et al., 2011; Miralles *et al.*, 2016). While solar radiation and temperature (which drive the Priestley and Taylor model) account for most of the variability in atmospheric demand, air humidity and wind speed also have some influence (Penman, 1948). This could be the cause of the mismatches in RMSD and VPD and $S_{\downarrow}$ sensitivities increasing with site energy-availability (Table 2). Our new method highlights these biases as potential targets for further model development. Such development is particularly significant considering the importance of ensuring these products capture extreme values of transpiration correctly, given the likelihood that extreme values of transpiration are likely to increase globally (Diffenbaugh *et al.*, 2017) and the fact that RS products are used to evaluate global climate models (Wang *et al.*, 2021).

Our tree-to-grid cell scaling approach does however have limitations – analysis is restricted to relative temporal trends rather than absolute values. Our work also assumes sap flow sensor data is equally accurate at different transpiration conditions, which may not be true (Flo *et al.* 2019). Using temporal trends of SF and T also cannot address issues of spatial mismatches between the products (often 0.25° for GLEAM-T versus one site/forest for SAPFLUXNET-T), which could be driving some of the disagreements between the products if site values are not representative of the broader landscape dynamics within that grid cell. Furthermore, it is possible that unmeasured trees have a different temporal dynamics compared to measured trees. All these sources of potential error should cause site-specific differences in temporal patterns. Given a sufficiently large number of sites however, such as used in this study, the differences are expected to be random, rather than creating the systematic mismatches we observe, which are instead related to climatic variables and GLEAM model parameterisation (Table 2). Consequently, with our approach confidence in conclusions reached for specific sites is limited, but cross-site analyses are likely to be robust.

5 Conclusions

Our work provides an initial template which could be expanded to evaluate other remote sensing based or T products, or T estimates from land surface and hydrological models. Other types of analyses, such as time lags between driver and T response and spatial correlations analysis, could provide valuable insights into evaluating other types of mismatches. A bridge between our approach, based on temporal trends, to an approach based on absolute SF values, such as done by Nelson *et al.* (2020), could be done by a joint comparison of both methods for those sites where sufficient data are available for this analysis. Future expansion of SF monitoring in a controlled and standardized way, particularly if paired with eddy-covariance towers, could greatly improve our capacity to utilize SF data to evaluate T products and optimize merging of different products (Jiménez *et al.*, 2018). Models behind global T products usually assume parameters are constant, which is an incorrect but necessary assumption, given the lack of data needed to monitor parameter stationarity (Stephens *et al.* 2021). Improved capabilities of evaluating T products, such as a global SF network, may also provide means to monitor how ongoing changes in vegetation structure and physiological acclimation to climate change may be shifting the parameters embedded in T products. We believe the initial steps we provide here can be the foundation for a wider SF based validation of T products, models and mechanisms.

Acknowledgements

This research was supported by the Newton Fund through the Met Office Climate Science for Service Partnership Brazil (CSSP Brazil) and a NERC independent fellowship grant NE/N014022/1 to LR. PRLB acknowledges support from NERC standard grant NE/V000071/1. RP was supported by the Spanish MICINN grant RTI2018-095297-J-I00 and by a Humboldt Fellowship for Experienced Researchers. DGM acknowledges support from the European Research Council (ERC) under grant agreement no. 715254 (DRY–2–DRY). We have no conflict of interest to declare.

Data Availability Statement

All data used in this work is freely available at the GLEAM (https://gleam.io/) and SAPFLUXNET (http://sapfluxnet.creaf.cat/) online repositories.

Author Contribution

PRLB, LR and MM conceived the research ideas, developed the project and wrote the manuscript. All authors contributed to manuscript preparation.

References

Bayat B, Camacho F, Nickeson J, Cosh M, Bolten J, Vereecken H, Montzka C. 2021. Toward operational validation systems for global satellite-based terrestrial essential climate variables. *International Journal of Applied Earth Observation and Geoinformation* 95: 102240.

Bright RM, Miralles DG, Poyatos R, Eisner S. 2022. Simple Models Outperform More Complex Big-Leaf Models of Daily Transpiration in Forested Biomes. *Geophysical Research Letters* 49.

Čermák J, Kučera J, Nadezhdina N. 2004. Sap flow measurements with some thermodynamic methods, flow integration within trees and scaling up from sample trees to entire forest stands. *Trees* 18: 529–546.

Dolman AJ, Miralles DG, de Jeu RAM. 2014. Fifty years since Monteith's 1965 seminal paper: the emergence of global ecohydrology: EMERGENCE GLOBAL ECOHYDROLOGY. *Ecohydrology* 7: 897–902.

Kelley D & Richards C. 2020. oce: Analysis of Oceanographic Data. R package version 1.2-0. https://CRAN.R-project.org/package=oce

Dembélé M, Ceperley N, Zwart SJ, Salvadore E, Mariethoz G, Schaefli B. 2020. Potential of satellite and reanalysis evaporation datasets for hydrological modelling under various model calibration strategies. *Advances in Water Resources* 143: 103667.

Diffenbaugh NS, Singh D, Mankin JS, Horton DE, Swain DL, Touma D, Charland A, Liu Y, Haugen M, Tsiang M, *et al.* 2017. Quantifying the influence of global warming on unprecedented extreme climate events. *Proceedings of the National Academy of Sciences* 114: 4881–4886.

DiMiceli, Charlene, Carroll, Mark, Sohlberg, Robert, Kim, Do-Hyung, Kelly, Maggi, Townshend, John. 2015. MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006.

Feng S, Liu J, Zhang Q, Zhang Y, Singh VP, Gu X, Sun P. 2020. A global quantitation of factors affecting evapotranspiration variability. *Journal of Hydrology* 584: 124688.

Fisher JB, Melton F, Middleton E, Hain C, Anderson M, Allen R, McCabe MF, Hook S, Baldocchi D, Townsend PA, *et al.* 2017. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources: THE FUTURE OF EVAPOTRANSPIRATION. *Water Resources Research* 53: 2618–2626.

Flo V, Martinez-Vilalta J, Steppe K, Schuldt B, Poyatos R. 2019. A synthesis of bias and uncertainty in sap flow methods. *Agricultural and Forest Meteorology* 271: 362–374.

Ford CR, Hubbard RM, Kloeppel BD, Vose JM. 2007. A comparison of sap flux-based evapotranspiration estimates with catchment-scale water balance. *Agricultural and Forest Meteorology* 145: 176–185.

Forster M. 2017. How Reliable Are Heat Pulse Velocity Methods for Estimating Tree Transpiration? *Forests* 8: 350.

Forzieri G, Miralles DG, Ciais P, Alkama R, Ryu Y, Duveiller G, Zhang K, Robertson E, Kautz M, Martens B, et al. 2020. Increased control of vegetation on global terrestrial energy fluxes. *Nature Climate Change* 10: 356–362.

Gash JHC. 1979. An analytical model of rainfall interception by forests. *Quarterly Journal of the Royal Meteorological Society* 105: 43–55.

Harris I, Osborn TJ, Jones P, Lister D. 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* 7.

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, *et al.* 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146: 1999–2049.

Hyndman RJ, Fan Y. 1996. Sample Quantiles in Statistical Packages. *The American Statistician* 50: 361.

Immerzeel WW, Lutz AF, Andrade M, Bahl A, Biemans H, Bolch T, Hyde S, Brumby S, Davies BJ, Elmore AC, *et al.* 2020. Importance and vulnerability of the world's water towers. *Nature* 577: 364–369.

Jiménez C, Martens B, Miralles DM, Fisher JB, Beck HE, Fernández-Prieto D. 2018. Exploring the merging of the global land evaporation WACMOS-ET products based on local tower measurements. *Hydrology and Earth System Sciences* 22: 4513–4533.

Jung M, Koirala S, Weber U, Ichii K, Gans F, Camps-Valls G, Papale D, Schwalm C, Tramontana G, Reichstein M. 2019. The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data* 6: 74.

Keenan TF, Hollinger DY, Bohrer G, Dragoni D, Munger JW, Schmid HP, Richardson AD. 2013. Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise. *Nature* 499: 324–327.

Kool D, Agam N, Lazarovitch N, Heitman JL, Sauer TJ, Ben-Gal A. 2014. A review of approaches for evapotranspiration partitioning. *Agricultural and Forest Meteorology* 184: 56–70.

Kumagai T, Aoki S, Otsuki K, Utsumi Y. 2009. Impact of stem water storage on diurnal estimates of whole-tree transpiration and canopy conductance from sap flow measurements in Japanese cedar and Japanese cypress trees. *Hydrological Processes* 23: 2335–2344.

Kumarathunge DP, Medlyn BE, Drake JE, Tjoelker MG, Aspinwall MJ, Battaglia M, Cano FJ, Carter KR, Cavaleri MA, Cernusak LA, et al. 2019. Acclimation and adaptation components of the temperature dependence of plant photosynthesis at the global scale. *New Phytologist* 222: 768–784.

Lemordant L, Gentine P, Swann AS, Cook BI, Scheff J. 2018. Critical impact of vegetation physiology on the continental hydrologic cycle in response to increasing CO 2. *Proceedings of the National Academy of Sciences* 115: 4093–4098.

Maes WH, Pagán BR, Martens B, Gentine P, Guanter L, Steppe K, Verhoest NEC, Dorigo W, Li X, Xiao J, *et al.* 2020. Sun-induced fluorescence closely linked to ecosystem transpiration as evidenced by satellite data and radiative transfer models. *Remote Sensing of Environment* 249: 112030.

Mackay DS, Ewers BE, Loranty MM, Kruger EL. 2010. On the representativeness of plot size and location for scaling transpiration from trees to a stand: SCALING CANOPY TRANSPIRATION FROM TREES. Journal of Geophysical Research: *Biogeosciences* 115: n/a-n/a.

Martens B, Miralles DG, Lievens H, van der Schalie R, de Jeu RAM, Fernández-Prieto D, Beck HE, Dorigo WA, Verhoest NEC. 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development* 10: 1903–1925.

Martens B, Waegeman W, Dorigo WA, Verhoest NEC, Miralles DG. 2018. Terrestrial evaporation response to modes of climate variability. *npj Climate and Atmospheric Science* 1: 43.

Michalsky JJ. 1988. The Astronomical Almanac’s algorithm for approximate solar position (1950–2050). *Solar Energy* 40: 227–235.

Michel D, Jiménez C, Miralles DG, Jung M, Hirschi M, Ershadi A, Martens B, McCabe MF, Fisher JB, Mu Q, *et al.* 2016. The WACMOS-ET project – Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms. *Hydrology and Earth System Sciences* 20: 803–822.

Miralles DG, van den Berg MJ, Gash JH, Parinussa RM, de Jeu RAM, Beck HE, Holmes TRH, Jiménez C, Verhoest NEC, Dorigo WA, *et al.* 2014. El Niño–La Niña cycle and recent trends in continental evaporation. *Nature Climate Change* 4: 122–126.

Miralles DG, Holmes TRH, De Jeu RAM, Gash JH, Meesters AGCA, Dolman AJ. 2011. Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences* 15: 453–469.

Miralles DG, Jiménez C, Jung M, Michel D, Ershadi A, McCabe MF, Hirschi M, Martens B, Dolman AJ, Fisher JB, *et al.* 2016. The WACMOS-ET project – Part 2: Evaluation of global terrestrial evaporation data sets. *Hydrology and Earth System Sciences* 20: 823–842.

Nelson JA, Pérez-Priego O, Zhou S, Poyatos R, Zhang Y, Blanken PD, Gimeno TE, Wohlfahrt G, Desai AR, Gioli B, *et al.* 2020. Ecosystem transpiration and evaporation: Insights from three water flux partitioning methods across FLUXNET sites. *Global Change Biology* 26: 6916–6930.

Penman, HL. 1948. 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.

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2020. nlme: Linear and Nonlinear Mixed Effects Models. URL https://CRAN.R-project.org/package=nlme.

Poyatos R, Granda V, Flo V, Adams MA, Adorján B, Aguadé D, Aidar MPM, Allen S, Alvarado-Barrientos MS, Anderson-Teixeira KJ, *et al.* 2021. Global transpiration data from sap flow measurements: the SAPFLUXNET database. *Earth System Science Data* 13: 2607–2649.

Poyatos R, Granda V, Molowny-Horas R, Mencuccini M, Steppe K, Martínez-Vilalta J. 2016. SAPFLUXNET: towards a global database of sap flow measurements (R Oren, Ed.). *Tree Physiology* 36: 1449–1455.

Priestley CHB, Taylor RJ. 1972. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Monthly Weather Review* 100: 81–92.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sabater AM, Ward HC, Hill TC, Gornall JL, Wade TJ, Evans JG, Prieto-Blanco A, Disney M, Phoenix GK, Williams M, *et al.* 2020. Transpiration from subarctic deciduous woodlands: Environmental controls and contribution to ecosystem evapotranspiration. *Ecohydrology* 13.

Smith RJ. 2009. Use and misuse of the reduced major axis for line-fitting. *American Journal of Physical Anthropology* 140: 476–486.

Stephens CM, Lall U, Johnson FM, Marshall LA. 2021. Landscape changes and their hydrologic effects: Interactions and feedbacks across scales. *Earth-Science Reviews* 212: 103466.

Stoy PC, El-Madany TS, Fisher JB, Gentine P, Gerken T, Good SP, Klosterhalfen A, Liu S, Miralles DG, Perez-Priego O, *et al.* 2019. Reviews and syntheses: Turning the challenges of partitioning ecosystem evaporation and transpiration into opportunities. *Biogeosciences* 16: 3747–3775.

Talsma CJ, Good SP, Jimenez C, Martens B, Fisher JB, Miralles DG, McCabe MF, Purdy AJ. 2018. Partitioning of evapotranspiration in remote sensing-based models. *Agricultural and Forest Meteorology* 260–261: 131–143.

Thomas R, Lello J, Medeiros R, Pollard A, Robinson P, Seward A, Smith J, Vafidis J, Vaughan I. 2017. Data analysis with R Statistical Software: a guidebook for scientists. Eco-Explore, United Kindgom.

Wang K, Dickinson RE. 2012. A REVIEW OF GLOBAL TERRESTRIAL EVAPOTRANSPIRATION : OBSERVATION ,.

Wang Z, Zhan C, Ning L, Guo H. 2021. Evaluation of global terrestrial evapotranspiration in CMIP6 models. *Theoretical and Applied Climatology* 143: 521–531.

Warton DI, Duursma RA, Falster DS, Taskinen S. 2012. smatr 3- an R package for estimation and inference about allometric lines: The smatr 3 - an R package. *Methods in Ecology and Evolution* 3: 257–259.

Wei Z, Yoshimura K, Wang L, Miralles DG, Jasechko S, Lee X. 2017. Revisiting the contribution of transpiration to global terrestrial evapotranspiration: Revisiting Global ET Partitioning. *Geophysical Research Letters* 44: 2792–2801.

Whelan ME, Lennartz ST, Gimeno TE, Wehr R, Wohlfahrt G, Wang Y, Kooijmans LMJ, Hilton TW, Belviso S, Peylin P, *et al.* 2018. Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles. *Biogeosciences* 15: 3625–3657.

Williams DG, Cable W, Hultine K, Hoedjes JCB, Yepez EA, Simonneaux V, Er-Raki S, Boulet G, de Bruin HAR, Chehbouni A, *et al.* 2004. Evapotranspiration components determined by stable isotope, sap flow and eddy covariance techniques. *Agricultural and Forest Meteorology* 125: 241–258.

Zuur AF, Ieno EN, Walker N, Saveliev AA, Smith GM. 2009. *Mixed effects models and extensions in ecology with R*. New York, NY: Springer New York.

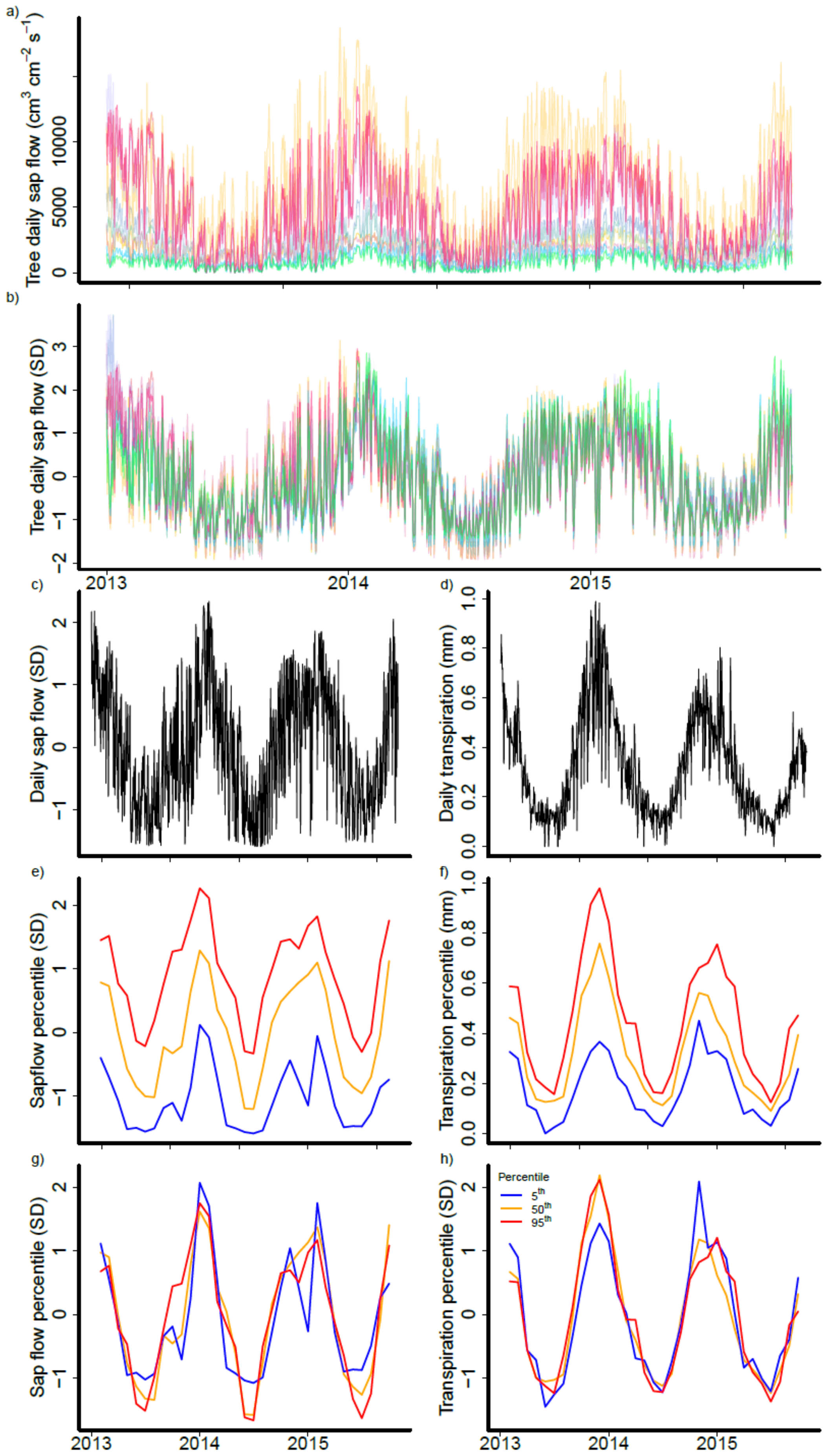
Table 1. Variable loadings and percentage contributions to the first and second axis of the principal component analysis (PC1 and PC2) of the climatic and model variables studied. Variables with high loading/contributions for each axis are highlighted in bold.

	PC1		PC2	
	Loading	Contribution	Loading	Contribution
VPD	**-0.49**	**24.0**	-0.11	1.1
Temp.	**-0.44**	**19.7**	0.26	7.0
$S_{\downarrow}$	**-0.50**	**24.6**	0.02	0.1
Prec.	0.20	3.9	**0.48**	**23.5**
ETp	-0.39	15.0	0.40	16.0
ET	-0.02	0.1	**0.60**	**35.8**
S	0.36	12.8	0.41	16.6

VPD – mean vapour pressure deficit; $S_{\downarrow}$– total monthly incoming net surface solar radiation (MJ m^{-2}); Temp – mean surface temperature; Prec. – mean precipitation; ET and ETp – GLEAM mean actual ET and potential ET; S – mean GLEAM evaporative stress factor (S equal to one equates to no stress). Site climatic data from ERA5 and CRUJRA products for the period 2001–2020.

Table 2. Results of the linear models of the first and second principal component analysis axes (PC1 and PC2) of the climatic and model variables studied as predictors of mismatches between GLEAM-T and SAPFLUXNET-SF: root mean squared difference (RMSD), bivariate correlation (r), VPD sensitivity mismatch (VPD_{sm}) and incoming solar radiation mismatch ($S_{\downarrow sm}$). The mismatch indices were scaled prior to analysis, thus the magnitude of their slopes is directly comparable. Blank cells for PC1 or PC2 indicates that predictor is not significant. Values in the PC1 and PC2 columns give the slope of the relationships, r^2 is percent of explained variance and p is probability value.

Index	Percentile	PC1	PC2	r^2	p
	P5		0.18	0.09	0.009
RMSD	P50		0.25	0.16	<0.001
	P95		0.24	0.15	<0.001
	P5	0.14		0.07	0.02
r	P50	0.24		0.21	<0.001
	P95	0.18	0.18	0.21	<0.001
	P5	0.23	0.21	0.3	<0.001
VPD_{sm}	P50		0.17	0.07	0.03
	P95	0.23	0.20	0.29	<0.001
	P5	0.12	0.27	0.24	<0.001
$S_{\downarrow sm}$	P50		0.26	0.17	<0.001
	P95				0.14



a)
Tree daily sap flow (cm^3 cm^{-2} s^{-1})
0
5000
10000
b)
Tree daily sap flow (SD)
2013
2014
2015
c)
Daily sap flow (SD)
d)
Daily transpiration (mm)
e)
Sapflow percentile (SD)
f)
Transpiration percentile (mm)
g)
Sap flow percentile (SD)
h)
Transpiration percentile (SD)
Percentile
5th
50th
95th

Figure 1. Example of processing of individual tree sap flow (SAPFLUXNET) and transpiration (GLEAM) to yield standardised ecosystem sap flow and standardised transpiration. For SAPFLUXNET site AUS_WOM (37.42^{o} S, 144.09^{o} E; Melbourne, Australia). a) Daily SF for eleven trees (each colour representing one tree) at the site; b) Standardized (Z-score) SF for the eleven trees. c) Site-level daily SF, calculated as the average of the standardized SF for the eleven trees; d) GLEAM daily tall vegetation T for the grid cell closest to site AUS_WOM; e-f) Monthly percentiles (5^{th}, 50^{th} and 95^{th}; blue, orange and red, respectively) of SF (e) and T (f), hereafter designated as SAPFLUXNET-SF and GLEAM-T, calculated from the monthly distribution of daily values in c) and d). The percentiles represent, in each month, conditions of days with low, median and high SF and T. g-h) Standardized (Z-scores) monthly SF and T percentiles (i.e. in number of standard deviations, SD).

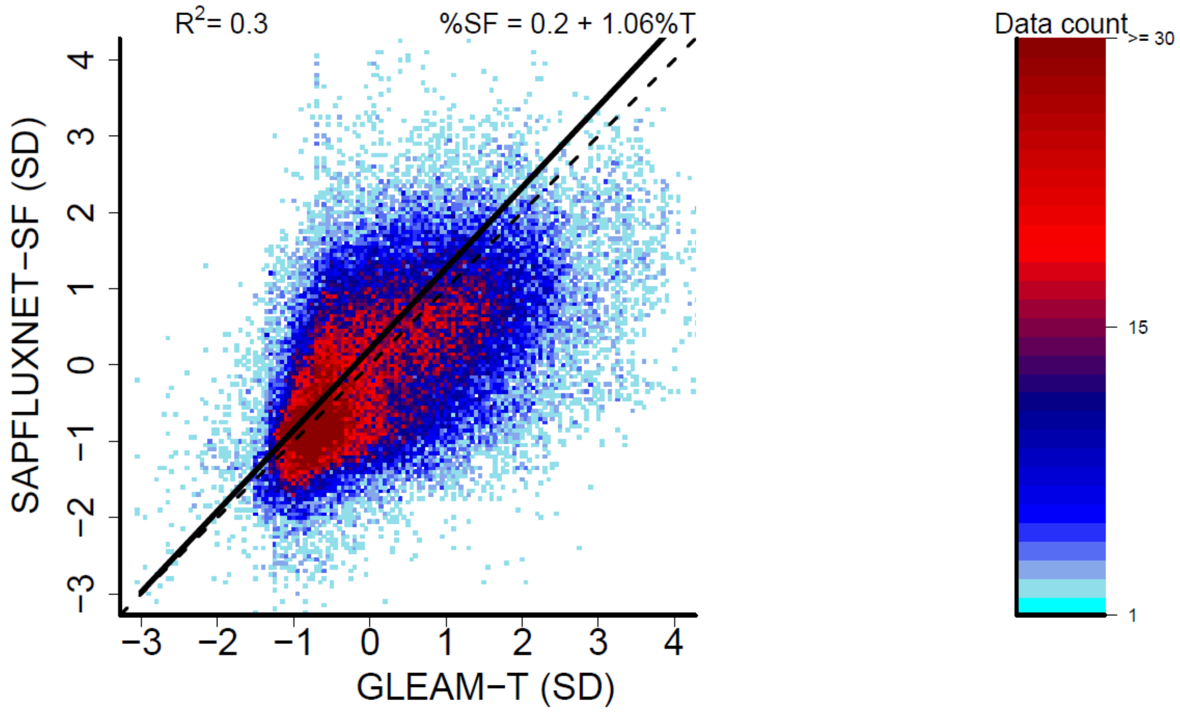


Figure 2. SAPFLUXNET-SF as a function of GLEAM-T variability for all daily points combined. Values are Z-scores for daily mean values of sap flow and transpiration; data point colour indicates the count of data point in each 0.05 bin. R^2 is the coefficient of determination of the standardized major axis regression model. The black line is the model fit and the dashed line marks the 1:1 relationship. The scaling slope of the relationship is 1.06 ± 0.007 (mean ± 95% confidence interval).

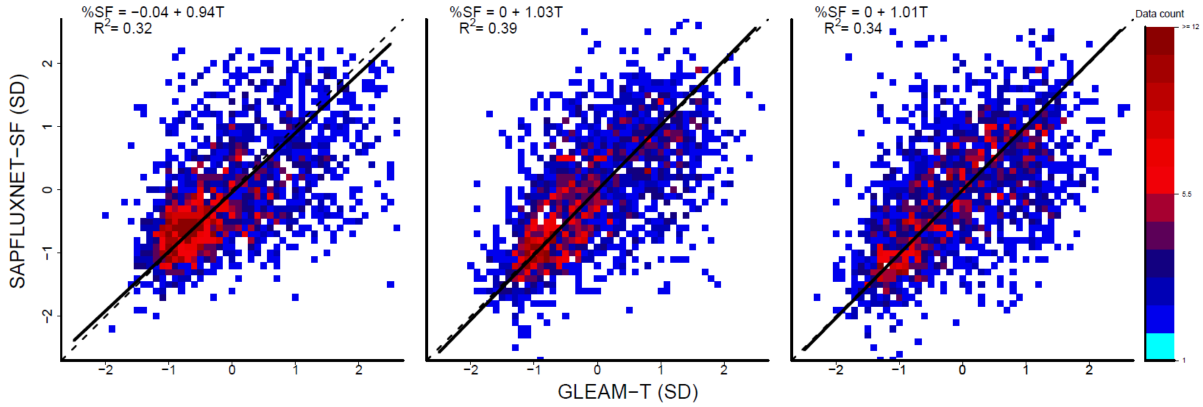


Figure 3. SAPFLUXNET-SF as a function of GLEAM-T. Graphs a, b and c are, respectively, low, median and high transpiration daily values within a month and site (i.e., the 5th, 50th and 95th monthly percentiles of daily values). Data point colour indicates the count of data point in each 0.1 bin. R^2 is the coefficient of determination of the standardized major axis regression model with sap flow scaling with transpiration and percentile as a covariate affecting the slope of the scaling. The black line is the model fit and the dashed line marks the 1:1 relationship.

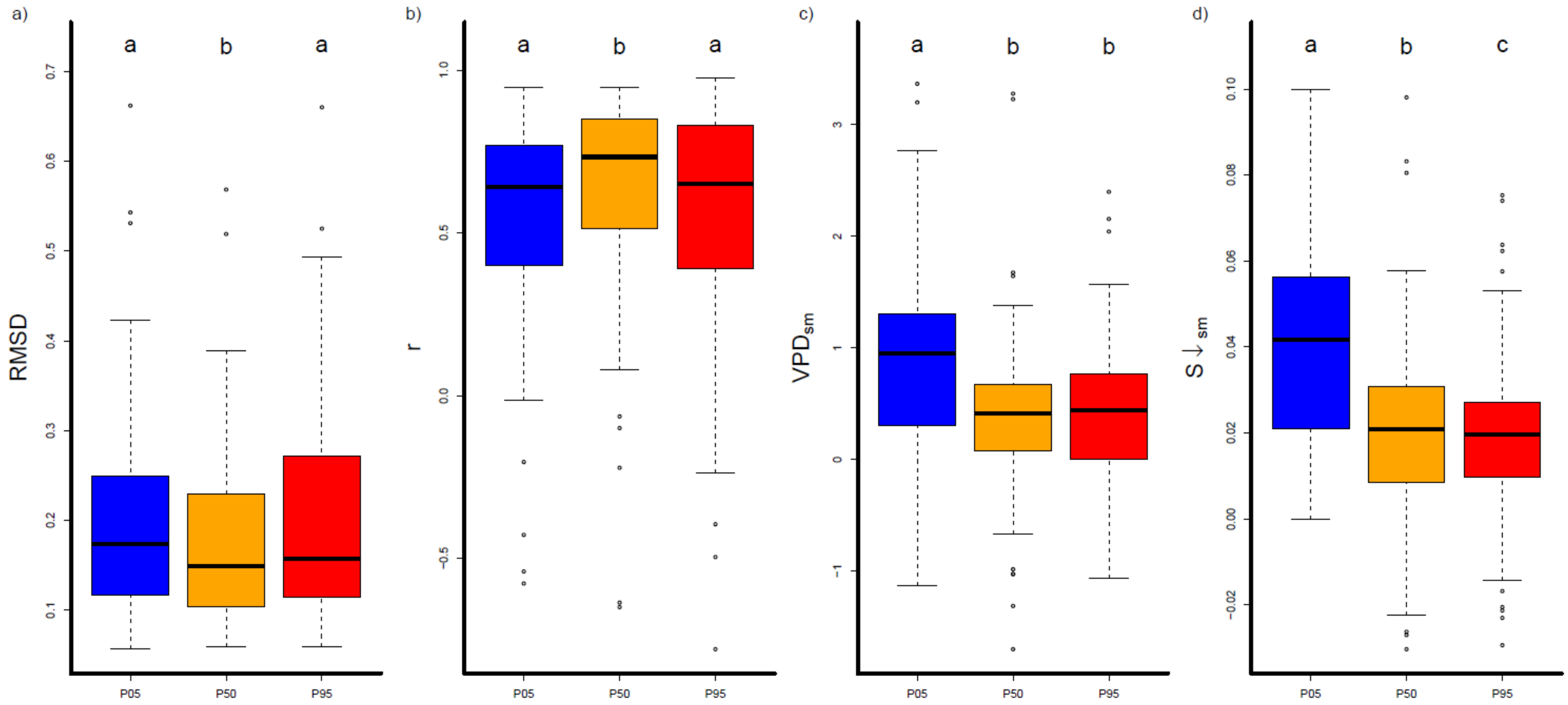
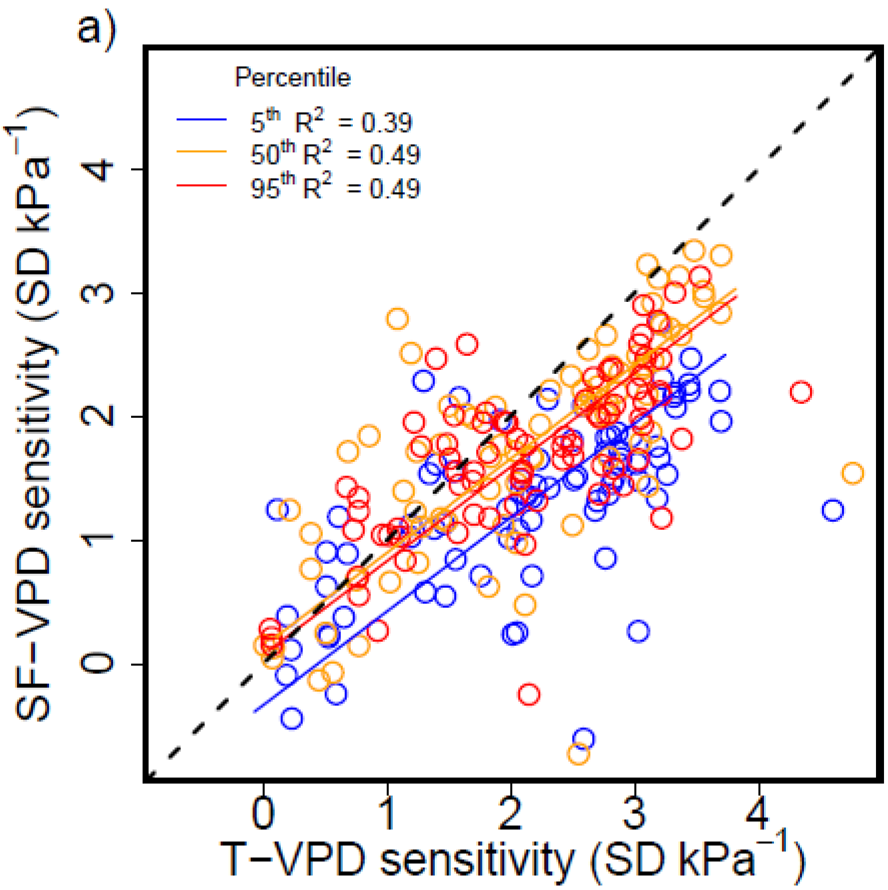
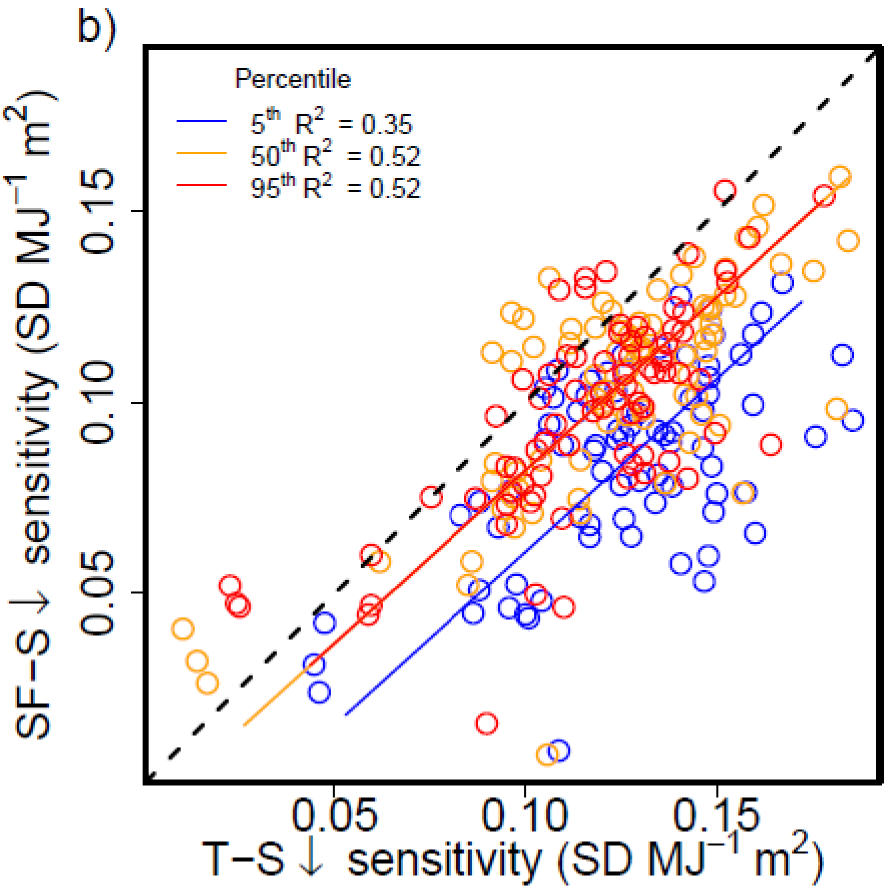


Figure 4. Site level mismatching indices between GLEAM-T and SAPFLUXNET-SF for the 5^{th}, 50^{th} and 95^{th} monthly percentiles (P5, P50 and P95; blue, orange and red, respectively): a) mean root squared difference (RMSD), b) bivariate correlation (r), c) VPD sensitivity mismatch (VPD_{sm}) and (d) and incoming solar radiation sensitivity mismatch ($S_{\downarrow sm}$). Groups with different letters in are significantly different from each other at least at $p < 0.05$ in a mixed model with site as random effect and percentile as fixed effect.


a)
Percentile
5th R2 = 0.39
50th R2 = 0.49
95th R2 = 0.49
SF−VPD sensitivity (SD kPa−1)
T−VPD sensitivity (SD kPa−1)
0
1
2
3
4


b)
Percentile
5th R2 = 0.35
50th R2 = 0.52
95th R2 = 0.52
SF−S↓ sensitivity (SD MJ−1 m2)
T−S↓ sensitivity (SD MJ−1 m2)
0.05
0.10
0.15

577

Figure 5. Relationships between GLEAM-T and SAPFLUXNET-SF sensitivities to vapour pressure deficit (VPD; a) and surface solar radiation ($S_{\downarrow}$; b). Blue, orange and red points indicate, respectively, daily conditions, within months, with low, median and high T (or SF) (i.e., 5th, 50th and 95th monthly percentiles of daily values, P5, P50 and P95, respectively). Each point is a different site. Sensitivity is the slope of the relationship between GLEAM-T (or SAPFLUXNET-SF) and site VPD (or $S_{\downarrow}$) (i.e., a value of 1 indicates T increases by one standard deviation per 1 kPa increase in VPD). Coloured lines are the standardized major axis fits for each percentile and the black dashed line is the 1:1 line.

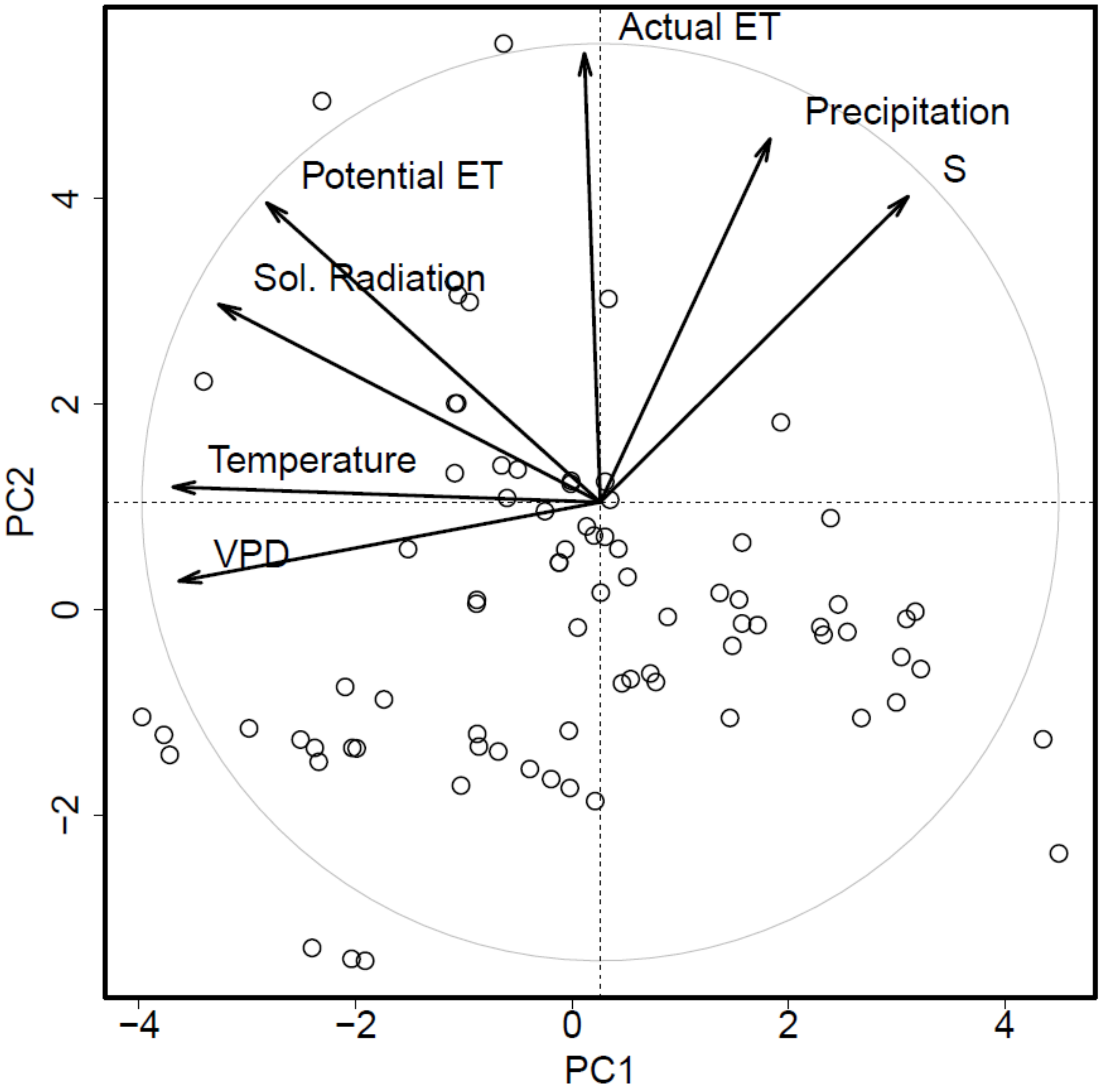


Figure 6. Principal component analysis of site climatic (vapour pressure deficit – VPD, incoming solar radiation, air temperature and precipitation) and model variables (potential and actual ET, and their ratio, i.e. S). The loadings of each variable into the PC1 and PC2 axis, as well as their contribution, are presented in Table 1. The grey circle is the correlation circle marking the correlation between variables and principal components.

JAMES Journal of Advances in Modeling Earth Systems

Supporting Information for

Bridging scales: a temporal approach to evaluate global transpiration products using tree-scale sap flow data

Paulo Bittencourt1, Lucy Rowland1, Stephen Sitch1, Rafael Poyatos2, Diego G. Miralles3, Maurizio Mencuccini2,4

1 College of Life and Environmental Sciences, University of Exeter, Exeter, EX4 4RJ, UK

2 CREAF, Bellaterra (Cerdanyola del Vallès), Catalonia, E08193, Spain

3 Hydro-Climate Extremes Lab (H-CEL), Ghent University, Coupure links 653, 9000 Ghent, Belgium

4 ICREA, Pg. Lluís Companys 23, Barcelona, 08010, Spain

Contents of this file

Tables S1

Introduction

This supplementary material presents a list of the SAPFLUXNET sites whose sap flow data was used in this work, including data availability dates and site meteorological summaries.

Table S1. Summary of the SAPFLUXNET sites used in this study (site codes given here correspond to those in SAPFLUXNET), temporal range of available data, total number of months of available data (n) and site summaries. P – mean precipitation (mm month^{-1}); Temp – mean surface temperature (°C); VPD – mean vapour pressure deficit (kPa); $S_{\downarrow}$– mean monthly incoming surface solar radiation (MJ m^{-2}); T, ET_p and ET – GLEAM mean tall vegetation T, potential ET and actual ET, respectively (mm month^{-1}); T/ET – mean tall vegetation T to total ET fraction; S – mean GLEAM evaporative stress factor (i.e. ET/ET_p). Site climatic data from ERA5 and CRUJRA for the period 2001–2020.

	Range			n	P	Temp.	VPD	$S_{\downarrow}$	T	ETp	ET	T/ET	S	Lat.	Long.
AUS_ELL_UNB	08/2010	to	02/2012	28	89	12.8	0.80	15.1	55.7	80.9	72.6	0.91	0.77	146.6	-36.8
AUS_MAR_UBD	02/2011	to	02/2012	24	92	12.9	0.72	13.9	58.5	82.5	79.7	0.97	0.73	145.6	-37.7
AUS_MAR_UBW	01/2011	to	02/2012	21	85	13.1	0.72	13.7	58.6	79.3	75.9	0.96	0.77	145.6	-37.9
AUS_RIC_EUC_ELE	08/2013	to	08/2014	24	74	17.1	0.73	14.2	57.9	79.5	67.0	0.86	0.86	150.7	-33.6
AUS_WOM	05/2014	to	10/2015	34	53	11.9	0.79	14.1	47.3	65.0	53.6	0.85	0.88	144.1	-37.4
AUT_PAT_FOR	07/2007	to	10/2007	6	143	4.1	0.20	9.4	28.8	49.9	48.9	0.97	0.59	11.5	47.3
AUT_TSC	06/2012	to	10/2012	8	135	2.5	0.18	9.3	28.1	50.2	49.5	0.98	0.57	10.8	47.2
BRA_CAM	07/2011	to	11/2011	9	152	16.6	0.40	15.1	69.9	85.8	81.6	0.95	0.86	-45.5	-22.7
BRA_CAX_CON	11/2015	to	11/2016	24	189	26.6	0.63	15.2	39.2	132.8	131.4	0.99	0.30	-51.4	-1.7
BRA_SAN	01/2009	to	09/2009	15	157	18.5	0.38	14.2	72.4	102.2	101.2	0.99	0.72	-45.2	-23.3
CHE_DAV_SEE	06/2010	to	12/2010	12	122	-0.6	0.19	9.6	29.6	41.8	40.5	0.97	0.73	9.9	46.8
CHE_LOT_NOR	01/2014	to	11/2015	34	147	-1.3	0.24	8.1	20.8	44.4	39.8	0.94	0.52	7.8	46.5
CHN_YUN_YUN	04/2011	to	10/2011	13	132	20.5	0.65	13.3	52.1	88.8	81.7	0.92	0.64	117.4	24.0
CRI_TAM_TOW	10/2015	to	07/2016	19	204	22.6	0.62	13.7	78.6	100.0	98.9	0.99	0.80	-84.6	10.4
CZE_LIZ_LES	08/2008	to	10/2009	28	91	6.6	0.35	10.1	32.2	49.9	48.6	0.97	0.66	13.7	49.1
CZE_STI	07/2016	to	10/2016	7	71	8.7	0.36	9.7	33.7	48.1	46.3	0.96	0.73	18.0	49.0
DEU_HIN_OAK	07/2013	to	09/2014	15	57	9.4	0.34	8.9	25.3	42.4	38.1	0.92	0.66	13.2	53.4
DEU_HIN_TER	07/2013	to	09/2014	15	57	9.4	0.34	8.9	25.3	42.4	38.1	0.92	0.66	13.2	53.4
ESP_ALT_ARM	05/2012	to	10/2014	54	50	11.6	0.66	14.6	28.5	59.0	37.1	0.72	0.77	-2.3	40.8
ESP_ALT_HUE	06/2011	to	04/2013	41	49	11.4	0.66	14.5	28.5	59.0	37.1	0.72	0.77	-2.3	40.8
ESP_ALT_TRI	05/2012	to	10/2014	58	49	11.3	0.66	14.5	28.3	58.1	38.1	0.74	0.74	-2.2	40.8
ESP_CAN	12/2011	to	12/2012	23	54	14.6	0.69	13.6	15.1	78.4	63.1	0.86	0.24	2.1	41.5

ESP_GUA_VAL	12/2012	to	10/2013	14	62	9.7	0.65	13.9	24.8	57.7	35.1	0.71	0.71	-4.0	40.9
ESP_MAJ_MAI	09/2016	to	05/2018	39	71	16.1	0.86	14.2	33.0	64.4	38.0	0.70	0.87	-5.8	39.9
ESP_MAJ_NOR_LM1	09/2016	to	05/2018	39	71	16.1	0.86	14.2	33.0	64.4	38.0	0.70	0.87	-5.8	39.9
ESP_MON_SIE_NAT	09/2011	to	07/2013	35	58	9.7	0.65	14.0	27.6	57.9	37.6	0.74	0.73	-3.5	41.1
ESP_RON_PIL	06/2012	to	11/2013	25	59	14.6	0.71	15.6	33.4	77.4	40.2	0.63	0.83	-5.0	36.8
ESP_TIL_MIX	01/2012	to	10/2013	42	47	13.7	0.74	14.0	28.3	65.8	37.2	0.66	0.76	1.0	41.4
ESP_TIL_OAK	01/2011	to	10/2011	17	47	13.7	0.74	14.0	28.3	65.8	37.2	0.66	0.76	1.0	41.4
ESP_TIL_PIN	03/2011	to	11/2011	15	47	13.6	0.74	14.0	28.3	65.8	37.2	0.66	0.76	1.0	41.4
ESP_VAL_BAR	06/2004	to	08/2005	17	73	8.2	0.46	13.8	43.8	65.2	56.9	0.89	0.77	1.8	42.2
ESP_VAL_SOR	07/2004	to	08/2005	27	72	8.1	0.46	13.6	43.8	65.2	56.9	0.89	0.77	1.8	42.3
ESP_YUN_C1	07/2012	to	10/2014	37	55	15.0	0.78	15.5	32.3	79.5	38.9	0.61	0.83	-5.0	36.8
ESP_YUN_C2	05/2013	to	10/2014	26	55	15.0	0.78	15.5	32.3	79.5	38.9	0.61	0.83	-5.0	36.8
ESP_YUN_T1_THI	09/2012	to	11/2014	36	59	14.5	0.78	15.9	32.3	79.5	38.9	0.61	0.83	-5.0	36.7
ESP_YUN_T3_THI	04/2012	to	11/2014	55	55	15.0	0.78	15.5	32.3	79.5	38.9	0.61	0.83	-5.0	36.8
FIN_HYY_SME	02/2015	to	11/2016	32	60	4.7	0.23	7.4	22.1	37.3	36.9	0.99	0.60	24.3	61.8
FRA_FON	08/2010	to	12/2014	77	60	11.7	0.45	9.9	30.6	47.8	41.4	0.88	0.74	2.8	48.5
FRA_HES_HE2_NON	01/2003	to	11/2005	24	86	10.1	0.40	9.9	35.2	48.6	46.6	0.96	0.76	7.1	48.7
FRA_PUE	12/2007	to	12/2015	156	91	13.3	0.67	13.2	37.8	64.7	47.5	0.80	0.80	3.6	43.7
GBR_GUI_ST1	07/2003	to	10/2003	6	124	5.8	0.16	6.6	22.1	37.5	37.2	0.99	0.59	-4.8	57.4
GUF_GUY_GUY	03/2015	to	05/2016	29	207	25.8	0.53	15.6	0.0	139.2	132.1	0.96	0.00	-52.9	5.3
GUF_GUY_ST2	10/2008	to	01/2009	7	208	25.8	0.53	15.6	0.0	139.2	132.1	0.96	0.00	-52.9	5.3
HUN_SIK	02/2015	to	11/2015	12	64	9.5	0.44	10.4	31.9	53.6	43.9	0.84	0.73	20.4	48.0
IDN_PON_STE	06/2008	to	12/2008	12	257	19.4	0.61	14.5	87.5	128.0	127.9	1.00	0.68	120.1	-1.5
ITA_KAE_S20	11/2013	to	11/2014	17	92	-0.9	0.15	8.0	24.9	42.6	39.0	0.93	0.64	10.6	46.8
ITA_MAT_S21	05/2013	to	11/2014	19	94	-0.5	0.15	7.2	24.9	42.6	39.0	0.93	0.64	10.7	46.7
ITA_REN	07/2016	to	10/2016	7	90	5.5	0.18	10.9	32.0	46.1	44.1	0.95	0.73	11.4	46.6
ITA_RUN_N20	07/2013	to	11/2014	25	93	1.5	0.15	9.1	24.9	42.6	39.0	0.93	0.64	10.6	46.7
ITA_TOR	01/2016	to	12/2016	22	109	3.5	0.29	9.2	15.6	45.9	40.1	0.92	0.39	7.6	45.8
MEX_VER_BSJ	08/2015	to	02/2016	6	135	16.4	0.52	15.9	81.5	105.2	100.5	0.96	0.81	-97.0	19.5
MEX_VER_BSM	08/2015	to	02/2016	6	138	15.8	0.52	15.5	81.5	105.2	100.5	0.96	0.81	-97.0	19.6
NZL_HUA_HUA	07/2013	to	09/2015	51	97	15.5	0.25	13.9	0.0	84.9	84.6	1.00	0.00	174.5	-36.8

PRT_LEZ_ARN	12/2007	to	09/2008	18	43	16.8	0.70	14.6	27.9	78.1	45.3	0.70	0.62	-8.8	38.8
PRT_MIT	06/2002	to	12/2003	12	48	16.5	0.72	14.1	35.5	71.7	39.6	0.67	0.90	-8.0	38.5
RUS_CHE_Y4	08/2014	to	08/2015	7	25	-9.6	0.14	6.5	8.1	21.2	19.1	0.96	0.42	161.4	68.7
RUS_FYO	11/2001	to	09/2004	13	68	5.4	0.24	8.0	27.2	43.5	42.2	0.97	0.64	32.9	56.6
SWE_NOR_ST1_AF1	01/2010	to	10/2010	12	54	6.7	0.22	8.1	21.6	39.7	37.1	0.96	0.58	17.5	60.1
SWE_NOR_ST1_AF2	03/2010	to	10/2010	11	54	6.7	0.22	8.1	21.6	39.7	37.1	0.96	0.58	17.5	60.1
SWE_NOR_ST1_BEF	02/2008	to	09/2008	8	54	6.7	0.22	8.1	21.6	39.7	37.1	0.96	0.58	17.5	60.1
SWE_NOR_ST3	12/2004	to	10/2007	19	54	6.7	0.22	8.1	21.6	39.7	37.1	0.96	0.58	17.5	60.1
SWE_SVA_MIX_NON	12/2016	to	06/2017	12	57	2.8	0.18	7.2	19.8	32.9	31.8	0.98	0.62	19.8	64.4
USA_BNZ_BLA	08/2014	to	09/2016	32	34	-1.7	0.27	7.9	15.0	29.4	25.6	0.94	0.58	-148.3	64.8
USA_DUK_HAR	12/2003	to	12/2005	29	91	14.9	0.62	13.1	56.8	74.7	69.8	0.94	0.81	-79.1	37.0
USA_HIL_HF1_POS	10/2013	to	09/2016	62	95	15.4	0.59	13.1	57.5	79.7	74.2	0.93	0.78	-78.9	36.3
USA_HIL_HF1_PRE	10/2010	to	01/2011	7	95	15.4	0.59	13.1	57.5	79.7	74.2	0.93	0.78	-78.9	36.3
USA_HIL_HF2	09/2013	to	12/2016	74	95	15.4	0.59	13.1	57.5	79.7	74.2	0.93	0.78	-78.9	36.3
USA_MOR_SF	07/2012	to	08/2013	6	98	12.2	0.47	12.3	45.5	62.6	58.2	0.93	0.78	-86.4	39.4
USA_PJS_P04_AMB	10/2012	to	12/2015	73	22	13.7	0.85	16.9	6.9	63.9	20.8	0.38	0.33	-106.5	34.4
USA_PJS_P08_AMB	10/2012	to	12/2015	73	22	13.7	0.85	16.9	6.9	63.9	20.8	0.38	0.33	-106.5	34.4
USA_PJS_P12_AMB	06/2010	to	12/2013	73	22	13.7	0.85	16.9	6.9	63.9	20.8	0.38	0.33	-106.5	34.4
USA_SIL_OAK_1PR	02/2010	to	11/2010	16	99	12.7	0.59	12.3	48.1	66.7	63.6	0.96	0.76	-74.6	40.0
USA_SIL_OAK_2PR	05/2007	to	11/2008	30	99	12.7	0.59	12.3	48.1	66.7	63.6	0.96	0.76	-74.6	40.0
USA_SIL_OAK_POS	07/2012	to	12/2013	34	99	12.7	0.59	12.3	48.1	66.7	63.6	0.96	0.76	-74.6	40.0
USA_SMI_SCB	03/2014	to	12/2014	15	88	12.2	0.55	12.4	49.8	67.5	63.4	0.94	0.79	-78.1	38.9
USA_SMI_SER	03/2015	to	12/2015	16	101	13.8	0.66	12.7	42.3	70.6	64.9	0.94	0.65	-76.6	38.9
USA_SYL_HL1	03/2003	to	12/2003	9	74	4.9	0.24	10.8	28.1	55.9	53.1	0.97	0.53	-89.3	46.3
USA_SYL_HL2	02/2016	to	12/2016	19	74	4.9	0.24	10.8	28.1	55.9	53.1	0.97	0.53	-89.3	46.3
USA_UMB_CON	10/2013	to	09/2016	30	75	6.6	0.24	11.7	24.5	57.2	55.4	0.98	0.44	-84.7	45.6
USA_UMB_GIR	11/2013	to	09/2016	29	75	6.6	0.24	11.7	24.5	57.2	55.4	0.98	0.44	-84.7	45.6
USA_WIL_WC2	03/2016	to	11/2016	17	73	5.2	0.27	11.0	33.4	54.7	52.7	0.98	0.63	-90.1	45.8
ZAF_FRA_FRA	08/2015	to	02/2016	13	66	15.3	0.68	16.6	37.7	81.7	48.5	0.68	0.78	19.1	-33.8
ZAF_WEL_SOR	05/2014	to	10/2014	11	51	17.4	0.88	16.1	29.7	78.4	37.2	0.58	0.80	19.0	-33.4