



## Supplement of

## Long-term relative decline in evapotranspiration with increasing runoff on fractional land surfaces

Ren Wang et al.

*Correspondence to:* Ren Wang (wangr67@mail2.sysu.edu.cn) and Pierre Gentine (pg2328@columbia.edu)

The copyright of individual parts of the supplement might differ from the article licence.



**Supplementary Figure S1.** Spatial distribution of (a) the flux towers in the FLUXNET2015 and (b) the weather stations used in this study. The plant function types of the flux towers include Croplands (CRO), Deciduous Needleleaf Forests (DNF), Evergreen Needleleaf Forest (ENF), Evergreen Broadleaf Forest (EBF), Deciduous Broadleaf Forest (DBF), Mixed Forest (MF), Grasslands (GRA), Savannas (SAV), Woody Savannas (WSA), Closed Shrublands (CSH), Open Shrublands (OSH), Wetlands (WET), and Snow and Ice (SNO).



**Supplementary Figure S2.** The performance of the ANN model using 2 hidden layers and different number of neurons.



**Supplementary Figure S3.** The performance of a well-trained random forest (RF) model for predicting daily  $\lambda E$  and H based on the same Fluxnet2015 data as the ANN model. The RF model is trained using 500 trees. The test samples are randomly selected from the training data using a ratio of 10%.



**Supplementary Figure S4.** Density scatter plot for the cross-validation in terms of values for evergreen needleleaf forest (ENF), evergreen broadleaf forest (EBF), deciduous broadleaf forest (DBF), mixed forest (MF), and open shrublands (OSH). Samples of the validation set are randomly composed of five flux towers from one plant function type.



**Supplementary Figure S5.** Density scatter plot for the cross-validation in terms of values for savannas (SAV), grasslands(GRA), croplands(CRO), and wetlands(WET). Samples of the validation set are randomly composed of five flux towers from one plant function type.



**Supplementary Figure S6.** Density scatter plot for the cross-validation in terms of trends for different samples from ENF, EBF, DBF, MF, and OSH, respectively. The validation set is randomly composed of one flux tower from one plant function type, and the trends are estimated for all time periods.



**Supplementary Figure S7.** Density scatter plot for the cross-validation in terms of trends for different samples from SAV, GRA, CRO, and WET, respectively. The validation set is randomly composed of one flux tower from one plant function type, and the trends are estimated for all time periods.



**Supplementary Figure S8.** Spatial distribution of correlation coefficient (R) and root mean squared error (RMSE) between the predicted latent heat flux ( $\lambda$ E) (sensible heat flux (H)) and observed  $\lambda$ E (H) at the daily scale for all flux towers.



**Supplementary Figure S9.** Spatial distribution of R and RMSE between the predicted  $\lambda E$  (H) and observed  $\lambda E$  (H) at the monthly scale for all flux towers.



Supplementary Figure S10. Spatial distribution of the mean annual latent heat and sensible heat fluxes.



Supplementary Figure S11. Mean annual evapotranspiration (ET) during 1982–2008.



Supplementary Figure S12. Regionally temporal changes of the  $\lambda E$  estimated by the ANN model in this study and the  $\lambda E$  estimated by the Model Tree Ensemble (MTE) over the regions of >30° North, 30° North-30° South, and global land, respectively.



Supplementary Figure S13. Trends in annual precipitation during different historical periods.



**Supplementary Figure S14.** Global patterns of agricultural irrigation. Global maps show percentage of the area equipped for irrigation that is (a) actually irrigated, (b) irrigated with groundwater and (c) irrigated with surface water. The base year for most countries is the 2000–2008 period. Data are collected from the Food and Agriculture Organization of the United Nations (http://www.fao.org/nr /water/aquastat/irrigationmap/index10.stm).



**Supplementary Figure S15.** Annual changes in (a) EF, (b) ET, and (c) P–ET in Phase 5 of the Coupled Model Intercomparison Project (CMIP5) under the RCP8.5 scenario with all anthropogenic forcing (e.g., land use/land cover changes, aerosols, and ozone). Changes are quantified by difference in years 2070–2099 of simulation and years 1941–1970. The data are collected from 6 models: bcc-csm1-1, CanESM2, CESM1-BGC, GFDL-ESM2M, HadGEM2-ES, NorESM1-ME. Most of the models only one ensemble member is available, so that we consider only one ensemble member per model.



**Supplementary Figure S16.** (a) Trends in observed runoff, (b) global distribution of large reservoirs, and (c) trends in runoff that are not too influenced by reservoir. Observed runoff data were collected from hydrologic stations with a controlled watershed area from 5 to 1000 km<sup>2</sup> with a data length over 20 years in the GRDC database. The right panel in (a) shows runoff trends at different latitudes, and the pink curve represents the median trend. The data of large reservoirs in (b) were collected from GRanD database, and the red dashed boxes show typical areas that the observed runoff presents a decrease trend which is opposite to the increase trend of P–ET. The observed runoff in (c) is selected by the following criteria: (1) 1 latitude/longitude degree away from the reservoirs with a capacity of <1 km<sup>3</sup> and (2) 5 latitude/longitude degree away from the reservoirs with a capacity of <1 km<sup>3</sup>.

## Model training using different variable combinations

Solar radiation is a key factor affecting surface energy and water cycle. However, there is currently a lack of long-term global ground shortwave radiation observation data. Thus, this study uses daily topof-atmosphere shortwave (SW IN POT) which can be estimated by a semi-empirical model as the input of the artificial neural networks (ANN) model. This study used different variable combinations as the input of ANN model to analyze the sensitivity of different variables to the estimation of latent heat flux  $(\lambda E)$  and sensible heat flux (H). To reduce time consumption, here we used a simplified ANN structure with 2 hidden layers and 6 neurons per layer to perform the test of different variable combinations. The results show that SW IN POT, relative humidity (RH), maximum temperature (Tmax), and minimum temperature (Tmin) are the key influencing variables, while mean temperature (Tmean) and mean wind speed (WS) are also important influencing variables (Table S1). Overall,  $\lambda E$  is sensitive to SW IN POT, RH, and WS, and H is sensitive to SW IN POT, Tmax, and Tmin. When the model training includes precipitation, the model shows a better performance in Table S1, but the improvement is limited and the signal of precipitation (P) can be reflected in the variations of relative humidity. In addition, latent heat flux measurements can be biased low during precipitation events (dew). To avoid overlapping information between the runoff assessed by P minus evapotranspiration (ET) and the ET, P is not used as a driving variable of the ANN model. Meanwhile, daily temperature range (DTR) has overlapping information with Tmax and Tmin, and thus it is not used to train the ANN model. All variables and their data sources for training ANN model are listed in Table S2.

	λΕ		Н	
Combination of different variables	R	RMSE (W m <sup>-2</sup> )	R	RMSE (W m <sup>-2</sup> )
{Tmax; Tmin}	0.60	36.22	0.52	43.75
{RH; Tmax; Tmin}	0.66	32.11	0.60	40.97
{RH; Tmean; Tmax; Tmin}	0.67	32.00	0.61	40.06
{RH; Tmax; Tmin; DTR}	0.67	30.89	0.62	39.48
{SW_IN_POT; Tmax; Tmin}	0.70	30.75	0.78	31.81
{SW_IN_POT; RH}	0.70	30.72	0.69	37.01
{SW_IN_POT; RH; Tmean; Tmax; Tmin}	0.74	28.80	0.72	35.8
{SW_IN_POT; RH; Tmean; Tmax; Tmin; WS}	0.75	28.65	0.74	34.72
{SW_IN_POT; RH; Tmax; Tmin; WS}	0.74	28.78	0.73	35.06
{SW_IN_POT; RH; Tmax; Tmin; WS; P}	0.75	28.90	0.75	33.01
{SW_IN_POT; RH; Tmax; WS; Tmin; P}	0.76	28.11	0.75	34.80
{SW_IN_POT; RH; Tmean; Tmax; Tmin; DTR;	0.77	27.12	0.74	34.34
WS; P}				

Table S1. Test results of model training using different variable combinations\*

\*Tmax, Tmin, and Tmean are maximum, minimum, and mean temperature, respectively. RH, DTR, and SW\_IN\_POT are relative humidity, daily temperature range, and top-of-atmosphere shortwave, respectively. WS and P are mean wind speed and total precipitation of the day.

Variables	Units	Data sources	Usage
SW_IN_POT	$W/m^2$	The daily integrated dataset	Input variable
Tmean	°C	The daily integrated dataset	Input variable
Tmax	°C	Half-hourly or hourly data	Input variable
Tmin	°C	Half-hourly or hourly data	Input variable
VPD	hPa	The daily integrated dataset	VPD was used to calculate RH
WS	m/s	The daily integrated dataset	Input variable
λΕ	$W/m^2$	The daily integrated dataset	Output variable
Н	$W/m^2$	The daily integrated dataset	Output variable

Table S2. Variables and data sources for training ANN model\*

\*Vapor pressure deficit (VPD) was used to calculate relative humidity.  $\lambda E$  and H are the latent heat flux and sensible heat flux, respectively.