Journal of Geophysical Research: Biogeosciences

RESEARCH ARTICLE
10.1029/2018JG004727

Coupling Water and Carbon Fluxes to Constrain Estimates of Transpiration: The TEA Algorithm

Jacob A. Nelson[1], Nuno Carvalhais[1,2], Matthias Cuntz[3], Nicolas Delpierre[4], Jürgen Knauer[1], Jérome Ogée[5], Mirco Migliavacca[1], Markus Reichstein[1,6], and Martin Jung[1]

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, Germany, [2]Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisbon, Portugal, [3]INRA, Université de Lorraine, UMR1137 Ecology et Ecophysiologie Forestières, Champenoux, France, [4]Ecologie Systématique Evolution, Université Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, Orsay, France, [5]INRA, UMR 1391 ISPA, Villenave d'Ornon, France, [6]Michael-Stifel-Center Jena for Data-Driven and Simulation Science, Jena, Germany

Key Points:
- TEA is a method for extracting water use efficiency (WUE) dynamics from flux data with minimal assumptions
- Validation shows that TEA is able to derive patterns of WUE and transpiration from three different models
- Method is applicable to eddy covariance data sets, opening the door to wide-scale transpiration estimates

Supporting Information:
- Supporting Information S1

Correspondence to:
J. A. Nelson,
jnelson@bgc-jena.mpg.de

Citation:
Nelson, J. A., Carvalhais, N., Cuntz, M., Delpierre, N., Knauer, J., Ogée, J., et al. (2018). Coupling water and carbon fluxes to constrain estimates of transpiration: The TEA algorithm. *Journal of Geophysical Research: Biogeosciences*, *123*, 3617–3632. https://doi.org/10.1029/2018JG004727

Received 7 AUG 2018
Accepted 15 NOV 2018
Accepted article online 23 NOV 2018
Published online 21 DEC 2018

Author Contributions
Writing - Original Draft: Jacob A. Nelson
Formal Analysis: Jacob A. Nelson
Visualization: Jacob A. Nelson

Abstract Plant transpiration (*T*), biologically controlled movement of water from soil to atmosphere, currently lacks sufficient estimates in space and time to characterize global ecohydrology. Here we describe the Transpiration Estimation Algorithm (TEA), which uses both the signals of gross primary productivity and evapotranspiration (ET) to estimate temporal patterns of water use efficiency (WUE, i.e., the ratio between gross primary productivity and *T*) from which *T* is calculated. The method first isolates periods when *T* is most likely to dominate ET. Then, a Random Forest Regressor is trained on WUE within the filtered periods and can thus estimate WUE and *T* at every time step. Performance of the method is validated using terrestrial biosphere model output as synthetic flux data sets, that is, flux data where WUE dynamics are encoded in the model structure and *T* is known. TEA reproduced temporal patterns of *T* with modeling efficiencies above 0.8 for all three models: JSBACH, MuSICA, and CASTANEA. Algorithm output is robust to data set noise but shows some sensitivity to sites and model structures with relatively constant evaporation levels, overestimating values of *T* while still capturing temporal patterns. The ability to capture between-site variability in the fraction of *T* to total ET varied by model, with root-mean-square error values between algorithm predicted and modeled *T*/*ET* ranging from 3% to 15% depending on the model. TEA provides a widely applicable method for estimating WUE while requiring minimal data and/or knowledge on physiology which can complement and inform the current understanding of underlying processes.

Plain Language Summary While it is widely known that plants need water to survive, exactly how much water plants in an ecosystem use is hard to quantify. However, many places have been measuring how much total water leaves an ecosystem, both the water plants use directly and the water that simply evaporates from the soil or the surfaces of leaves, using eddy covariance towers. These eddy covariance towers also measure the coming and going of carbon, such as the total amount of carbon taken up by photosynthesis. Here we present the idea that by using the signals from both photosynthesis and total water losses together, we can capture the water signal related to plants, namely, transpiration, using an algorithm called Transpiration Estimation Algorithm (TEA). To verify that TEA is working the way we expect, we test it out using artificial ecosystem simulations where transpiration and photosynthesis come from mathematical models. By thoroughly testing TEA, we have a better idea of how it will work in a real world situation, hopefully opening the door for a better understanding on how much water ecosystems are using and how it might affect our changing planet.

1. Introduction

At current state, transpiration (*T*) is a key ecosystem process that lacks the widespread and consistent estimates necessary to study ecohydrological processes globally. For example, a recent meta-analysis by Wei et al. (2017) analyzed an aggregation of ecosystem level *T* estimates resulting in a data set of only 64 studies conducted between 1941 and 2014, a relatively sparse data set when attempting to capture global variability. As such, demand for *T* data sets that can encompass the variety of ecosystem responses to water availability has been highlighted as a key need, from the perspectives of both the water (Fisher et al., 2017) and carbon

(Rogers et al., 2017) cycle communities. Though transpiration and evaporation (*E*) processes are built into most ecosystem and land surface models, resulting estimates are poorly constrained, as can be seen in the spread of global *T*/*ET* estimates from the Coupled Model Intercomparison Project Phase 5 which ranged from 22% to 58% (Wei et al., 2017). Here we present an approach for estimating *T* which is applicable to eddy covariance (EC) networks and is data driven providing an alternative perspective to current process-based approaches.

The difficulty in partitioning evapotranspiration (ET) into the biotic component (transpiration, *T*) and the abiotic component (here evaporation, *E*) is partially due to equifinality, as *E* and *T* share the same primary environmental drivers making the problem difficult to constrain. From the view point of physics, transpiration is an evaporation which is then modulated by stomatal resistance, making the task of distinguishing the two fluxes particularly challenging. However, a key distinction of *T* lies in that it is regulated by an active process via stomatal control, which is linked to plant photosynthesis. To this end, the method we propose aims to utilize this link between water and carbon cycles as the key differentiating process between *E* and *T* in an effort to distinguish the two.

As reviewed in Kool et al. (2014), many approaches to partition ET attempt to pair a separate *E* and/or *T* distinguishing estimate, such as measurements of sap flux, isotope fractionation, or carbonyl sulfide (OCS) flux, in tandem with an ET estimate. Sap flux measurements, which estimate the flow of water through a stem (Granier, 1987), is currently the most widespread method. Though sap flux measurements have proven to be effective at measuring tree water fluxes, estimating ecosystem *T* relies on upscaling point source sap flow estimates based on an approximation of sapwood area, which can be problematic in ecosystems with high plant diversity, hampering suitability for universal application (Oishi et al., 2008; Poyatos et al., 2016). Isotopic methods take advantage of the isotopic fractionation of water oxygen ($^{18}O/^{16}O$) and hydrogen ($^{2}H/^{1}H$) which occurs from evaporation but not root uptake, producing isotopic signatures related to the *T*:*E* ratio. Isotopic methods have been used both at global scales (Good et al., 2015; Jasechko et al., 2013) and at high temporal resolutions (Good et al., 2014) yet are limited in the number of sites and length of time series. In general, global isotopic estimates of *T*/*ET* tend to be higher than site estimates (Wei et al., 2017), some even controversially so (Coenders-Gerrits et al., 2014). The OCS method attempts to use the flux of OCS uptake by leaves to estimate ecosystem canopy conductance directly, as the pathways of CO_2 and OCS are similar (Sandoval-Soto et al., 2005; Whelan et al., 2017). The calculation of conductance is simplified when using OCS, as it does not have the complication of having a large source component such as is the case with respiration and CO_2 (Wehr et al., 2017). While the OCS method is promising, the novelty and potential complications due to alternate sources/sinks of OCS (Gimeno et al., 2017; Wohlfahrt, 2017) have resulted in limited applications in practice.

Due to the limits of current *T* estimates, and shortfalls in understanding ecosystem water dynamics, data-driven approaches can provide an alternate perspective. Widespread monitoring of both water and CO_2 fluxes provide rich data sets that can inform *T* estimates by utilizing concepts of water use efficiency (WUE), here defined as the ratio of gross primary productivity (GPP) to *T*. At present, data-driven approaches to estimate ecosystem WUE and *T* do exist, such as the method proposed by Zhou et al. (2016; hereafter referred to as the *Zhou method*) which is based on estimates of annual underlying WUE from GPP and vapor pressure deficit (VPD), calculated as

$$uWUE_t = \frac{GPP_t \cdot \sqrt{VPD_t}}{ET_t}, \tag{1}$$

where the $\sqrt{VPD}$ term represents an approximate stomatal response which is broadly applied to many ecosystems. *uWUE* is related to the carbon cost of water which is assumed to be constant in light-limited leaves over time scales of days to weeks. By incorporating the $\sqrt{VPD}$ term, the carbon:water relationship becomes linear and *uWUE* values can be estimated using linear regression where *uWUE* is the slope parameter. The Zhou method makes a *T*/*ET* estimate by taking the ratio of a normal linear regression of *uWUE* estimated within 1 day (which would include the *E* component), and the 95th percentile regression of annual *uWUE* which is assumed to contain only transpiration. The key assumptions are then that *uWUE* is constant within a year and that the 95th percentile of *uWUE* corresponds to conditions where $E \approx 0$. *E* is most likely to be 0 at high percentiles because these points correspond to periods with the highest ratio of *GPP*/*ET*, whereas points with a high *E* component would increase ET with no added *T* causing the *uWUE* to decrease. Points over the 95th percentile are assumed not to be representative of *uWUE*, possibly due to noise.

While Zhou et al. (2016) and other data-driven WUE estimates (Beer et al., 2009; Scott & Biederman, 2017) have laid the foundation for globally useful WUE and *T* estimates, they have yet to be rigorously validated, likely

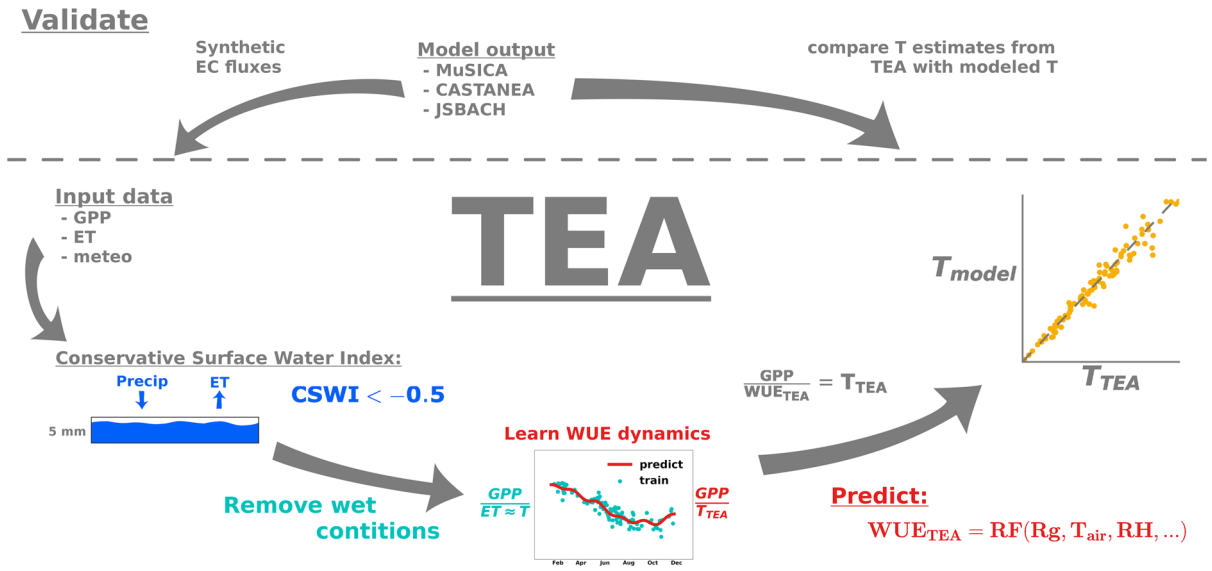

Figure 1. Theoretical outline of both the TEA algorithm (lower section) and validation steps (upper section). TEA consists of a filtration step to isolate dry periods when vegetation is active using the CSWI, then a Random Forest Regressor is trained on the filtered data set to characterize WUE dynamics, which are then predicted for the entire time series. The validation scheme involves using model output as a synthetic flux data set, to evaluate if TEA is able to replicate the WUE dynamics encoded in the models. TEA = Transpiration Estimation Algorithm; CSWI = conservative surface wetness index; WUE = water use efficiency.

in part due to the limited availability of verification data sets as described in the previous section. Notably, assumptions on ecosystem WUE dynamics which are not fully understood must be taken into consideration. In particular, nonlinearities in the GPP to *T* relationship must be addressed such as the known effects from stomatal response (VPD) (Beer et al., 2009; Katul et al., 2009; Zhou et al., 2014). Though the Zhou method does attempt to account for VPD effects, the resulting *uWUE* estimate is tied to annual time periods and does not allow for seasonal or diurnal variations in plant and ecosystem responses, only accounting for the VPD response. Boese et al. (2017) concluded that the *uWUE* framework could be outperformed by empirical models that included incoming radiation, suggesting that only incorporating VPD may not be sufficient to characterize the carbon:water relationships at ecosystem level. As such, the method proposed here attempts to derive WUE dynamics from a data-driven perspective, using a nonlinear, machine learning method to characterize the carbon:water relationship and thus make few assumptions on the ecosystem WUE dynamics.

1.1. Method Outline and Objectives

We identify two key limitations of the current methods outlined: (1) restricted applicability or spatiotemporal scope, particularly with direct *T* measurements; (2) strong assumptions of carbon:water relationship, particularly with EC-dependent methods, which have the potential to bias WUE and *T* estimates.

We aim to overcome the first limitation by basing the method only on water, energy, and carbon EC fluxes with associated meteorological data to make predictions at half-hourly to hourly scales with minimal data requirements.

To address the second limitation, we validated the presented ET partitioning method against model output in an effort to assess sensitivities and limitations. The use of artificial data sets has proven useful both in the field of biogeochemistry (Jung et al., 2009), as well as adjacent fields (Ishizaki et al., 2014; Jasechko et al., 2014). We used three separate models in an effort to reduce the influence of any one set of model assumptions.

Here we introduce the Transpiration Estimation Algorithm (TEA), which uses ecosystem WUE ($eWUE = GPP/ET$) to predict transpiration in two steps (see Figure 1): (1) a data filtration step to isolate the signal of ET for periods where *E* is minimized and ET is likely dominated by the signal of *T*, that is, during periods of the growing season with dry surfaces; and (2) a step that predicts the WUE using meteorological variables, as well as information derived from the carbon and energy fluxes. This prediction of WUE translates to a novel transpiration estimate which aims to be capable of capturing seasonal and diurnal dynamics with wide application potential.

Table 1
Overview of Filters Used to Isolate Conditions Where the Signal of ET is Dominated by T

Variable	Long name	Half-hourly limit	Daily limit
GPP	Gross primary productivity	> 0.05 μmol C·m^{-2}·s^{-1}	> 0.5 g C·m^{-2}·day^{-1}
T_{air}	Air temperature	>5 °C	—
Rg	Incoming radiation	> 0 W/m^2	—
CSWI	Conservative surface wetness index	<−3 to 2 mm	—

Note. GPP and T_{air} filters were designed to ensure plants are active, while *Rg* filters remove nighttime values. The *CSWI* filter attempts to remove periods where the surface is likely to be wet, a sensitivity analysis of which can be found in section 3.2.

The key hypothesis to be tested here is as follows: Does the TEA algorithm capture the dynamics of WUE and T encoded in the models? If the method cannot capture WUE dynamics from the three different models, we can assume it will not capture real world WUE dynamics; thus, the exercise is a sanity check on whether TEA is capable of extracting physiological patterns of ecosystem WUE. Furthermore, we explore scenarios when a key assumption is broken, that is, evaporation is persistent at every point in time, as well as how evaporation can bias the results and how to mitigate this bias using percentile regression.

2. Methods

2.1. Isolating Training Periods

To accomplish the characterization of T in time, we used the assumption that the signal of $T/ET \approx 1$ under conditions where the ecosystem has minimal surface moisture and the plants are photosynthetically active, as manifest via the set of filters outlined in Table 1. Filters were constructed from half hourly flux and meteorological data which excluded periods that did not meet filter criteria (individual half hours were removed). Periods likely to have no or low photosynthetic activity were removed, such as nighttime values or periods with low temperatures, as well as full days that did not reach a minimum threshold of total daily GPP. Periods expected to have high surface moisture were removed using the conservative surface wetness index (CSWI), a shallow bucket model where the bucket represents the surface water storage (S) for each half-hourly time step t (S_t) relative to the last precipitation event, or,

$$S_t = \min\left(S_{t-1} + P_t - ET_t, S_{\max}\right), \tag{2}$$

where P_t is the precipitation at time t and S_{max} is the maximum allowable storage (size of bucket). S_{max} was set to 5 mm, and values from 3 to 9 mm showed no difference in filter utility (data not shown, further discussed in section 4.2). The CSWI is then calculated as

$$CSWI = \max\left(S_t, \min\left(P_t, S_{\max}\right)\right). \tag{3}$$

Periods were considered sufficiently dry based on a CSWI limit, that is, periods where $CSWI < limit$are assumed to have dry surfaces. As opposed to other methods of identifying wet and dry conditions, such as removing periods after rain events, the CSWI accounts for the amount of rain evaporated and therefore compensates for small rain events which may evaporate relatively quickly as well as for periods of low ET after rain events such as persistent clouds reducing radiation inputs where surfaces may stay wet longer. As the appropriate limit for CSWI was unknown, this limit then becomes an input parameter to the algorithm which is not optimized, or hyperparameter; hence, a sensitivity analysis was conducted across a range of limits from 2 to −3 mm (see section 3.2). The CSWI limits were not extended past −3 mm, as lower limits resulted in fewer than 500 half hours remaining in the training data set at some sites, which was considered too few to properly characterize site variability. Note that the limit of 500 half hours is arbitrary and possibly conservative; however, results indicate stricter limits (i.e., CSWI < −3 mm) may cause the training data set to only include periods of water stress and decrease prediction performance (see Figure 7). Similarly for CSWI, periods when daily GPP was too low were also filtered in an effort to remove periods when the plants are relatively inactive, such as transition periods from winter to spring. A minimum daily threshold of 0.5 g C·m^{-2}·day^{-1} was found to give a good performance, and a sensitivity analysis to daily GPP filter can be found in supporting information Figure S1.

Each individual filter was combined (logical AND), resulting in a filtered time series that was then used to calculate half-hourly values of *eWUE* to be used as a training data set in the next section, hereafter referred to as the training data set.

2.2. Modeling WUE and Predicting *T*

Using a set of features **X**, we trained a random forest regressor (RF; RandomForestRegressor from Pedregosa et al., 2011, based on Breiman, 2001) on *eWUE* within the training data set (for each site) made with the filters outlined in Table 1. Features consisted of four meteorological variables: incoming radiation (Rg), air temperature (T_{air}), relative humidity (RH), wind speed (u); four derived variables: the derivative of a Gaussian filtered GPP (GPP'), the Rg normalized diurnal centroid of ET (C^*_{ET}), the diurnal water:carbon index (DWCI), CSWI; and daily potential radiation ($Rg_{pot,daily}$), the derivative of daily potential radiation ($Rg'_{pot,daily}$), and year. The resulting feature vector **X** is

$$\mathbf{X} = [Rg, T_{air}, RH, u, Rg_{pot,daily}, Rg'_{pot,daily}, CSWI, GPP', C^*_{ET}, DWCI, year]. \quad (4)$$

Note that C^*_{ET} measures the morning shift of diurnal ET, and DWCI measures the degree of correlation in 1 day between GPP and ET; a detailed explanation of C^*_{ET}, DWCI, and CSWI can be found in Nelson et al. (2017), code for which can be found at Nelson (2017). The set of features **X** was designed to give the RF regressor information on processes that may impact WUE.

The full time series of WUE was then predicted for all half hours (unfiltered data) using the resulting model as

$$WUE_{t,pred} = RF_P(\mathbf{X_t}, P), \quad (5)$$

where P is the percentile used from each resulting predictive leaf, or prediction percentile (Meinshausen, 2006). Quantile random forest regression is analogous to the linear quantile regression use by the Zhou method but makes no assumptions on linearity.

The RF utilized 100 trees which were fully grown, and each splitting node consisted of a maximum number of features equal to one third the total number of features, rounded up. A sensitivity analysis of the number of trees and maximum number of feature parameters can be found in supporting information Figure S2.

As ET in the training data set is assumed to be only a proxy of T, there is likely E still present even after filtering. For example, when making a prediction for a particular half hour the process would work as follows: Features of the half hour would be fed to the RF (Rg, T_{air}, RH, etc …); in turn the RF will return a number of WUE values which it has identified as associated with the particular features of that half hour; this set of returned values can then be summarized, which is typically via the mean, but can also be a median or any other quantile such as the percentiles used here. If one assumes that all these WUE values from the RF for a half hour represents a single *true WUE* (GPP/T) that is contaminated by some residual evaporation (GPP/($T + E$)), the best summary statistic to use would be the maximum, as that would be the point most likely to have minimal residual evaporation. However, because the assumption that the WUE values returned from the RF likely do not represent a single true WUE, and instead variability comes both from residual evaporation and variability in WUE, the most appropriate percentile is not known. Therefore, the magnitude of predicted WUE can be adjusted using the percentile of prediction from the random forest and the optimum percentile, another hyperparameter which is not known a priori. A sensitivity analysis of prediction percentiles can be found in section 3.2. Note that extraction of percentiles from 50 to 100 is the result of a single prediction step with a single trained RF regressor; that is, the RF was not retrained for each percentile.

Given an estimate of WUE, the prediction of transpiration at time t was calculated as

$$T_t = \frac{GPP_t}{WUE_{t,pred}} \quad (6)$$

for each half hour, where nighttime values of T are considered 0. The evaporation component at time t was then estimated as

$$E_t = ET_t - T_t. \quad (7)$$

All code for processing and partitioning, as well as interactive examples, can be found in Nelson (2018).

2.3. Model Output Used for Method Evaluation

To test the predictive performance of the method, WUE and T estimates were compared to output from three separate models with different underlying carbon:water coupling mechanisms: CASTANEA (Delpierre et al., 2012; Dufrêne et al., 2005), JSBACH (Knauer et al., 2015; Reick et al., 2013), and MuSICA (Ogée et al., 2003;

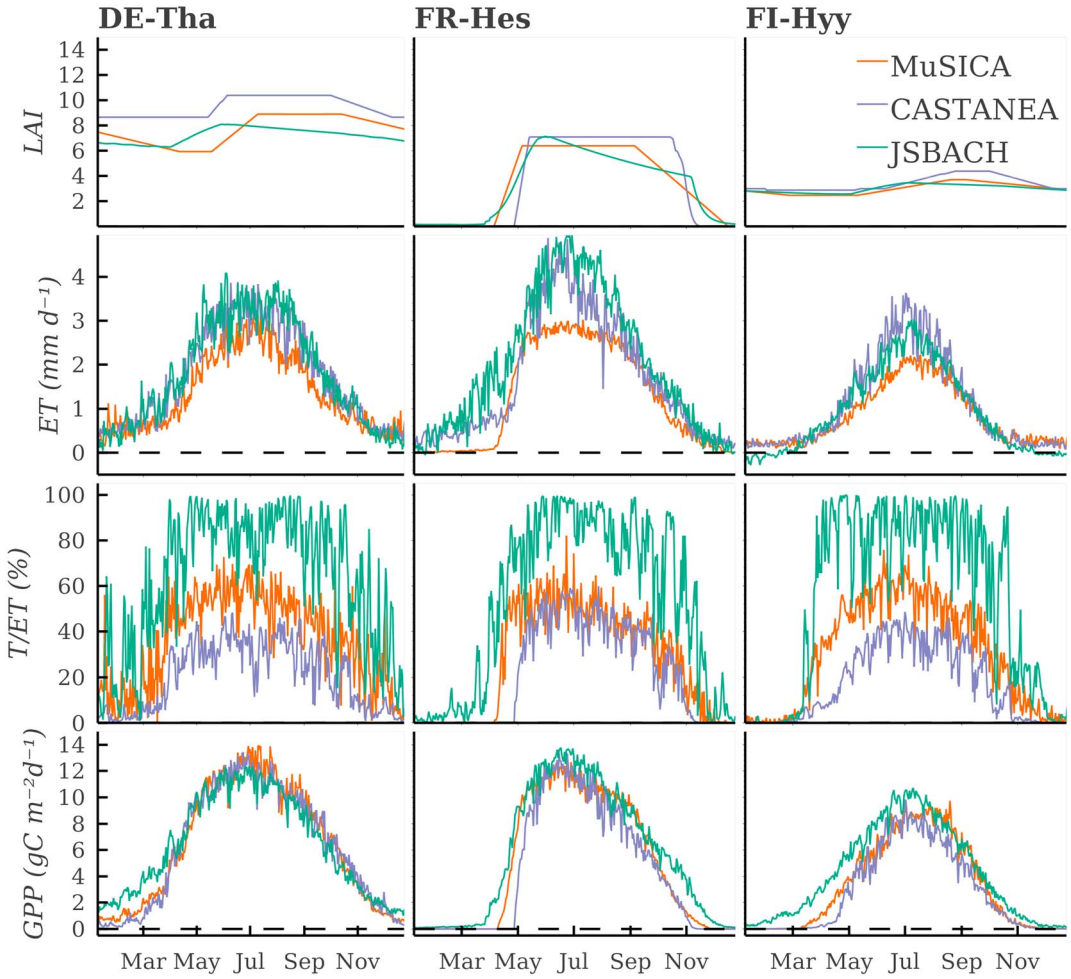

Figure 2. Average seasonal cycle of four output variables from JSBACH, CASTANEA, and MuSICA, all driven with the same meteorological forcing from three forest sites. Modeled years were 1997–2011 for FI-Hyy and FR-Hes, and 1997–2010 for DE-Tha. Data presented correspond to daily data averaged across all years and are intended to show seasonal trends. FI-Hyy = Hyytiälä, Finland; FR-Hes = Hesse beech forest in France; DE-Tha = Anchor Station Tharandt, Germany.

Potier et al., 2015; Wilkinson et al., 2015). Each model comes from a slightly different perspective, characterized by different model structures and ways of dealing with carbon-water relationships. JSBACH differs from the other two models in that it is a land surface model designed to be integrated into a global climate model, which was run off-line for this study. MuSICA separates the canopy into multiple layers, with each layer containing various plant components each with their own water status, light regime, and age. CASTANEA focuses on the growth, carbon allocation, and water budget of a monospecific forest stand.

Models were run using meteorological forcing data from 73 different sites, with 85 model runs in total (see full list in File S6 in the supporting information). Meteorological forcing data for the models came directly from the flux towers. This exercise was designed to test whether the method is capable of extracting a known carbon:water relationship even when the underlying assumptions are different. The ability of the algorithm to infer the complex formulations from these process-based models gives credence to the capability of the method to estimate these processes in real data. Therefore, the method was applied to the modeled GPP and ET fluxes paired with the respective forcing meteorological data, with the resulting TEA algorithm transpiration estimates compared to the modeled T. An intercomparison of the three models which used the same meteorological forcing data set can be seen for three sites in Figure 2. Key distinctions between the models can be seen in the leaf area index (LAI) and T/ET, with highest values of LAI from CASTANEA and MuSICA and highest T/ET values being from JSBACH. An overview of key model features can be found in Table 2.

Table 2
Overview of Key Processes in the Three Models Used for Validation

Key processes	Remarks
MuSICA	
Interception	Canopy rain interception and water storage on leaf surfaces are computed in each vegetation layer using a water balance equation and the concept of maximum storage capacity, scaled by the leaf area of each layer (Rutter et al., 1971). Evaporation from the interception storage is taken at the potential rate in each layer. More details are provided in Potier et al. (2015).
Water stress	Stomatal conductance, leaf photosynthetic capacity, and/or root hydrauilc conductivity downregulated based on instantaneous (Tuzet) or predawn (Ball, Leuning) leaf water potential.
Stomatal conductance	Ball et al. (1987), Leuning (1995), or Tuzet et al. (2003) depending on parameterizations available for individual sites.
Soil evaporation	zlitter acts as a separate, insulating layer.
Phenology	Dates of phenology events (bud burst, senescence) and minimum/maximum leaf area are constant throughout the simulation and supplied by the user.
JSBACH	
Interception	Water storage for the whole canopy, scaled by LAI, with evaporation from interception storage at the potential rate.
Water stress	Nonlinear reduction of g1 (stomatal slope parameter) and photosynthetic capacity (V_{cmax} and J_{max}) based on available soil moisture.
Stomatal conductance	Medlyn et al. (2011)
Soil evaporation	Soil evaporation coming from top soil layer (of 5).
Phenology	Logistic Growth Phenology model (LoGro-P); calculation depends on the phenotype, dependent on temperature, soil moisture, and NPP; for evergreen and deciduous forests (described in Böttcher et al., 2016): heat sum approach in combination with a critical number of chill days.
CASTANEA	
Interception	Water storage for the whole canopy, function of WAI and LAI.
Water stress	Linear reduction of g1 based on extractable soil water content.
Stomatal conductance	Ball et al. (1987)
Soil evaporation	Evaporation coming from both litter and top soil layer, soil moisture levels updated daily.
Phenology	LAI dynamics based on degree-days (Delpierre et al., 2009); for coniferous trees, winter regulation of photosynthetic transpiration activity further modulated by thermal acclimation (Delpierre et al., 2012).

Note. LAI = leaf area index; NPP = net primary production; WAI = wood area index.

Comparisons between model output and TEA estimations were focused on two key aspects: replication of patterns and minimizing bias. The ability to capture patterns was assessed using the modeling efficiency (MEF; Nash & Sutcliffe, 1970), calculated as

$$MEF = 1 - \frac{\sum (T_{t,\text{model}} - T_{t,\text{TEA}})^2}{\sum (T_{t,\text{model}} - \overline{T_{\text{model}}})^2}. \quad (8)$$

As this metric is meant to identify only patterns so as to differentiate bias due to consistent overestimations/underestimations and inability to capture temporal patterns, the mean values of T_{TEA} and T_{model} are removed prior to calculating the MEF. Quantification of bias was calculated as a relative bias,

$$bias = \frac{\sum T_{t,\text{TEA}} - \sum T_{t,\text{model}}}{\sum T_{t,\text{model}}}. \quad (9)$$

2.4. Noise and Evaporation Sensitivity Experiments

To isolate the effects of noise and training set *E*, two artificial experiments were conducted where the data from each model run were used to create a series of new experimental data sets. The first case attempted to assess the sensitivity of the TEA algorithm hyperparameters to the presence of noise, which is likely to be present in real EC data and is not present in the model output. The second experiment aimed to isolate the effect of *E* on prediction bias, with the aim of understanding how a persistent fraction of *E* may potentially bias *T* estimates.

To test the effects of noise, Random Gaussian noise was added to the original modeled GPP and ET values with a standard deviation corresponding to a scaling factor (s) according to percentages of the original value:

5%, 10%, 15%, 20%, and 25%. The experimental GPP and ET fluxes were then calculated as

$$flux_{t,\text{noise}} = N\left(flux_{t,\text{original}}, \left(flux_{t,\text{original}} \cdot s\right)^2\right), \tag{10}$$

with the resulting *eWUE* containing noise in both the GPP and ET components. Noise levels were designed to encompass the range expected in real EC data (Hollinger & Richardson, 2005).

To isolate the effect of E on prediction bias, it is important to distinguish between total E and E which is persistent within the training data set (E_{train}, see Figure 3 for a conceptual overview). As the RF is trained on GPP/ET within the filtered periods, only E which is present in the training data set can bias WUE predictions. To quantify how sensitive the method is to E_{train}, experimental ET data were calculated from simulated model T data to give a consistent E_{train} for the entire time series which could not be filtered via CSWI. E_{train} levels ranged from 0% to 50% of ET, with some added noise to give some uncharacterizable variability. Calculations utilized a multiplier (e_{factor}) which was centered on the desired E_{train}, with standard deviation equal to 25% of E_{train}:

$$e_{t,\text{factor}} = N(E_{\text{train}}, \left(E_{\text{train}} \cdot 0.25\right)^2), \tag{11}$$

from which ET was calculated as

$$ET_t = \frac{T_t}{1 - e_{t,\text{factor}}}. \tag{12}$$

The resulting ET data set had a consistent fraction of E_{train} in ET which was independent of the magnitude of ET and which the random forest was unable to characterize. The range of E in the experiments encompassed the E levels in the original model training data sets, which reached values up to 37%.

These two experimental data sets were then partitioned using the exact same procedure as the original data set.

2.5. Application to Real EC Data

The TEA algorithm was used as described above to partition the real EC data from three sites: Hesse beech forest in France (FR-Hes, Granier et al., 2008), a Scots pine forest in Hyytiälä, Finland (FI-Hyy, Mammarella et al., 2009), and a spruce forest at Anchor Station Tharandt, Germany (DE-Tha, Grünwald and Bernhofer (2007)). Flux data were flagged as good or bad quality as per Papale et al. (2006), and gap filling and net ecosystem exchange partitioning were performed as per Reichstein et al. (2005). TEA estimates from real flux data can be found as part of the discussion.

3. Results

Figure 4 shows the average annual cycle of T/ET for the three sites run by each model with the same forcing data set, as well as the total T/ET values. For clarity, TEA estimated values are reported here using hyperparameter values of CSWI limit of −0.5 mm and the 75th prediction percentile. The 75th prediction percentile corresponds to the median of the 50th–100th percentiles and offers a value useful to quantify spatial and temporal patterns (see section 3.2 for sensitivity analysis of hyperparameters and section 4.2 for a discussion on their use). It can seem surprising that the TEA algorithm, which is trained on periods when E is assumed to be 0, performs so well given that the mean daily evaporation across years is often over 40% (Figure 2). This is because this mean daily evaporation includes nighttime periods (when all ET is likely from evaporation) and rainy or postrain periods, while the TEA algorithm training data set excludes all those periods (via the radiation and CSWI filters). Thus, the amount of evaporation in the training data set is much lower than seen in these two plots. The ability of TEA to extract the WUE dynamics can be seen in Figure 5, with seasonal and diurnal WUE patterns of T_{TEA}/ET matching those of T_{model}/ET, including during periods where ET/GPP shows the obvious effects of E during the wet winter periods.

3.1. Resulting Partitioning Performance

The predictive performance of the TEA algorithm applied to the model outputs across both aggregated time scales, as well as across all sites, is shown in Figure 6. The median MEF values between T_{TEA} and T_{model} are greater than 0.9 for all models across all time aggregations up to quarterly, with a slight decrease at annual aggregation. This decrease in performance at annual scale may be due to the limited variability at these time scales, as well as the limited number of years at some sites. Model bias varies between sites and particularly between models, indicating that the optimal prediction percentile for minimizing bias may vary for each

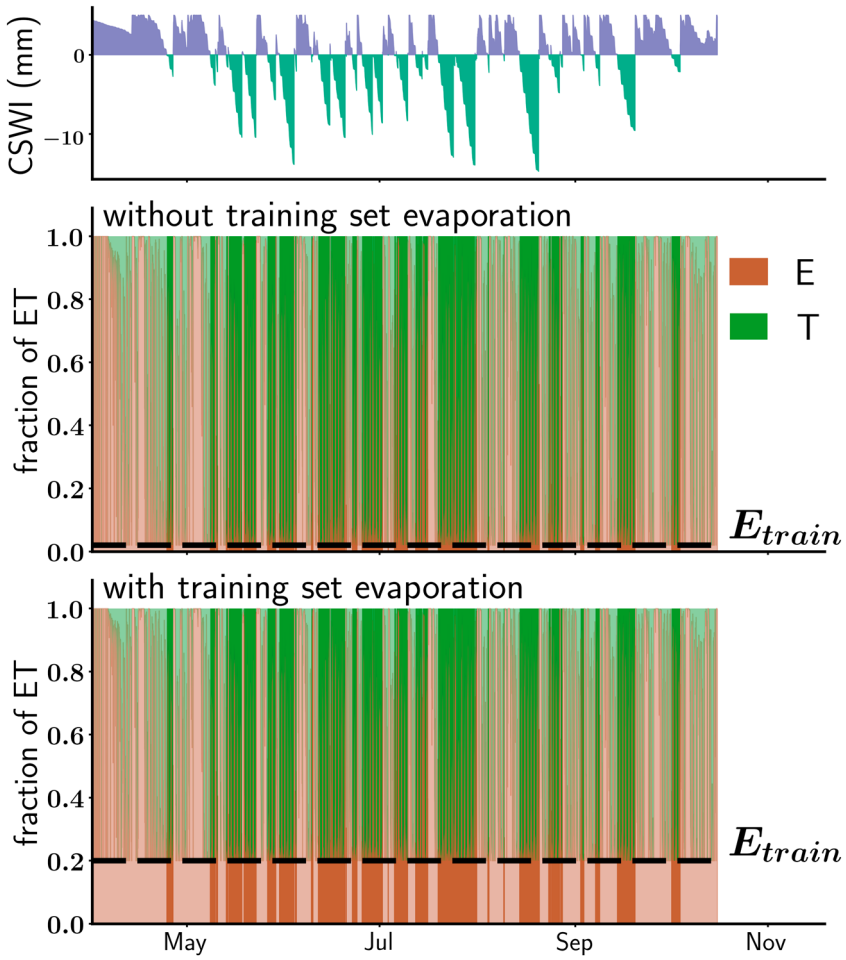

Figure 3. Conceptual diagram showing persistent evaporation in the training data set (E_{train}). In the case without persistent evaporation, many periods of the training data set contain little E_{train}, meaning the algorithm can likely find periods where $eWUE = WUE$, $(GPP/ET) = (GPP/T)$, and $ET = T$. In the case with high E_{train}, every period contains significant E, which is likely to cause a bias in WUE estimates and ultimately an overestimation of T. WUE = water use efficiency; GPP = gross primary productivity; ET = evapotranspiration.

model (see section 3.2 for analysis of prediction percentiles). The method performed well spatially, that is, across sites, with slopes between predicted and modeled T/ET varying between 1.02 and 1.12.

Using the wider range of sites afforded by JSBACH, site characteristics such as aridity index, mean annual temperature, max LAI, and plant functional type (PFT) were shown to have no significant effect on MEF or bias (see supporting information Figure S3).

3.2. Sensitivity to Hyperparameters: CSWI Limit and Prediction Percentile

The TEA algorithm provides two key hyperparameters to tune the resulting WUE and transpiration estimates: the CSWI limit which controls the amount of required accumulated ET after a rain event before data are included in the training data set, and the prediction percentile which adjusts the predictions higher or lower, that is, closer to the limits of the training data set *eWUE* values. As the CSWI limit attempts to remove *E*-contaminated data, the limit should be optimized to maximize the MEF between T_{TEA} and T_{model} while maintaining a high number of points in the training data set. As seen in Figure 7a, MEF values improve with limits below 0 mm and stabilize below a limit of −0.5 mm. Furthermore, spatial correlations (Figure 7b) show a similar improvement at a limit of −0.5 mm, followed with a sharp decline from the JSBACH runs below a limit of −2.0 mm. A CSWI limit of −0.5 mm was used for all further analysis (see section 4.2 for further discussion).

Figure 8 shows sensitivity and model performance with respect to prediction percentile. MEF for each prediction percentile was generally above 0.7 for all models and sites, with some MEF values less than 0.7 at the highest (P_{100}) and lowest (P_{50}) prediction percentiles (Figure 8a). In accordance with the hypothesis laid out in

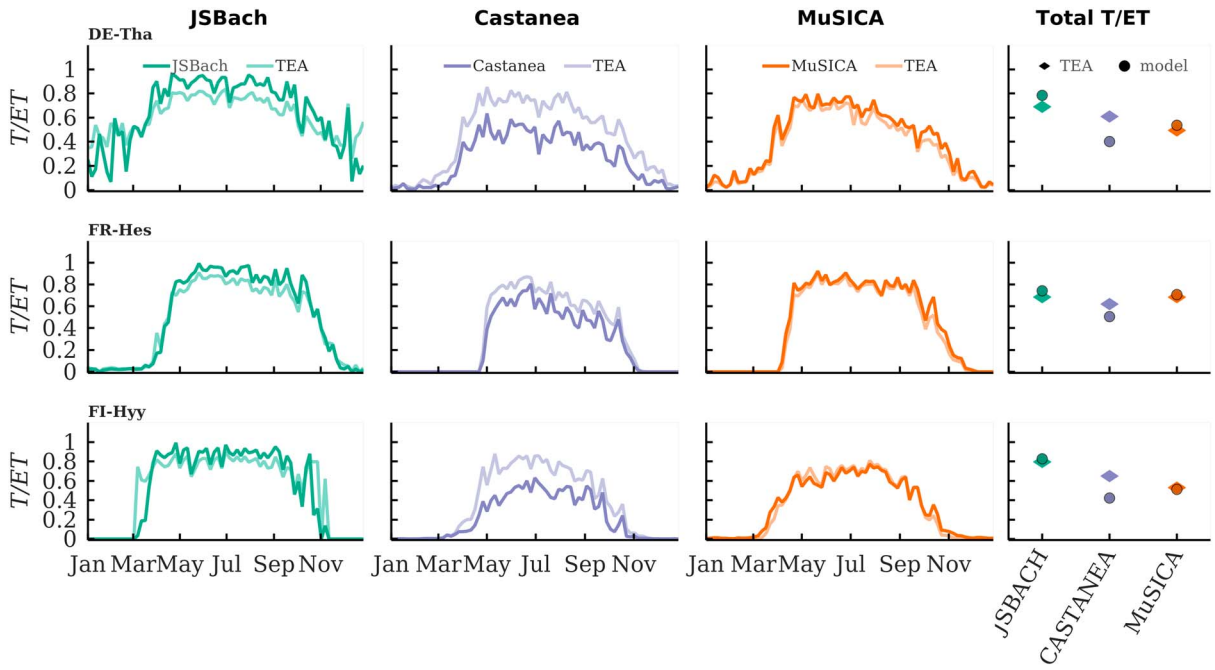

Figure 4. Average seasonal cycle of T/ET both from model output and the TEA estimate from model output, as well as total T/ET from the entire time series. TEA algorithm was run on each individual site and model independently. Modeled years were 1997–2011 for FI-Hyy and FR-Hes, and 1997–2010 for DE-Tha. Data presented correspond to daily data averaged across all years and are intended to show seasonal trends. TEA = Transpiration Estimation Algorithm; ET = evapotranspiration; FI-Hyy = Hyytiälä, Finland; FR-Hes = Hesse beech forest in France; DE-Tha = Anchor Station Tharandt, Germany.

Zhou et al. (2016), prediction percentiles closer to the limit which maximizes the GPP:ET ratio should be associated with periods where E contamination is minimized. In contrast, looking at the relative bias between T_{TEA} and T_{model} (Figure 8b), we found that the prediction percentile which minimizes bias varied depending on the model, with optimal prediction percentiles to minimize bias for CASTANEA being around P_{95}, compared to P_{70} for MuSICA and P_{60} for JSBACH. This difference in optimal prediction percentiles may be due to the differences in the inherent residual E predicted by the models (see Figure 2), with CASTANEA having the highest level of E throughout the growing season (and highest bias-minimized prediction percentile) and JSBACH having the lowest. Supporting information Figure S4, which shows the relationship of E/ET in prediction points at various prediction percentiles, further indicates that indeed the E component from TEA predictions is minimized at different percentiles for the three models.

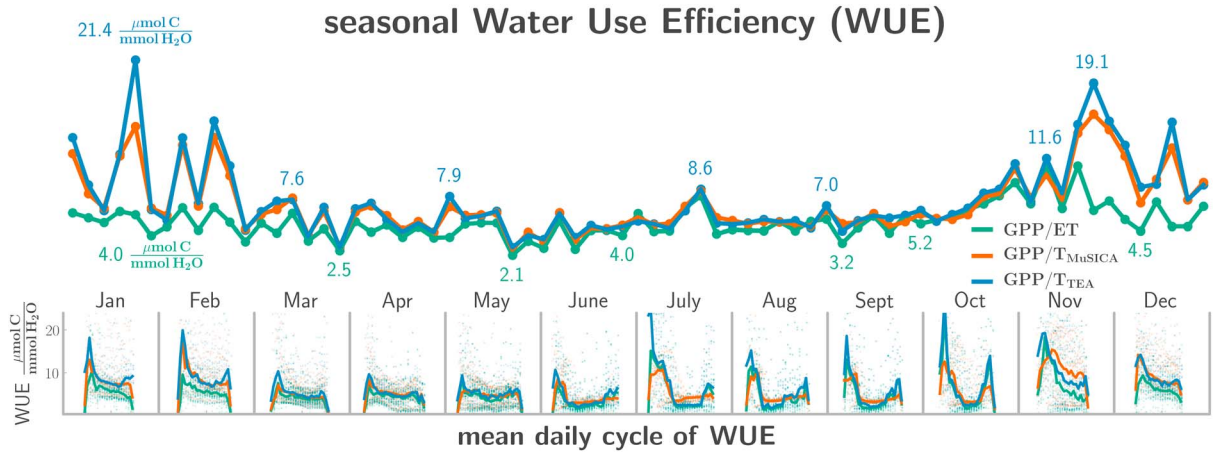

Figure 5. The TEA algorithm captures mean seasonal and daily cycles of WUE from MuSICA output forced with data from Yatir Forest from 2001. Mean daily cycles are based on half hourly data for each month, whereas the seasonal cycles are an average from 5 days. TEA = Transpiration Estimation Algorithm; WUE = water use efficiency;

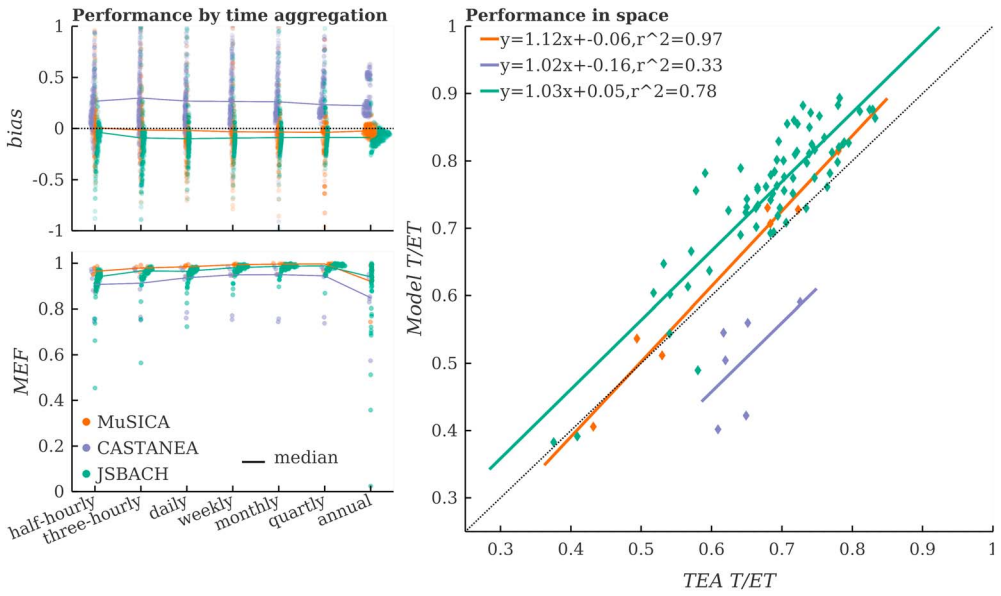

Figure 6. Overall model performance in space (right) and time (left). All models show high correlation across time scales, with some degradation at annual scale. In space (i.e., across sites), the TEA algorithm shows the highest agreement with the MuSICA model runs ($r^2 \approx 0.97$), and root-mean-square error for JSBACH, CASTANEA, and MuSICA was 8.4%, 15.3%, and 3.2%, respectively.

3.2.1. Sensitivity to Training Set Evaporation and Noise

As seen in Figure 8d, the TEA algorithm shows a response in bias to E_{train} (see section 2.4 for experimental outline), with the slope between bias and E_{train} being between 1 and 2 for the 50th–90th prediction percentiles. These slopes correspond to a worst case scenario, representing a situation where a site would have a constant E component, for example, a site where E never goes below 15% of ET at any time. So if a site is estimated to have at least 15% E_{train} at every half hour, the transpiration rates may be 22% overestimated using the 75th prediction percentile. An overestimation of 150% of E_{train} is consistent with the CASTANEA model runs, which across sites has a mean E_{train} of 20% and a mean total bias of 32%, which translates to a mean overestimation of total T/ET of 14%. For context the percentage of training data set half hours with less than 15% E/ET was on average 67%, 33%, and 95% for MuSICA, JSBACH, and CASTANEA, respectively, with the lowest percentage, 6%, for the Hyytiälä Forest simulation from CASTANEA.

Though the highest prediction percentiles show the lowest sensitivity to E_{train} and could thus mitigate this bias, high prediction percentiles also show large sensitivity to noise (Figure 8c; see section 2.4 for experimental outline), indicating that directly using prediction percentiles above P_{95} is not suitable. Prediction percentiles below P_{90} show less sensitivity to noise, with slopes between MEF and the noise-to-signal ratio (inverse of signal-to-noise ratio used to simplify sign convention) generally being between −0.1 and 0 for the majority of sites. To put a slope of −0.1 into context, if a site had an MEF of 0.9 and a noise-to-signal ratio of 1:10, the same site would have an MEF of 0.83 if noise was then added making the noise-to-signal ratio 1:2 (see section 4.2 for further discussion).

4. Discussion

4.1. Broadly Applicable WUE and *T* Estimates

The validation experiment presented here indicates that while ET is composed of two signals (E and T), by pairing the ET signal with GPP, TEA is able to extract the WUE dynamics and thus the biologically controlled T signal. Figures 5 and 6 demonstrate that TEA-estimated WUE captures variability in transpiration from subdaily to interannual scales and between sites, particularly when comparing only simulations from an individual model. The relatively fine temporal resolution of TEA provides the possibility of exploring the dynamics of carbon:water interactions such as seasonal and diurnal cycles. In general, the method as outlined here can be directly applied to real EC data with minimal alteration, allowing for potential global application, with the limitations and cautionary remarks described in the following sections. As a demonstration of TEA using real data, Figure 9 shows a comparison of modeled T/ET at three sites compared to the TEA-estimated T/ET

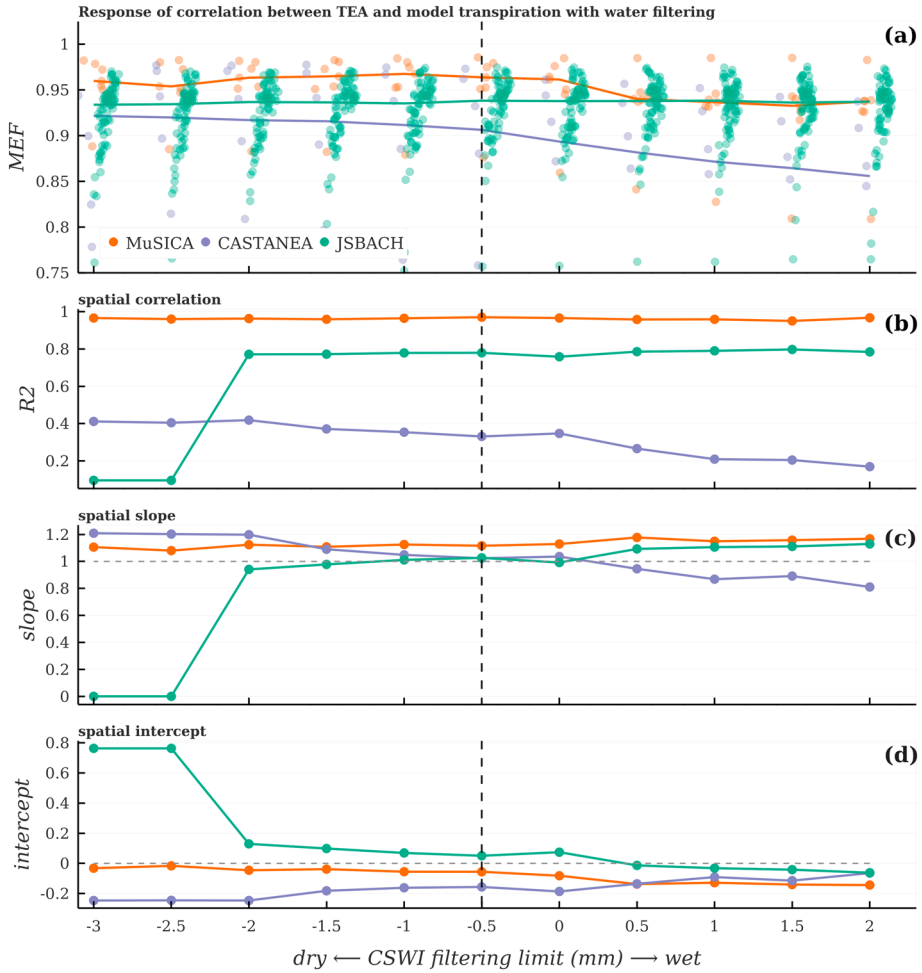

Figure 7. Sensitivity of TEA performance to CSWI limit hyperparameter. Though low CSWI limits help reduce the impact of E_{train}, strict filters also decrease the number of points available in the training data set, which can exclude some wet sites entirely. A filter of −0.5 shows good MEF between T_{TEA} and T_{model} across sites (a), with little improvement using stricter filtering. The negative effect of strict filtering can be seen in the spatial correlation (b) of the JSBACH models, which significantly decreases at limits less than −2, as well as in the spatial slopes (c) and spatial intercepts (d). TEA = Transpiration Estimation Algorithm; CSWI = conservative surface water index; MEF = modeling efficiency.

using actual EC data. *T* estimates from TEA using real EC data fall between the process model *T* estimates, all while requiring no parameterizations nor having any assumptions on the underlying biological processes. Importantly, TEA does not rely on the model data in any way, as model runs were only used as a validation experiment; thus, TEA is purely data driven and represents the statistical prediction of WUE and *T* based on input data of GPP, ET, and meteorological data. To see the value that these widely applicable methods provide, one needs to look no further than the partitioning of carbon EC fluxes, which have provided a wealth of information despite having known limitations (Reichstein et al., 2012). Combining such widely applicable methods, such as TEA, with the unconstrained processed-based models and the sparse independent *T* measurements provides a multifaceted and complementary view of ecosystem *T*.

Though methods for estimating *T* from independent measurements such as upscaled sap flux methods have existed for decades, there are still relatively few published values that coincide with EC sites. One set of estimates at Hesse forest gives a seasonal *T*/*ET* (2 May to 27 October) from sap flux upscaling of 0.72, 0.82, and 0.86 for 1997, 1998, and 1999, respectively (Vainshtein, 2010), which is in relative agreement with the TEA estimates from EC data at Hesse forest being 0.75, 0.82, and 0.73. Though an in-depth comparison to independent *T* measurements is beyond the scope of this analysis, initiatives such as SAPFLUXNET, which aims to aggregate sap flux data sets from around the world (Poyatos et al., 2016), as well as aggregations of isotope-based mea-

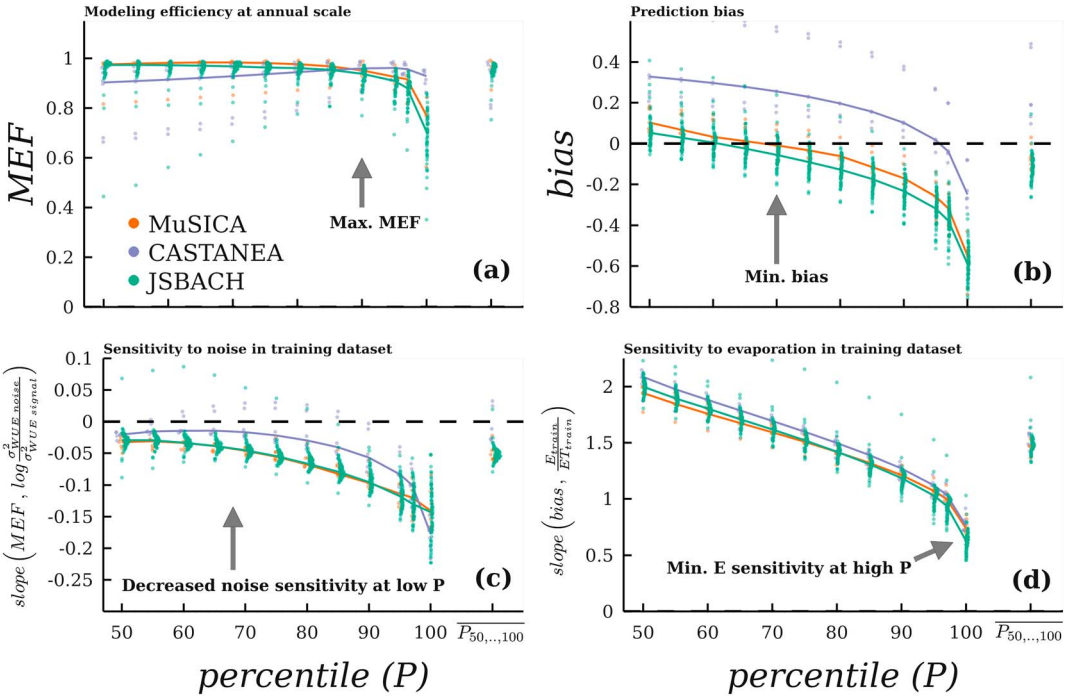

Figure 8. Sensitivity analysis of prediction percentile (P) used in TEA. MEF (a), while generally stable, shows the highest values at relatively high percentiles ($P \approx 80-90$), whereas bias (b) is minimized at around P_{70} for MuSICA and JSBACH. Similarly, percentiles near the maximum (P_{100}) shows the lowest sensitivity to evaporation content in the training data set (d), yet these high percentiles are also very sensitive to noise in the training data set (c). Given that residual evaporation will likely be present in the training data set, causing predictions to be overestimated, percentiles below the median (P_{50}) can be discounted. As such, by treating each percentile above the median ($P \geq 50$) as an equally likely estimate, we can calculate the mean of $T_{P50}, \ldots, T_{P100}$, the results of which are seen as the far right points in each plot or the median which corresponds to P_{75}. TEA = Transpiration Estimation Algorithm; MEF = modeling efficiency.

surements and the continued aggregation of EC data set will help constrain ecosystem transpiration estimates within the next few years.

4.2. Sensitivities and Limitations

As seen in Figure 7, the CSWI limit hyperparameter should always be less than 0. However, prediction performance did not improve at increasingly negative values and may actually deteriorate performance due to declines in sample size. This lack of improvement indicates that CSWI, though likely an improvement to time-based methods, does not do well at indicating degrees of moisture levels past simple wet and dry. Therefore, a CSWI value of around −0.5 or −1.0 mm is warranted, as it creates the largest sample size while still being below 0. It is possible that the TEA algorithm could be improved with a filter that better minimizes the amount of evaporation left in the training data set, E_{train}, such as using surface soil moisture data. As such, the TEA algorithm would likely benefit from site-specific information on water status, both as a means to filter the training data set and as a predictor variable. Additionally, it should be noted that as the filtering step removes all periods during and immediately after rain, these periods will not be represented in the training data set and therefore any response of WUE during these rainy periods will not be captured. However, as none of the filters was based on humidity levels, periods of high relative humidity are included in the training data set, so both stressed and unstressed conditions will be included in the training data set. As rain specifically should not have a dramatic influence on WUE, indeed, we did not observe any error increase in the validation data set during wet periods.

Overall, the method tends to be more precise than accurate, that is, it robustly produces precise patterns but with a propensity for systematic overestimation or underestimation. In particular, the method is sensitive to E_{train} such as is the case with the CASTANEA model runs, producing an overestimation of transpiration while still capturing the temporal patterns across time scales. The CASTANEA simulations here provide an important test as to how E can impact the TEA estimates, as the simulations have not only relatively high E/ET throughout the year, but, due to the fact that soil moisture levels are updated daily, E/ET is also relatively consistent

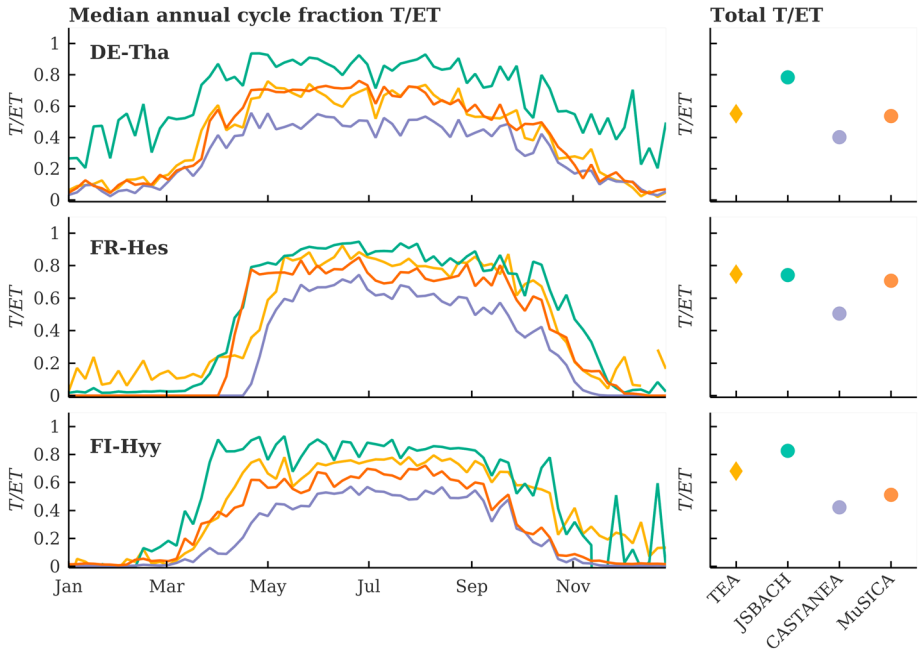

Figure 9. Comparison of mean seasonal cycle of T/ET (5-day aggregation) results from model simulations (JSBACH, CASTANEA, and MuSCIA) and TEA algorithm partitioning of original EC data (ET, GPP, and meteorological variables). Modeled years were 2001–2006 for FI-Hyy and FR-Hes, and 1998–2003 for DE-Tha. Seasonal cycles are an average of 5 days. TEA = Transpiration Estimation Algorithm; EC = eddy covariance; ET = evapotranspiration; GPP = gross primary productivity. FI-Hyy = Hyytiälä, Finland; FR-Hes = Hesse beech forest in France; DE-Tha = Anchor Station Tharandt, Germany.

throughout the day. Therefore, the training data set from CASTANEA simulations is always contaminated by E, in contrast to MuSICA which can have daily E but still have periods when T dominates within that day, and JSBACH which has very low E throughout the growing season (see Figure S4). That being said, it is possible to predict an accurate average T/ET using a higher prediction percentile even when E is always present as long as there is variability in WUE (see Figure S5), though the highest WUE values will be underestimated.

Given that the optimal prediction percentile for minimizing bias differed among the three models, all prediction percentiles above the 50th can be considered equally likely predictors, with P_{50} representing the case with no E_{train} and a constant WUE, and P_{100} representing the maximum *eWUE*. The result is a distribution of estimates for WUE and T, which can be translated into an average and uncertainty. As this distribution tends to be rather skewed, the median of this distribution (or P_{75}) is likely a more robust estimator. While the lower bound of the distribution is well bounded, the maximum (P_{100}) case could still systematically underestimate WUE if E_{train} is significantly higher than 0. In contrast, P_{100} could also grossly overestimate WUE as it can correspond to conditions which are not at all representative, for example, conditions of high humidity when WUE tends to infinity, which can be further complicated with the added effects of noise.

Ecosystems with sparse vegetation coverage are likely most at risk of having high E_{train} levels, and therefore overestimation, as the canopy is potentially not the key control on ET. The risk of overestimation is especially high at wetland sites with exposed open water. Therefore, site-specific estimations are warranted to determine if TEA estimations would benefit from hyperparameter adjustments such as using a higher prediction percentile, improved training set filtering, or other improvements based on site knowledge (e.g., filtering periods during irrigation). Another important consideration when applying the method to actual data is the existence of noise which is not present in the synthetic validation data sets. This is particularly pertinent due to very large or small values of WUE (a ratio) during mornings and evenings when the fluxes are low. In this case, a filter for small values of either ET, GPP, or Rg will likely be warranted, even though the method was shown to be relatively insensitive to noise for most prediction percentiles (Figure 7c). Given the considerations outlined above, a general framework for implementing TEA for EC data would be to use a CSWI limit of −0.5 mm and the 75th percentile for prediction, which corresponds to the median of predictions from the 50th to 100th percentiles.

5. Conclusion

In its current state, ecosystem transpiration is far a more concrete physiological concept than it is actually quantifiable, as one can isolate transpiration in relatively controlled leaf- or plant-scale experiments in contrast to the difficulties of isolating soil and interception evaporation components from the transpiration of each needle and leaf at a field site. However, by utilizing the carbon cycle, transpiration dynamics can be extracted from the overall evapotranspiration signal. As such, the TEA algorithm is a novel evapotranspiration partitioning method designed for EC data sets which is able to capture dynamics in WUE and transpiration across spatial and temporal scales. The method is the first such evapotranspiration partitioning approach to attempt such an extensive validation exercise, utilizing a synthetic experiment of process model output, which demonstrates the ability of the method to replicate the carbon:water relationship across three model frameworks. Furthermore, we outline the biases and uncertainties of the approach with particular respect to effect of persistent evaporation fluxes, with the prospect that by thoroughly scrutinizing and testing the limits of TEA we can open the door to wide scale application.

Acknowledgments

The authors would like to thank those who gave thoughtful input and feedback on the manuscript such as Sven Boese, Sujan Koirala, and Gustau Camps-Valls. Jacob Nelson would like to thank Bernhard Ahrens for initial editing work and Tiana Hammer for personal and scientific support. The authors would like to thank the broader eddy covariance community, whose work and expertise are fundamental for this analysis. This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization was carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux, and AsiaFlux offices. FLUXNET data can be found at http://fluxnet.fluxdata.org/. Data and code used in this analysis can be found in the associated repository (Nelson, 2018).

References

Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In J. Biggins (Ed.), *Progress in photosynthesis research* (pp. 221–224). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-017-0519-6_48

Beer, C., Ciais, P., Reichstein, M., Baldocchi, D., Law, B. E., Papale, D., et al. (2009). Temporal and among-site variability of inherent water use efficiency at the ecosystem level: Variability of inherent WUE. *Global Biogeochemical Cycles*, *23*, GB2018. https://doi.org/10.1029/2008GB003233

Boese, S., Jung, M., Carvalhais, N., & Reichstein, M. (2017). The importance of radiation for semi-empirical water-use efficiency models. *Biogeosciences Discussions*, *14*(12). https://doi.org/10.5194/bg-2016-524

Böttcher, K., Markkanen, T., Thum, T., Aalto, T., Aurela, M., Reick, C., et al. (2016). Evaluating biosphere model estimates of the start of the vegetation active season in boreal forests by satellite observations. *Remote Sensing*, *8*(7), 580. https://doi.org/10.3390/rs8070580

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Coenders-Gerrits, A. M. J., van der Ent, R. J., Bogaard, T. A., Wang-Erlandsson, L., Hrachowitz, M., & Savenije, H. H. G. (2014). Uncertainties in transpiration estimates. *Nature*, *506*(7487), E1–E2.

Delpierre, N., Dufrêne, E., Soudani, K., Ulrich, E., Cecchini, S., Boé, J., & François, C. (2009). Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France. *Agricultural and Forest Meteorology*, *149*(6-7), 938–948. https://doi.org/10.1016/j.agrformet.2008.11.014

Delpierre, N., Soudani, K., François, C., Le Maire, G., Bernhofer, C., Kutsch, W., et al. (2012). Quantifying the influence of climate and biological drivers on the interannual variability of carbon exchanges in European forests through process-based modelling. *Agricultural and Forest Meteorology*, *154-155*, 99–112. https://doi.org/10.1016/j.agrformet.2011.10.010

Dufrêne, E., Davi, H., François, C., Le Maire, G., Le Dantec, V., & Granier, A. (2005). Modelling carbon and water cycles in a beech forest: Part I: Model description and uncertainty analysis on modelled NEE. *Ecological Modelling*, *185*(2), 407–436.

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*, 2618–2626. https://doi.org/10.1002/2016WR020175

Gimeno, T. E., Ogée, J., Royles, J., Gibon, Y., West, J. B., Burlett, R., et al. (2017). Bryophyte gas-exchange dynamics along varying hydration status reveal a significant carbonyl sulphide (COS) sink in the dark and COS source in the light. *New Phytologist*, *215*(3), 965–976. https://doi.org/10.1111/nph.14584

Good, S. P., Noone, D., & Bowen, G. (2015). Hydrologic connectivity constrains partitioning of global terrestrial water fluxes. *Science*, *349*(6244), 175–177.

Good, S. P., Soderberg, K., Guan, K., King, E. G., Scanlon, T. M., & Caylor, K. K. (2014). δ^2H isotopic flux partitioning of evapotranspiration over a grass field following a water pulse and subsequent dry down. *Water Resources Research*, *50*, 1410–1432. https://doi.org/10.1002/2013WR014333

Granier, A. (1987). Evaluation of transpiration in a Douglas-fir stand by means of sap flow measurements. *Tree Physiology*, *3*(4), 309–320. https://doi.org/10.1093/treephys/3.4.309

Granier, A., Bréda, N., Longdoz, B., Gross, P., & Ngao, J. (2008). Ten years of fluxes and stand growth in a young beech forest at Hesse, north-eastern France. *Annals of Forest Science*, *65*(7), 704–704. https://doi.org/10.1051/forest:2008052

Grünwald, T., & Bernhofer, C. (2007). A decade of carbon, water and energy flux measurements of an old spruce forest at the Anchor Station Tharandt. *Tellus B*, *59*(3), 387–396. https://doi.org/10.1111/j.1600-0889.2007.00259.x

Hollinger, D. Y., & Richardson, A. D. (2005). Uncertainty in eddy covariance measurements and its application to physiological models. *Tree Physiology*, *25*(7), 873–885.

Ishizaki, Y., Yokohata, T., Emori, S., Shiogama, H., Takahashi, K., Hanasaki, N., et al. (2014). Validation of a pattern scaling approach for determining the maximum available renewable freshwater resource. *Journal of Hydrometeorology*, *15*(1), 505–516. https://doi.org/10.1175/JHM-D-12-0114.1

Jasechko, S., Birks, S. J., Gleeson, T., Wada, Y., Fawcett, P. J., Sharp, Z. D., et al. (2014). The pronounced seasonality of global groundwater recharge. *Water Resources Research*, *50*, 8845–8867. https://doi.org/10.1002/2014WR015809

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., & Fawcett, P. J. (2013). Terrestrial water fluxes dominated by transpiration. *Nature*, *496*(7445), 347–350. https://doi.org/10.1038/nature11983

Jung, M., Reichstein, M., & Bondeau, A. (2009). Towards global empirical upscaling of FLUXNET eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, *6*(10), 2001–2013.

Katul, G. G., Palmroth, S., & Oren, R. (2009). Leaf stomatal responses to vapour pressure deficit under current and CO_2-enriched atmosphere explained by the economics of gas exchange. *Plant, Cell & Environment*, *32*(8), 968–979. https://doi.org/10.1111/j.1365-3040.2009.01977.x

Knauer, J., Werner, C., & Zaehle, S. (2015). Evaluating stomatal models and their atmospheric drought response in a land surface scheme: A multibiome analysis: Multibiome stomatal model evaluation. *Journal of Geophysical Research: Biogeosciences*, *120*, 1894–1911. https://doi.org/10.1002/2015JG003114

Kool, D., Agam, N., Lazarovitch, N., Heitman, J., Sauer, T., & 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

Leuning, R. (1995). A critical appraisal of a combined stomatal-photosynthesis model for C3 plants. *Plant, Cell & Environment*, *18*(4), 339–355. https://doi.org/10.1111/j.1365-3040.1995.tb00370.x

Mammarella, I., Launiainen, S., Gronholm, T., Keronen, P., Pumpanen, J., Rannik, U., & Vesala, T. (2009). Relative humidity effect on the high-frequency attenuation of water vapor flux measured by a closed-path eddy covariance system. *Journal of Atmospheric and Oceanic Technology*, *26*(9), 1856–1866. https://doi.org/10.1175/2009JTECHA1179.1

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., et al. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance: Reconciling optimal and empirical stomatal models. *Global Change Biology*, *17*(6), 2134–2144. https://doi.org/10.1111/j.1365-2486.2010.02375.x

Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, *7*(Jun), 983–999.

Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nelson, J. A. (2017). Jnelson18/Fluxnettools: Initial release. https://doi.org/10.5281/zenodo.1010483

Nelson, J. A. (2018). Jnelson18/TranspirationEstimationAlgorithm: Release for publication review. https://doi.org/10.5281/zenodo.1305058

Nelson, J. A., Carvalhais, N., Migliavacca, M., Reichstein, M., & Jung, M. (2017). Water stress induced breakdown of carbon-water relations: Indicators from diurnal fluxnet patterns. *Biogeosciences*, *15*, 2433–2447. https://doi.org/10.5194/bg-15-2433-2018

Ogée, J., Brunet, Y., Loustau, D., Berbigier, P., & Delzon, S. (2003). MuSICA, a CO_2, water and energy multilayer, multileaf pine forest model: Evaluation from hourly to yearly time scales and sensitivity analysis. *Global Change Biology*, *9*(5), 697–717.

Oishi, A. C., Oren, R., & Stoy, P. C. (2008). Estimating components of forest evapotranspiration: A footprint approach for scaling sap flux measurements. *Agricultural and Forest Meteorology*, *148*(11), 1719–1732. https://doi.org/10.1016/j.agrformet.2008.06.013

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., et al. (2006). Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: Algorithms and uncertainty estimation. *Biogeosciences*, *3*(4), 571–583.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*(Oct), 2825–2830.

Potier, E., Ogée, J., Jouanguy, J., Lamaud, E., Stella, P., Personne, E., et al. (2015). Multilayer modelling of ozone fluxes on winter wheat reveals large deposition on wet senescing leaves. *Agricultural and Forest Meteorology*, *211-212*, 58–71. https://doi.org/10.1016/j.agrformet.2015.05.006

Poyatos, R., Granda, V., Molowny-Horas, R., Mencuccini, M., Steppe, K., & Martínez-Vilalta, J. (2016). SAPFLUXNET: Towards a global database of sap flow measurements. *Tree Physiology*, *36*(12), 1449–1455. https://doi.org/10.1093/treephys/tpw110

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., et al. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm. *Global Change Biology*, *11*(9), 1424–1439. https://doi.org/10.1111/j.1365-2486.2005.001002.x

Reichstein, M., Stoy, P. C., Desai, A. R., Lasslop, G., & Richardson, A. D. (2012). Partitioning of net fluxes. In M. Aubinet, T. Vesala, & D. Papale (Eds.), *Eddy covariance* (pp. 263–289). Dordrecht, Netherlands: Springer Netherlands. https://doi.org/10.1007/978-94-007-2351-1_9

Reick, C. H., Raddatz, T., Brovkin, V., & Gayler, V. (2013). Representation of natural and anthropogenic land cover change in MPI-ESM: Land cover in MPI-ESM. *Journal of Advances in Modeling Earth Systems*, *5*, 459–482. https://doi.org/10.1002/jame.20022

Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., Caemmerer, S., Dietze, M. C., et al. (2017). A roadmap for improving the representation of photosynthesis in Earth system models. *New Phytologist*, *213*(1), 22–42.

Rutter, A., Kershaw, K., Robins, P., & Morton, A. (1971). A predictive model of rainfall interception in forests, 1. Derivation of the model from observations in a plantation of Corsican pine. *Agricultural Meteorology*, *9*, 367–384. https://doi.org/10.1016/0002-1571(71)90034-3

Sandoval-Soto, L., Stanimirov, M., von Hobe, M., Schmitt, V., Valdes, J., Wild, A., & Kesselmeier, J. (2005). Global uptake of carbonyl sulfide (COS) by terrestrial vegetation: Estimates corrected by deposition velocities normalized to the uptake of carbon dioxide (CO_2). *Biogeosciences*, *2*(2), 125–132. https://doi.org/10.5194/bg-2-125-2005

Scott, R. L., & Biederman, J. A. (2017). Partitioning evapotranspiration using long-term carbon dioxide and water vapor fluxes: New approach to ET partitioning. *Geophysical Research Letters*, *44*, 6833–6840. https://doi.org/10.1002/2017GL074324

Tuzet, A., Perrier, A., & Leuning, R. (2003). A coupled model of stomatal conductance, photosynthesis and transpiration. *Plant, Cell and Environment*, *26*(7), 1097–1116. https://doi.org/10.1046/j.1365-3040.2003.01035.x

Wehr, R., Commane, R., Munger, J. W., McManus, J. B., Nelson, D. D., Zahniser, M. S., et al. (2017). Dynamics of canopy stomatal conductance, transpiration, and evaporation in a temperate deciduous forest, validated by carbonyl sulfide uptake. *Biogeosciences*, *14*(2), 389–401. https://doi.org/10.5194/bg-14-389-2017

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*, 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. (2017). Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles. *Biogeosciences Discussions*, *15*(12). https://doi.org/10.5194/bg-2017-427

Wilkinson, S., Ogee, J., Domec, J.-C., Rayment, M., & Wingate, L. (2015). Biophysical modelling of intra-ring variations in tracheid features and wood density of Pinus pinaster trees exposed to seasonal droughts. *Tree Physiology*, *35*(3), 305–318. https://doi.org/10.1093/treephys/tpv010

Wohlfahrt, G. (2017). Bi-directional COS exchange in bryophytes challenges its use as a tracer for gross primary productivity. *New Phytologist*, *215*(3), 923–925.

Zhou, S., Yu, B., Huang, Y., & Wang, G. (2014). The effect of vapor pressure deficit on water use efficiency at the subdaily time scale: Underlying water use efficiency. *Geophysical Research Letters*, *41*, 5005–5013. https://doi.org/10.1002/2014GL060741

Zhou, S., Yu, B., Zhang, Y., Huang, Y., & Wang, G. (2016). Partitioning evapotranspiration based on the concept of underlying water use efficiency: ET partitioning. *Water Resources Research*, *52*, 1160–1175. https://doi.org/10.1002/2015WR017766