## *New Phytologist* Supporting Information

Article title: Deep learning to diagnose evapotranspiration responses to water deficit across biomes

Authors: Francesco Giardina, Pierre Gentine, Alexandra G. Konings, Sonia I. Seneviratne, Benjamin D. Stocker

Article acceptance date: Click here to enter a date.

The following Supporting Information is available 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.

**Fig. S5** Seasonality of predicted and observed evapotranspiration (ET) at two sample dry sites.

**Fig. S6** Performance of the deep-learning model at predicting evapotranspiration (ET) and potential evapotranspiration (PET), using the four predictors retained in the study vs a set of seven predictors.

**Fig. S7** Evolution of the fractional reduction in evapotranspiration (fET) vs transpiration (fT) with the cumulative water deficit (CWD) and performance of the deep-learning model at predicting evapotranspiration (ET) vs transpiration (T).

**Fig. S8** Range and distribution of EVI for 'moist' and 'dry' days.

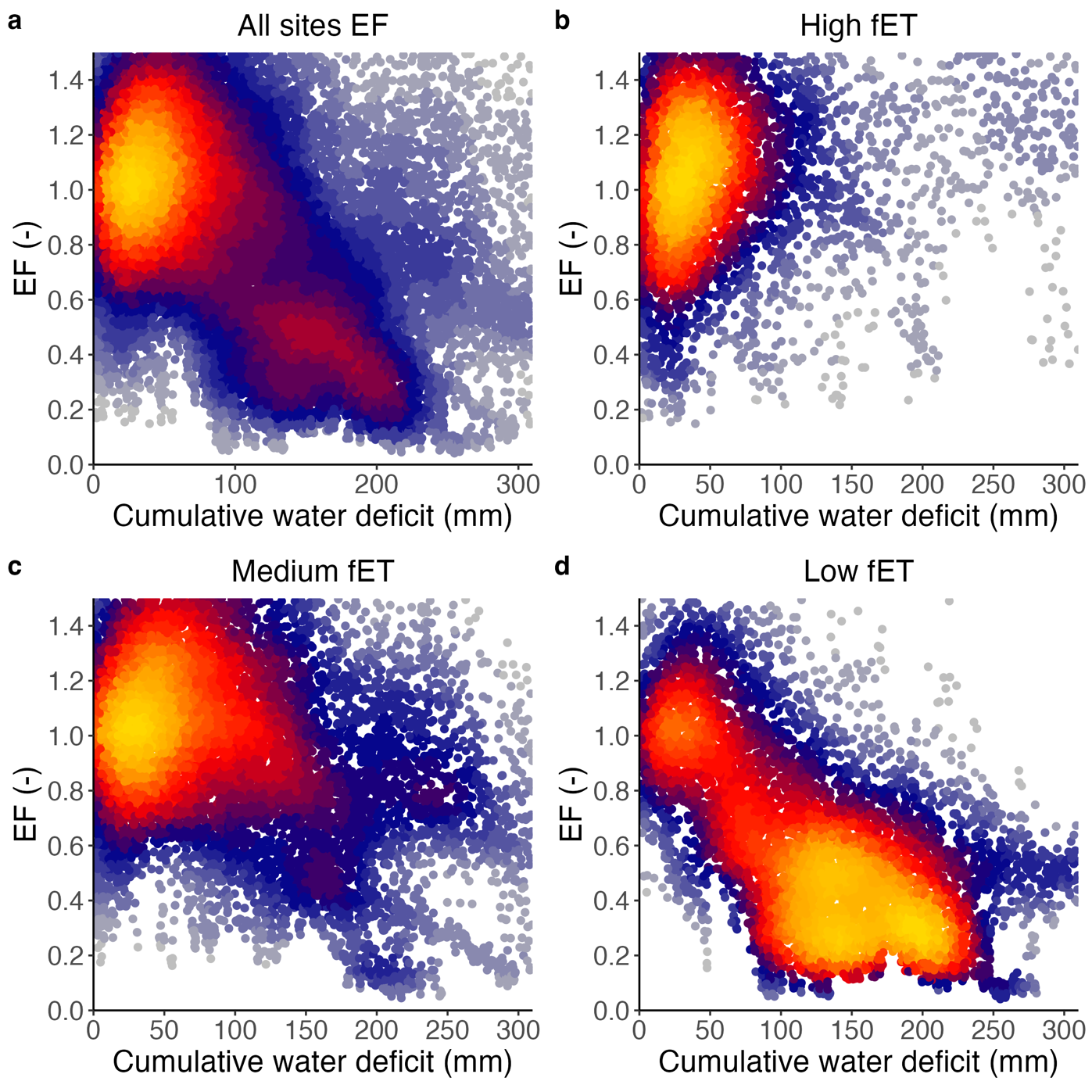
**Table S1** FLUXNET Tier 1 sites included in the analysis.

**Methods S1** Extended description of the deep learning models.

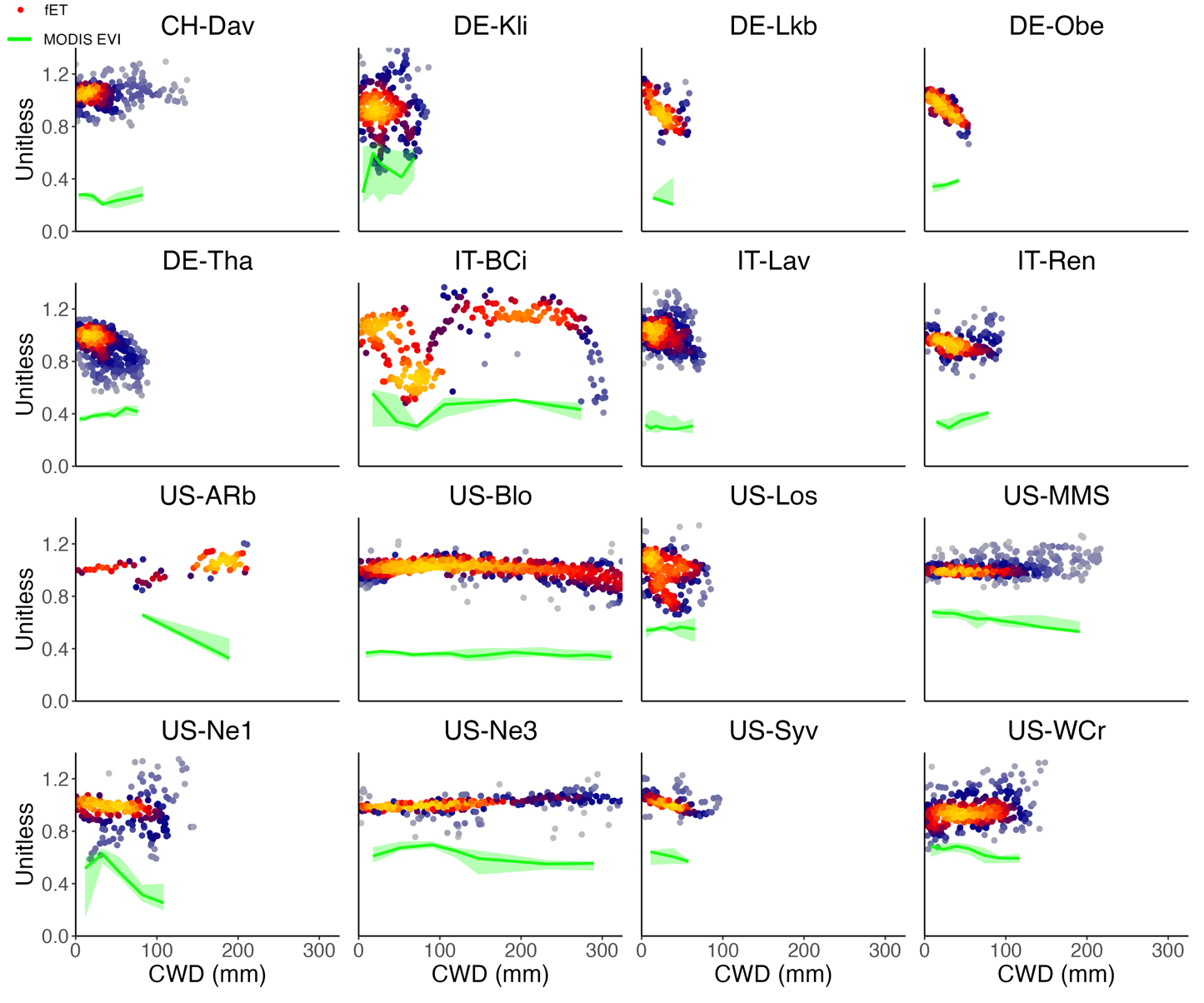
**Figures  
Fig. S1** Performance of the deep-learning model at predicting evapotranspiration (ET) across sites.ETNN and PETNN are respectively ET and PET predicted with our deep learning model. ETobs corresponds to observational ET from FLUXNET2015. **a**,PETNN vs ETNN, evaluated on moist days. **b**, PETNN vs observational ETobs, evaluated on dry days. Red line: regression line between modelled and observed data. Dashed black line: *y = x* line. RMSE, root-mean-square error. R package 'LSD' was used to plot the point density (Schwalb et al., 2020).

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

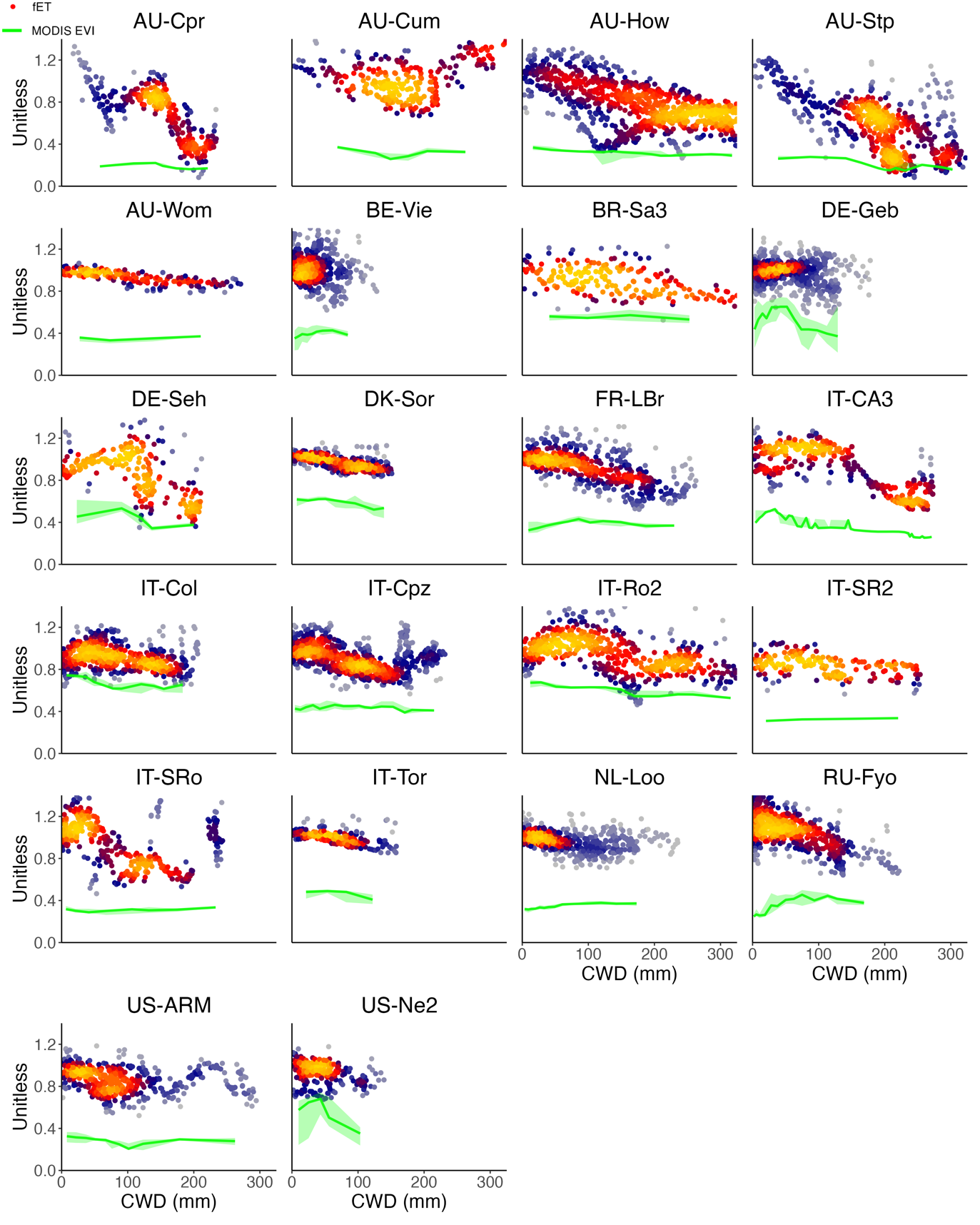
**Fig. S2** EF vs CWD for sites grouped according to their median fET. a, All sites. b, High fET. c, Medium fET. d, Low fET. The clustering is consistent with Figure 3. EF was scaled by dividing it by the median in its lower CWD bin (CWD < 20 mm). EF was calculated as latent heat divided by net radiation, two quantities directly downloaded from the FLUXNET2015 dataset that do not depend on any model. R package 'LSD' was used to plot the point density (Schwalb et al., 2020).

****

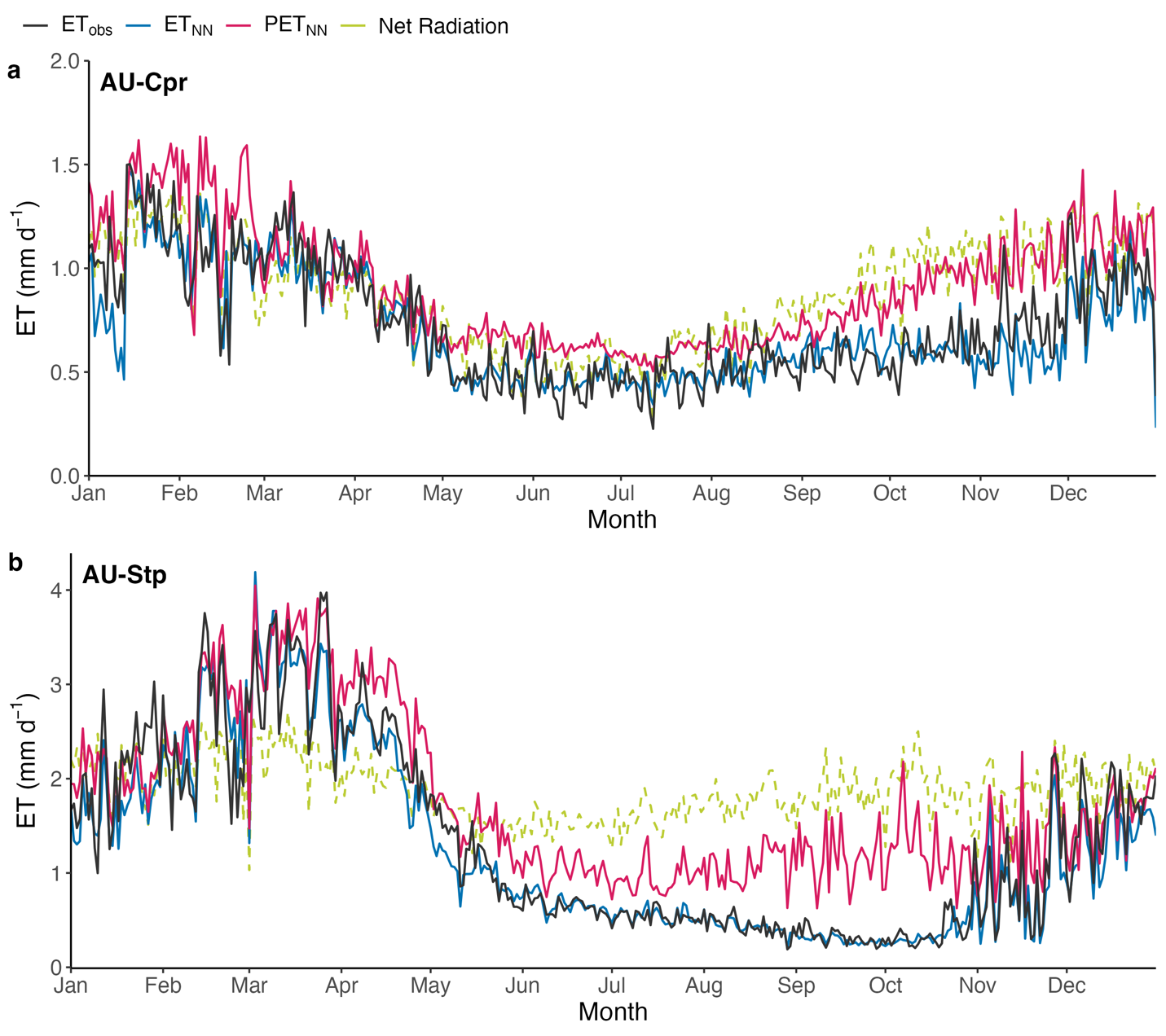
**Fig. S3** Evolution of the fractional reduction in evapotranspiration (fET) with the cumulative water deficit (CWD) for sites classified in the **'**high fET' group. Colored dots: fET. Green line: MODIS Enhanced Vegetation Index (EVI). EVI was binned by CWD intervals of 50 points. Shading represents the lower and upper quartiles, and the solid line the median in every bin. R package 'LSD' was used to plot the point density (Schwalb et al., 2020).



**Fig. S4** Evolution of the fractional reduction in evapotranspiration (fET) with the cumulative water deficit (CWD) for sites classified in the **'**medium fET' group. Colored dots: fET. Green line: MODIS Enhanced Vegetation Index (EVI). EVI was binned by CWD intervals of 50 points. Shading represents the lower and upper quartiles, and the solid line the median in every bin. R package 'LSD' was used to plot the point density (Schwalb et al., 2020).

****

**Fig. S5** Seasonality of predicted and observed evapotranspiration (ET) at two sample dry sites. **(a)** AU-Cpr. (**b)** AU-Stp. ETNN and PETNN are respectively ET and PET predicted with our deep learning model. ETobs corresponds to observational ET from FLUXNET2015. Blue line: ETNN. Red line: PETNN. Black line: ETobs. Dashed green 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.



**Fig. S6** Performance of the deep-learning model at predicting evapotranspiration (ET) and potential evapotranspiration (PET), using the four predictors retained in the study vs a set of seven predictors at US-Ton. *obs* corresponds to observational ET from FLUXNET2015. *nn\_act* and *nn\_pot* are respectively ET and PET predicted with our deep learning model.

|  |  |
| --- | --- |
| Four predictors: net radiation, VPD, air temperature, EVI | Seven predictors: net radiation, VPD, air temperature, EVI, wind speed, soil temperature, U\* |
|  |  |
|  |  |

**Fig. S7** Evolution of the fractional reduction in evapotranspiration (fET) vs transpiration (fT) with the cumulative water deficit (CWD) and performance of the deep-learning model at predicting evapotranspiration (ET) vs transpiration (T) at US-Ton, US-Var and DK-Sor. Transpiration data (Tobs) are from a from a published dataset (Li et al., 2019b). ETobs corresponds to observational ET from FLUXNET2015. ETNN and TNN are respectively ET and T predicted with our deep learning model.

|  |  |
| --- | --- |
| **US-Ton** | |
| **ET** | **T** |
| Chart, scatter chart  Description automatically generated | Chart, scatter chart  Description automatically generated |
| Chart, scatter chart  Description automatically generated | Chart, scatter chart  Description automatically generated |

|  |  |
| --- | --- |
| **US-Var** | |
| **ET** | **T** |
|  |  |
|  |  |

|  |  |
| --- | --- |
| **DK-Sor** | |
| **ET** | **T** |
|  | Chart, scatter chart  Description automatically generated |
|  | Chart, scatter chart  Description automatically generated |

**Fig. S8** Range and distribution of EVI for 'moist' and 'dry' days. **(a)** Boxplot of EVI for all sites. Blue: 'moist days. Red: 'dry days'. **(b)** Histogram of EVI for all sites. Blue: 'moist days. Red: 'dry days'. **(c)** Boxplots of EVI at individual sites showing all data**.** **(d)** Boxplots of EVI at individual sites showing 'moist days' data only**.**

**Chart

Description automatically generated**

**Tables**

**Table S1** FLUXNET Tier 1 sites included in the analysis. Group refers to the grouping of sites according to their median fET (see Methods). Coordinates in decimal degrees. IGBP is the vegetation class (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). MAT, mean annual temperature (°C). MAP, mean annual precipitation (mm).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Site** | **Coordinates** | **Years** | **IGBP** | **fET group** | **MAT (**°C) | **MAP (mm)** |
| **AU-Cpr** | 140.59, -34 | 2010 - 2014 | SAV | medium fET | 17.35 | 269 |
| **AU-Cum** | 150.72, -33.61 | 2012 - 2014 | EBF | medium fET | 17.28 | 850 |
| **AU-DaP** | 131.32, -14.06 | 2007 - 2013 | GRA | low fET | 27.22 | 1175 |
| **AU-DaS** | 131.39, -14.16 | 2008 - 2014 | SAV | low fET | 27.13 | 1134 |
| **AU-Gin** | 115.71, -31.38 | 2011 - 2014 | WSA | low fET | 18.53 | 697 |
| **AU-How** | 131.15, -12.49 | 2001 - 2014 | WSA | medium fET | 27.03 | 1640 |
| **AU-RDF** | 132.48, -14.56 | 2011 - 2013 | WSA | low fET | 26.98 | 958 |
| **AU-Stp** | 133.35, -17.15 | 2008 - 2014 | GRA | medium fET | 26.39 | 639 |
| **AU-Wom** | 144.09, -37.42 | 2010 - 2012 | EBF | medium fET | 10.70 | 1071 |
| **BE-Vie** | 6, 50.31 | 1996 - 2014 | MF | medium fET | 8.01 | 1085 |
| **BR-Sa3** | -54.97, -3.02 | 2000 - 2004 | EBF | medium fET | 25.50 | 1856 |
| **CH-Dav** | 9.86, 46.82 | 1997 - 2014 | ENF | high fET | 3.53 | 1053 |
| **DE-Geb** | 10.91, 51.1 | 2001 - 2014 | CRO | medium fET | 8.96 | 496 |
| **DE-Kli** | 13.52, 50.89 | 2004 - 2014 | CRO | high fET | 7.61 | 839 |
| **DE-Lkb** | 13.3, 49.1 | 2009 - 2013 | ENF | high fET | 4.20 | 1364 |
| **DE-Obe** | 13.72, 50.78 | 2008 - 2014 | ENF | high fET | 6.09 | 820 |
| **DE-Seh** | 6.45, 50.87 | 2007 - 2010 | CRO | medium fET | 10.17 | 709 |
| **DE-Tha** | 13.57, 50.96 | 1996 - 2014 | ENF | high fET | 8.26 | 754 |
| **DK-Sor** | 11.64, 55.49 | 1996 - 2014 | DBF | medium fET | 8.48 | 614 |
| **FR-LBr** | -0.77, 44.72 | 1996 - 2008 | ENF | medium fET | 12.52 | 908 |
| **FR-Pue** | 3.6, 43.74 | 2000 - 2014 | EBF | low fET | 13.35 | 683 |
| **IT-BCi** | 14.96, 40.52 | 2004 - 2014 | CRO | high fET | 16.10 | 1035 |
| **IT-CA3** | 12.02, 42.38 | 2011 - 2014 | DBF | medium fET | 14.52 | 328 |
| **IT-Col** | 13.59, 41.85 | 1996 - 2014 | DBF | medium fET | 6.95 | 789 |
| **IT-Cpz** | 12.38, 41.71 | 1997 - 2009 | EBF | medium fET | 15.54 | 757 |
| **IT-Lav** | 11.28, 45.96 | 2003 - 2014 | ENF | high fET | 6.79 | 502 |
| **IT-Noe** | 8.15, 40.61 | 2004 - 2014 | CSH | low fET | 15.85 | 748 |
| **IT-Ren** | 11.43, 46.59 | 1998 - 2013 | ENF | high fET | 3.98 | 664 |
| **IT-Ro2** | 11.92, 42.39 | 2002 - 2012 | DBF | medium fET | 14.54 | 380 |
| **IT-SR2** | 10.29, 43.73 | 2013 - 2014 | ENF | medium fET | 14.02 | 888 |
| **IT-SRo** | 10.28, 43.73 | 1999 - 2012 | ENF | medium fET | 14.02 | 888 |
| **IT-Tor** | 7.58, 45.84 | 2008 - 2014 | GRA | medium fET | 1.56 | 1317 |
| **NL-Loo** | 5.74, 52.17 | 1996 - 2013 | ENF | medium fET | 9.38 | 839 |
| **RU-Fyo** | 32.92, 56.46 | 1998 - 2014 | ENF | medium fET | 4.51 | 693 |
| **US-AR2** | -99.6, 36.64 | 2009 - 2012 | GRA | low fET | 14.22 | 600 |
| **US-ARb** | -98.04, 35.55 | 2005 - 2006 | GRA | high fET | 15.29 | 793 |
| **US-ARM** | -97.49, 36.61 | 2003 - 2012 | CRO | medium fET | 14.90 | 861 |
| **US-Blo** | -120.63, 38.9 | 1997 - 2007 | ENF | high fET | 11.04 | 1510 |
| **US-Los** | -89.98, 46.08 | 2000 - 2014 | WET | high fET | 4.12 | 833 |
| **US-MMS** | -86.41, 39.32 | 1999 - 2014 | DBF | high fET | 11.13 | 1097 |
| **US-Ne1** | -96.48, 41.17 | 2001 - 2013 | CRO | high fET | 10.24 | 799 |
| **US-Ne2** | -96.47, 41.16 | 2001 - 2013 | CRO | medium fET | 10.21 | 799 |
| **US-Ne3** | -96.44, 41.18 | 2001 - 2013 | CRO | high fET | 10.20 | 793 |
| **US-SRG** | -110.83, 31.79 | 2008 - 2014 | GRA | low fET | 17.57 | 537 |
| **US-SRM** | -110.87, 31.82 | 2004 - 2014 | WSA | low fET | 18.53 | 459 |
| **US-Syv** | -89.35, 46.24 | 2001 - 2014 | MF | high fET | 3.76 | 844 |
| **US-Ton** | -120.97, 38.43 | 2001 - 2014 | WSA | low fET | 16.14 | 656 |
| **US-Var** | -120.95, 38.41 | 2000 - 2014 | GRA | low fET | 16.20 | 639 |
| **US-WCr** | -90.08, 45.81 | 1999 - 2014 | DBF | high fET | 4.23 | 828 |

**Methods**

**Methods S1** Extended description of the deep learning models.

## 1 Data processing

In accordance with previous studies, we filtered the data to reduce biases in ET prediction (Li et al., 2019a; 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., 2019a). We further removed data with relative humidity higher than the 95% quantile to exclude the impact of dew evaporation on ET (Knauer et al., 2018). Please note that this is a simplification, given that dew formation is technically associated with cooler surface temperature compared to air temperature. However, this filter, together with the filter below on shortwave radiation, should be robust enough to exclude conditions with dewfall. 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.

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 (soil moisture >> ET − P or soil moisture << ET − P) calculated using precipitation and latent heat flux, from the same FLUXNET2015 database.

## 2 Derivation of the soil moisture threshold

The threshold to divide data into ‘moist’ and ‘dry’ days was determined by building site-specific models for a sequence of soil moisture thresholds. The sequence of thresholds was sampled from 0.1 to 0.7 with an increment of 0.05. 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. The method was only applied for sites where sufficient data above and below the soil moisture threshold were available.

## 3 Hyperparameter tuning of the deep neural networks

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). Before training the machine learning models, we excluded all NAs and centered and scaled all variables with the R package ‘caret’ (Kuhn 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.

In hyperparameter tuning, 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. We then retained five sets of hyperparameters that achieved the lowest validation loss. Among the five models with the lowest validation loss, we chose the simplest model (i.e., the model with the lowest number of total features). We repeated this procedure for every site.

Once we had defined the site-specific set of hyperparameters, we trained the feed-forward DNNs performing a five-fold cross-validation, with a 75%-25% split between training and testing 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.

## 4 Gap-filling of ET for the calculation of the cumulative water deficit

We 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 section 1 of Methods S1. We trained the DNN model with the ET timeseries cleaned as described in section 2.2 with no additional gap-filling.

## References

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

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

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). https://keras.rstudio.com/

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

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

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. (2019a). 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

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

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

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

Schwalb, B., Tresch, A., Torkler, P., Duemcke, S., Demel, C., Ripley, B., & Venables, B. (2020). *LSD: Lots of Superior Depictions* (4.1-0). https://CRAN.R-project.org/package=LSD

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

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