**Plant access to belowground moisture allows sustained evapotranspiration during drought**

Francesco Giardina1,2

Corresponding author: fgiardina@ethz.ch /

ORCID: 0000-0002-8359-809X

Pierre Gentine3,4

pg2328@columbia.edu

ORCID: 0000-0002-0845-8345

Alexandra G. Konings5

konings@stanford.edu

ORCID: 0000-0002-2810-1722

Benjamin D. Stocker1,2

benjamin.stocker@giub.unibe.ch

ORCID: 0000-0003-2697-9096

1Department of Environmental Systems Science, ETH Zürich, CH-8092 Zürich, Switzerland

2Swiss Federal Institute for Forest, Snow and Landscape Research WSL, CH-8903 Birmensdorf, Switzerland

3Department of Earth and Environmental Engineering, Columbia University, New York, New York 10027, USA

4Center for Learning the Earth with Artificial intelligence and Physics (LEAP), Columbia University, New York, New York 10027, USA

5Department of Earth System Science, Stanford University, Stanford, California 94305, USA

konings@stanford.edu

**Summary**

* Accounting for subsurface water limitation is key to determining vegetation sensitivity to drought. There is evidence that water sources deeper than soil moisture exert a control on evapotranspiration (ET). However, these effects are difficult to quantify.
* Here, we train deep neural networks using flux measurements to study the impact of belowground water limitation on ET in progressing droughts. We determine a belowground water stress factor (fET) that isolates ET reductions from effects by atmospheric aridity and other co-varying drivers. We regress fET against the cumulative water deficit (CWD), which reveals the control of belowground moisture availability.
* We find rapid declines of fET down to 10% of its water-unlimited rate in savannahs and grasslands, and negligible fET reductions in forests, in spite of substantial water deficits. In both cases, ET is almost never completely shut off during droughts. The response across sites is regulated by climate type.
* Access to belowground water reservoirs could explain the different behaviors observed across sites. This variety of responses is not captured by a state-of-the-art land surface model, likely reflecting simplifications in its representation of belowground water storage. Our findings illuminate the understanding of vegetation sensitivity to drought and provide insight into modelling water stress effects.

**Keywords**: climate change, data science, deep learning, groundwater, rock moisture, root zone water storage capacity, soil moisture, vapor pressure deficit (VPD)

# Introduction

Water availability controls evapotranspiration (ET) over most of the vegetated land surface (Ahlström et al., 2015; Beer et al., 2010; Schwalm et al., 2010; Seneviratne et al., 2010). Under prolonged rain-free periods, plants can depend on belowground water reservoirs (Hahm et al., 2019; McCormick et al., 2021; Milly, 1994) or rely on very resistant xylem (Mackay et al., 2015; Plaut et al., 2012) to sustain their function and growth. As droughts may become more severe and frequent in the future (Seneviratne et al., 2021), it is crucial to understand how plant function links ecosystem evapotranspiration (ET) and subsurface water limitation. Here, we identify 'droughts' by the effects of water balance deficits on vegetation productivity (Seneviratne et al., 2021). Quantifying belowground water limitation effects on ET is challenged by the heterogeneity of the soil and bedrock structure, both vertically along the rooting zone and laterally across the landscape (Dralle et al., 2020; Gao et al., 2014; Thompson et al., 2011).

The root zone water storage capacity (S0) defines the maximum amount of soil moisture that plants can access for transpiration (and thus, the amount of soil moisture that affects ET). A large S0 has been linked to a higher plant resistance to drought, sustained ET, and persistent vegetation cover during rain-free periods (Gao et al., 2014; Teuling et al., 2006). Recent studies have identified locally important contributions to S0 by water reservoirs below the soil layer, such as in fractured bedrock (Dawson et al., 2020; McCormick et al., 2021; Rempe & Dietrich, 2018; Stocker et al., 2021) or groundwater (Fan et al., 2017; Hain et al., 2015; Thompson et al., 2011). In settings with pronounced topographical gradients and local subsurface water flow convergence, or in the proximity of water bodies, water-saturated zones can be within the reach of roots (Fan et al., 2013, 2017). This implies large variations in plant access to belowground water stores at multiple scales, as also suggested by inverse modelling (Fan et al., 2017; Hain et al., 2015). However, direct observations of the contributions of subsurface moisture to land-atmosphere exchanges are sparse and limited to a small number of individual field sites (Rempe & Dietrich, 2018).

The role of plant access to belowground water stores and its importance for ecosystem fluxes across vegetation types and climate zones has remained poorly understood. This is partly due to challenges in separating partial effects on ecosystem fluxes from multiple covarying drivers. In particular, soil moisture (which is correlated to subsurface moisture (Qiu et al., 2016; Salvucci & Entekhabi, 1994)) and vapor pressure deficit (VPD) jointly affect ET, but tend to covary (Giardina et al., 2018; Liu et al., 2020; Novick et al., 2016; Zhou et al., 2019). This limits what can be inferred from flux measurements to characterize the role of belowground water stores across biomes and climates, and whether generalizations can be found to inform models across a wide range of conditions.

The complexity of subsurface water storage contrasts with its typical representation in land surface models (LSMs) (Fan et al., 2019). With a few exceptions, LSMs represent the land surface as a flat surface (Fan et al., 2019; Prentice et al., 2015), assume free drainage from the bottom soil layer which prevents the formation of water-saturated zones (Liang et al., 1994; Schlemmer et al., 2018), and do not account for the variety of bedrock lithology and its role as a moisture storage component (McCormick et al., 2021). Spatial variations of S0 are typically represented in models based on variations in soil type and in plant rooting depth that are assumed fixed and assigned to plant functional types (PFTs) (Drewniak, 2019; Tumber-Dávila et al., 2022). These simplifications affect the accuracy at predicting water limitation effects on ET, particularly under drought conditions (Green et al., 2017; Kennedy et al., 2019; K. Zhang et al., 2016). Taken together, this highlights the need to develop observational benchmarks for the role of belowground moisture limitation, separated from atmospheric aridity and other co-varying drivers (Giardina et al., 2018; Novick et al., 2016; Zhou et al., 2019), for reliable ET modelling under a wide range of water limitation levels.

Here, using site-specific deep neural network (DNN) models, we estimate the fractional reduction in ET (fET) at eddy-covariance sites compared to its potential rate. We use 88,000 site-days of ecosystem-scale ET measurements, complemented by meteorological measurements, multiple soil moisture datasets and a remotely sensed greenness index. Our method isolates the control of belowground water availability on ET, factoring out effects of VPD and other drivers, i.e., vegetation greenness, air temperature and net radiation (Rn). We evaluate fET reductions due to belowground water availability by regressing it against the cumulative water deficit (CWD, the cumulative difference between observed ET and precipitation). In contrast to evaluating fET against surface soil moisture, assessment against CWD reveals patterns of water stress effects below the relatively shallow soil layers.

# Description

## Overview

We started by estimating potential and actual evapotranspiration (PET and ET, respectively) across a large set of sites. PET is defined here as equal to ET in the absence of belowground water limitations. For each site we defined two separate deep neural network models (DNNPET and DNNET, respectively). The key difference between the two models is that DNNPET was trained using data from days with high soil moisture only, whereas DNNET was trained using all available data (see section 2.2). We defined a normalized measure of belowground moisture effects on ET (referred to as 'fET') by dividing the neural network estimate of ET (ETNN) by the neural network estimate of PET (PETNN). We then analyzed how fET evolves with increasing water stress by regressing it against the observed cumulative water deficit (CWD), a measure of whole-column water availability. Since the CWD is the integration of the observed water balance over time (see section 2.5), it constitutes a proxy for the depletion of total belowground moisture. By definition, the CWD is independent from assumptions regarding soil depth or the total water storage capacity. It therefore implicitly includes contributions from both soil and subsurface water storage. We then grouped sites based on their fET-CWD relationship and analyzed how the site groups vary with soil texture, vegetation classes, aridity index, topographical context of the site, mean annual precipitation and mean annual temperature within each group.

## Estimating potential ET

Our approach was based on a published method that separates soil moisture effects on light-use efficiency using FLUXNET2015 data (Stocker et al., 2018). Here, ETNN was estimated using observational ET (ETobs) as the target variable and soil moisture, Rn, VPD, air temperature (T) and the enhanced vegetation index (EVI) as predictors:

PETNN was estimated using ETobs as target variable and again Rn, VPD, T and EVI as predictors, but no longer considering soil moisture:

The choice of predictors was limited to a small number representing known environmental controls on ET and PET (Maes et al., 2019; Miralles et al., 2019). PETNN was derived by training the model with data from days when soil moisture was relatively high (i.e., above a site-specific threshold) at the specific site (‘moist days’). We thus defined PETNN as a soil-moisture-unlimited ET. The method was only applied for sites where sufficient data above and below the soil moisture threshold were available. In contrast, the model for predicting ETNN was trained using all data and with soil moisture as an additional predictor.

The threshold to divide data into ‘moist’ and ‘dry’ days was determined by running the site-specific models for a sequence of soil moisture thresholds. For each threshold, we calculated the median of the ratio in moist and dry days and determined the three models with the highest difference between these two ratios. Among these three models, the one with the smallest variance in fET during moist days was chosen.

We derived the belowground water stress factor (fET) as:

|  |  |
| --- | --- |
|  | Eq. 3 |

Using this method, we quantified the control of belowground water reservoirs on ET, separated from other predictors (net radiation, VPD, vegetation greenness and air temperature) without needing explicit information about the amount or distribution of root-accessible water. The only water availability information used was the surface soil moisture values (which are correlated to deeper soil moisture (Qiu et al., 2016; Salvucci & Entekhabi, 1994)) used to determine the 'moist' days, i.e. the days when ET = PET.

The use of this neural-network-based approach avoided the necessity to determine PET *a priori* based on theoretical estimations that are difficult to parameterize (Maes et al., 2019). Using ETNN instead of ETobs in Eq. 3 resulted in less noisy fET estimates, as ETNN and PETNN were affected by similar prediction errors, such as varying footprint and incorrect instantaneous energy balance closure. By definition, PETNN should agree with ETobs during ‘moist days’. On the contrary, PETNN was expected to overestimate ETobs during ‘dry days’, as the former was trained on ‘moist days’ data only (Fig. S1).

We evaluated PETNN against an empirical Priestley-Taylor estimate of PET (PETPT) obtained from the SPLASH model (Davis et al., 2017) and against a linear model (lm) estimate of PET (PETlm). The site-specific linear regression models were defined as PETlm=k\*Rn, where Rn was converted to mass units (mm d-1) and k is a site-specific constant that scales Rn, calibrated against ETobs.

## Deep neural network models architecture

The site-specific DNN models were built as feed-forward deep neural networks, implemented using R packages Tensorflow (Falbel et al., 2022) and Keras (Falbel, Allaire, Chollet, et al., 2021). The hyperparameter tuning was performed on the number of neurons per hidden layer (sampled among 8, 16, 32, 64), the number of hidden layers (sampled from 1 to 5), the optimizer (sampled among ‘adam’ or ‘rmsprop’), the activation function (“relu”, “leaky\_relu”, “linear”), the batch size (sampled among 16, 32, 64, 128, 256) and the learning rate (sampled among 0.01, 0.001, 0.0001) using the R packages ‘tfruns’ and ‘tfestimators’ (Allaire et al., 2018; Falbel, Allaire, Bostock, et al., 2021). We first ran the same tuning algorithm for a limited set of representative sites. To reduce the number of hyperparameters of the DNN, the ones that were giving consistent results from the beginning were removed from the final tuning across all sites, e.g., we retained only the activation function “relu” as it was clearly outperforming “linear” and “leaky\_relu”; we similarly kept a learning rate of 0.01.

We used the mean square error (MSE) as loss function during model training and the mean absolute percentage error as the error metric during model validation. Other error metrics were tested (e.g., mean absolute error), but resulted in a lower performance. To monitor the validation loss, we used the “EarlyStopping” callback function. This function stopped the training process should the validation loss not improve after five epochs. This limited the number of training epochs and avoided overfitting.

We ran hyperparameter tuning for a subset of 5% of the total hyperparameter combinations and chose the simplest model (i.e., the model with the lowest number of total parameters) among the five models with the lowest validation loss. We built one model per site, and the hyperparameters were tuned on a site-by-site basis. We then trained the feed-forward DNNs performing a five-fold cross-validation, with a 75%-25% split between training and validation data, respectively. The model with the lowest root-mean-square error was selected and the same procedure was repeated five times. We retained the mean prediction across the five ensemble neural network members. This was done to consider the variability caused by the inherent randomness of the initialization of the weights of the neural network neurons.

## Data

Half-hourly data were downloaded from the FLUXNET2015 website and filtered to keep only measured values (\*\_QC = 0) or values gap-filled with high confidence (\*\_QC = 1) (Pastorello et al., 2020). The latent heat flux was converted from energy units to mass units (ET), dividing it by the latent heat of vaporization as a function of air temperature, using the R package ‘Bigleaf’ (Knauer, El-Madany, et al., 2018)*.*

To reduce biases in ET predictions, we applied additional filters to the data, in accordance with other previous studies (Li et al., 2019; Liu et al., 2022; Medlyn et al., 2017; Zhou et al., 2016). We first applied a rainfall filter with a buffer of 6 hours after each rain event to exclude interception evaporation and to avoid sensor saturation with high relative humidity (Li et al., 2019). We further removed data with relative humidity higher than the 95% quantile to exclude the impact of dew evaporation on ET (Knauer, Zaehle, et al., 2018). To avoid stable boundary layer conditions, we excluded data where the sensible heat flux was smaller than 5 W m-2 and incoming shortwave radiation was smaller than 50 W m-2. Finally, only daytime data (GPP, ET and VPD > 0) were considered. Half-hourly data were aggregated into daily data to reduce noise and to avoid the ET-VPD hysteresis effect, observed at sub-daily timescales (Q. Zhang et al., 2014). While aggregating to the daily level, the daily mean was calculated for all variables, except for VPD (for which we calculated the daily maximum), ET and precipitation (for which we used the daily sum). We only retained daily estimates with at least 8 measured half-hourly points, as in (Li et al., 2019).

The evaporative fraction (EF) (which was used as a consistency check on typical fET values in our analysis) was calculated using daily aggregates as .

We used observational soil moisture data as available from the FLUXNET2015 dataset. To be consistent across all sites, we only used soil moisture data from the uppermost layer, as multiple depths were available only for a very limited number of sites. For many FLUXNET2015 sites, we found that observational soil moisture data was unavailable, incomplete, or inconsistent with ET observations. 'Incomplete' meant that the soil moisture timeseries had large data gaps (on the scale of >50% of the data, which made the calculation of CWD not possible). 'Inconsistent with ET observations' referred to the fact that soil moisture was not consistent with the soil water balance calculated using precipitation and latent heat flux, from the same FLUXNET2015 database. For these sites, we simulated soil moisture with SPLASH, a bucket-type soil water balance model (Davis et al., 2017). This model was based on a Priestley-Taylor formulation for ET estimation. We set the water-holding capacity (‘bucket depth’) to 220 mm (Davis et al., 2017; Orth et al., 2013). Using a modelled soil moisture represented a potential source of circularity in our analysis, since we were assuming a water-holding capacity specific to a rooting depth. However, we mostly used soil moisture to separate the training data into 'moist' and 'dry' days (see 'Estimating potential ET'), so that the exact interpolated values should have only a minimal influence on our results.

We downloaded MODIS EVI (MOD13Q1, Collection 5) with the ‘MODISTools’ R package (Tuck et al., 2014). Images with clouds, snow, ice or shadows were excluded. To get daily values and remove noise, we applied a Savitzky–Golay smoothing filter (‘SIGNAL’ R package) with a 3rd order polynomial and frame length of 31 days.

Before running the machine learning model, we excluded all NAs. Soil moisture was normalized between 0 and 1 on a site-by-site basis for better comparison across sites. We centered and scaled all variables as part of the resampling process with the R package ‘caret’ (Kuhn et al., 2021).

We downloaded precipitation, ET and PET estimates of the GLDAS\_NOAH025\_3H product from the NASA Global Land Data Assimilation System Version 2 (GLDAS-2) (Beaudoing et al., 2020; Rodell et al., 2004). The product is based on an assimilation of ground- and space-based observations to model energy and water exchanges. We referred to ET and PET from this data product as ETGLDAS and PETGLDAS. We defined fETGLDAS = ETGLDAS/PETGLDAS to compare it with our deep learning estimate of fET. The data was originally available at a 3-hour time resolution and at a grid resolution of 0.25°x0.25°. After extracting the variables at the site level, we calculated daily means. PETGLDAS displays good correlation with in-situ measurements, although it was found to be biased relative to some satellite products (Zhan et al., 2019). For comparison with PETNN, PETGLDAS was therefore scaled by dividing it by its median value in the lower CWD bin (CWD < 20 mm). This way, fETGLDAS was visually comparable with fET (roughly comprised between 0 and 1).

We extracted soil texture distribution data at FLUXNET2015 locations from the Regridded Harmonized World Soil Database v1.2 (Wieder et al., 2014). We downloaded mean annual temperature (MAT) and mean annual precipitation (MAP) from WorldClim version 2.1 and extracted their values at FLUXNET2015 sites (Fick & Hijmans, 2017).

The aridity index was calculated with FLUXNET2015 data as the ratio of annual precipitation (P) over potential evapotranspiration (PET), for all years in which data were available for the respective sites. Precipitation data are from the FLUXNET 2015 Tier 1 dataset; PET was calculated following the Priestley–Taylor equation (PETPT), as implemented in the SPLASH model (Davis et al., 2017).

To investigate the role of topography across sites, Global topographic index (GTI) values were downloaded from a high-resolution dataset and extracted at FLUXNET2015 locations (Marthews et al., 2015a, 2015b).

## Derivation of the cumulative water deficit

We derived the cumulative water deficit (CWD) as the cumulative difference of the actual evapotranspiration (ETobs) and precipitation (P), considered over continuous dry periods, i.e., periods where the difference *P – ETobs* was negative. The summation was stopped when the rain had compensated the water loss due to ET, i.e., the cumulative sum across days was zero (). We defined a ‘CWD event’ as the period between the start and the end of the summation, i.e., a dry-down event. We assumed that water stress was already mitigated when new precipitation was re-wetting the topsoil layers, even before the CWD was fully offset. For this reason, we removed data after rain had reduced the CWD to below 90% of its maximum value within the same ‘CWD event’. To eliminate the noise caused by smaller CWD events and focus on extreme events, we only retained the single largest CWD event each year. We used latent heat flux and precipitation timeseries from FLUXNET2015 at the daily time scale.

When calculating CWD, it was important to focus on high quality observations and use a continuous time series of ET to avoid gaps. We therefore processed the data in a slightly different way than the data used for the DNN models. We only retained measured values (\*\_QC = 0) or values gap-filled with high confidence (\*\_QC = 1) (Pastorello et al., 2020). We then gap-filled ET based on site-specific single-layer neural networks, using temperature, PAR, VPD and ET simulated by the SPLASH model as predictors (Davis et al., 2017). To build this model, we used the R packages ‘NNET’ (Venables & Ripley, 2002) and ‘CARET’ (Kuhn et al., 2021), and used a neural network with a single hidden layer, 20 nodes, 10-fold cross-validated. We only used the gap-filled ET to calculate the CWD. Note that this single hidden layer neural network used for gap filling ET is different from the DNN model defined in 2.3. We trained the DNN model with the ET timeseries cleaned as described in section 2.4 with no additional gap-filling.

## Site selection and binning

We evaluated fET for 135 sites of the FLUXNET Tier 1 dataset, where modelled soil moisture (or observational soil moisture where available) gave consistent results (as defined in section 2.4). The sites were filtered according to the final number of days after data cleaning (> 300 d) and to the performance of the DNN model. Mean PETNN had to be greater than or equal to ET during ‘dry days’ and the R2 between modelled ET and ETobs had to be > 0.5, retaining 58 sites. Nine sites were excluded upon visual inspection of the fET vs CWD relationship, which was not giving physically consistent results. The remaining 49 sites were used in this study and are listed in Table S1. We removed fET outliers, defined as any value that fell outside of the interval To categorize the behavior of different sites, we calculated the median of fET in the CWD interval comprised between 125 mm and 175 mm for each site. We grouped all sites along this single dimension using a k-means algorithm, with predefined k = 3. We conducted a sensitivity analysis with k = 2 and k = 4. We retained the grouping with k = 3, as it identified clearly the dominant fET distributions across sites. The resulting groups were: high fET (8 sites), medium fET (22 sites) and low fET (11 sites), with each group showing a similar pattern in the decline of fET during drought (based on visual inspect). The CWD interval centered at 150 mm was a trade-off between using the largest possible CWD, thus capturing the most divergent responses across sites, and choosing a CWD value attained and exceeded in the largest number of sites. Sites that had a maximum CWD lower than 125 mm were manually assigned to the 'high fET' group (8 sites), as their behavior was visually deemed to be consistent with other sites in the same group (Fig. S3).

# Results

## Reliability of the deep neural network

Across all sites and days, ETNN achieved consistent results against ETobs (R2 = 0.78, Fig. 1a). To evaluate the accuracy of PETNN, we compared it against ETobs during moist days only (R2 = 0.77, Fig. 1b). This method gave better results compared to PETPT (R2 = 0.45, Fig. 1c) or PETlm (R2 = 0.40, Fig. 1d), both evaluated for moist days only. Although the general patterns are robust across the three different methods, they emerge clearer when using DNNs for modelling PET. We thus retained PETNN for further analyses.

|  |  |
| --- | --- |
|  |  |
|  |  |

**Fig. 1** Performance of the deep-learning model at predicting evapotranspiration (ET). ETNN and PETNN are respectively ET and PET predicted with our deep learning model. ETobs corresponds to observational ET from FLUXNET2015. (**a**)ETNN vs ETobs, evaluated on all days. (**b**)PETNN vs ETobs, evaluated on moist days only. (**c)** PETPT is from the SPLASH model, based on a Priestley-Taylor formulation of evapotranspiration, vs ETobs. (**d**) PETlm is based on a linear model (lm), defined as PET=k\*Rn, where *Rn* was converted to mass units (mm d-1) and *k* is a site-specific constant that scales *Rn*, calibrated to ETobs.

To further understand the timing and magnitude of water limitation on ETobs, we evaluated its seasonality compared to ETNN and PETNN (Fig. 2). We chose two sample sites with contrasting behavior: DK-Sor, a humid deciduous broadleaf forest, and US-Ton, a dry woody savannah site. At DK-Sor, PETNN and ETobs almost perfectly overlap, suggesting that belowground moisture limitation has little effect on ET at this site, allowing it to be almost always at its potential rate, i.e., energy-limited and without belowground moisture limitations (Fig. 2a, red and black lines). At US-Ton, PETNN departs substantially from ETobs during the dry season, indicating significant impacts of water stress on ET at this site (Fig. 2b, red and black lines). In both cases, ETNN is consistent with ETobs (black and blue lines) and the ET seasonality is confirmed by Rn converted to mass units (Fig. 2, dashed green line).

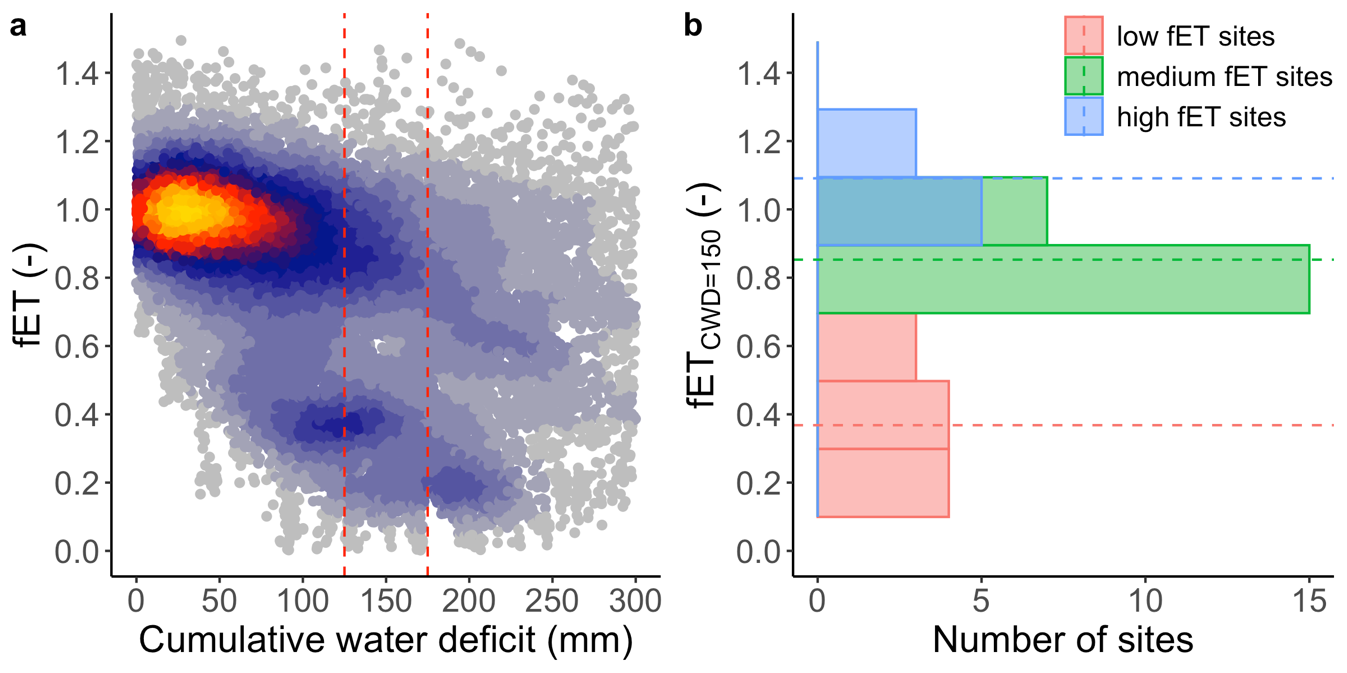
**Chart, histogram

Description automatically generated**

**Fig. 2** Seasonality of predicted and observed ET for sample sites. **(a)** DK-Sor. (**b)** US-Ton. ETNN and PETNN are respectively ET and PET predicted with our deep learning model. ETobs corresponds to observational ET from FLUXNET2015. Blue line: ETNN. Green line: PETNN. Red line: ETobs. Dashed gray line: Net radiation converted to mass units (mm d-1). We derived the seasonality by calculating the mean across all years for every day of the year.

## Binning of fET vs CWD responses

When the density of fET is plotted against CWD across all sites, we observe a variety of responses (Fig. 3a). For CWD values up to approximately 100 mm, there is an accumulation of points centered at fET equal to one. At higher CWD values, fET declines gradually. We can distinguish two other smaller clouds of points, one centered at around fET equal to 0.4 and another one at 0.2. The response of fET vs. CWD seems to be most variable around CWD = 150 mm (red dotted lines in Fig. 3a). To further investigate commonalities of the fET-CWD relationship across sites, we divided them into three groups based on the fET median in this interval (Fig. 3b, see 2.6). Sites were thus grouped into low, medium, and high fET. Each group exhibits a different fET vs CWD behavior (Fig. 5 and Figs. S3-S4).

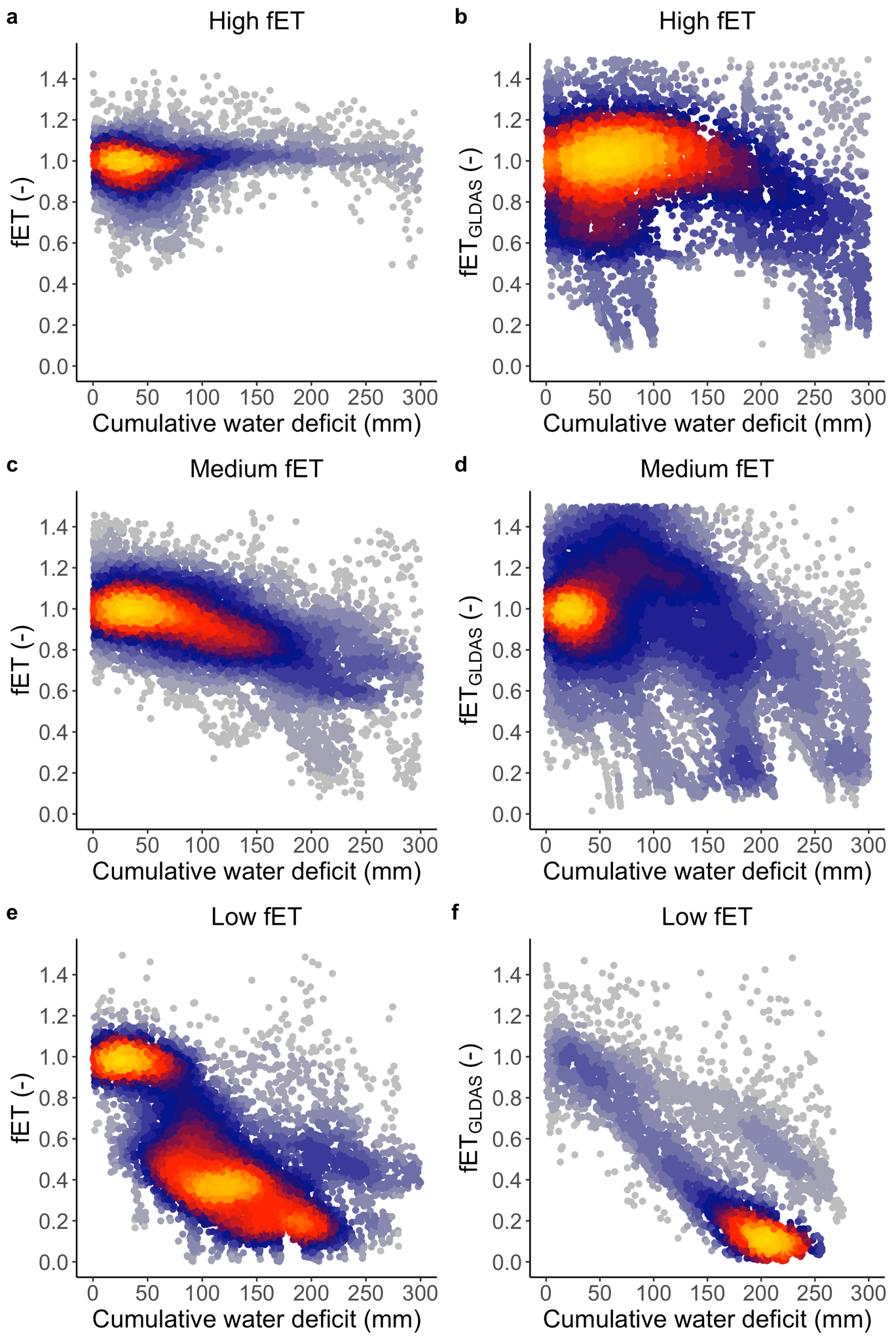


**Fig. 3** Partition of sites in three groups according to their median fET at CWD = 150 mm.

**(a)** fET for all sites plotted against the cumulative water deficit (CWD). Dashed lines at CWD = 125 mm and CWD = 175 mm delimit the interval (fET150) in which the median of fET was calculated for each site to define the three fET groups (low fET, medium fET and high fET). **(b)** Number of sites per each fET group: low fET sites (red), medium fET sites (green), high fET sites (blue). Dashed lines represent the average fET inside each group.The width of each bin was determined with the Freedman-Diaconis rule, that considers not only sample size but also the spread of each sample.

## Patterns of fET vs CWD responses across sites

At 'high fET' sites, there is almost no effect of water stress on plants, as fET is almost always near one, in spite of substantial water deficits (Fig. 4a and Fig. S3). In contrast, when predicted by a standard land surface model, after a CWD threshold of around 150 mm, fETGLDAS is decreasing linearly with progressing drought at these sites (Fig. 4b). At 'medium fET' sites (Fig. 4c), the bulk of fET values is equal to one up to a CWD of around 100 mm. At a CWD greater than 100 mm, fET slowly decreases with progressing CWD, reaching an fET of around 0.5 at a corresponding CWD of 250 mm. In contrast, fETGLDAS displays several tails, which decrease linearly with progressing CWD, down to zero (Fig. 4d). The observed fETGLDAS behavior reflects different responses at different sites. At 'low fET' sites, fET stays equal to one until a CWD of 50 mm. After that, fET values drop off abruptly with increasing CWD, followed by a relationship that largely levels-off, slowly approaching but never reaching fET = 0 (Fig. 4e). In the same fET group, fETGLDAS is decreasing almost linearly with progressing drought, reaching values around zero (Fig. 4f). This drop in fET with CWD followed by a levelling-off is also seen for each of the individual sites in this group (Fig. 5). This confirms that the trend observed in Fig. 4e is not simply due to plotting several sites pooled together, and adds further support to the site grouping used here. Across the three fET groups, fETGLDAS reduces more quickly with increasing CWD than fET does, suggesting GLDAS overestimates water stress. The general trends observed for fET vs CWD are confirmed by EF, a model-independent indicator of ET stress, plotted as a function of CWD (Fig. S2).



**Fig. 4** fET vs CWD for sites grouped according to their median fET. (**a**) High fET, fET predicted from our observations-driven model. (**b**) High fET, fET extracted as ET/PET from the GLDAS data product. (**c)** Medium fET, fET predicted from our observations-driven model. (**d)** Medium fET, fET extracted as ET/PET from the GLDAS data product. (**e)** Low fET, fET predicted from our observations-driven model. (**f)** Low fET, fET extracted as ET/PET from the GLDAS data product. PET from GLDAS was scaled by dividing it by its median in the lower CWD bin (CWD < 20 mm).

**Chart, scatter chart

Description automatically generated**

**Fig. 5** fET vs CWD for sites grouped according to their median fET.Low fET group.

## Relationship to other soil and climate variables

To gain insight into what factors drive the different fET behaviors, we investigated how the different site groups vary with soil and climate variables (Fig. 6). The mean sand fraction increases slightly across 'low fET' to 'high fET' sites, and the silt fraction decreases slightly (Fig. 6a). Sites in the 'high fET' group are usually found in humid climates (Fig. 6c). They consist mostly of forests, and in particular evergreen needleleaf forest (ENF, 7 sites out of 16, Fig. 6b). We found that sites with intermediate fET reductions are found in mesic regions (Fig. 6c). They are mostly characterized by forests (14 sites out of 22) and croplands (4 sites, Fig. 6b). The 'low fET' sites are found in the driest climates (Fig. 6c). These sites are composed mostly of savannahs (5 sites over 11) and grasslands (4 sites over 11, Fig. 6b). Topographic effects and mean annual precipitation do not vary significantly across fET groups (Fig. 6d,e). By contrast, the mean annual temperature increases from high to low fET groups (Fig. 6f).

Chart, box and whisker chart

Description automatically generated

**Fig. 6** Analysis of soil and climate variables per fET group. (**a)** Topsoil soil texture in percentage weight. The mean was calculated across all sites within a certain fET group. (**b)** IGBP vegetation classes (GRA, grasslands; SAV, savannah; WSA, woody savannah; ENF, evergreen needleleaf forest; EBF, evergreen broadleaf forest; DBF, deciduous broadleaf forest; CSH, closed shrubland; WET, wetland; CRO, cropland; MF, mixed forest). (**c**) Aridity index, defined as the ratio of annual precipitation (P) over potential evapotranspiration (PET), calculated for all available years on a site-by-site basis. Precipitation data was taken from the FLUXNET2015 dataset, whereas PET was calculated with the SPLASH model, based on a Priestley-Taylor formulation of evapotranspiration. (**d**) Global topographic index (GTI), defining the tendency of the soil to become saturated with water because of its topographic position. (**e**) Mean Annual Precipitation (MAP). (**f**) Mean Annual Temperature (MAT).

We can observe a clear correlation between the median fET (as defined above) and the aridity index (Fig. 7). At the same time, fET is also dependent on soil texture. In particular, the clay fraction exerts a strong control on the fET vs aridity index relationship (Fig. 7a). The higher the silt fraction, the stronger the relationship between fET and the aridity index (slopes of the regression lines, Fig. 7c). The opposite can be said for the sand fraction: the higher the sand fraction, the weaker the relationship between fET and the aridity index (Fig. 7b).

Chart, scatter chart

Description automatically generated

**Fig. 7** Median fET as a function of the aridity index per site. Each dot represents one site. Circles size indicates soil texture (in percentage weight). The colored lines show the regression between the site-specific median fET and the aridity index, divided in three equal-sized groups of soil texture fractions (in ascending order, from lower to higher soil fraction: blue, green and red lines). Slopes values were tested to be statistically significant (t-test with a significance level of 0.05). (**a)** Clay fraction (**b)** Sand fraction. (**c**) Silt fraction. All soil fractions are from the topsoil.

# Discussion

Even at high CWD values, vegetation activity is never completely shut off. As CWD increases and fET decreases in low and medium fET sites (Fig. 4c,e), the fET vs CWD pattern suggests that water continues to be supplied to maintain a base ET level without further reductions in plant access to moisture, a pattern not reflected by a standard land surface model (Fig. 4b,d,f). This suggests that plants may have access to deep water or have a low xylem potential cavitation point that allows them to sustain ET even during the unfolding of a drought event. Field studies could be devised to discriminate between these alternative hypotheses using measurements that are linked to ET and can directly sample water in the soil-plant-atmosphere continuum (SPAC, i.e., sap flow and leaf water potential measurements) (Novick et al., 2022).

## Drivers of the abruptness of the fET decrease with CWD

Our findings are consistent with the notion that forests (which are more common at high fET sites) are more resistant to drought than grasslands and can support vegetation activity over longer dry periods (Konings & Gentine, 2017; Martínez-Vilalta & Garcia-Forner, 2017; Teuling et al., 2010). Grasslands tend to have shallower roots and thus are more likely to experience water stress when the topsoil dries out. Forests tend to have deeper roots that can access deeper water stores, and therefore are more resilient to high CWD (Fan et al., 2017; Tumber-Dávila et al., 2022).

The different behavior observed in grasslands and forests could explain the drop in fET after an intermediate CWD threshold at low fET sites (Fig. 4e). In this group, the dominant PFT is woody savannah, a tree-grass ecosystem characterized by a herbaceous understory (grassland) scattered with sufficiently spaced trees, so that the canopy is never continuous (Fig. 6b) (El-Madany et al., 2020; Luo et al., 2018). After a certain CWD value (around 50 mm), we hypothesize that the herbaceous layer loses access to water, and thus stops contributing to ecosystem ET. In turn, trees can have a more resistant xylem, allowing them to pull water at lower water potentials or, alternatively, they can rely on deep roots to access deeper water reservoirs. These strategies allow trees to keep transpiring even at higher CWD. This could explain why we observe an abrupt change in fET, followed by a levelling-off which never reaches zero (Fig. 4e). In this framework, the levelling-off corresponds to a period where activity of the understory ceases and mostly trees contribute to a base evapotranspiration.

The cavitation resistance found in arid plant communities could also explain the levelling-off of the fET vs CWD relationship in this group of sites (Jacobsen et al., 2007). In particular, when the belowground moisture availability decreases in such a way that the root water potential decreases, one would also expect the xylem potential to decrease, and in turn the leaf water potential and the stomatal conductance. The relatively flat fET curve suggests that either *a*) plants access deep belowground moisture and maintain relatively high water potentials along the SPAC while water is continuously consumed (thus continuously increasing the cumulative water deficit), and/or *b*) conductance along the SPAC is relatively insensitive to the range of water potentials experienced by plants under the conditions investigated here. That insensitivity would thus correspond to a specific range in the curve of the stomatal conductance as a function of leaf water potential, where the water potential is varying without dramatically affecting the conductance (Wolf et al., 2016).

The correlation between fET groups and the aridity index (Fig. 6c) indicates different adaptive plant strategies to water stress. At intermediate to low aridity, woody vegetation invests in biomass to be more drought-resistant, i.e. deeper rooting depth, and lower minimum leaf water potentials (van der Molen et al., 2011). At high aridity, carbon uptake is limited so that investments into structures for maintaining activity (e.g., deep roots, evergreen foliage) is not possible (Christiansen et al., 1987; Stamp, 2003). This is also reflected by the relatively high prevalence of drought-deciduous vegetation at arid sites. These plants have a more resistant xylem and typically have tracheid rather than vessels, allowing a smaller leaf area index (LAI) which in turn reduces the maximum transpiration rate (McDowell et al., 2008). Semi-arid regions are a key driver of the interannual variability of the terrestrial carbon cycle (Ahlström et al., 2015; Poulter et al., 2014). The seasonal reductions in ET found at arid sites suggest that a more accurate accounting of drought conditions of these areas in global models could improve the prediction of the variability of the carbon cycle (Biederman et al., 2017). (van der Molen et al., 2011)

The relationship between fET and aridity index is regulated by soil texture, which modulates the relationship between water availability and soil water potential, thereby affecting plant water stress (Novick et al., 2022). Nevertheless, the aridity effects mentioned above are dominant. Additionally, there was no significant difference in topographic index across fET groups (Fig. 6d). This is consistent with the fact that most flux towers are located in flat areas and valleys, so that most sites cluster into similar topographic index values (Thompson et al., 2011).

## Possible explanations of the discrepancy between models and observations

The hypothesis that trees access deeper water reservoirs is consistent with recent findings, which highlight the often-neglected importance of deep, non-soil water stores for root water uptake (Dawson et al., 2020). Rock moisture is often accessible to plants (McCormick et al., 2021), but is essentially not represented by models, and is especially important during drought, after soil moisture is depleted (Rempe & Dietrich, 2018). This could explain why the lower tail of the fET vs CWD relationship in the 'low fET' group never reaches zero, even at very high CWD (i.e. under ongoing drought, Fig. 4e). The same behaviour is not captured by models, which cannot quantify rock moisture (Fig. 4f).

Groundwater contributions to ET could also explain the discrepancy between models and observations. Groundwater is generally not taken into account by global models (Condon et al., 2021; Hain et al., 2015), but it has been shown to have a pivotal effect in sustaining transpiration during drought (Mu et al., 2021). The neglect of groundwater in many land surface models could thus explain why they tend to overestimate water stress effects compared to observations (Fig. 4b,d,f). This is consistent with a previous finding that quantifying plant access to groundwater can improve ET prediction (Thompson et al., 2011).

In our anaysis, we compared our results to a standard land surface model, the GLDAS\_NOAH025\_3H product. GLDAS uses vegetation tiling to represent sub-grid heterogeneity of vegetation types, so the fluxes in a grid box are a weighted average of land cover tiles at 1 km resolution. NOAH assigns a rooting depth to each vegetation type (Beaudoing et al., 2020; Rodell et al., 2004). This could explain why GLDAS data does not capture some of the water stress responses highlighted by our approach, as the rooting depth has been shown to vary strongly even within the same species and climate (Fan et al., 2017). NOAH also does not account for heterogeneity in rooting depth within the same vegetation type. Similarly, NOAH does not account for heterogeneity in stomatal or xylem traits within the same vegetation type (Konings and Gentine, 2017).

The almost linear signal found in 'high fET' sites when using GLDAS data (Fig. 4f) could be linked to missing spatial heterogeneity in the model simulations. There is an intrinsic problem in ET prediction related to the fact that soil moisture, and by extension belowground water availability, varies in nature at scales in the order of 1-10 m, while models have grid cell sizes at scales as big as one to several hundred kilometers. The upscaling would naturally smooth out the stress response (Baker et al., 2017). The fact that the GLDAS model uses vegetation tiles to simulate low-scale heterogeneity could entail an ecosystem response more linear than what it really is. In other words, large scale averaging could lead to a less abrupt regulation of fET in reality than in models (Baker et al., 2017). More research is needed to resolve the apparent model-observation bias in light of the role of belowground water availability.

An observation-driven statistical approach is used here to evaluate the effects of belowground water stress on ET, separated from the contribution of other drivers, including radiation, VPD and vegetation greenness. We demonstrated that it is crucial to account for belowground water availability effects to diagnose ET responses to drought. We highlighted substantial differences in plant responses to water stress across vegetation types, mean site climate, and soil texture. Most forests tend to show little sensitivity to water stress, whereas most savannahs, shrublands and grasslands sites show an abrupt drop in ET after an initial stress-free phase. In both cases and in contrast with a representative land surface model, ET is almost never completely shut off, even during progressing drought conditions. Access to belowground water reserves could explain the muted effects of belowground water limitation, observed here also in sites where seasonal water deficits are substantial. The fact that most models do not account for sub-grid heterogeneity in belowground water availability and ignore moisture supplied by the saturated zone and weathered bedrock leads to a biased quantification of water stress effects on plants. Future research should address this bias focusing on the role of deep unquantified water stores, the quantification of sub-grid heterogeneity, and the representation of rooting strategies and plant conductance across aridity gradients.

**Data availability**

All datasets used in this study are freely available from the following sources:

* Ecosystem fluxes and meteorological data: https://fluxnet.org/data/fluxnet2015-dataset/
* MODIS EVI (from MOD13Q1, 16 days, 250 m): https://lpdaac.usgs.gov/products/mod13q1v006/
* Precipitation, ET and PET estimates from the GLDAS\_NOAH025\_3H product: https://disc.gsfc.nasa.gov/datasets/GLDAS\_NOAH025\_3H\_2.1/summary?keywords=gldas
* Soil texture distribution data: https://daac.ornl.gov/SOILS/guides/HWSD.html
* Mean annual temperature (MAT) and mean annual precipitation (MAP): https://www.worldclim.org/data/worldclim21.html
* Global topographic index (GTI): https://data-search.nerc.ac.uk/geonetwork/srv/api/records/6b0c4358-2bf3-4924-aa8f-793d468b92be
* All other intermediate data and computer codes that support this study are available from the Zenodo Digital Repository: https://doi.org/10.5281/zenodo.6897699 (Giardina et al. 2022).
* Final modelled fET output is also made available as a separate dataset: https://zenodo.org/record/6885164

**Acknowledgements**The authors thank the providers of the data sets used in this study. We also thank group members Koen Hufkens and Laura Marqués for providing feedback. This project was supported by the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program. F.G. and B.D.S. were funded by the Swiss National Science Foundation grant no. PCEFP2\_181115. A.G.K. was supported by NSF DEB-1942133 and by the NASA MAP program under award 80NSSC21K1523. P.G. was supported by the European research Council grant SMILE: Understanding and modelling the Earth System with Machine learning, by the NASA funding Remote Sensing Estimate of Evapotranspiration Partitioning to Transpiration and (with A.G.K.) and the NASA SMAP Understanding memory effects and climatic drivers of net primary productivity and respiration enabled by SMAP vegetation optical depth.

**Conflict of interest**

The authors declare no conflicts of interest.

**Author contributions**

F.G., B.D.S., P.G. and A.G.K. designed the study and wrote the manuscript text. F.G. performed the research and prepared figures.

**List of Figures and Tables**

**Fig. 1** Performance of the deep-learning model at predicting evapotranspiration (ET).

**Fig. 2** Seasonality of predicted and observed ET for sample sites.

**Fig. 3** Partition of sites in three groups according to their median fET at CWD = 150 mm.

**Fig. 4** fET vs CWD for sites grouped according to their median fET.

**Fig. 5** fET vs CWD for sites grouped according to their median fET.Low fET group.

**Fig. 6** Analysis of soil and climate variables per fET group.

**Fig. 7** Median fET as a function of the aridity index per site.

**Fig. S1** Performance of the deep-learning model at predicting evapotranspiration (ET) across sites.

**Fig. S2** EF vs CWD for sites grouped according to their median fET

**Fig. S3** fET vs CWD for sites grouped according to their median fET. High fET group.

**Fig. S4** fET vs CWD for sites grouped according to their median fET. Medium fET group.

**Table S1** FLUXNET Tier 1 sites included in the analysis.

**References**

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J. G., Friedlingstein, P., Jain, A. K., Kato, E., Poulter, B., Sitch, S., Stocker, B. D., Viovy, N., Wang, Y. P., Wiltshire, A., Zaehle, S., & Zeng, N. (2015). The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink. *Science*, *348*(6237), 895–899.

Allaire, J., Tang, Y., Ushey, K., Kuo, K., & Falbel, D. (2018). *tfestimators: Interface to “TensorFlow” Estimators* (1.9.1).

Baker, I. T., Sellers, P. J., Denning, A. S., Medina, I., Kraus, P., Haynes, K. D., & Biraud, S. C. (2017). Closing the scale gap between land surface parameterizations and GCMs with a new scheme, SiB3-Bins. *Journal of Advances in Modeling Earth Systems*, *9*(1), 691–711. https://doi.org/10.1002/2016MS000764

Beaudoing, H., Rodell, M., & NASA/GSFC/HSL. (2020). *GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.1*. Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC). https://disc.gsfc.nasa.gov/datasets/GLDAS\_NOAH025\_3H\_2.1/summary

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., … Papale, D. (2010). *Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate*. *329*(August), 834–839.

Biederman, J. A., Scott, R. L., Bell, T. W., Bowling, D. R., Dore, S., Garatuza-Payan, J., Kolb, T. E., Krishnan, P., Krofcheck, D. J., Litvak, M. E., Maurer, G. E., Meyers, T. P., Oechel, W. C., Papuga, S. A., Ponce-Campos, G. E., Rodriguez, J. C., Smith, W. K., Vargas, R., Watts, C. J., … Goulden, M. L. (2017). CO2 exchange and evapotranspiration across dryland ecosystems of southwestern North America. *Global Change Biology*, *23*(10), 4204–4221. https://doi.org/10.1111/gcb.13686

Christiansen, E., Waring, R. H., & Berryman, A. A. (1987). Resistance of conifers to bark beetle attack: Searching for general relationships. *Forest Ecology and Management*, *22*(1–2), 89–106. https://doi.org/10.1016/0378-1127(87)90098-3

Condon, L. E., Kollet, S., Bierkens, M. F. P., Fogg, G. E., Maxwell, R. M., Hill, M. C., Fransen, H. J. H., Verhoef, A., Van Loon, A. F., Sulis, M., & Abesser, C. (2021). Global Groundwater Modeling and Monitoring: Opportunities and Challenges. In *Water Resources Research* (Vol. 57, Issue 12, p. e2020WR029500). John Wiley and Sons Inc. https://doi.org/10.1029/2020WR029500

Davis, T. W., Prentice, I. C., Stocker, B. D., Thomas, R. T., Whitley, R. J., Wang, H., Evans, B. J., Gallego-Sala, A. V., Sykes, M. T., & Cramer, W. (2017). Simple process-led algorithms for simulating habitats (SPLASH v.1.0): Robust indices of radiation, evapotranspiration and plant-available moisture. *Geoscientific Model Development*, *10*(2), 689–708. https://doi.org/10.5194/gmd-10-689-2017

Dawson, T. E., Hahm, W. J., & Crutchfield-Peters, K. (2020). Digging deeper: what the critical zone perspective adds to the study of plant ecophysiology. *New Phytologist*, *226*(3), 666–671. https://doi.org/10.1111/nph.16410

Dralle, D. N., Jesse Hahm, W., Rempe, D. M., Karst, N., Anderegg, L. D. L., Thompson, S. E., Dawson, T. E., & Dietrich, W. E. (2020). Plants as sensors: Vegetation response to rainfall predicts root-zone water storage capacity in Mediterranean-type climates. *Environmental Research Letters*, *15*(10). https://doi.org/10.1088/1748-9326/abb10b

Drewniak, B. A. (2019). Simulating Dynamic Roots in the Energy Exascale Earth System Land Model. *Journal of Advances in Modeling Earth Systems*, *11*(1), 338–359. https://doi.org/10.1029/2018MS001334

El-Madany, T. S., Carrara, A., Martín, M. P., Moreno, G., Kolle, O., Pacheco-Labrador, J., Weber, U., Wutzler, T., Reichstein, M., & Migliavacca, M. (2020). Drought and heatwave impacts on semi-arid ecosystems’ carbon fluxes along a precipitation gradient: Drought and Heatwave Impacts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1810). https://doi.org/10.1098/rstb.2019.0519

Falbel, D., Allaire, J., Bostock, M., Tanaka, M., Bowe, S., You, Y., Decker, K., Fernandes, R., Sagalaev, I., & Pakala, Y. (2021). *tfruns: Training Run Tools for “TensorFlow”* (1.5.0).

Falbel, D., Allaire, J., Chollet, F., RStudio, Google, Tang, Y., Bijl, W. Van Der, Studer, M., & Keydana, S. (2021). *keras: R Interface to “Keras”* (2.4.0).

Falbel, D., Allaire, J., RStudio, Yuan Tang, Dirk Eddelbuettel, Golding, N., Kalinowski, T., & Inc., G. (2022). *Package ‘tensorflow’* (2.8.0).

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., … Yamazaki, D. (2019). Hillslope Hydrology in Global Change Research and Earth System Modeling. *Water Resources Research*, *55*(2), 1737–1772. https://doi.org/10.1029/2018WR023903

Fan, Y., Li, H., & Miguez-Macho, G. (2013). Global patterns of groundwater table depth. *Science*, *339*(6122), 940–943. https://doi.org/10.1126/science.1229881

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(40), 10572–10577. https://doi.org/10.1073/pnas.1712381114

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1‐km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., & Savenije, H. H. G. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research Letters*, *41*(22), 7916–7923. https://doi.org/10.1002/2014GL061668

Giardina, F., Konings, A. G., Kennedy, D., Alemohammad, S. H., Oliveira, R. S., Uriarte, M., & Gentine, P. (2018). Tall Amazonian forests are less sensitive to precipitation variability. *Nature Geoscience*, *11*(6), 405–409. https://doi.org/10.1038/s41561-018-0133-5

Green, J. K., Konings, A. G., Alemohammad, S. H., Berry, J., Entekhabi, D., Kolassa, J., Lee, J.-E., & Gentine, P. (2017). Regionally strong feedbacks between the atmosphere and terrestrial biosphere. *Nature Geosci*, *advance on*(May). https://doi.org/10.1038/ngeo2957

Hahm, W. J., Dralle, D. N., Rempe, D. M., Bryk, A. B., Thompson, S. E., Dawson, T. E., & Dietrich, W. E. (2019). Low Subsurface Water Storage Capacity Relative to Annual Rainfall Decouples Mediterranean Plant Productivity and Water Use From Rainfall Variability. *Geophysical Research Letters*, *46*(12), 6544–6553. https://doi.org/10.1029/2019GL083294

Hain, C. R., Crow, W. T., Anderson, M. C., & Tugrul Yilmaz, M. (2015). Diagnosing neglected soil moisture source-sink processes via a thermal infrared-based two-source energy balance model. *Journal of Hydrometeorology*, *16*(3), 1070–1086. https://doi.org/10.1175/JHM-D-14-0017.1

Jacobsen, A. L., Pratt, R. B., Davis, S. D., & Ewers, F. W. (2007). Cavitation resistance and seasonal hydraulics differ among three arid Californian plant communities. *Plant, Cell and Environment*, *30*(12), 1599–1609. https://doi.org/10.1111/j.1365-3040.2007.01729.x

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C., & Gentine, P. (2019). Implementing Plant Hydraulics in the Community Land Model, Version 5. *Journal of Advances in Modeling Earth Systems*, *11*(2), 485–513. https://doi.org/10.1029/2018MS001500

Knauer, J., El-Madany, T. S., Zaehle, S., & Migliavacca, M. (2018). Bigleaf - An R package for the calculation of physical and physiological ecosystem properties from eddy covariance data. *PLoS ONE*, *13*(8), 1–26. https://doi.org/10.1371/journal.pone.0201114

Knauer, J., Zaehle, S., Medlyn, B. E., Reichstein, M., Williams, C. A., Migliavacca, M., De Kauwe, M. G., Werner, C., Keitel, C., Kolari, P., Limousin, J. M., & Linderson, M. L. (2018). Towards physiologically meaningful water-use efficiency estimates from eddy covariance data. *Global Change Biology*, *24*(2), 694–710. https://doi.org/10.1111/gcb.13893

Konings, A. G., & Gentine, P. (2017). Global variations in ecosystem-scale isohydricity. *Global Change Biology*, *23*(2), 891–905. https://doi.org/10.1111/gcb.13389

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Team, R. C., Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., & Hunt, T. (2021). *caret: Classification and Regression Training* (6.0-88). https://doi.org/10.1887/0750303123/b365c43

Li, X., Gentine, P., Lin, C., Zhou, S., Sun, Z., Zheng, Y., Liu, J., & Zheng, C. (2019). A simple and objective method to partition evapotranspiration into transpiration and evaporation at eddy-covariance sites. *Agricultural and Forest Meteorology*, *265*(May 2018), 171–182. https://doi.org/10.1016/j.agrformet.2018.11.017

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

Liu, Y., Flournoy, O., Zhang, Q., Novick, K. A., Koster, R. D., & Konings, A. G. (2022). Canopy Height and Climate Dryness Parsimoniously Explain Spatial Variation of Unstressed Stomatal Conductance. *Geophysical Research Letters*, *49*(15), e2022GL099339. https://doi.org/10.1029/2022GL099339

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

Luo, Y., El-Madany, T. S., Filippa, G., Ma, X., Ahrens, B., Carrara, A., Gonzalez-Cascon, R., Cremonese, E., Galvagno, M., Hammer, T. W., Pacheco-Labrador, J., Martín, M. P., Moreno, G., Perez-Priego, O., Reichstein, M., Richardson, A. D., Römermann, C., & Migliavacca, M. (2018). Using Near-Infrared-Enabled Digital Repeat Photography to Track Structural and Physiological Phenology in Mediterranean Tree–Grass Ecosystems. *Remote Sensing*, *10*(8), 1293. https://doi.org/10.3390/rs10081293

Mackay, D. S., Roberts, D. E., Ewers, B. E., Sperry, J. S., McDowell, N. G., & Pockman, W. T. (2015). Interdependence of chronic hydraulic dysfunction and canopy processes can improve integrated models of tree response to drought. *Water Resources Research*, *51*(8), 6156–6176. https://doi.org/10.1002/2015WR017244

Maes, W. H., Gentine, P., Verhoest, N. E. C., & Miralles, D. G. (2019). Potential evaporation at eddy-covariance sites across the globe. *Hydrology and Earth System Sciences*, *23*(2), 925–948. https://doi.org/10.5194/hess-23-925-2019

Marthews, T. R., Dadson, S. J., Lehner, B., Abele, S., & Gedney, N. (2015a). *High-resolution global topographic index values*. NERC Environmental Information Data Centre. (Dataset).

Marthews, T. R., Dadson, S. J., Lehner, B., Abele, S., & Gedney, N. (2015b). High-resolution global topographic index values for use in large-scale hydrological modelling. *Hydrology and Earth System Sciences*, *19*(1), 91–104. https://doi.org/10.5194/hess-19-91-2015

Martínez-Vilalta, J., & Garcia-Forner, N. (2017). Water potential regulation, stomatal behaviour and hydraulic transport under drought: deconstructing the iso/anisohydric concept. *Plant Cell and Environment*, *40*(6), 962–976. https://doi.org/10.1111/pce.12846

McCormick, E. L., Dralle, D. N., Hahm, W. J., Tune, A. K., Schmidt, L. M., Chadwick, K. D., & Rempe, D. M. (2021). Widespread woody plant use of water stored in bedrock. *Nature*, *597*(7875), 225–229. https://doi.org/10.1038/s41586-021-03761-3

McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D. G., & Yepez, E. A. (2008). Mechanisms of Plant Survival and Mortality during Drought: why do survive while others succumb plants drought ? to. *New Phytologist*, *178*(4), 719–739.

Medlyn, B. E., De Kauwe, M. G., Lin, Y. S., Knauer, J., Duursma, R. A., Williams, C. A., Arneth, A., Clement, R., Isaac, P., Limousin, J. M., Linderson, M. L., Meir, P., Martin-Stpaul, N., & Wingate, L. (2017). How do leaf and ecosystem measures of water-use efficiency compare? *New Phytologist*, *216*, 758–770. https://doi.org/10.1111/nph.14626

Milly, P. C. D. (1994). Climate, soil water storage, and the average annual water balance. *Water Resources Research*, *30*(7), 2143–2156. https://doi.org/10.1029/94WR00586

Miralles, D. G., Gentine, P., Seneviratne, S. I., & Teuling, A. J. (2019). Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. *Annals of the New York Academy of Sciences*, *1436*(1), 19–35. https://doi.org/10.1111/nyas.13912

Mu, M., De Kauwe, M. G., Ukkola, A. M., Pitman, A. J., Guo, W., Hobeichi, S., & Briggs, P. R. (2021). Exploring how groundwater buffers the influence of heatwaves on vegetation function during multi-year droughts. *Earth System Dynamics*, *12*(3), 919–938. https://doi.org/10.5194/esd-12-919-2021

Novick, K. A., Ficklin, D. L., Baldocchi, D., Davis, K. J., Ghezzehei, T. A., Konings, A. G., MacBean, N., Raoult, N., Scott, R. L., Shi, Y., Sulman, B. N., & Wood, J. D. (2022). Confronting the water potential information gap. *Nature Geoscience*, *15*(3), 158–164. https://doi.org/10.1038/s41561-022-00909-2

Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., Papuga, S. A., Blanken, P. D., Noormets, A., Sulman, B. N., Scott, R. L., Wang, L., & Phillips, R. P. (2016). The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. *Nature Climate Change*, *6*(11), 1023–1027. https://doi.org/10.1038/nclimate3114

Orth, R., Koster, R. D., & Seneviratne, S. I. (2013). Inferring soil moisture memory from streamflow observations using a simple water balance model. *Journal of Hydrometeorology*, *14*(6), 1773–1790. https://doi.org/10.1175/JHM-D-12-099.1

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Ribeca, A., van Ingen, C., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., … Papale, D. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1), 225. https://doi.org/10.1038/s41597-020-0534-3

Plaut, J. A., Yepez, E. A., Hill, J., Pangle, R., Sperry, J. S., Pockman, W. T., & Mcdowell, N. G. (2012). Hydraulic limits preceding mortality in a piñon-juniper woodland under experimental drought. *Plant, Cell & Environment*, *35*(9), 1601–1617. https://doi.org/10.1111/j.1365-3040.2012.02512.x

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y. Y., Running, S. W., Sitch, S., & Van Der Werf, G. R. (2014). Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature*, *509*(7502), 600–603. https://doi.org/10.1038/nature13376

Prentice, I. C., Liang, X., Medlyn, B. E., & Wang, Y. P. (2015). Reliable, robust and realistic: The three R’s of next-generation land-surface modelling. *Atmospheric Chemistry and Physics*, *15*(10), 5987–6005. https://doi.org/10.5194/acp-15-5987-2015

Qiu, J., Crow, W. T., & Nearing, G. S. (2016). The impact of vertical measurement depth on the information content of soil moisture for latent heat flux estimation. *Journal of Hydrometeorology*, *17*(9), 2419–2430. https://doi.org/10.1175/JHM-D-16-0044.1

Rempe, D. M., & Dietrich, W. E. (2018). Direct observations of rock moisture, a hidden component of the hydrologic cycle. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2664–2669. https://doi.org/10.1073/pnas.1800141115

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Salvucci, G. D., & Entekhabi, D. (1994). Equivalent steady soil moisture profile and the time compression approximation in water balance modeling. *Water Resources Research*, *30*(10), 2737–2749.

Schlemmer, L., Schär, C., Lüthi, D., & Strebel, L. (2018). A Groundwater and Runoff Formulation for Weather and Climate Models. *Journal of Advances in Modeling Earth Systems*, *10*(8), 1809–1832. https://doi.org/10.1029/2017MS001260

Schwalm, C. R., Williams, C. A., Schaefer, K., Arneth, A., Bonal, D., Buchmann, N., Chen, J., Law, B., Lindroth, A., Luyssaert, S., Reichstein, M., & Richardson, A. D. (2010). Assimilation exceeds respiration sensitivity to drought: A FLUXNET synthesis. *Global Change Biology*, *16*(2), 657–670. https://doi.org/10.1111/j.1365-2486.2009.01991.x

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

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A. Di, Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., & Zhou, B. (2021). Weather and Climate Extreme Events in a Changing Climate. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, *Cambridge*(In Press.).

Stamp, N. (2003). Out Of The Quagmire Of Plant Defense Hypotheses. *Source: The Quarterly Review of Biology*, *78*(1), 23–55. https://doi.org/10.1086/367580

Stocker, B. D., Tumber-d, S. J., Konings, A. G., Anderson, M. B., Hain, C., & Jackson, R. B. (2021). Global distribution of the rooting zone water storage capacity reflects plant adaptation to the environment. *BioRxiv*. https://doi.org/https://doi.org/10.1101/2021.09.17.460332

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Peñuelas, J., & Seneviratne, S. I. (2018). Quantifying soil moisture impacts on light use efficiency across biomes. *New Phytologist*, *218*(4), 1430–1449. https://doi.org/10.1111/nph.15123

Teuling, A. J., Seneviratne, S. I., Stöckli, R., Reichstein, M., Moors, E., Ciais, P., Luyssaert, S., Van Den Hurk, B., Ammann, C., Bernhofer, C., Dellwik, E., Gianelle, D., Gielen, B., Grünwald, T., Klumpp, K., Montagnani, L., Moureaux, C., Sottocornola, M., & Wohlfahrt, G. (2010). Contrasting response of European forest and grassland energy exchange to heatwaves. *Nature Geoscience*. https://doi.org/10.1038/NGEO950

Teuling, A. J., Seneviratne, S. I., Williams, C., & Troch, P. A. (2006). Observed timescales of evapotranspiration response to soil moisture. *Geophysical Research Letters*, *33*(23), 0–4. https://doi.org/10.1029/2006GL028178

Thompson, S. E., Harman, C. J., Konings, A. G., Sivapalan, M., Neal, A., & Troch, P. A. (2011). Comparative hydrology across AmeriFlux sites: The variable roles of climate, vegetation, and groundwater. *Water Resources Research*, *47*(7), 1–17. https://doi.org/10.1029/2010WR009797

Tuck, S. L., Phillips, H. R. P., Hintzen, R. E., Hudson, J. P. W. S., Purvis, A., & N., L. (2014). MODISTools - downloading and processing MODIS remotely sensed data in R. *Ecology and Evolution*, *4*(24), 4658–4668. https://doi.org/10.1002/ece3.1273

Tumber-Dávila, S. J., Schenk, H. J., Du, E., & Jackson, R. B. (2022). Plant sizes and shapes above and belowground and their interactions with climate. *New Phytologist*, *235*(3), 1032–1056. https://doi.org/10.1111/nph.18031

van der Molen, M. K., Dolman, A. J., Ciais, P., Eglin, T., Gobron, N., Law, B. E., Meir, P., Peters, W., Phillips, O. L., Reichstein, M., Chen, T., Dekker, S. C., Doubková, M., Friedl, M. A., Jung, M., van den Hurk, B. J. J. M., de Jeu, R. A. M., Kruijt, B., Ohta, T., … Wang, G. (2011). Drought and ecosystem carbon cycling. In *Agricultural and Forest Meteorology* (Vol. 151, Issue 7, pp. 765–773). Elsevier. https://doi.org/10.1016/j.agrformet.2011.01.018

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth). Springer.

Wieder, W. R., Boehnert, J., Bonan, G. B., & Langseth., M. (2014). *Regridded Harmonized World Soil Database v1.2*. ORNL DAAC, Oak Ridge, Tennessee, USA.

Wolf, A., Anderegg, W. R. L., & Pacala, S. W. (2016). Optimal stomatal behavior with competition for water and risk of hydraulic impairment. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), E7222–E7230. https://doi.org/10.1073/pnas.1615144113

Zhan, S., Song, C., Wang, J., Sheng, Y., & Quan, J. (2019). A Global Assessment of Terrestrial Evapotranspiration Increase Due to Surface Water Area Change. *Earth’s Future*, *7*(3), 266–282. https://doi.org/10.1029/2018EF001066

Zhang, K., Kimball, J. S., & Running, S. W. (2016). A review of remote sensing based actual evapotranspiration estimation. *Wiley Interdisciplinary Reviews: Water*, *3*(6), 834–853. https://doi.org/10.1002/wat2.1168

Zhang, Q., Manzoni, S., Katul, G., Porporato, A., & Yang, D. (2014). The hysteretic evapotranspiration—Vapor pressure deficit relation. *Journal of Geophysical Research: Biogeosciences*, *119*(2), 125–140. https://doi.org/doi:10.1002/ 2013JG002484

Zhou, S., Yu, B., Zhang, Y., Huang, Y., & Guangqian, W. (2016). Partitioning evapotranspiration based on the concept of underlying water use efficiency. *Water Resources Research*, *52*, 1160– 1175. https://doi.org/10.1111/j.1752-1688.1969.tb04897.x

Zhou, S., Zhang, Y., Williams, A. P., & Gentine, P. (2019). Projected increases in intensity, frequency, and terrestrial carbon costs of compound drought and aridity events. *Science Advances*, *5*(1), 1–9. https://doi.org/10.1126/sciadv.aau5740

**Supporting Information**

Additional Supporting Information may be found online in the Supporting Information tab for this article:

**Fig. S1** Performance of the deep-learning model at predicting evapotranspiration (ET) across sites.

**Fig. S2** EF vs CWD for sites grouped according to their median fET

**Fig. S3** fET vs CWD for sites grouped according to their median fET. High fET group.

**Fig. S4** fET vs CWD for sites grouped according to their median fET. Medium fET group.

**Table S1** FLUXNET Tier 1 sites included in the analysis.