JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2022JG007308

Key Points:

- Transpiration products are vital for understanding land-atmosphere processes, but their validation is limited by lack of suitable data sets
- 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

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

P. Bittencourt,
paulo09d@gmail.com

Citation:

Bittencourt, P., Rowland, L., Sitch, S., Poyatos, R., Miralles, D. G., & Mencuccini, M. (2023). Bridging scales: An approach to evaluate the temporal patterns of global transpiration products using tree-scale sap flow data. *Journal of Geophysical Research: Biogeosciences*, *128*, e2022JG007308. https://doi.org/10.1029/2022JG007308

Received 23 NOV 2022
Accepted 9 FEB 2023

Author Contributions:

Conceptualization: Paulo Bittencourt, Lucy Rowland, Maurizio Mencuccini
Data curation: Paulo Bittencourt, Rafael Poyatos
Formal analysis: Paulo Bittencourt, Diego G. Miralles, Maurizio Mencuccini
Funding acquisition: Lucy Rowland, Maurizio Mencuccini
Investigation: Paulo Bittencourt, Lucy Rowland, Stephen Sitch, Rafael Poyatos, Diego G. Miralles, Maurizio Mencuccini

Bridging Scales: An Approach to Evaluate the Temporal Patterns of 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], and Maurizio Mencuccini[2,4]

[1]College of Life and Environmental Sciences, University of Exeter, Exeter, UK, [2]CREAF, Catalonia, Spain, [3]Hydro-Climate Extremes Lab (H-CEL), Ghent University, Ghent, Belgium, [4]ICREA, Barcelona, Spain

Abstract Transpiration is a key process driving energy, water and thus carbon dynamics. Global transpiration products are fundamental for understanding and predicting vegetation processes. However, validation of these transpiration products is limited, mainly due to lack of suitable data sets. We propose a method to use SAPFLUXNET, the first quality-controlled global tree sap flow (SF) 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 SF 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 data set 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 (SF) data from SAPFLUXNET, the first global SF 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.

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) (Stoy et al., 2019; Wei et al., 2017). T is regulated by a complex combination of energy availability and soil moisture and atmospheric demand (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 (Keenan et al., 2013; Lemordant et al., 2018). Additionally, ongoing global changes are causing plants to acclimate and communities

Methodology: Paulo Bittencourt, Lucy Rowland, Stephen Sitch, Rafael Poyatos, Diego G. Miralles, Maurizio Mencuccini
Project Administration: Paulo Bittencourt, Lucy Rowland, Maurizio Mencuccini
Resources: Paulo Bittencourt, Lucy Rowland, Rafael Poyatos, Diego G. Miralles, Maurizio Mencuccini
Software: Paulo Bittencourt, Rafael Poyatos, Diego G. Miralles
Supervision: Lucy Rowland, Maurizio Mencuccini
Validation: Paulo Bittencourt, Rafael Poyatos
Visualization: Paulo Bittencourt, Maurizio Mencuccini
Writing – original draft: Paulo Bittencourt, Lucy Rowland, Stephen Sitch, Rafael Poyatos, Diego G. Miralles, Maurizio Mencuccini
Writing – review & editing: Paulo Bittencourt, Lucy Rowland, Stephen Sitch, Rafael Poyatos, Diego G. Miralles, Maurizio Mencuccini

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, for example, quantification of water resources (Immerzeel et al., 2020), driving basin hydrological models (Dembélé et al., 2020), studying global climate (Martens et al., 2018; Miralles et al., 2014) and benchmarking climate models, such as those from CMIP6 (Z. Z. 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 modeling 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 (Feng et al., 2020; Miralles et al., 2016; Talsma et al., 2018). Model improvement is limited by a lack of suitable data sets 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 data sets at the necessary spatial and temporal scales (Bayat et al., 2021; Kool et al., 2014; Talsma et al., 2018).

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 modeling 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 (Poyatos et al., 2021; Stoy et al., 2019; K. K. Wang & Dickinson, 2012). At daily or longer time scales, average SF can be equated to T with minimal errors (Kool et al., 2014; Kumagai et al., 2009). 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 data set at a global scale. SAPFLUXNET opens new opportunities to validate T products directly (Bright et al., 2022). However, new generalized 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 (Flo et al., 2019; Forster, 2017). 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 kilometers) to global data sets spatial scales (10–50 km), requires further consideration of landscape heterogeneity, which increases uncertainty (Ford et al., 2007; Mackay et al., 2010). Consequently, the use of SF 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 data set 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 focusing on temporal mismatches rather than absolute values. We use SF data from >80 sites across the globe and analyze 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 analyzing 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 behavior of ground SF data and modeled T data. We assess, for days with low, median, and high transpiration values, (a) how GLEAM and SAPFLUXNET compare over time, (b) whether GLEAM and SAPFLUXNET sensitivity to vapor pressure deficit and radiation match, and (c) 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 RS 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 Data Sets

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 (a) had non-native vegetation, (b) were affected by experimental manipulations or recent fire, or (c) 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 1,195 (Table S1 in Supporting Information S1).

We use the outputs from the GLEAM model (Martens et al., 2017; Miralles et al., 2011). GLEAM uses remote sensing data to calculate potential ET based on the Priestley and 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° (~25 km at equator) grid cell (water, soil, short, and tall vegetation); these fractions are derived based on the Moderate Resolution Imaging Spectroradiometer product MOD44B (DiMiceli 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 vapor 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; 0.36° resolution) at the monthly time scale. We calculate VPD from the CRUJRA monthly data set of air vapor pressure and air temperature (Harris et al., 2020; 0.5° resolution) 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; Figures 1a and 1b). 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

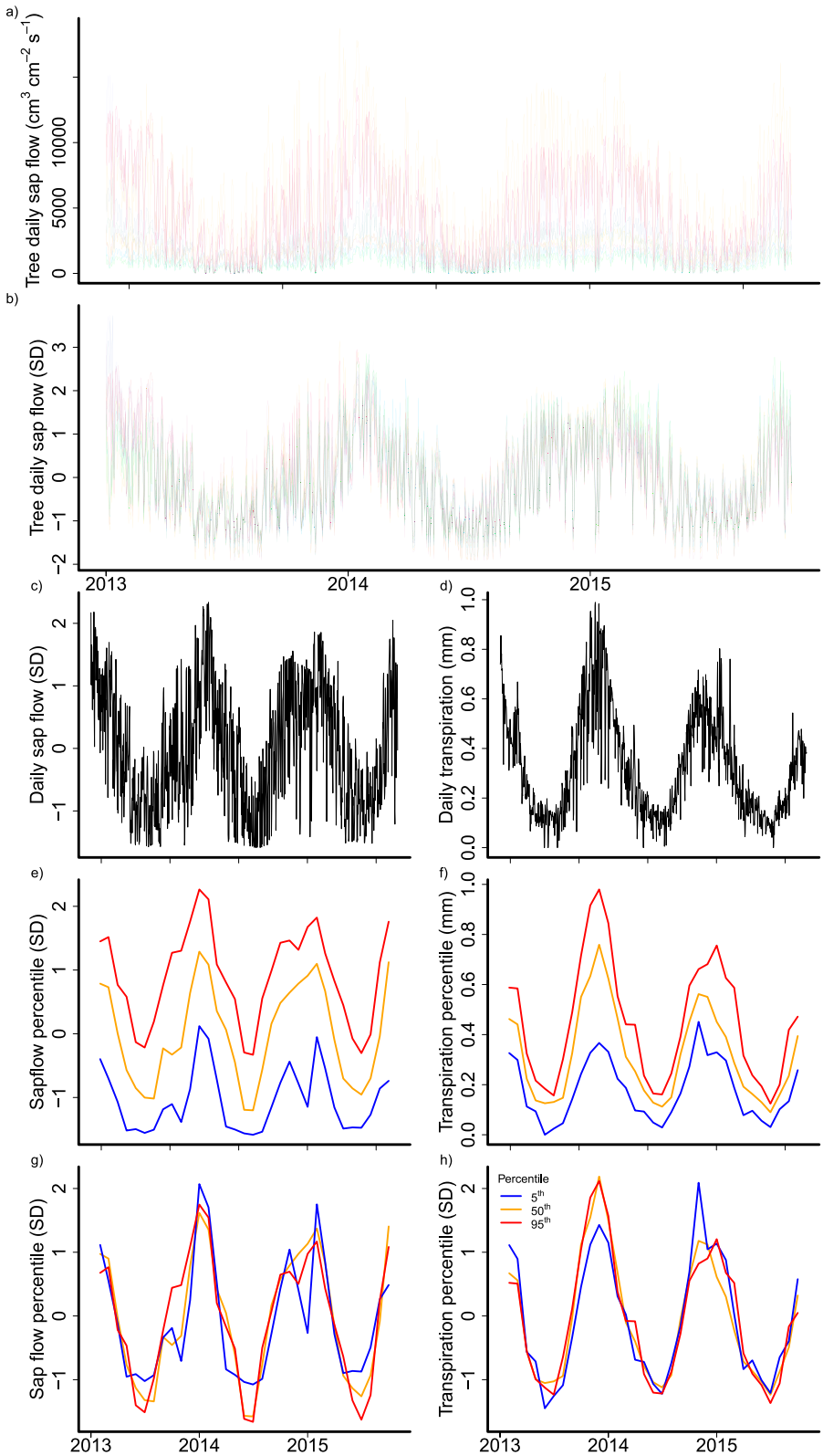

Figure 1.

site (Figure 1c). We performed the same 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 (Figures 1c–1f). 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 (Figures 1e–1h). 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. A flowchart with the data processing steps above is presented in Figure S1 of the Supporting Information S1.

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: (a) the root mean squared difference (RMSD) of T in relation to SF (Figure S2c in Supporting Information S1) and (b) the bivariate correlation between T and SF (r—the Pearson's correlation). Both indices were calculated for each of the time series (i.e., low, median and high T and SF percentiles).

2.6. Sensitivity to Vapor 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 1 and 2, 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):

1. T or SF = a + b*VPD + (1 + VPD|site)
2. 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 centered 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 and whether the scaling is different among days with low, median and high transpiration (i.e., whether the scaling relationship changes with the percentile analyzed) using standardized major axis regression (SMA; Smith, 2009). We focus particularly on whether the scaling relationship is consistent with a 1:1 relationship (slope of 1) as deviations from this

Figure 1. Example of processing of individual tree sap flow (SF) (SAPFLUXNET) and transpiration (Global Land Evaporation Amsterdam Model (GLEAM)) to yield standardized ecosystem SF and standardized transpiration. For SAPFLUXNET site AUS_WOM (37.42°S, 144.09°E; Melbourne, Australia). (a) Daily SF for 11 trees (each color representing one tree) at the site; (b) Standardized (Z-score) SF for the 11 trees. (c) Site-level daily SF, calculated as the average of the standardized SF for the 11 trees; (d) GLEAM daily tall vegetation T for the grid cell containing the site AUS_WOM; (e and f) Monthly percentiles (5th, 50th and 95th; 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 and h) Standardized (Z-scores) monthly SF and T percentiles (i.e., in number of standard deviations, SD). A flowchart with the steps used in the data processing are presented on Figure S1 in Supporting Information S1.

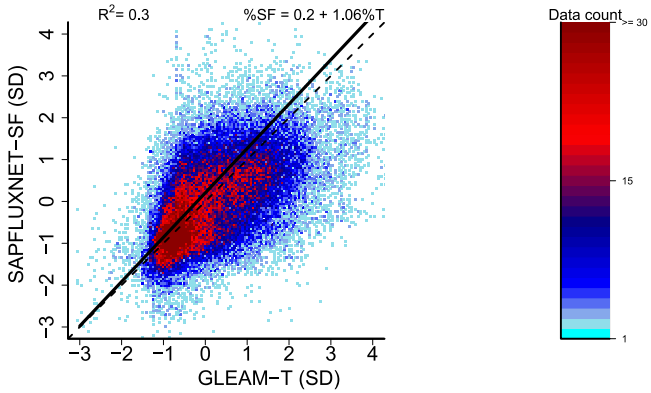

Figure 2. SAPFLUXNET-SF as a function of GLEAM-T variability for all daily points combined. Values are Z-scores for daily mean values of sap flow and transpiration; data point color indicates the count of data point in each 0.05 bin. R^2 is the coefficient of determination of the standardized major axis 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).

relationship would indicate biases at low or high T. 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. In this model, the mismatching indices are the response variable, the percentile is the fixed effect and site is a random effect on the intercept. The random effect on the intercept allows pairing percentiles by site and controlling for site effects.

We use the same approach as above, SMA, 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. For site climatology indexes, 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 interpret each PCA axis based on the weight of the individual variables composing it.

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

Analyzing 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 ($p < 0.001$; Figure 2). This indicates a reasonable 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 5th, 50th and 95th percentiles—P05, P50, and P95) differed across percentiles ($p < 0.001$; Figure 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 accounting for the influence of trees of different sizes on site SF, by using a weighted mean instead of simple mean to calculate site SF (i.e., from Figures 1b and 1c; 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 (Figures 4a–4c). 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 ($p <= 0.03$; Figure 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; $p <= 0.01$; Figure 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 analyzed 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,

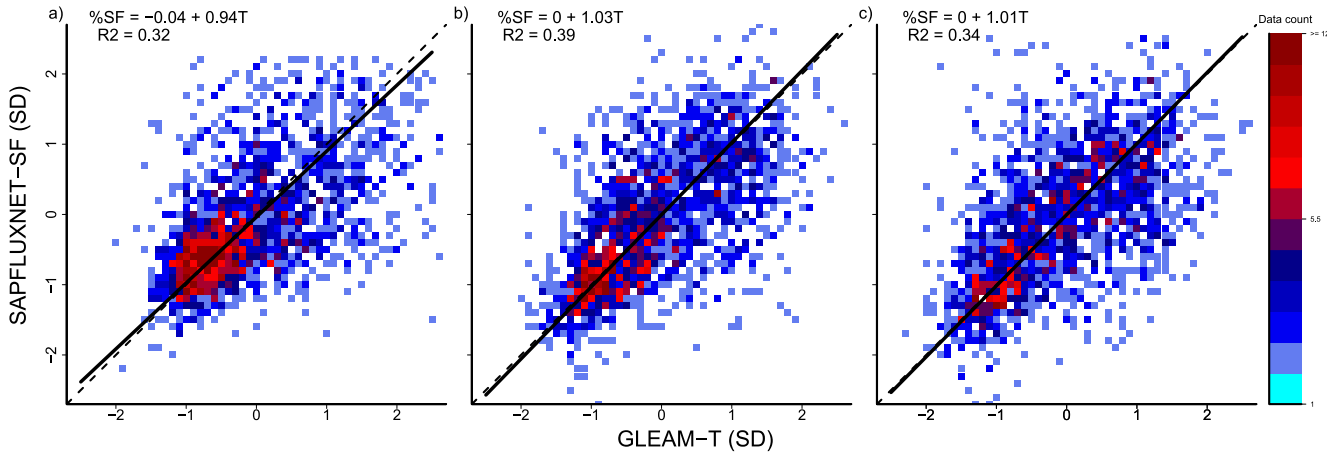

Figure 3. SAPFLUXNET-SF as a function of GLEAM-T. Graphs (a, b, and c) are, respectively, low, median and high transpiration daily values within a month and site (i.e., the 5th, 50th and 95th monthly percentiles of daily values). Data point color indicates the count of data point in each 0.1 bin. R^2 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.

using SMA. 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; Figure 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 ($p < 0.001$; Figure 4c) 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 ($p = 0.87$; Figure 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 ($p < 0.01$; Figure 4d).

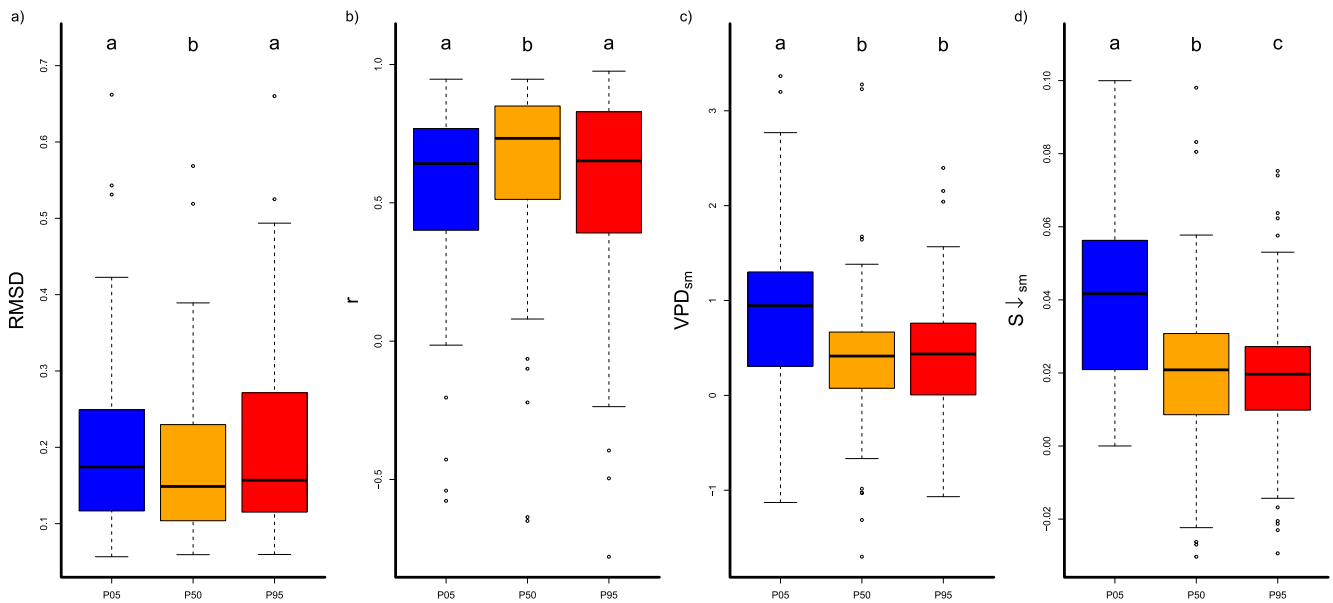

Figure 4. Site level mismatching indices between GLEAM-T and SAPFLUXNET-SF for the 5th, 50th, and 95th monthly percentiles (P5, P50, and P95; blue, orange, and red, respectively): (a) mean root squared difference (RMSD), (b) bivariate correlation (r), (c) vapor pressure deficit sensitivity mismatch (VPD_{sm}) 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.

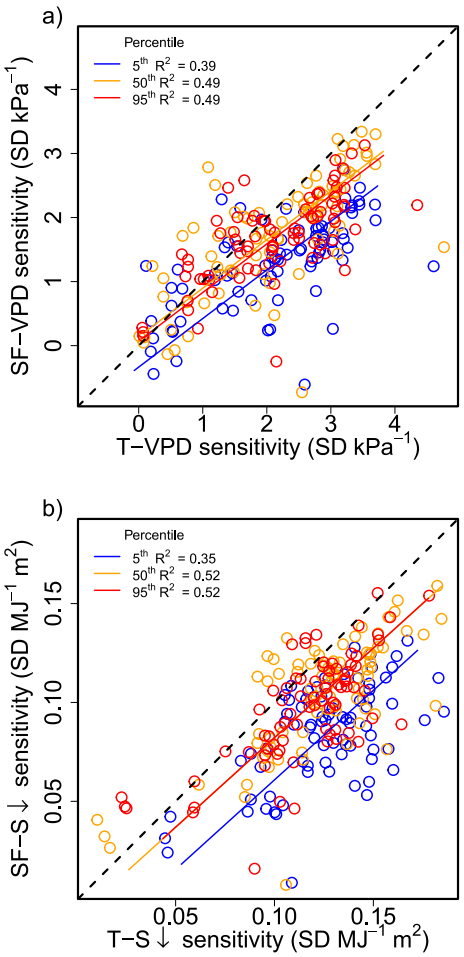

Figure 5. Relationships between GLEAM-T and SAPFLUXNET-SF sensitivities to vapor pressure deficit (VPD; a) and surface solar radiation ($S_{\downarrow}$; b). Blue, orange, and red points indicate, respectively, daily conditions, within months, with low, median, and high T (or sap flow) (i.e., 5th, 50th, 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). Colored lines are the standardized major axis fits for each percentile and the black dashed line is the 1:1 line.

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 PCA axes (Figure 6). The first and second axis of the PCA (PC1 and PC2) explained most of the data set 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 versus SAPFLUXNET -SF mismatches varied depending on the percentile analyzed (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 SF has been proposed as a mean to evaluate T data sets, 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 (Figures 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 (Feng et al., 2020; Miralles et al., 2016; Talsma et al., 2018).

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 (Figures 4d and 4e), especially for low transpiration conditions. Lower agreement between GLEAM and eddy-covariance data in arid conditions has been reported previously (Michel et al., 2016) and are often related to inaccurate plant stomatal control mechanisms or improper

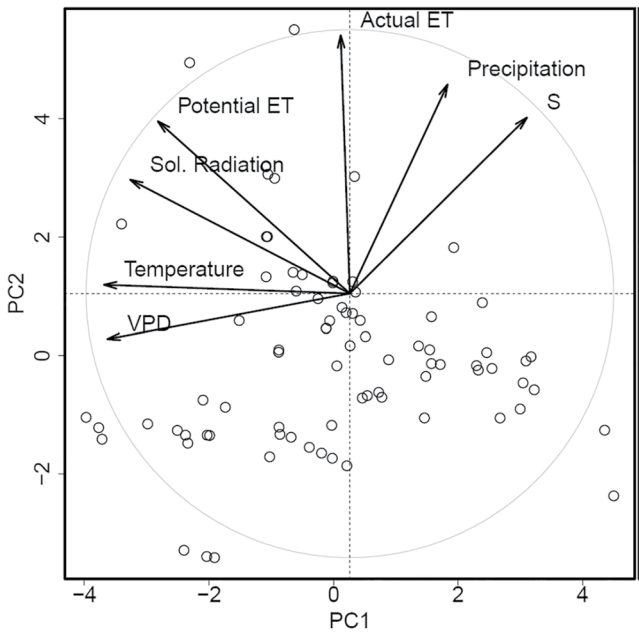

Figure 6. Principal component analysis of site climatic (vapor pressure deficit, incoming solar radiation, air temperature and precipitation) and model variables (potential and actual ET, and their ratio, i.e., S). The loadings of each variable into the PC1 and PC2 axis, as well as their contribution,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

root zone water availability (Feng et al., 2020; Liu et al., 2020). However, 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 (Ford et al., 2007; 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 (Z. Z. 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 SF 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 vs. 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 parameterization (Table 2). Consequently, with our approach confidence in conclusions reached for specific sites is limited, but cross-site analyses are likely to be robust.

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

	PC1		PC2	
	Loading	Contribution	Loading	Contribution
VPD	**−0.49**	**24.0**	−0.11	1.1
Temperature	**−0.44**	**19.7**	0.26	7.0
$S_{\downarrow}$	**−0.50**	**24.6**	0.02	0.1
Precipitation	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

Note. VPD, mean vapor 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. Variables with high loading/contributions for each axis are highlighted in bold.

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.

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, Bivariate Correlation (r), VPD Sensitivity Mismatch (VPD_{sm}) and Incoming Solar Radiation Mismatch ($S_{\downarrow sm}$)

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

Note. 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.

We believe the initial steps we provide here can be the foundation for a wider SF based validation of T products, models and mechanisms.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

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.

Acknowledgments
This work and its contributors Paulo Bittencourt and Lucy Rowland were 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).

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. https://doi.org/10.1016/j.jag.2020.102240

Bright, R. M., Miralles, D. G., Poyatos, R., & Eisner, S. (2022). Simple models outperform more complex big-leaf models of daily transpiration in forested biomes. *Geophysical Research Letters*, *49*(18), e2022GL100100. https://doi.org/10.1029/2022gl100100

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*(5), 529–546. https://doi.org/10.1007/s00468-004-0339-6

Dembélé, M., Ceperley, N., Zwart, S. J., 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. https://doi.org/10.1016/j.advwatres.2020.103667

Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., et al. (2017). Quantifying the influence of global warming on unprecedented extreme climate events. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(19), 4881–4886. https://doi.org/10.1073/pnas.1618082114

DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D.-H., Kelly, M., & Townshend, J. (2015). MOD44B MODIS/terra vegetation continuous fields yearly L3 global 250m SIN grid V006.

Dolman, A. J., Miralles, D. G., & de Jeu, R. A. M. (2014). Fifty years since Monteith's 1965 seminal paper: The emergence of global ecohydrology: Emergence global ecohydrology. *Ecohydrology*, *7*(3), 897–902. https://doi.org/10.1002/eco.1505

Feng, S., Liu, J., Zhang, Q., Zhang, Y., Singh, V. P., Gu, X., & Sun, P. (2020). A global quantitation of factors affecting evapotranspiration variability. *Journal of Hydrology*, *584*, 124688. https://doi.org/10.1016/j.jhydrol.2020.124688

Fisher, J. B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., 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*(4), 2618–2626. https://doi.org/10.1002/2016wr020175

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. https://doi.org/10.1016/j.agrformet.2019.03.012

Ford, C. R., Hubbard, R. M., Kloeppel, B. D., & Vose, J. M. (2007). A comparison of sap flux-based evapotranspiration estimates with catchment-scale water balance. *Agricultural and Forest Meteorology*, *145*(3–4), 176–185. https://doi.org/10.1016/j.agrformet.2007.04.010

Forster, M. (2017). How reliable are heat pulse velocity methods for estimating tree transpiration? *Forests*, *8*(9), 350. https://doi.org/10.3390/f8090350

Forzieri, G., Miralles, D. G., Ciais, P., Alkama, R., Ryu, Y., Duveiller, G., et al. (2020). Increased control of vegetation on global terrestrial energy fluxes. *Nature Climate Change*, *10*(4), 356–362. https://doi.org/10.1038/s41558-020-0717-0

Gash, J. H. C. (1979). An analytical model of rainfall interception by forests. *Quarterly Journal of the Royal Meteorological Society*, *105*(443), 43–55. https://doi.org/10.1002/qj.49710544304

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*(1), 109. https://doi.org/10.1038/s41597-020-0453-3

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hyndman, R. J., & Fan, Y. (1996). Sample quantiles in statistical packages. *The American Statistician*, *50*(4), 361. https://doi.org/10.2307/2684934

Immerzeel, W. W., Lutz, A. F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., et al. (2020). Importance and vulnerability of the world's water towers. *Nature*, *577*(7790), 364–369. https://doi.org/10.1038/s41586-019-1822-y

Jiménez, C., Martens, B., Miralles, D. M., Fisher, J. B., Beck, H. E., & 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*(8), 4513–4533. https://doi.org/10.5194/hess-22-4513-2018

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., et al. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, *6*(1), 74. https://doi.org/10.1038/s41597-019-0076-8

Keenan, T. F., Hollinger, D. Y., Bohrer, G., Dragoni, D., Munger, J. W., Schmid, H. P., & Richardson, A. D. (2013). Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise. *Nature*, *499*(7458), 324–327. https://doi.org/10.1038/nature12291

Kelley, D., & Richards, C. (2020). OCE: Analysis of oceanographic data. R package version 1.2-0. Retrieved from https://CRAN.R-project.org/package=oce

Kool, D., Agam, N., Lazarovitch, N., Heitman, J. L., Sauer, T. J., & Ben-Gal, A. (2014). A review of approaches for evapotranspiration partitioning. *Agricultural and Forest Meteorology*, *184*, 56–70. https://doi.org/10.1016/j.agrformet.2013.09.003

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*(16), 2335–2344. https://doi.org/10.1002/hyp.7338

Kumarathunge, D. P., Medlyn, B. E., Drake, J. E., Tjoelker, M. G., Aspinwall, M. J., Battaglia, M., et al. (2019). Acclimation and adaptation components of the temperature dependence of plant photosynthesis at the global scale. *New Phytologist*, *222*(2), 768–784. https://doi.org/10.1111/nph.15668

Lemordant, L., Gentine, P., Swann, A. S., Cook, B. I., & 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 of the United States of America*, *115*(16), 4093–4098. https://doi.org/10.1073/pnas.1720712115

Liu, Y., Kumar, M., Katul, G. G., Feng, X., & Konings, A. G. (2020). Plant hydraulics accentuates the effect of atmospheric moisture stress on transpiration. *Nature Climate Change*, *10*(7), 691–695. https://doi.org/10.1038/s41558-020-0781-5

Mackay, D. S., Ewers, B. E., Loranty, M. M., & Kruger, E. L. (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*, *115*(G2), G02016. https://doi.org/10.1029/2009jg001092

Maes, W. H., Pagán, B. R., Martens, B., Gentine, P., Guanter, L., Steppe, K., 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. https://doi.org/10.1016/j.rse.2020.112030

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

Martens, B., Waegeman, W., Dorigo, W. A., Verhoest, N. E. C., & Miralles, D. G. (2018). Terrestrial evaporation response to modes of climate variability. *NPJ Climate and Atmospheric Science*, *1*, 43. https://doi.org/10.1038/s41612-018-0053-5

Michalsky, J. J. (1988). The astronomical Almanac's algorithm for approximate solar position (1950–2050). *Solar Energy*, *40*(3), 227–235. https://doi.org/10.1016/0038-092x(88)90045-x

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., 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*(2), 803–822. https://doi.org/10.5194/hess-20-803-2016

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, *15*(2), 453–469. https://doi.org/10.5194/hess-15-453-2011

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., McCabe, M. F., et al. (2016). The WACMOS-ET project—Part 2: Evaluation of global terrestrial evaporation data sets. *Hydrology and Earth System Sciences*, *20*(2), 823–842. https://doi.org/10.5194/hess-20-823-2016

Miralles, D. G., van den Berg, M. J., Gash, J. H., Parinussa, R. M., de Jeu, R. A. M., Beck, H. E., et al. (2014). El Niño–La Niña cycle and recent trends in continental evaporation. *Nature Climate Change*, *4*(2), 122–126. https://doi.org/10.1038/nclimate2068

Nelson, J. A., Pérez-Priego, O., Zhou, S., Poyatos, R., Zhang, Y., Blanken, P. D., et al. (2020). Ecosystem transpiration and evaporation: Insights from three water flux partitioning methods across FLUXNET sites. *Global Change Biology*, *26*(12), 6916–6930. https://doi.org/10.1111/gcb.15314

Penman, H. L. (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. Retrieved from https://CRAN.R-project.org/package=nlme

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

Poyatos, R., Granda, V., Flo, V., Adams, M. A., Adorján, B., Aguadé, D., et al. (2021). Global transpiration data from sap flow measurements: The SAPFLUXNET database. *Earth System Science Data*, *13*, 2607–2649.

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

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Sabater, A. M., Ward, H. C., Hill, T. C., Gornall, J. L., Wade, T. J., Evans, J. G., et al. (2020). Transpiration from subarctic deciduous woodlands: Environmental controls and contribution to ecosystem evapotranspiration. *Ecohydrology*, *13*(3). https://doi.org/10.1002/eco.2190

Smith, R. J. (2009). Use and misuse of the reduced major axis for line-fitting. *American Journal of Physical Anthropology*, *140*(3), 476–486. https://doi.org/10.1002/ajpa.21090

Stephens, C. M., Lall, U., Johnson, F. M., & Marshall, L. A. (2021). Landscape changes and their hydrologic effects: Interactions and feedbacks across scales. *Earth-Science Reviews*, *212*, 103466. https://doi.org/10.1016/j.earscirev.2020.103466

Stoy, P. C., El-Madany, T. S., Fisher, J. B., Gentine, P., Gerken, T., Good, S. P., et al. (2019). Reviews and syntheses: Turning the challenges of partitioning ecosystem evaporation and transpiration into opportunities. *Biogeosciences*, *16*(19), 3747–3775. https://doi.org/10.5194/bg-16-3747-2019

Talsma, C. J., Good, S. P., Jimenez, C., Martens, B., Fisher, J. B., Miralles, D. G., et al. (2018). Partitioning of evapotranspiration in remote sensing-based models. *Agricultural and Forest Meteorology*, *260–261*, 131–143. https://doi.org/10.1016/j.agrformet.2018.05.010

Thomas, R., Lello, J., Medeiros, R., Pollard, A., Robinson, P., Seward, A., et al. (2017). *Data analysis with R statistical software: A guidebook for scientists*. Eco-Explore.

Wang, K., & Dickinson, R. E. (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*(1–2), 521–531. https://doi.org/10.1007/s00704-020-03437-4

Warton, D. I., Duursma, R. A., Falster, D. S., & 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*(2), 257–259. https://doi.org/10.1111/j.2041-210x.2011.00153.x

Wei, Z., Yoshimura, K., Wang, L., Miralles, D. G., Jasechko, S., & Lee, X. (2017). Revisiting the contribution of transpiration to global terrestrial evapotranspiration: Revisiting Global ET Partitioning. *Geophysical Research Letters*, *44*(6), 2792–2801. https://doi.org/10.1002/2016gl072235

Whelan, M. E., Lennartz, S. T., Gimeno, T. E., Wehr, R., Wohlfahrt, G., Wang, Y., et al. (2018). Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles. *Biogeosciences*, *15*(12), 3625–3657. https://doi.org/10.5194/bg-15-3625-2018

Williams, D. G., Cable, W., Hultine, K., Hoedjes, J. C. B., Yepez, E. A., Simonneaux, V., et al. (2004). Evapotranspiration components determined by stable isotope, sap flow and eddy covariance techniques. *Agricultural and Forest Meteorology*, *125*(3–4), 241–258. https://doi.org/10.1016/j.agrformet.2004.04.008

Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Springer.