

Responses to the comments on hess-2020-590

Dear Editor,

We greatly appreciate you for your efforts and time on our manuscript “Long-term relative decline in evapotranspiration with increasing runoff on fractional land surfaces” (#hess-2020-590), and we also thank Dr. René Orth and the other reviewer for providing valuable comments and suggestions on our manuscript. Following those comments and suggestions, we have revised the manuscript carefully. We are pleased to submit the revised version for your consideration with the main revisions listed below:

1. We revised each sentence according to the points raised by the reviewer, and we also carefully double-checked the language of the entire manuscript.
2. We improved our manuscript after seriously considering each comment, and replied to the questions/doubts with providing more experimental information.

We hope that the revised version of our manuscript will satisfy both you and the reviewers. Should you have any queries on our manuscript, please don't hesitate to contact us. Thank you again for time and efforts on our manuscript.

Yours sincerely,

Dr. Ren Wang and Dr. Pierre Gentine (on behalf of the authors' team)

Email: wangr67@mail2.sysu.edu.cn (R.W.); pg2328@columbia.edu (P.G.)

No.1, Wenyuan Road, Xianlin University District, Nanjing 210023, China (R.W.);

New York, NY 10027, USA (P. G.)

Response to the reviewer

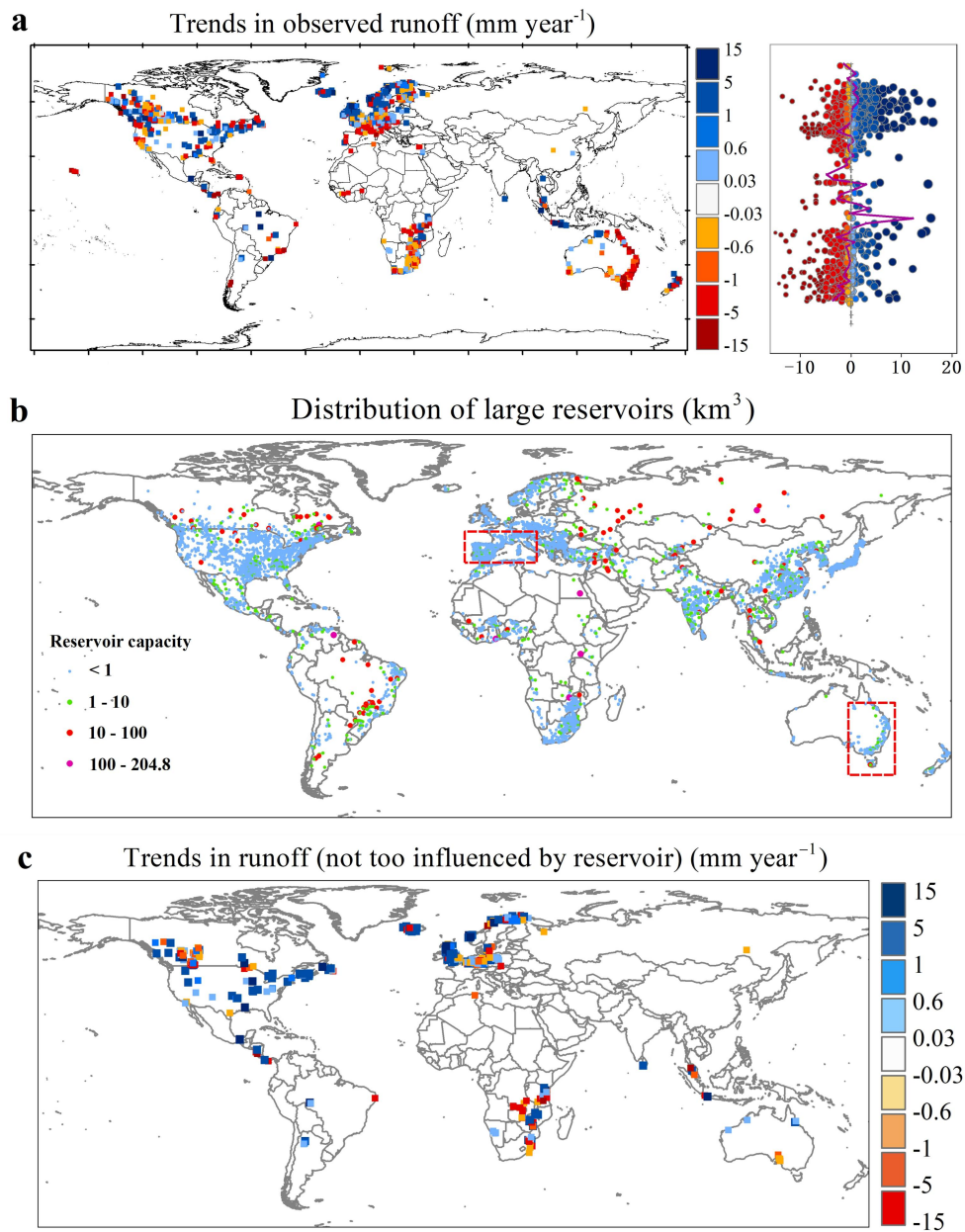
The paper has overall improved as the authors have addressed some of the concerns raised by me and the other reviewer. However, at the same time important issues remain unresolved:

Comment (1): As pointed out in main comment (1) of my previous review, the validation of the inferred time series against independent data such as the Fluxnet-MTE product, and observed streamflow is not convincing. The differences are not satisfactorily explained by the authors in their rebuttal. Figure S11 clearly shows diverging trends in the tropics, and the author's explanation about "different models, forcing data and time scales" does not provide evidence why their product (and trends therein) is more reliable than the Fluxnet-MTE reference product. Moreover, the different runoff trends between Figures 4c and S14 cannot be entirely explained with human influence on the observed streamflow, I think, as even the large-scale patterns differ clearly, for example in Europe and Australia. Given that the differences between the products here are partly so remarkable I feel that they need to be understood better in order to illustrate the reliability of the newly derived dataset.

Reply: We thank the reviewer for this valuable comment. (1) We believe that the more important significance of our study is that it can provide an alternative choice of latent heat and sensible heat fluxes, which have its unique feature that it is purely driven by observations and can meet the needs for long-term trend analysis. Given that the Fluxnet-MTE assumes stationarity in the relationship between fluxes and environmental drivers (e.g. CO₂ effects are not taken into account), a direct comparison is difficult. In fact, this gap in the study of long-term trends with current products was the very reason behind our new product. Yet, we acknowledge that the well-trained ANN model and the retrieved results of our study can be an important supplement to the existing product.

(2) We provided more experimental information on the global distribution of large reservoirs which is related to the key impacts from human activities, e.g., reservoir scheduling and agriculture irrigation (Lehner et al., 2011) (Supplementary Fig. S16b). As suggested by the reviewer, Australia and Europe are regions where the observed runoff shows a decrease trend and is contrary to the retrieved P-ET trend. These regions are also distributed with a

number of large reservoirs. When we only considered the stations that are not too influenced by large reservoirs, we found that the direction and the spatial pattern of P-ET trend (Fig. 4c) are more obvious consistent with observed runoff trend, including the upward trend in northern Australia and the downward trend in southern Australia, the upward trend in western Europe and the downward trend in eastern Europe (Supplementary Fig. S16c). The spatial pattern of P-ET trends and observed runoff trends are also generally consistent in other regions including North and South America, Southern Africa, East Asia, and Southeast Asia.



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

km² 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 km³ and (2) 5 latitude/longitude degree away from the reservoirs with a capacity of >1 km³.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rodel, R., Sindorf, N., and Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Front. Ecol. Environ.*, 9: 494–502. <https://doi.org/10.2307/23034466>, 2011.

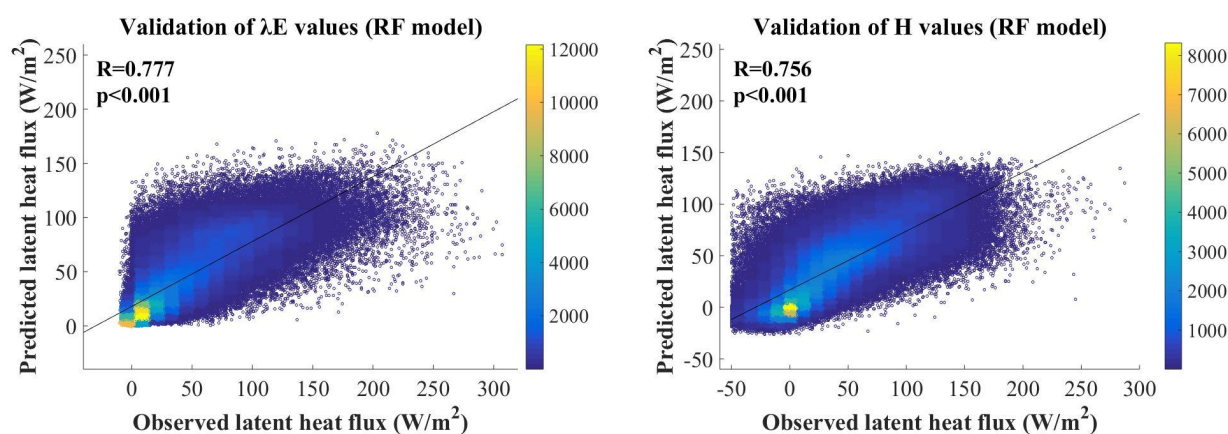
Comment (2): As pointed out in main comment (2) of my previous review, different time periods are considered during which trends from the data derived in this study are determined versus the trends from the CMIP5 data. The authors acknowledge that matching time periods would be "better, but difficult to achieve", and state that this might not be so important as the evaluation focuses on directions of change rather than the actual magnitude. I do not agree with this. Even the sign of change may be different in different time periods, as for example introduced by decreasing and later increasing radiation within the global dimming and brightening periods. Further, decadal climate variability can introduce varying trends over time. Therefore, I think that with using different time periods, the comparison of trends between the data derived in this study and the CMIP5 data is not meaningful.

Reply: Thanks for the considerable comment. The prediction results of CMIP5 are auxiliary material for analyzing the phenomenon found in this study. After considering this comment, we placed the Fig. 6 to the supplementary material in the revised manuscript. We think the more important purpose here is to check the trend of EF changes, and the meaning of these comparisons lies in examining whether ET or EF show the similar trends in history and in the future CO₂ rising scenario.

Comment (3): Even though I requested this in main comment (4) of my previous review, the

authors did not explain why they chose to employ artificial neural networks over alternative machine learning approaches.

Reply: We thank the reviewer for this valuable comment. We added an comparison with another machine learning algorithm, i.e., random forest (RF). The RF model was trained for predicting daily λE and H based on the same Fluxnet2015 data as the ANN model. The RF model showed very similar performance to the ANN model. The correlation coefficients of RF model in predicting daily λE and daily H are 0.777 ($p < 0.001$) and 0.756 ($p < 0.001$), respectively (Supplementary Fig. S3). The correlation coefficients of ANN model in predicting daily λE and daily H are 0.849 ($p < 0.001$) and 0.743 ($p < 0.001$), respectively (Fig. 2). We want to point out that our retrievals are based on daily scale and the existing studies are based on monthly scale, and thus the values of correlation analysis can be different. Therefore, it is feasible to use the neural network algorithm to retrieve the latent heat and sensible heat fluxes.



Supplementary Figure S3. The performance of a well-trained random forest (RF) model for predicting daily λE and daily 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%.

The following explanations have been added in the revised manuscript and were copied here:

“The artificial neural networks (ANN) have been shown to be powerful regression algorithm, and unlike other machine learning algorithms, ANN can build multi-layer and multi-node network models to achieve deep learning of a complex simulation. Pure ANN

model has been proven to show good performance in retrieving surface fluxes (Chen et al., 2020; Haughton et al., 2018; Zhao et al., 2019).” (In the section 2.4)

“Moreover, we trained a random forest (RF) model for predicting daily λE and H based on the same Fluxnet2015 dataset as the ANN model. The RF model shows very similar performance to the ANN model. The correlation coefficients of RF model in predicting daily λE and daily H are 0.777 ($p < 0.001$) and 0.756 ($p < 0.001$), respectively (Supplementary Fig. S3). Therefore, it is feasible to use the neural network algorithm to retrieve the latent heat and sensible heat fluxes.” (In the section 3.1)

Comment (4): Even though I mentioned this in main comment (5) of my previous review, there are still many small language errors throughout the manuscript, for example in lines 71, 77, 131, 155, 159, 196/197, 205, 224, and 249.

Reply: Thanks for the careful review. We have carefully double-checked the language in those lines and all expressions in the manuscript.

Comment (5): One additional issue caught my attention this time, sorry for realizing this only now. The ANN model is trained with target data from flux tower sites which typically cover 5-15 years. I doubt that CO₂ in such short time periods would change sufficiently to lead to an emerging influence of CO₂ fertilization on ET or EF. Therefore I think the reasoning of the CO₂ fertilization effect on stomatal conductance potentially explaining differences in the partitioning between ET and runoff over time, which can be found throughout the manuscript, is not valid. I do not wish to remain anonymous - René Orth.

Reply: We appreciate this valuable point raised here. The effect of CO₂ fertilization dynamically affects the leaf area and physiology of vegetation, thereby changing the partitioning of latent heat and sensible heat fluxes. We completely agree with the reviewer for FLUXNET-MTE but not for our model. A major advantage of our method compared to other methods (e.g. FLUXNET-MTE) is that it does not rely on any assumption on a CO₂ effect on the link between environmental variables and fluxes. Indeed, we flipped the strategy around its head by diagnosing the diurnal changes in temperature and humidity in the boundary layer. As such this diurnal cycle reflects naturally any change in CO₂. For instance, if stoma were to

substantially close they would increase the sensible heat flux and reduce latent heat flux. This would in turn lead to increased temperature diurnal range and reduced atmospheric humidity in the boundary layer (Salvucci and Gentine 2013, Gentine et al. 2016). Therefore, this CO₂ effect is completely detectable with our method. We believe this is a major advantage of our method based on a boundary layer energy budget, as the physics of the boundary layer does not change (fluid dynamics).

Besides, according to Wang et al. (2020), which also used 22 flux towers with a time period over 14 years, the CO₂ fertilization effect has changed. This means that the observation period of the Fluxnet data used in this study is long enough.

Salvucci, G. D. and Gentine, P.: Emergent relation between surface vapor conductance and relative humidity profiles yields evaporation rates from weather data. *Proc. Natl. Acad. Sci. USA*, 110, 6287–6291, <https://doi.org/10.1073/pnas.1215844110>, 2013.

Gentine, P., Chhang, A., Rigden, A., and Salvucci, G.: Evaporation estimates using weather station data and boundary layer theory. *Geophys. Res. Lett.*, 43, 11661–11670, <https://doi.org/10.1002/2016GL070819>, 2016.

Wang, S. H., Zhang, Y. G., Ju, W. M., et al: Recent global decline of CO₂ fertilization effects on vegetation photosynthesis. *Science*, <https://doi.org/10.1126/science.abb7772>, 2020.

Comment (6): lines 66-67: not clear what is meant with "boundary layer budget" and "non-linearity"

Reply: The “boundary layer budget” has been revised to “boundary layer energy budget” and the “non-linearity” has been revised to “non-linear effects of changing environment factors on surface energy fluxes”. (Line 65-67).

Comment (7): lines 70-74: The authors state that existing datasets are somewhat biased towards the (flux) data availability which is largely concentrated in the northern hemisphere. This is true, but it is also true or the data derived in this study even if the weather stations are more distributed, because the model training is confined to the flux tower sites, as for the existing products.

Reply: Thanks for the comment. We agree with the point raised by the reviewer here, but the purpose here is to emphasize the weather station network are more widely distributed, ground-based observation, and long-term availability. One thing we want to emphasize again is that the physics of the boundary layer is consistent across regions unlike the flux to environmental response (which varies for instance with plant functional type, soil type etc). As such, the dataset we have is more representative of generic and global conditions. After considering this comment, we have modified the related expressions that existing products are based on highly localized data.

Comment (8): lines 156-157: Why was exactly this combination of input variables chosen while other combinations showed even better performance in Table S1, for example after including precipitation?

Reply: The combination of variables is based on the various environmental factors that reflect the changes in latent heat and sensible heat fluxes and the test results of the model training. When the training includes precipitation, the model shows a better performance in Table S1, but the improvement is limited and the signal of precipitation can be reflected in the variation of air humidity. In addition, latent heat flux measurements can be biased low during precipitation events (dew). As clarified in the Supplementary Materials (Page 13), to avoid overlapping information between the runoff assessed by P-ET and the retrieved ET, P is not included.

Comment (9): lines 165-166: What do you mean with initial tests and how did that influence the final decision? How was this formula applied to guide these decisions?

Reply: We meant that this empirical formula can provide a reference for us to determine a good neuron number when training our ANN model, and it can reduce the possibility of overfitting.

Comment (10): lines 204-213: Why are there different p-value thresholds used here?

Reply: The p-value is determined by the significance level of the correlation, and thus p-value of correlation analysis can be different. The statistical significance levels 0.001 and 0.05 are commonly used critical values.

Comment (11): line 220: Why >12000?

Reply: This is the magnitude of the sample number for correlation analysis. The value of the correlation coefficient can be affected by the number of samples.

Comment (12): lines 220-225: How are these trends computed? As linear trends? Over which time period?

Reply: The trends are computed using linear trend estimation. The time period is determined by the availability of the flux tower observed latent heat and sensible heat fluxes.

Comment (13): line 241: Why are you using the 11-year old Fluxnet-MTE product from Jung et al. 2010 rather than the more recent Fluxcom product from Jung et al. 2019?

Reply: The Fluxnet-MTE is a mature and widely applied machine learning product that can be used as a bench work. The MTE is based on only one machine learning method trained on monthly flux data and may be regarded as a precursor to FLUXCOM which comprises 147 products in two setups: (1) 0.0833° resolution using MODIS remote sensing data (RS) and (2) 0.5° resolution using remote sensing and meteorological data (RS + METEO).

As our study only use one machine learning method and one data-driven source (weather station observations), which is more similar to the data-driven methods of Fluxnet-MTE, so we use Fluxnet-MTE as a reference. We also want to point out that when we started this analysis FLUXCOM was not readily available (2018).

Comment (14): line 298: Which CMIP5 models are used in the considered ensemble, and how were their simulations aggregated?

Reply: We used data from 6 models: bcc-csm1-1, CanESM2, CESM1-BGC, GFDL-ESM2M, HadGEM2-ES, NorESM1-ME. Most of the models only have one ensemble member

available, so that we consider only one ensemble member per model. These detailed information have been provided in the revised manuscript.

Comment (15): Figures 3-7: How did you determine the area of the Sahara? References: Jung, M., et al. The FLUXCOM ensemble of global land-atmosphere energy fluxes, *Scientific Data* 6, 74 (2019). Jung, M., et al. Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature* 467, 951–954 (2010).

Reply: As stated in the revised manuscript, the area of the Sahara is based on Vicente-Serrano et al., (2015). We have provided this information when it first appeared (Fig. 3).

Vicente-Serrano, S. M., Van Gerard, V. D. S., Beguería, S., Azorin-Molina, C., and Lopez-Moreno, J. I.: Contribution of precipitation and reference evapotranspiration to drought indices under different climates. *J. Hydrol.*, 526, 42–54. <http://dx.doi.org/10.1016/j.jhydrol.2014.11.025>, 2015.

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

Ren Wang^{1,2,3}, Pierre Gentine^{4,5}, Jiabo Yin⁶, Lijuan Chen^{1,2}, Jianyao Chen^{7,8}, and Longhui Li^{1,2,3}

¹Key Laboratory of Virtual Geographical Environment (Nanjing Normal University), Ministry of Education, Nanjing, 210023, China

²School of Geographical Sciences, Nanjing Normal University, Nanjing, 210023, China

³Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, 210023, China

⁴Earth and Environmental Engineering Department, Columbia University, New York, NY 10027, USA

10 ⁵Earth Institute, Columbia University, New York, NY 10025, USA

⁶State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, 430072, China

⁷School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510275, China

⁸Guandong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-sen University, Guangzhou, 510275, China

15

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

Abstract. Evapotranspiration (ET) accompanied by water and heat transport in the hydrological cycle is a key component in regulating surface aridity. Existing studies documenting changes in surface aridity have typically estimated ET using semi-empirical equations or parameterizations of land surface processes, which are based on the assumption that the parameters in the equation are stationary. However, plant physiological effects and its responses to a changing environment are dynamically modifying ET, thereby challenging this assumption and limiting the estimation of long-term ET. In this study, the latent heat flux (ET in energy units) and sensible heat flux were retrieved for recent decades on a global scale using machine learning approach and driven by ground observations from flux towers and weather stations. This study resulted in several findings, that is, the evaporative fraction (EF)—the ratio of latent heat flux to available surface energy—exhibited a relatively decreasing trend on fractional land surfaces; in particular, the decrease in EF was accompanied by an increase in long-term runoff as assessed by precipitation (P) minus ET, accounting for 27.06% of the global land areas. The signs are indicative of reduced surface conductance, which further emphasizes that ~~land~~-surface vegetation has major impacts in regulating water and energy cycles, as well as aridity variability.

20

25

30 1 Introduction

Evapotranspiration (ET) mainly includes two processes: (1) evaporation from soil and plant surfaces and (2) transpiration from plants to the atmosphere (Miralles et al., 2020). These processes connect the transfer of moisture and energy in soil, vegetation, and atmospheric systems (Salvucci et al., 2013; Yang et al., 2020). Quantifying changes in the exchange of moisture and heat between the land and atmosphere is very important for understanding and characterizing water and energy cycles, which has implications in various fields such as hydrology, climatology, and agronomy (Hoek van Dijke et al., 2020; Gentine et al., 2016; Komatsu and Kume, 2020).

ET is expected to intensify with the warming climate, thereby contributing to the increase in surface aridity stress (Baruga et al., 2020; Berg et al., 2016; Fu et al., 2014; Trenberth et al., 2014). However, quantification of changes in aridity/wetness is usually derived from traditional drought indices such as the Standardized Precipitation Evapotranspiration Index (Vicente-Serrano et al., 2015), which is embedded with a semi-empirical equation, such as the Thornthwaite equation or Penman–Monteith equation, for ET estimation (Dai et al., 2013; Sheffield et al., 2012). Using potential evaporation rather than actual ET or calculating offline ET using meteorological variables from climate model outputs in traditional drought indices, the calculation implicitly assumes that soil can always supply moisture to meet the atmospheric evaporation demand, which is an incorrect assumption for most land surfaces (Greve et al., 2014; Milly and Dunn, 2016; Yang et al., 2020). Moreover, when using a semi-empirical equation for ET estimation, some parameters such as soil surface resistance and stomatal resistance, are assumed to be stationary over time; however, we know that these parameters are dynamically changing with environmental conditions (Miralles et al., 2011; Yang et al., 2019; Zhou et al., 2016).

Why are the soil surface resistance and stomatal resistance not stationary? Changes in plant stomata and leaf area, with increasing CO₂ concentrations in particular, reshape the allocation of surface energy fluxes and affect plant transpiration (Forzieri et al., 2020; Sorokin et al., 2017; Mallick et al., 2016; Williams and Torn, 2015). With increasing CO₂ concentrations, the density and opening degree of leaf stomata decrease, while the water-use efficiency and biomass production of plant increase, which can modify vegetation transpiration and even affect soil moisture or surface runoff (Keenan et al., 2013; Massmann et al., 2019; Orth and Destouni, 2018; Rigden et al., 2016; van Der Sleen et al., 2015; Wagle et al., 2015). Vegetation transpiration occupies most of ET amount, so vegetation control effects can greatly alter the variability of land surface ET (Costa et al., 2010; Jaramillo et al., 2018; Wei et al., 2017; Williams et al., 2012). Moreover, human activities including agricultural irrigation and land use management, are constantly altering the exchange of water and heat between terrestrial ecosystems and the atmosphere (Padrón et al., 2020; Teuling et al., 2019). When these effects are taken into account, the semi-empirical equations for estimating ET and traditional drought indices also face challenges (Yang et al., 2020). Existing studies with respect to global surface fluxes inferred from flux tower observations, remote sensing products, and reanalysis data, e.g., the [resultssurface fluxes estimateddriven](#) by model tree ensemble ([Fluxnet-MTE](#)), rely on the satellite era and instantaneous meteorological observations (Jung et al., 2010; Jung et al., 2011; Miralles et al.,

2013). Thus, the existing products cannot be used for long-term trends as they cannot represent the long-term effects of confounders such as CO₂ or species composition changes. This is why we use an opposite view – we use in essence a boundary layer energy budget (~~based on~~ Salvucci and Gentine, (2013) ~~and~~; Gentine et al., (2016) except that we lump non-linear effects of changing environment factors on surface energy fluxes in a neural network. Indeed, the diurnal cycle of temperature is directly related to ~~the~~ sensible heat flux, ~~and~~. Similarly, the course of specific humidity related to the rate of latent heat flux variation (Gentine et al., 2011). If there are changes in latent heat flux due to vegetation in response to higher CO₂, this is still captured by the change in the specific humidity.

In this study, we propose a new strategy for estimating latent heat flux (λE) (ET in energy units) and sensible heat flux (H) using machine learning approach and ground observations from flux towers and weather stations. This strategy utilizes daily observations of meteorological variables such as temperatures, humidity, and solar radiation. A major advantage of such retrieval is that it does not rely on any assumption on a CO₂ effect on the link between environmental variables and fluxes. Indeed, we flipped the strategy around its head by diagnosing the diurnal changes in temperature and humidity in the boundary layer. As such this diurnal cycle reflects naturally any change in CO₂. For instance, if stomata were to substantially close they would increase H and reduce λE . This would in turn lead to increased temperature diurnal range and reduced air humidity in the boundary layer (Salvucci and Gentine 2013, Gentine et al. 2016). Therefore, this CO₂ effect is completely detectable. This is a major advantage of our method based on a boundary layer energy budget, as the physics of the boundary layer does not change (fluid dynamics). Moreover, the observational record of the weather station network is not only longer, but also extends to more remote places, such as the tropics. This study also employed the evaporative fraction (EF), i.e., the ratio of λE to the sum of λE and H, and a proxy for long-term runoff, i.e., the difference of precipitation (P) and ET (P-ET), to quantify the change in aridity/wetness.

2 Observational data and methodology

2.1 Flux tower observational data

We collected the half-hourly/hourly observational data and the integrated daily product from the FLUXNET2015 FULLSET dataset (Pastorello et al., 2020). To control the quality of the observational dataset, this study only used measurements and good-quality gap-filled data from 212 globally distributed flux towers (Supplementary Fig. S1a). The flux towers used in this study across various climate regions and land cover types (Fig. 1). The longest period of data availability is 22 years. This study intended to build a machine learning model for retrieving latent heat and sensible heat fluxes on a daily scale. Therefore, daily-scale data of top-of-atmosphere shortwave radiation, vapor pressure deficit (VPD), mean temperature, and surface wind speed were collected from the integrated daily product. VPD was used to calculate relative humidity. Daily maximum and minimum temperatures were obtained from the half-hourly/hourly flux tower measurement data. Moreover, daily-scale λE and H were also collected from the integrated daily product. The underlying surfaces of the flux towers

covered different plant function types (PFTs). According to the classification scheme of International Geosphere-Biosphere Programme, the PFTs 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). These flux tower observation data across different ecosystems are used to train and build machine learning model for the retrieval_s of latent heat and sensible heat fluxes.

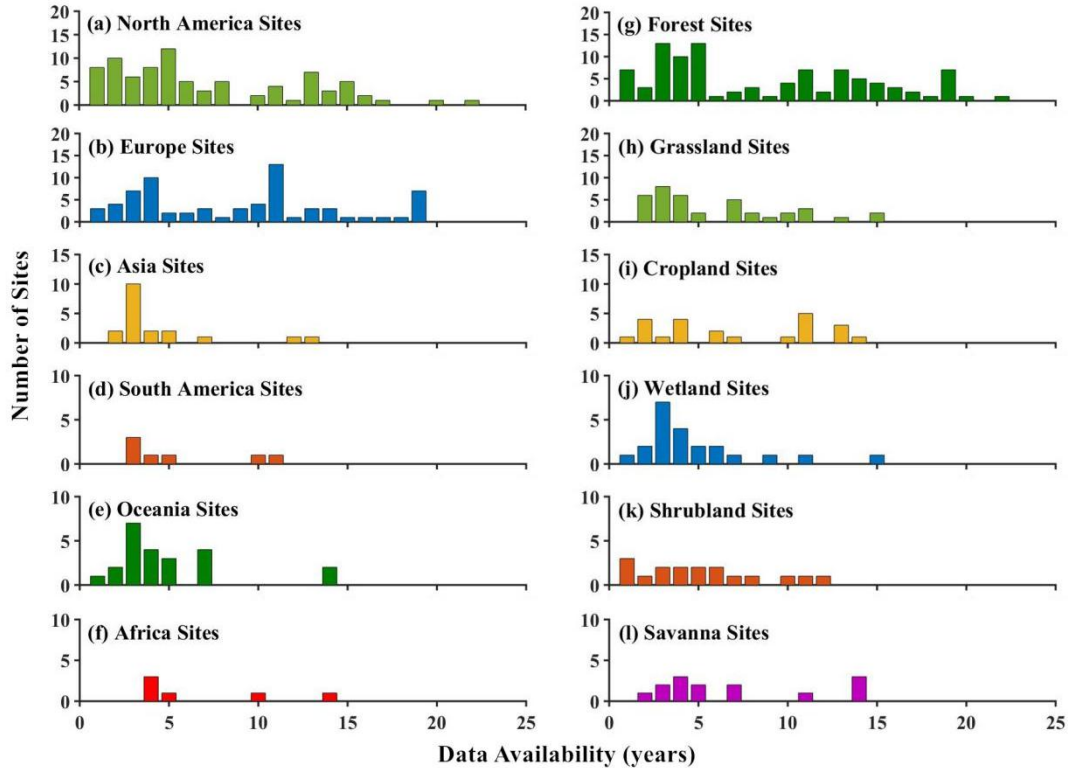


Figure 1. Data summary of the flux towers used in this study.

2.2 Weather station observation data

105 Daily observational records of precipitation (P), temperature (mean, maximum, and minimum temperatures), dew point temperature, and wind speed at weather stations were collected from the Global Summary of the Day (GSOD) during the 1950–2017 period. Dew point temperature data were used to calculate the relative humidity, and the daily weather station data on the global land were used to drive a well-trained machine learning model to retrieve [surface](#) fluxes. The quality of the data was controlled through several procedures (Durre et al., 2010; Matsuura et al., 2009; Yin et al., 2018). First, we

110 divided the weather stations into two groups: the original stations and the target stations. We used 20 048 sites in total as the original station group (Supplementary Fig. S1b). The target station group was obtained according to the following steps: (1)

The stations with a time series spanning less than ten years in length were excluded; (2) if the stations had the same geographic coordinates, we used the stations with a long observation record to replace the stations with a short observation record; (3) if there were multiple stations having different coordinates in a 0.1-degree grid, we removed the stations with a short observation record. After filtering, the target station group which were determined to be used to estimate long-term trends were obtained.

Other procedures for controlling data quality were also implemented. Any implausible values, such as negative precipitation or maximum temperature lower than the minimum temperature on that day, were excluded. Monthly mean, maximum, and minimum temperatures, as well as monthly precipitation were derived from daily observational data at the original stations. Considering the large uncertainty in the observational data of precipitation, we also compiled the daily precipitation records with precipitation records in another archives, i.e., the Global Historical Climatology Network (GHCN-Daily). The daily records of ~~all~~ weather stations in the GSOD that had the same coordinates as the GHCN-Daily were compared, and the missing daily records were supplemented using the GHCN-Daily archives. Monthly precipitation, temperatures (mean, maximum, and minimum temperatures), relative humidity, and surface wind speed were calculated when the number of missing days within a month was no more than seven days. Additionally, missing monthly data from the target stations were spatially interpolated from the original weather stations using the Kriging method.

2.3 Top-of-atmosphere shortwave radiation model

Solar shortwave radiation is a key factor affecting surface energy and water cycles. Since there is no reliable long-term observational solar radiation data, this study uses shortwave radiation at the top of the atmosphere (top-of-atmosphere shortwave radiation) as a replacement. Cloud effects are inherently captured by the diurnal cycle of temperature and humidity (Gentine et al., 2013a,b). Daily top-of-atmosphere shortwave radiation converted from the hourly top-of-atmosphere shortwave radiation was forced to drive the ~~ANN~~ model for predicting the daily λE (H) at the target weather stations. The amount of incoming shortwave radiation at any location/time at the top of atmosphere is a function of Earth–Sun geometry, which is defined as: i) latitude (i.e., location); ii) hour of day (due to the rotation of the earth); and iii) day of year (due to the tilted axis of the earth and its elliptical orbit around the sun). Several models for the top-of-atmosphere fluxes based on these inputs are available at varying levels of precision. The time-location model (Margulis, 2017) used in this study is shown as follows.

$$R_{s,0} = \begin{cases} I_0 \frac{\cos \theta_0}{d^2}, & \text{daytime : } |\theta_0| \leq 90^\circ \\ 0, & \text{nighttime} \end{cases} \quad (1)$$

where the cosine of the solar zenith angle is as follows:

$$\cos \theta_0 = \sin \delta \sin \lambda + \cos \delta \cos \lambda \cos \tau \quad (2)$$

$$\delta = \frac{23.45\pi}{180} \cos\left[\frac{2\pi}{365}(172 - DOY)\right] \quad (3)$$

$$\tau = 2\pi \frac{T_h - 12}{24} \quad (4)$$

$$d = 1 + 0.017 \cos\left[\frac{2\pi}{365}(186 - DOY)\right] \quad (5)$$

145 Here, θ_0 is the solar zenith angle, δ is the declination angle, λ is latitude, τ is the hour angle, DOY represents the day of year, d represents the distance between the sun and Earth normalized by the mean distance, and T_h represents solar hour of the day.

2.4 Artificial neural network model training

The artificial neural networks (ANN) have been shown to be powerful non-linear regression algorithm, and unlike other machine learning algorithms, ANN can build multi-layer and multi-node network models to achieve deep learning of a complex simulation. Pure ANN model has been proven to show good performance in retrieving surface fluxes (Chen et al., 2020; Haughton et al., 2018; Zhao et al., 2019). In this study, we trained a multi-layer feedforward neural network model that consisted of an input layer, hidden layers, and an output layer to predict daily λE and H at the globally distributed weather stations. To identify the sensitivities of latent heat and sensible heat fluxes to different variables in the retrieval, we used different variable combinations to train ANN model and tested the changes in the model performance (Supplementary Table S1). Top-of-atmosphere shortwave radiation, relative humidity, wind speed, and the mean, maximum, and minimum temperatures were determined to be the inputs of the neural network (Supplementary Table S2).

In the process of training ANN model, input data were randomly divided into three subsets using the percentages of 80%, 10%, and 10% for training, validation, and testing, respectively. Mean squared error (MSE) was used to evaluate the performance of the neural network in the training process of adjusting weight. The rRMSE and Pearson correlation coefficient (R) between the ANN predicted λE (H) and the observed λE (H) in the validation set were used to evaluate the retrieval performance of the well-trained ANN model. A neural network with 2 hidden layers can achieve the same performance as with a large number of hidden layers, so we used the lowest complexity model and enhanced its nonlinear ability by adding neurons. As for the optimal number of neurons, we initially tested it according to an empirical formula, i.e., $h = \sqrt{(n + m)} + a$ (n is the number of input neurons, m is the number of output neurons, and a is a constant ranging from 0 to 10). This empirical formula can provide a reference for us to choose the number of neurons when training neural network, and it can reduce the possibility of overfitting. The neural network was determined to have two hidden layers and 15 neurons per hidden layer, and the ANN model showed good performance and appropriate training time (Supplementary Fig. S2). A tangent sigmoid transfer function was used in the hidden layers, and a linear transfer function

was used in the output layer. To avoid over-fitting, the early stopping method was used, that is, we recorded the best validation accuracy during the training process, and the training was stopped when the MSE was no longer reduced after going through additional epochs. The maximum number of training epochs and training accuracy goal were set to 500 epochs and 0.0001, respectively. Once one of the parameters exceeded the threshold, the ~~model~~-training was stopped.

175 2.5 EF linked to surface resistance (r_s) and aerodynamic resistance (r_a)

Here, we show that a long-term decline in EF can be strongly impacted by an increase in surface resistance (r_s). The latent heat flux ($L_v E$) is expressed by the ~~following~~ formula:-

$$L_v E = L_v \rho \frac{e_{sat}(T_s) - e_a}{r_a + r_s} \quad (6)$$

where L_v is the latent heat of vaporization, E is evaporation flux, ρ is air density, T_s is near-surface air temperature, $e_{sat}(T_s)$ is saturated vapor pressure at the surface, e_a is actual vapor pressure, r_a is aerodynamic resistance, and r_s is surface resistance. EF can be expressed as follows.

$$EF = \frac{L_v E}{L_v E + H} = \frac{L_v \rho \frac{e_{sat}(T_s) - e_a}{r_a + r_s}}{L_v \rho \frac{e_{sat}(T_s) - e_a}{r_a + r_s} + H} \quad (7)$$

We used the linearized Clausius–Clapeyron relation (Eq. (8) and Eq. (9)) to simplify Eq. (7).

$$e_{sat}(T_s) = e_{sat}(T_a) + \Delta(T_s - T_a) \quad (8)$$

$$185 \quad T_s - T_a = \frac{H}{\rho c_p} \quad (9)$$

Where T_a is the air temperature, $e_{sat}(T_a)$ is saturated vapor pressure of the air, and $\Delta = \frac{L_v}{R_v} \frac{e_s}{T^2}$, R_v is the gas constant

for water vapor. Furthermore, c_p is the specific heat capacity, which is 4216 J kg⁻¹ K⁻¹ when the temperature is 0 °C.

$$EF = \frac{\frac{L_v \rho}{r_a + r_s} ((e_{sat}(T_a) - e_a)) + \frac{\Delta H}{\rho c_p}}{\frac{L_v \rho}{r_a + r_s} ((e_{sat}(T_a) - e_a)) + H} \quad (10)$$

$$= \frac{\frac{L_v \rho}{r_a + r_s} \left\{ VPD + \frac{\Delta H}{\rho c_p} \right\}}{\frac{L_v \rho}{r_a + r_s} \left\{ VPD + \frac{\Delta H}{\rho c_p} \right\} + H} \quad (11)$$

$$= \frac{1}{1 + \frac{r_a + r_s}{L_v \rho \left(\frac{\Delta}{\rho c_p} + \frac{VPD}{H} \right)}} \quad (12)$$

The incremental variation of $\frac{VPD}{H}$ is small because both variations of VPD and H are proportional to the temperature variation. EF can be expressed as follows:

$$\frac{1}{EF} = 1 + \frac{r_a + r_s}{\frac{L_v}{c_p} \Delta} \quad (13)$$

Hence, r_s is a function of EF.

$$r_s = \frac{L_v}{c_p} \Delta \left(\frac{1}{EF} - 1 \right) - r_a \quad (14)$$

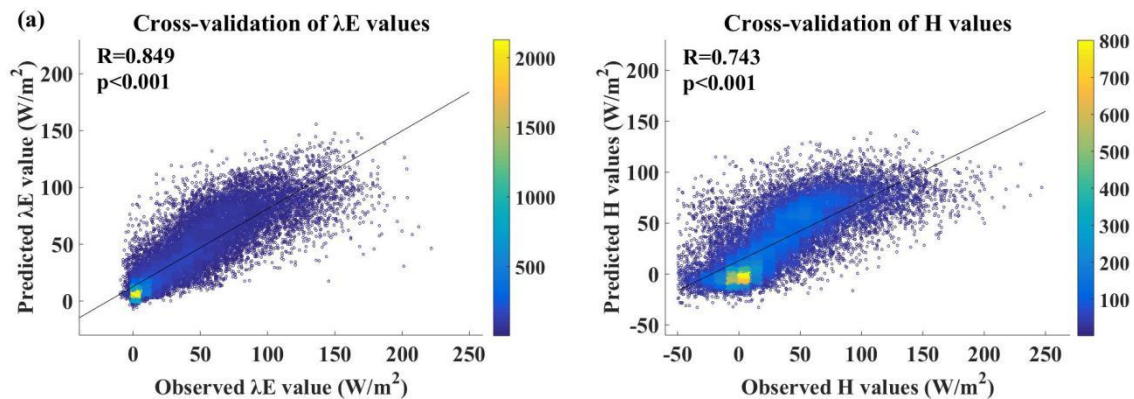
Annual EF ranges from 0 to 1, and EF is closely connected with surface resistance and aerodynamic resistance. A decline in EF can be induced by an increase in surface resistance. r_a is a function of wind speed, and the variation in r_a is with a relatively small ~~variations~~ while the variations in r_s can be strong.

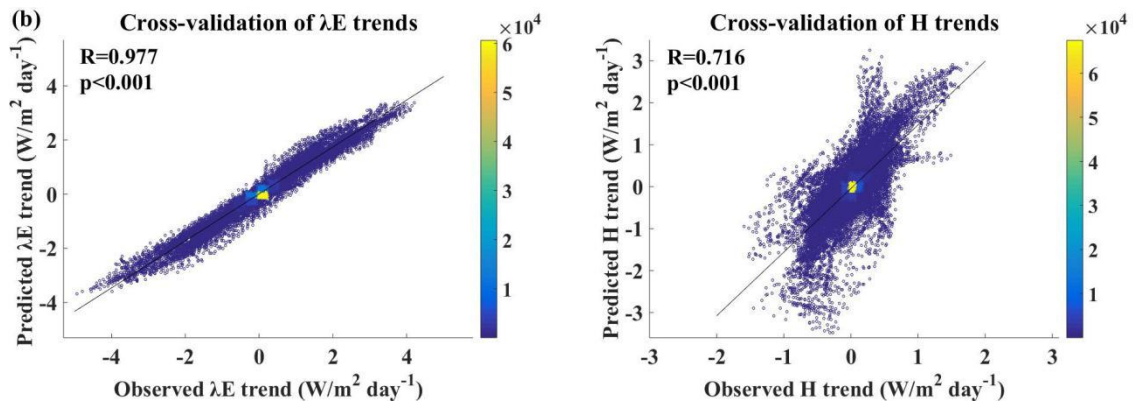
3 Results and discussion

200 3.1 ANN model retrievals

Cross-validations of the ANN model were performed in terms of values and trends. We randomized samples of 10 randomly chosen flux towers from different PFTs as the validation set, and then used the remaining samples to train the ANN model. The ~~ANN~~-predicted daily λE (H) values of the validation set were compared with their observed values (Fig. 2a). The R between predicted daily λE and ~~the~~-observed daily λE is 0.849 and the R between predicted daily H and observed daily H is 205 0.743, and both correlations are significant at the $p < 0.001$ level. Moreover, we trained a random forest (RF) model for predicting daily λE and H based on the same Fluxnet2015 dataset as the ANN model. The RF model shows very similar performance to the ANN model. The correlation coefficients of RF model in predicting daily λE and daily H are 0.777 ($p < 0.001$) and 0.756 ($p < 0.001$), respectively (Supplementary Fig. S3). Therefore, it is feasible to use the neural network algorithm to retrieve the surface fluxes. Cross-validations were also performed in different land covers (Supplementary Fig.

210 S4). The abilities of the trained ANN model for predicting latent heat and sensible heat fluxes were different for various PFTs. With the exception of OSH ($R=0.680$, $p<0.05$), the R of ~~the~~ daily λE of DBF, MF, SAV, GRA, CRO, and WET were all greater than 0.80, and all correlations were ~~statistically~~ significant at the $p<0.001$ level. A common feature of these ~~land-cover types~~ PFTs is that they belong to the ecosystems with relatively open water bodies or high vegetation coverage, while the OSH is mixed with vegetation and bare soil and thus the vegetation coverage is highly heterogeneous. Therefore, the R at
215 OSH was relatively low ($R=0.680$), but the correlation was ~~also~~ significant at the $p<0.05$ level. With respect to ~~the~~ daily H , the correlation coefficients for all PFTs were greater than 0.716 with the exception of R for CRO ($R=0.656$, $p<0.05$), and all were statistically significant at the $p<0.001$ level. In addition, the trained ANN model also shows good simulation ability under some other ecosystems with relatively sparse vegetation cover such as savannas (SAV), grasslands (GRA), croplands (CRO), and wetlands (WET) (Supplementary Fig. S54). In summary, in addition to OSH, the accuracy of retrieving λE is
220 relatively high in GRA, CRO, WET, and various forest ecosystems, and these ecosystems were characterized by sufficient water supply or dense vegetation ~~cover~~. For the estimation of H , except for the estimation of H in GRA, the correlations of predicted and observed H at all ecosystems are correlated at the $p<0.001$ level, especially in forest ~~ecosystems~~. It needs to be emphasized that the magnitude of R could be affected by the number of samples, and the sample numbers of each verification set in those cross-validations are large. As for the prediction of trends in λE and H , the ANN model also shows
225 good performance (Fig. 2b). All correlation coefficients between the estimated λE (H) trends and the observed λE (H) trends exceeded 0.90 ($p<0.001$) ~~over~~ ENF, DBF, GRA, and WET, and the correlations ~~over~~ MF, OSH, and CRO exceeded 0.80 ($p<0.001$), and the ~~correlations~~ correlation coefficients are greater than ~~surpassed~~ 0.70 ($p<0.001$) ~~in~~ EBF and SAV (Supplementary Fig. S65 and Fig. S76). In most causes, the estimations of λE (H) trends are more reliable than the retrieved λE (H) values.





230 **Figure 2.** Density scatter plot for (a) the cross-validation in terms of values and (b) the cross-validation in terms of trends. The validation set of values cross-validation is composed of 10 flux towers randomly selected from different plant function types, and the validation set of trends cross-validation is composed of the trends calculated from all time periods of the availability of the flux tower observations. The trends are calculated using linear trend estimation.

235 The uncertainty and bias characteristics of the ANN model retrievals were further analyzed on both daily and monthly scales. At the daily ~~time~~ scale, the RMSE of λE (H) ranged from 26.05 to 26.32 $W m^{-2}$ (28.61 to 29.15 $W m^{-2}$), and more than 80% of the 212 flux towers had a correlation greater than 0.70. As for the RMSE, 85% and 89% of the daily λE and (H) were less than 30 $W m^{-2}$, respectively (Supplementary Fig. S87). It was obvious that flux towers with large biases were mainly located on the coast of Australia and the west coast and Great Lakes region of the United States, as well as the Mediterranean region,

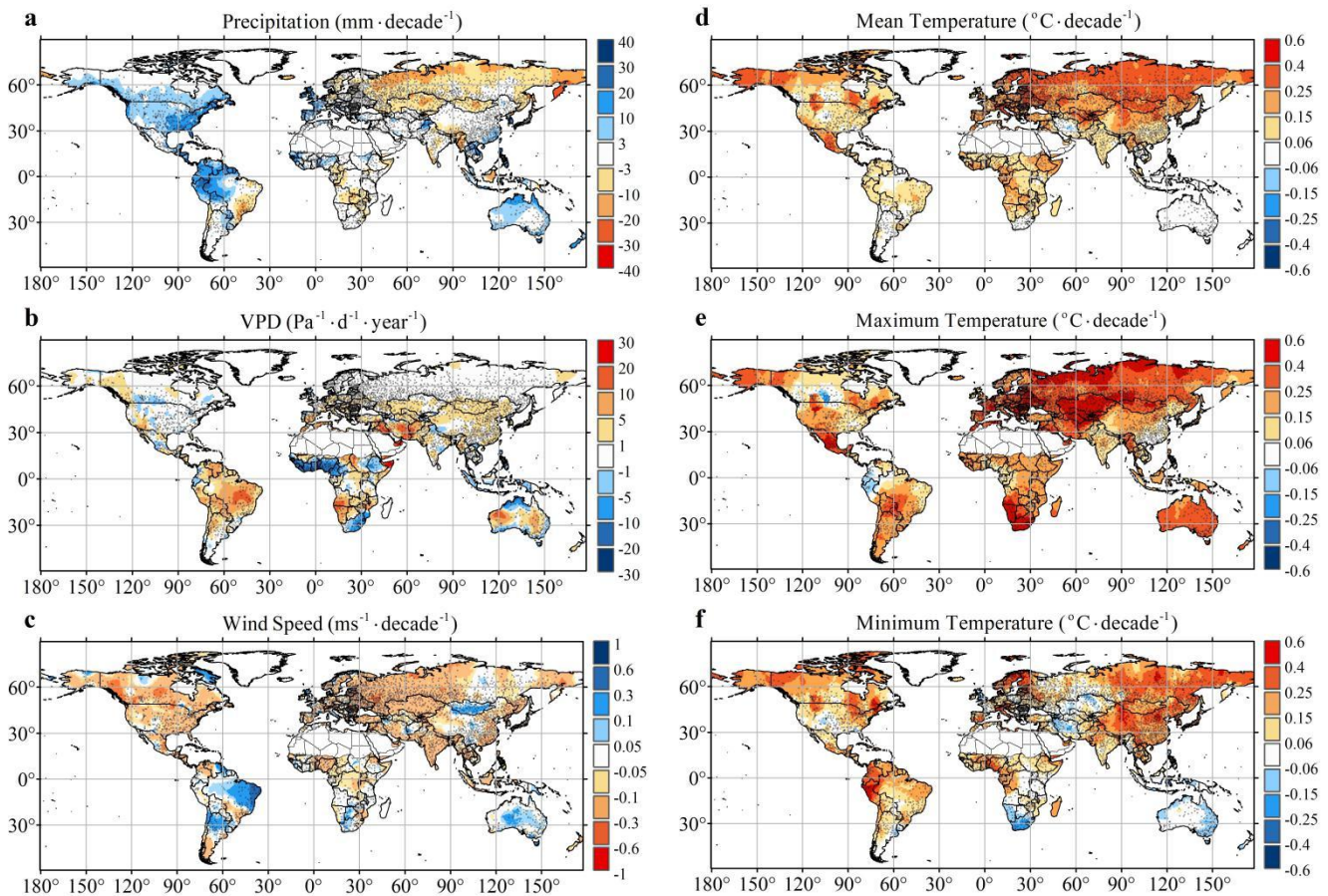
240 all of which are strongly impacted by advection from neighboring open water bodies. The biases of the monthly λE (H)-estimate were smaller than the biases of the daily λE (H)-estimate. More than 89% and 90% of the sites had an R greater than 0.70 (Supplementary Fig. S98). Meanwhile, the λE estimation at more than 88% of the sites and the H estimation at more than 89% of the sites showed an RMSE less than 30 $W m^{-2}$. Finally, the daily λE and H of aHof each weather stations over the past few decades were predicted by the well-trained ANN model. The spatial distribution patterns of mean annual λE and

245 H are consistent with the results in the results predicted by the MTE model Fluxnet-MTE (Jung et al., 2011) (Supplementary Fig. S109). The Fluxnet-MTE is a mature and widely applied machine learning product that can be used as a bench work. This ensemble of statistical estimates of λE were obtained from the Department of Biogeochemical Integration (BGI) of the Max Planck Institute (MPI) (<https://www.bgc-jena.mpg.de/geodb/projects/Data.php>). The mean annual ET of the MET model ranged from 0 to 1400 mm (Jung et al., 2010), while the mean annual ET of this study ranged from 0 to 1416 mm

250 during the 1982–2008 period (Supplementary Fig. S110). In different large-scale latitude intervals, the temporal changes of the ANN model estimated λE and the temporal changes of the MET model estimated λE are significantly correlated at the $p < 0.05$ level, which further emphasizes the reliability of the ANN model retrieval results (Supplementary Fig. S121).

3.2 Attribution of trends in climate variables

The attribution of trends in climate variables were estimated for two reasons: (1) to quantify the changes in the atmospheric water supply, and (2) to estimate the long-term trends in atmospheric evaporative demand factors including VPD, air temperature, and surface wind speed. The annual annual precipitation exhibited an increasing trend ranging from 3 to 40 mm per decade in western Europe, the United States, Southeast Asia, and Australia. Conversely, the annual precipitation exhibited a decreasing trend ranging from -3 to -30 mm per decade in northern Eurasia, the savanna region of Brazil, and South Africa (Fig. 3a). In particular, annual precipitation showed a more obvious upward trend than before in a large area of land in recent period, i.e., 2001–2017 (Supplementary Fig. S132). Rising air temperature and the associated increasing/increased water holding capacity of the atmosphere were the primary causes for the substantial increase in precipitation (Byrne et al., 2015), except for some regions (e.g., Russia) with insufficient moisture advection from an ocean or regional evaporation.



265 **Figure 3.** Long-term trends in annual precipitation, vapor pressure deficit (VPD), surface wind speed, mean temperature, maximum temperature, and minimum temperature. Values are not shown for the Greenland and/or Sahara region as there are

scarce weather stations, ~~and the range of the Sahara is referred to the existing study (Vicente-Serrano et al., 2015).~~ Small gray squares show locations of the weather stations used to interpolate global patterns.

270 With respect to the atmospheric water demand sides, ~~the~~-VPD primarily presented an increasing trend because of an increase in air temperature and a decrease in relative humidity, especially in the subtropics (Fig. 3b); this was consistent with the expectations of atmospheric dynamics and the influence of free-tropospheric warming (Held and Soden, 2006). Additional meteorological variables influencing the evaporative demand, such as the mean, maximum, and minimum temperatures, mostly presented increasing trends on the global scale, with the exception of a few areas, such as the United States/Canadian Corn Belt, and Mexico, which showed signs of cooling due to ~~land-use management such as~~ agricultural irrigation (Thiery et al., 2017) (Fig. 3d–f). Therefore, both rising air temperatures and increased VPD indicate that the driving forces of soil evaporation and plant transpiration are increasing under the climate warming trend. In addition, ~~the~~-mean surface wind speed—a meteorologic factor associated with evaporation—showed an overall decreasing trend (i.e., global stilling) except in the Amazon, Argentina, Australia, and Mongolia (Fig. 3c).

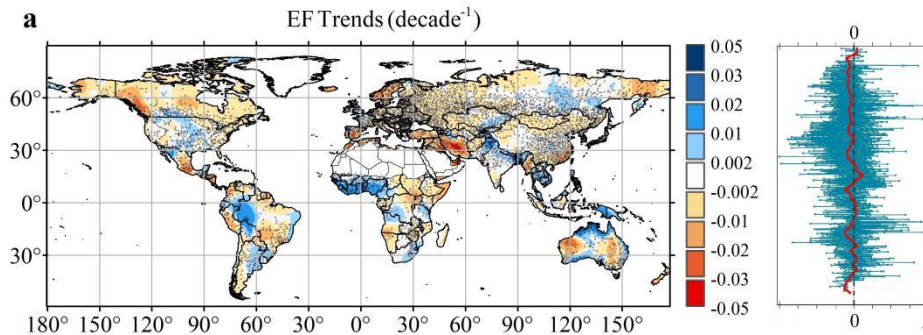
3.3 Long-term trends in EF, ET, and P–ET

280 ~~The annual~~Annual EF ranges from 0 (full aridity stress) to 1 (no aridity stress), and it is an indicator of surface aridity linked to soil moisture availability and vegetation phenology, as well as the physiological effects of atmospheric CO₂ concentrations on vegetation (Francesco et al., 2014; Lemordant et al., 2018; Swann et al., 2016). The decreasing trend in the EF varied from 0 to 0.05 per decade and was prevalent in several land areas (Fig. 4a), except in the most humid areas of tropical rainforest (e.g., the Amazon, West Africa, New Guinea Island, and Southeast Asia) and dense agricultural irrigation areas, including central North America and Punjab in India (Supplementary Fig. S143). Changes in the EF at different latitudinal intervals were consistent with the “dry gets drier, wet gets wetter” paradigm in the tropical areas (Chou et al., 2009; Liu et al., 2013). Moreover, the observed increase in EF further suggested a wet trend in ~~the~~-western Sahel, where ~~an increase in~~increasing rainfall was reported recently (Biasutti, 2019; Dong et al., 2015). It was systematically determined that the EF declined across large swaths of the globe and exhibited different spatial patterns in different periods of the past few decades, which ~~emphasized~~emphasized that this is not a short-term phenomenon (Fig. 5a–c). As the climate has warmed, ~~decline in a~~decrease in the EF reflected an increase in surface resistance (see Methodology), which can be controlled by one of two factors—either an increase in stomatal resistance associated with the physiological effects of CO₂ or a decrease in soil moisture. Therefore, if ~~we find that~~ soil moisture or surface runoff increases while EF decreases, it is a sign of increased surface resistance impacting the water balance.

295 The evolution of El Niño–Southern Oscillation (ENSO) can greatly influence ~~the~~ global hydrological cycle and ~~the~~ patterns of aridity/wetness (Fu et al., 2012; Miralles et al., 2013; Nalley et al., 2019), and thus we analyzed the patterns of EF in different ENSO phases based on the multivariate ENSO index (MEI). However, no significant changes in EF trends were

300 detected between different ENSO phases, with the exception of La Niña ~~having~~ showing a significant impact on the aridity in East Asia (Fig. 5d–f). In addition, the EF simulated using an ensemble from Phase 5 of the Coupled Model Intercomparison Project (CMIP5) under the future Representative Concentration Pathway (RCP) 8.5 scenario (the warming scenario with the highest CO₂ emissions) also presented a decreasing trend in most global land areas, even if the trend in the CMIP5 model appeared to be stronger than that of the trend estimated by the data-driven EF ([Supplementary Fig. S15a](#)). Although the trend magnitudes vary across different periods, ~~we can determine whether it indicated~~ the direction of EF decline ~~may be~~ is a long-term existing phenomenon in historical ~~years~~ periods and simulated future scenarios. The ~~climate model Earth System Model~~ simulation results further suggested that ~~the~~ increasing CO₂ concentrations can affect the allocation of surface energy ~~fluxes~~ and may cause a decrease in EF on large land surface areas. However, the climate model missed several areas where agricultural practices led to an increase in EF, as well as in the tropical rainforest where there was a deviation in peak precipitation (Yin et al., 2013).

310 As the climate warmed, ET showed a significant upward trend ranging from 0 to 0.03 mm per day per year (Fig. 4b), especially in the core regions of tropical rainforest ~~climate zones~~ (e.g., the Amazon, West Africa, and Southeast Asia), the coast of Australia, and the areas with a high density of agricultural irrigation (e.g., northern India, Central Asia, and Central America). The increase in ET was primarily induced by the radiative effect of a warming climate, which can compensate for
315 the observed decrease trend in EF ($ET = EF \times R_n$). Moreover, the observation-driven results showed a declining trend in ET at a rate of 0 to -0.03 mm per day per year on fractional land surfaces, such as North America, South Africa, Australia, Southeast Asia, and the Mediterranean region, which was consistent with the ET declining trends simulated by the CMIP5 climate models ~~at~~ RCP8.5 scenario ([Supplementary Fig. S15b](#)).



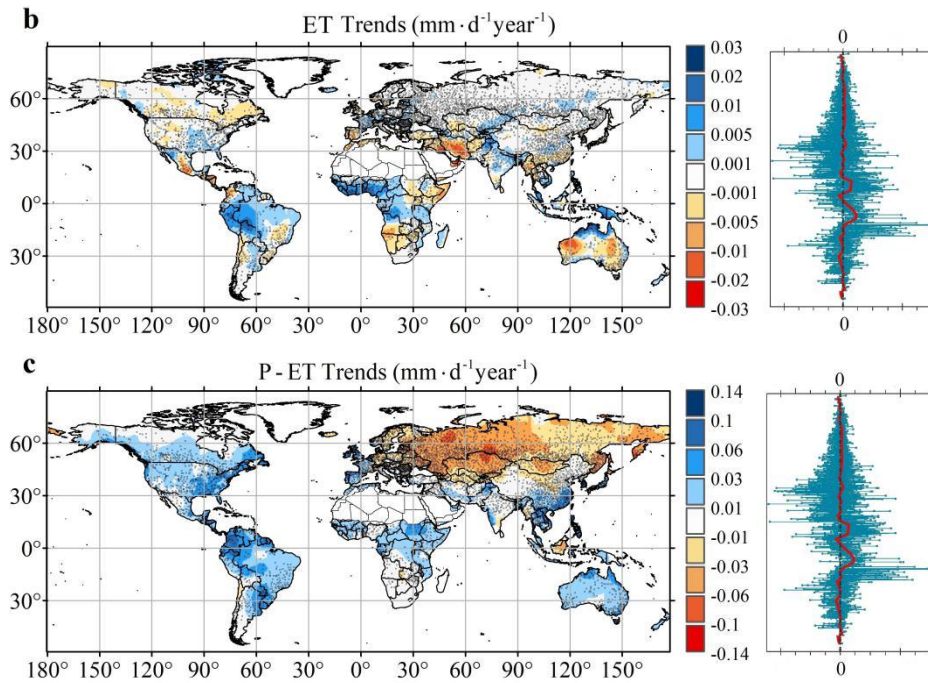
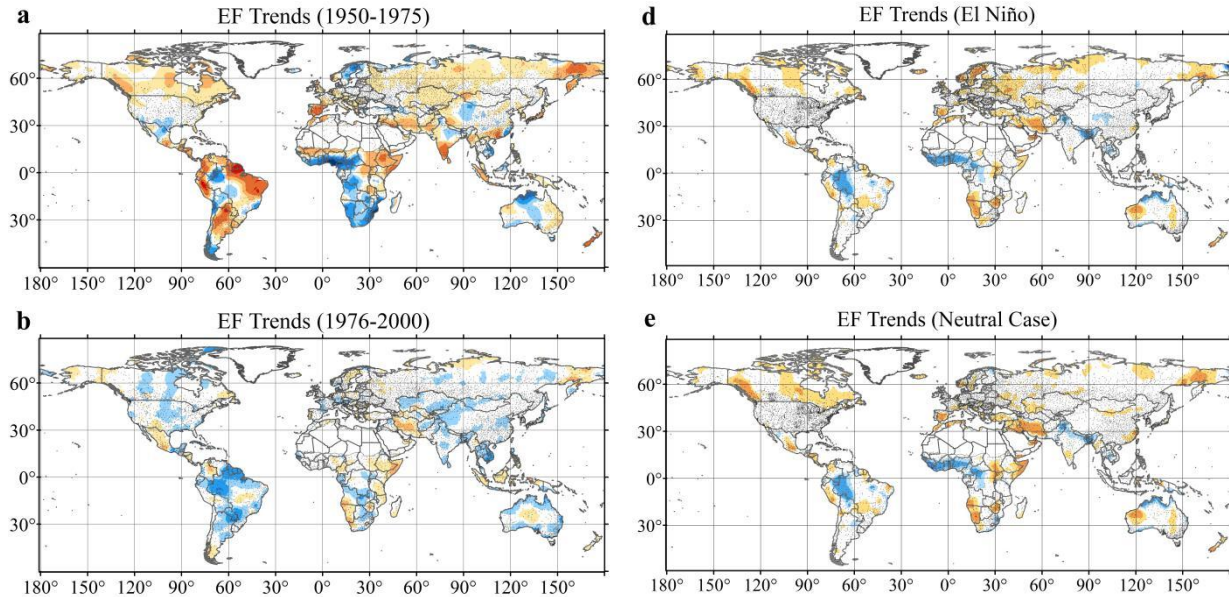


Figure 4. Long-term trends in evaporative fraction (EF), evapotranspiration (ET), and precipitation (P) minus ET (P-ET).

320

ET was converted from the ANN retrieved latent heat flux. The red curve represents median trends at different latitudes.



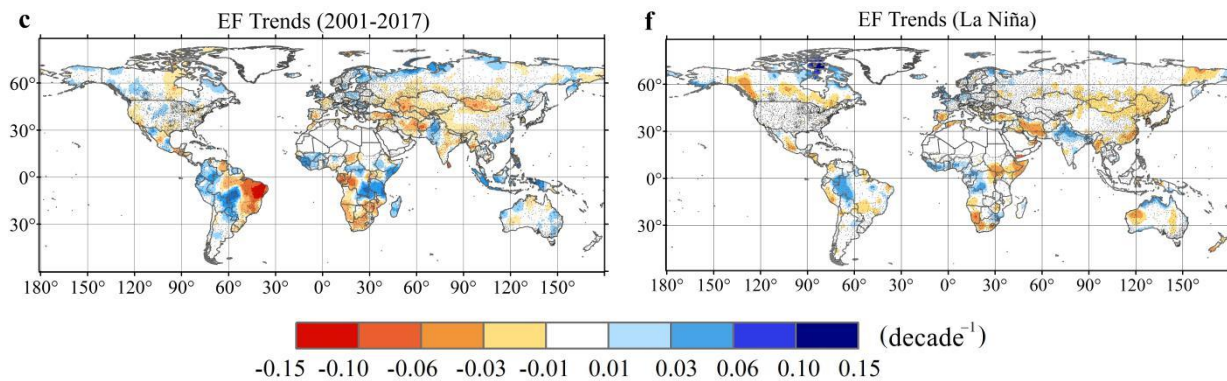


Figure 5. Spatial patterns of EF trends during different periods. (a–c) The spatial patterns show EF trends during different historical periods, and (d–f) the spatial patterns show EF trends during El Niño period, a neutral case period, and La Niña period, respectively.

325

P–ET, a proxy for long-term runoff, assumes that changes in storage due to human activity are negligible and are closely linked to water availability and soil moisture trends (Alkama et al., 2013; Sophocleous et al., 2002). Therefore, long-term runoff mainly presented an increasing trend on most of the global land, with the exception of a decrease in northern Eurasia (Fig. 4c).

330

To verify the retrieved P–ET trend, we ~~compared~~made an comparison between the P–ET trend ~~with~~and the observed runoff trend during the same periods in small- and medium-sized watersheds (5~1000 km²) (Supplementary Fig. S16+14). The P–ET and observed runoff presented different trends in eastern Australia, which can be attributed to a decrease in runoff caused by human activities such as reservoir scheduling and agriculture irrigation (Bosmans et al., 2017; Lehner et al., 2017).

335

When we only considered the stations that are not too influenced by large reservoirs, we found that the direction and the spatial pattern of P–ET trend (Fig. 4c) are more obvious consistent with observed runoff trend, including the upward trend in northern Australia and the downward trend in southern Australia, the upward trend in western Europe and the downward trend in eastern Europe (Supplementary Fig. S16c). The spatial pattern of P–ET trends and observed runoff trends are also generally consistent in other regions including North and South America, Southern Africa, East Asia, and Southeast Asia.

340

We ~~should~~do not expect the P–ET to be completely consistent with observed streamflow, because in addition to measurement errors, the observed streamflow is strongly affected by human activities especially over long-term period on long time scales. The CMIP5 models under RCP8.5 also predicted an overall increasing trend in P–ET (Supplementary Fig. S15c), while a decrease was predicted (but it has not been observed) in the western United States and western Europe and P–ET was predicted to increase in northern Eurasia, but it has not been widely observed.

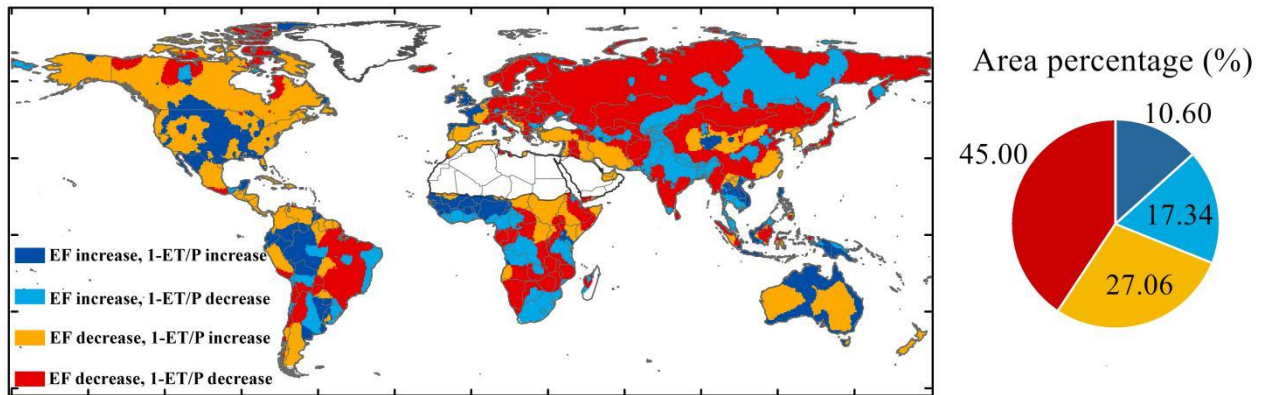
3.4 Signs of covariations in long-term EF and runoff

345

The signs of covariations in normalized ET, i.e., EF, and normalized P–ET, i.e., $1-ET/P$, were further investigated to determine the patterns of surface aridity. We superimposed the EF trend, indicative of changes in aridity stress (e.g.,

temperature and soil moisture) or plant physiological effects (see Methodology), and the 1-ET/P trend, which was indicative of changes in long-term runoff. Land areas with a decreased EF and an increased 1-ET/P were indicative of a dominance of CO₂ plant physiological effects, because a long-term decline in ET with increasing runoff mainly attributable to surface vegetation control. A decline in the EF caused by a decrease in surface conductance can be offset by an increase in the EF caused by the effects of climate warming. Nevertheless, in 27.06% of the global land areas, the EF has declined and has been accompanied by an increase in long-term runoff, which has been observed in most of North America, parts of South America, the Mediterranean, Africa, Australia, and Southeast China (Fig. 6). These signals further emphasized that surface vegetation controls and its response to changing environment have a great influence on water cycle and surface aridity variability.

Covariations in EF and 1-ET/P



355

Figure 6. Signs of covariation in EF and 1-ET/P. Right panel shows area percentage of different signs, and the area fractions are calculated by the spherical area.

In addition, the signs of increase in EF with decreasing runoff accounted for 17.34% of the global land areas, which was mainly due to agricultural irrigation and land use management, such as in the Punjab region of India, Central Asia, and downstream Amazon where there is a high density of irrigation (Supplementary Fig. S14). The land areas showing increase trend in both EF and runoff were typically located in humid regions and accounted for 10.60% of the global land surface. With the increase of EF and 1-ET/P, the humid areas of the Amazon, West Africa, Southeast Asia, and the coast of Australia are getting wetter (Fig. 6). Particularly, the previously reported wet trend in western Sahel was captured by the increase trends in both EF and 1-ET/P. Additionally, 45.00% of the global land areas experienced a decreasing trend in EF and 1-ET/P, and thus aridity stress posed a relatively larger risk to these regions. EF and 1-ET/P both exhibited a decreasing trend in the arid regions of the Amazon (e.g., the savanna region of Brazil), and thus those areas are getting drier. Moreover, the Mediterranean region, northern Eurasia, and South Africa also experienced a decrease trend in EF and 1-ET/P, which was consistent with existing observed or predicted [increasing drought](#) (Padrón et al., 2020; Samaniego et al., 2018; Zhou et al., 2019).

370

4 Concluding Remarks

This study for the first time provides the ~~first~~ strategy for retrieving consistent latent heat and sensible heat fluxes on a global scale over many decades, based on boundary layer energy budget perspective and using machine learning approach driven by the ground observations of globally distributed flux towers and weather stations. After that, we quantified the attributions of long-term changes in surface aridity/wetness. The latent heat and sensible heat fluxes retrieved in this study can be an important supplement to the existing product. ~~The results~~Our study had important implications for understanding variability of land surface aridity under changing environment and provided constraints for model predictions. Although we attempted to infer latent heat and sensible heat fluxes from in situ ground observations and used various data quality control methods to reduce uncertainty, the quality of the observational data from the flux towers and weather stations can influence the retrievals.

In the absence of surface regulation of plant physiological effect and changes in biomass, a warming climate was expected to intensify ET at a rate roughly governed by the Clausius–Clapeyron relation. However, a long-term relative decrease in normalized ET accompanied by increasing runoff was found in 27.06% of the global land areas, which was indicative of a reduction in surface conductance. The findings further emphasized that vegetation controls have strong impacts in regulating the water cycle and surface aridity variability. Climate models have captured some of these changes; however, they have also exhibited large regional discrepancies. Therefore, representations of land use management and plant physiological effects are essential for the improvement of future predictions with respect to the water, energy, and carbon cycles.

Data/code availability. The data, results and Matlab codes in this study are available upon request. The eddy-covariance data are available at <http://fluxnet.fluxdata.org/>. The Global Summary of the Day and the Global Historical Climatology Network datasets are collected from the NOAA at <https://www.ncdc.noaa.gov/data-access>. The data of Global Runoff Data Center are available at https://www.bafg.de/GRDC/EN/01_GRDC/13_dtbse/database_node.html. The reservoirs data in the GRanD database are available at <http://sedac.ciesin.columbia.edu/pfs/grand.html>. Global irrigation data are available at <http://www.fao.org/nr/water/aquastat/irrigationmap/index10.stm>. The Multivariate ENSO Index (MEI) is available at <https://www.esrl.noaa.gov/psd/enso/mei/>.

Author contributions. RW and PG designed the study. RW performed the experiment and wrote the manuscript. PG designed the methodology, analyzed the results, and revised the manuscript. JY and LC contributed to data collection and validation. JC and LL contributed to discussion and supervision.

Competing interests. The authors declare that they have no conflicts of interest.

405 *Acknowledgments.* We acknowledge the members of FLUXNET community for sharing flux tower observational data, and
Ren Wang acknowledge Dr. Léo Lemordant for helping with the simulations of Earth System Models.

Financial support. This work was financially supported by the National Key Research and Development Program of China
(2017YFA0603603), the China Postdoctoral Science Foundation (2020M681656) and the China Scholarship Council
410 scholarship (201706380063). JB Yin acknowledge the National Natural Science Foundation of China (52009091).

References

- Alkama, R., Marchand, L., Ribes, A., and Decharme, B.: Detection of global runoff changes: results from observations and
CMIP5 experiments. *Hydrol. Earth Syst. Sci.*, 17, 2967–2979, <https://doi.org/10.5194/hess-17-2967-2013>, 2013.
- 415 Baruga, C. K., Kim, D., and Choi, M.: A national-scale drought assessment in Uganda based on evapotranspiration deficits
from the Bouchet hypothesis. *J. Hydrol.*, 580, 124348, <https://doi.org/10.1016/j.jhydrol.2019.124348>, 2020.
- Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S. I., van den Hurk, B., Lorenz, R., Pitman, A., Hagemann, S.,
Meier, A., Cheruy, F., Ducharne, A., Malyshev, S., and Milly, P. C. D.: Land-atmosphere feedbacks amplify aridity
increase over land under global warming. *Nat. Clim. Change*, 6, 869–874, <https://doi.org/10.1038/nclimate3029>, 2016.
- 420 Biasutti, M.: Rainfall trends in the African Sahel: Characteristics, processes, and causes. *WIREs Clim. Change*, 10, e591,
<https://doi.org/10.1002/wcc.591>, 2019.
- Byrne, M. P. and O’Gorman, P. A.: The response of precipitation minus evapotranspiration to climate warming: Why the
“wet-get-wetter, dry-get-drier” scaling does not hold over land. *J. Clim.*, 28, 8078–8092, <https://doi.org/10.1175/JCLI-D-15-0369.1>, 2015.
- 425 Chen, Z. J., Zhu, Z. C., Jiang, H., and Sun S. J.: Estimating daily reference evapotranspiration based on limited
meteorological data using deep learning and classical machine learning methods. *J. Hydrol.*, 591, 125286,
<https://doi.org/10.1016/j.jhydrol.2020.125286>, 2020.
- Chou, C., Neelin, J. D., Chen, C. A., and Tu, J. Y.: Evaluating the “rich-get-richer” mechanism in tropical precipitation
change under global warming. *J. Clim.*, 22, 1982–2005, <https://doi.org/10.1175/2008JCLI2471.1>, 2009.
- 430 Cook, B. I., Smerdon, J. E., Seager, R., and Coats, S.: Global warming and 21st century drying. *Clim. Dyn.*, 43, 2607–2627,
<https://doi.org/10.1007/s00382-014-2075-y>, 2014.
- Costa, M. H., Biajoli, M. C., Sanches, L., Malhado, A. C. M., Hutyra, L. R., da Rocha, H. R., Aguiar, R. G., and de Araújo,
A. C.: Atmospheric versus vegetation controls of Amazonian tropical rain forest evapotranspiration: Are the wet and
seasonally dry rain forests any different?, *J. Geophys. Res.: Biogeo.*, 115, G04021,
435 <https://doi.org/10.1029/2009jg001179>, 2010.
- Dai, A.: Increasing drought under global warming in observations and models. *Nat. Clim. Change*, 3, 171–171,
<https://doi.org/10.1038/nclimate1633>, 2013.

- Durre, I., Menne, M. J., Gleason, B. E., Houston, T. G., and Vose, R. S.: Comprehensive automated quality assurance of daily surface observations. *J. Appl. Meteorol. Clim.*, 49, 1615–1633, <https://doi.org/10.1175/2010JAMC2375.1>, 2010.
- 440 Dong, B. and Sutton, R.: Dominant role of greenhouse-gas forcing in the recovery of Sahel rainfall. *Nat. Clim. Change*, 5, 757–760, <https://doi.org/10.1038/nclimate2664>, 2015.
- Forzieri, G., Miralles, D., Ciais, P., Alkama, R., Ryu, Y., Duveiller, G., Zhang, K., Robertson, E., Kautz, M., Martens, B., Jiang, C., Arneth, A., Georgievski, G., Li, W., Ceccherini, G., Anthoni, P., Lawrence, P., Wiltshire, A., Pongratz, J., Piao, S., Sitch, S., Goll, D. S., Arora, V. K., Lienert, S., Lombardozzi, D., Kato, E., Nabel, J. E. M. S., Tian, H., Friedlingstein, P., and Cescatti, A.: Increased control of vegetation on global terrestrial energy fluxes. *Nat. Clim. Change*, 10, 356–362, <https://doi.org/10.1038/s41558-020-0717-0>, 2020.
- 445 Fu, C., James, A.L., and Wachowiak, M.P. Analyzing the combined influence of solar activity and El Niño on streamflow across southern Canada. *Water Resour. Res.*, 48, W05507. <https://doi.org/10.1029/2011WR011507>, 2012.
- Fu, Q. and Feng, S.: Responses of terrestrial aridity to global warming. *J. Geophys. Res.: Atmos.* 119, 7863–7875, <https://doi.org/10.1002/2014JD021608>, 2014.
- 450 Francesco, N., Mirco, B., Gabriele, C., Stefano, B., and Pietro, B.: Evaporative fraction as an indicator of moisture condition and water stress status in semi-arid range land ecosystems. *Remote Sens.*, 6, 6300–6323, <https://doi.org/10.3390/rs6076300>, 2014.
- Gentine, P., Chhang, A., Rigden, A., and Salvucci, G.: Evaporation estimates using weather station data and boundary layer theory. *Geophys. Res. Lett.*, 43, 11661–11670, <https://doi.org/10.1002/2016GL070819>, 2016.
- 455 Gentine, P., Entekhabi, D., and Polcher, J.: The diurnal behavior of evaporative fraction in the soil-vegetation-atmospheric boundary layer continuum. *J. Hydrometeorol.*, 12, 1530–1546, <https://doi.org/10.1175/2011JHM1261.1>, 2011.
- Gentine, P., Holtslag, A.A., D'Andrea, F. and Ek, M. Surface and atmospheric controls on the onset of moist convection over land. *Journal of Hydrometeorology*, 14(5), 1443-1462., 2013a.
- 460 Gentine, P., Ferguson, C.R. and Holtslag, A.A., Diagnosing evaporative fraction over land from boundary-layer clouds. *Journal of Geophysical Research: Atmospheres*, 118(15), 8185-8196, 2013b.
- Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., and Seneviratne, S. I.: Global assessment of trends in wetting and drying over land. *Nat. Geosci.*, 7, 716–721, <https://doi.org/10.1038/ngeo2247>, 2014.
- Held, I. M. and Soden, B. J. Robust Responses of the Hydrological Cycle to Global Warming. *J. Clim.*, 19, 5686–5699, <https://doi.org/10.1175/JCLI3990.1>, 2006.
- 465 Hoek van Dijke, A. J., Mallick, K., Schlerf, M., Machwitz, M., Herold, M., and Teuling, A. J. Examining the link between vegetation leaf area and land–atmosphere exchange of water, energy and carbon fluxes using FLUXNET data. *Hydrol. Earth Syst. Sci.*, 17, 4443–4457, <https://doi.org/10.5194/bg-17-4443-2020>, 2020.
- Jaramillo, F., Cory, N., Arheimer, B., Laudon, H., van der Velde, Y., Hasper, T. B., Teutschbein, C., and Uddling, J.: Dominant effect of increasing forest biomass on evapotranspiration: interpretations of movement in Budyko space. *Hydrol. Earth Syst. Sci.*, 22, 567–580, <https://doi.org/10.5194/hess-22-567-2018>, 2018.
- 470

- 475 Jung, M., Reichstein, M., Ciais, P., Seneviratne, S.I., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., de Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Rouspard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nat.*, 467, 951–954, <https://doi.org/10.1038/nature09396>, 2010.
- 480 Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneeth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res.: Biogeo.*, 116, G00J07, <http://dx.doi.org/10.1029/2010JG001566>, 2011.
- Keenan, T. F., Hollinger, D. Y., Bohrer, G., Dragoni, D., Munger, J. W., Schmid, H. P., and Richardson, A. D.: Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise. *Nat.*, 499, 324–327, <https://doi.org/10.1038/nature12291>, 2013.
- 485 Komatsu, H. and Kume, T.: Modeling of evapotranspiration changes with forest management practices: A genealogical review. *J. Hydrol.*, 585, 124835, <https://doi.org/10.1016/j.jhydrol.2020.124835>, 2020.
- [Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rodel, R., Sindorf, N., and Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Front. Ecol. Environ.*, 9: 494–502, <https://doi.org/10.2307/23034466>, 2011.](#)
- 490 Lemordant, L., Gentine, P., Swann, A. S., Cook, B. I., and Scheff, J.: Critical impact of vegetation physiology on the continental hydrologic cycle in response to increasing CO₂. *Proc. Natl. Acad. Sci. USA*, 16, 4093–4098, <https://doi.org/10.1073/pnas.1720712115>, 2018.
- Liu, C. and Allan, R. P.: Observed and simulated precipitation responses in wet and dry regions 1850–2100. *Environ. Res. Lett.*, 8, 034002, <https://doi.org/10.1088/1748-9326/8/3/034002>, 2013.
- Mallick, K., Trebs, I., Boegh, E., Giustarini, L., Schlerf, M., Drewry, D. T., Hoffmann, L., Von Randow, C., Kruijt, B., Araùjo, A., Saleska, S., Ehleringer, J. R., Domingues, T. F., Ometto, J. P. H. B., Nobre, A. D., Luiz Leal De Moraes, O., Hayek, M., William Munger, J., and Wofsy, S. C.: Canopy-scale biophysical controls of transpiration and evaporation in the Amazon Basin, *Hydrol. Earth Syst. Sci.*, 20, 4237–4264, <https://doi.org/10.5194/hess-204237-2016>, 2016.
- 500 Margulis, S. A., Introduction to hydrology, 2017. <https://margulis-group.github.io/teaching/>.
- Massmann, A., Gentine, P., and Lin, C.: When does vapor pressure deficit drive or reduce evapotranspiration?. *J. Adv. Model. Earth Syst.*, 11, 3305–3320, <https://doi.org/10.1029/2019MS001790>, 2019.
- Miralles, D. G., DeJeu, R. A. M., Gash, J. H., Holmes, T. R. H., and Dolman, A. J.: Magnitude and variability of land evaporation and its components at the global scale, *Hydrol. Earth Syst. Sci.*, 15, 967–981, [https://doi.org/10.5194/hess-](https://doi.org/10.5194/hess-15-967-2011)
- 505 15-967-2011, 2011.

- Miralles, D. G., Van, d. B. M. J., Gash, J. H., Parinussa, R. M., de Jeu, R. A. M., Beck, H. E., Holmes, T. R. H., Carlos Jiménez, C., Verhoest, N. E. C., Dorigo, W. A., Teuling, A. J., and Dolman, A. J.: El Niño–La Niña cycle and recent trends in continental evaporation. *Nat. Clim. Change*, 4, 122–126, <https://doi.org/10.1038/nclimate2068>, 2013.
- Miralles, D. G., Brutsaert, W., Dolman, A. J., and Gash, J. H.: On the use of the term “evapotranspiration”. *Water Resour. Res.*, 56, e2020WR028055, <https://doi.org/10.1029/2020WR028055>, 2020.
- Milly, P. C. and Dunne, K. A.: Potential evapotranspiration and continental drying. *Nat. Clim. Change*, 6, 946–949, <https://doi.org/10.1038/nclimate3046>, 2016.
- Nalley, D., Adamowski, J., Biswas, A., Gharabaghi, B., and Hu, W.: A multiscale and multivariate analysis of precipitation and streamflow variability in relation to ENSO, NAO and PDO. *J. Hydrol.*, 574, 288–307, <https://doi.org/10.1016/j.jhydrol.2019.04.024>, 2019.
- Naumann, G., Alfieri, L., Wyser, K., Mentaschi, L., Betts, R. A., Carrao, H., Spinoni, J., Vogt, J., and Feyen, L.: Global changes in drought conditions under different levels of warming. *Geophys. Res. Lett.*, 45, 3285–3296, <https://doi.org/10.1002/2017GL076521>, 2018.
- Ned, H., Gab, A., and Pitman, A. J.: On the predictability of land surface fluxes from meteorological variables. *Geosci. Model Dev.*, 11, 195–212, <https://doi.org/10.5194/bg-15-4495-2018>, 2018.
- Orth, R., and Destouni, G.: Drought reduces blue-water fluxes more strongly than green-water fluxes in Europe. *Nat. Commun.*, 9, 3602, <http://dx.doi.org/10.1038/s41467-018-06013-7>, 2018.
- Padrón, R. S., Gudmundsson, L., Decharme, B., Ducharne, A., Lawrence, D. M., Mao, J., Peano, D., Krinner, G., Hyungjun Kim, H., and Seneviratne, S. I.: Observed changes in dry-season water availability attributed to human-induced climate change. *Nat. Geosci.*, 2020, 13(7): 477–481, <https://doi.org/10.1038/s41561-020-0594-1>, 2020.
- Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Frank, J., Massman, W., and Urbanski, S.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Sci. Data*, 7, 225, <https://doi.org/10.1038/s41597-020-0534-3>, 2020.
- Rigden, A. J. and Salvucci, G. D.: Evapotranspiration based on equilibrated relative humidity (ETRHEQ): Evaluation over the continental US. *Water Resour. Res.*, 51, 2951–2973, <https://doi.org/10.1002/2014WR016072>, 2015.
- Rigden, A. J. and Salvucci, G. D.: Stomatal response to humidity and CO₂ implicated in recent decline in US evaporation. *Global Change Biol.*, 23, 1140–1150, <https://doi.org/10.1111/gcb.13439>, 2016.
- Salvucci, G. D. and Gentine, P.: Emergent relation between surface vapor conductance and relative humidity profiles yields evaporation rates from weather data. *Proc. Natl. Acad. Sci. USA*, 110, 6287–6291, <https://doi.org/10.1073/pnas.1215844110>, 2013.
- Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E. F., and Marx, A.: Anthropogenic warming exacerbates European soil moisture droughts. *Nat. Clim. Change*, 8, 421–428, <https://doi.org/10.1038/s41558-018-0138-5>, 2018.

- 540 Sheffield, J., Wood, E. F., and Roderick, M. L.: Little change in global drought over the past 60 years. *Nat.*, 491, 435–438, <https://doi.org/10.1038/nature11575>, 2012.
- Sorokin, Y., Jane Zelikova, T., Blumenthal, D., Williams, D. G., and Pendall, E.: Seasonally contrasting responses of evapotranspiration to warming and elevated CO₂ in a semi-arid grassland. *Ecohydrol.*, 10, e1880, <https://doi.org/10.1002/eco.1880>, 2017.
- 545 Sophocleous, M.: Interactions between ground water and surface water: the state of the science. *Hydrogeol. J.*, 10, 348–348, <https://doi.org/10.1007/s10040-001-0170-8>, 2002.
- Swann, A. L., Hoffman, F. M., Koven, C. D., and Randerson, J. T.: Plant responses to increasing CO₂ reduce estimates of climate impacts on drought severity. *Proc. Natl. Acad. Sci. USA*, 113, 10019–10024, <https://doi.org/10.1073/pnas.1604581113>, 2016.
- 550 Thiery, W., Davin, E. L., Lawrence, D. M., Hirsch, A. L., and Seneviratne, S. I.: Present-day irrigation mitigates heat extremes. *J. Geophys. Res.: Atmos.*, 122, 1403–1422, <https://doi.org/10.1002/2016JD025740>, 2017.
- Trenberth, K. E., Dai, A., Van Der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., and Sheffield, J.: Global warming and changes in drought. *Nat. Clim. Change*, 4, 17–22, <https://doi.org/10.1038/NCLIMATE2067>, 2014.
- 555 Teuling, A. J., De Badts, E. A. G., Jansen, F. A., Fuchs, R., Buitink, J., Hoek Van Dijke, A. J., and Sterling, S. M.: Climate change, reforestation/afforestation, and urbanization impacts on evapotranspiration and streamflow in Europe. *Hydrol. Earth Syst. Sci.*, 23, 3631–3652, <https://doi.org/10.5194/hess-23-3631-2019>, 2019.
- Van der Schrier, G., Jones, P. D., and Briffa, K. R.: The sensitivity of the PDSI to the Thornthwaite and Penman-Monteith parameterizations for potential evapotranspiration. *J. Geophys. Res.: Atmos.* 116, D03106, <https://doi.org/10.1029/2010JD015001>, 2011.
- 560 Vicente-Serrano, S. M., Van Gerard, V. D. S., Beguería, S., Azorin-Molina, C., and Lopez-Moreno, J. I.: Contribution of precipitation and reference evapotranspiration to drought indices under different climates. *J. Hydrol.*, 526, 42–54. <http://dx.doi.org/10.1016/j.jhydrol.2014.11.025>, 2015.
- van der Sleen, P., Groenendijk, P., Vlam, M., Anten, N. P. R., Boom, A., Bongers, F., Pons, T. L., Terburg, G., and Zuidema, P. A.: No growth stimulation of tropical trees by 150 years of CO₂ fertilization but water-use efficiency increased. *Nat. Geoscience*, 8, 24–28, <https://doi.org/10.1038/NGEO2313>, 2015.
- 565 Wagle, P., Xiao, X., Scott, R. L., Kolb, T. E., Cook, D. R., Brunsell, N., Baldocchi, D. D., Basara, J., Matamala, R., Zhou, Y., and Bajgain, R.: Biophysical controls on carbon and water vapor fluxes across a grassland climatic gradient in the United States, *Agr. Forest Meteorol.*, 214/215, 293–305, <https://doi.org/10.1016/j.agrformet.2015.08.265>, 2015.
- 570 Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C., Wohlfahrt, G., Hasler, N., Bernhofer, C., Foken, T., Papale, D., Schymanski, S., and Schaefer, K.: Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers, *Water Resour. Res.*, 48, W06523, <https://doi.org/10.1029/2011WR011586>, 2012.

- Williams, I. N. and Torn, M. S.: Vegetation controls on surface heat flux partitioning, and land-atmosphere coupling, *Geophys. Res. Lett.*, 42, 9416–9424, <https://doi.org/10.1002/2015gl066305>, 2015.
- 575 Wei, Z., Yoshimura, K., Wang, L., Miralles, D. G., Jasechko, S., and Lee, X. H.: Revisiting the contribution of transpiration to global terrestrial evapotranspiration, *Geophys. Res. Lett.*, 44, 2792–2801, <https://doi.org/10.1002/2016GL072235>, 2017.
- Yang, Y. T., Roderick, M. L., Zhang, S., McVicar, T. R., and Donohue, R. J.: Hydrologic implications of vegetation response to elevated CO₂ in climate projections. *Nat. Clim. Change*, 9, 44–48. <https://doi.org/10.1038/s41558-018-0361-0>, 2019.
- 580 Yang, Y. T., Zhang, S. L., Roderick, M. L., McVicar, T. R., Yang, D. W., Liu W. B., and Li X. Y.: Comparing Palmer Drought Severity Index drought assessments using the traditional offline approach with direct climate model outputs. *Hydrol. Earth Syst. Sci.*, 24, 2921–2930. <https://doi.org/10.5194/hess-24-2921-2020>, 2020.
- Yin, J. B., Gentine, P., Zhou, S., Sullivan, S. C., Wang, R., Zhang, Y., and Guo S. L.: Large increase in global storm runoff extremes driven by climate and anthropogenic changes. *Nat. Commun.*, 22, 9, 4389, <https://doi.org/10.1038/s41467-018-06765-2>, 2018.
- 585 Yin, L., Fu, R., Shevliakova, E., and Dickinson, R. E.: How well can CMIP5 simulate precipitation and its controlling processes over tropical south America?. *Clim. Dynam.*, 41, 3127–3143, <https://doi.org/10.1007/s00382-012-1582-y>, 2013.
- Zhao, W. L., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S., Wen, Y., Lin, C., Li, X., and Qiu, G.Y.: Physics-constrained machine learning of evapotranspiration. *Geophys. Res. Lett.*, 46, 14496–14507, <https://doi.org/10.1029/2019GL085291>, 2019.
- 590 Zhou, C. and Wang, K.: Biological and environmental controls on evaporative fractions at AmeriFlux sites. *J. Appl. Meteorol. Clim.*, 55, 145–161. <https://doi.org/10.1175/JAMC-D-15-0126.1>, 2016.
- Zhou, S., Williams, A. P., Berg, A. M., Cook, B. I., Zhang, Y., Stefan, H., Ruth, L., Seneviratne, S. I., and Gentine, P.: Land-atmosphere feedbacks exacerbate concurrent soil drought and atmospheric aridity. *Proc. Natl. Acad. Sci. USA*, 595 116, 18848–18853, <https://doi.org/10.1073/pnas.1904955116>, 2019.

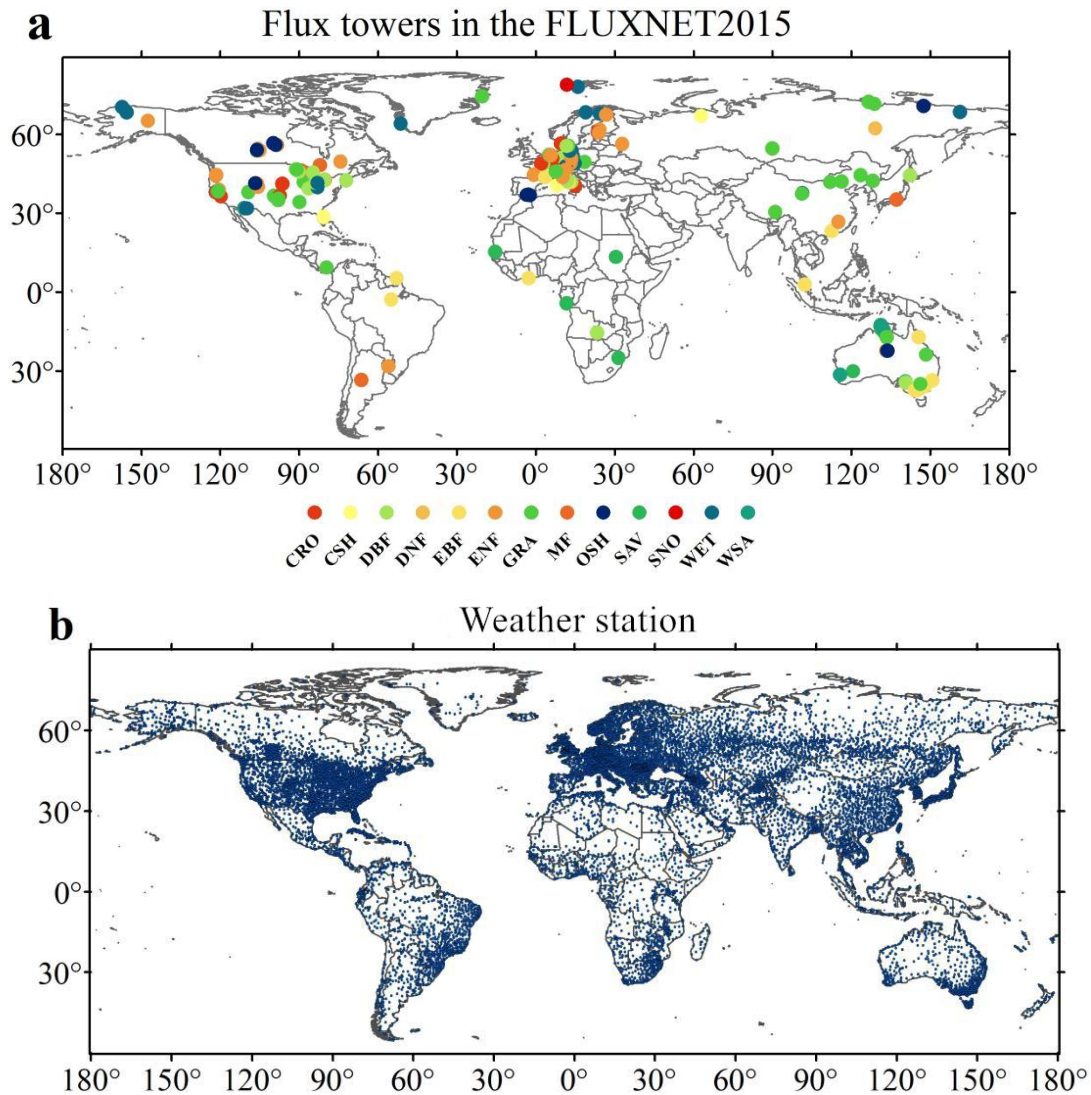
Supplement of

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

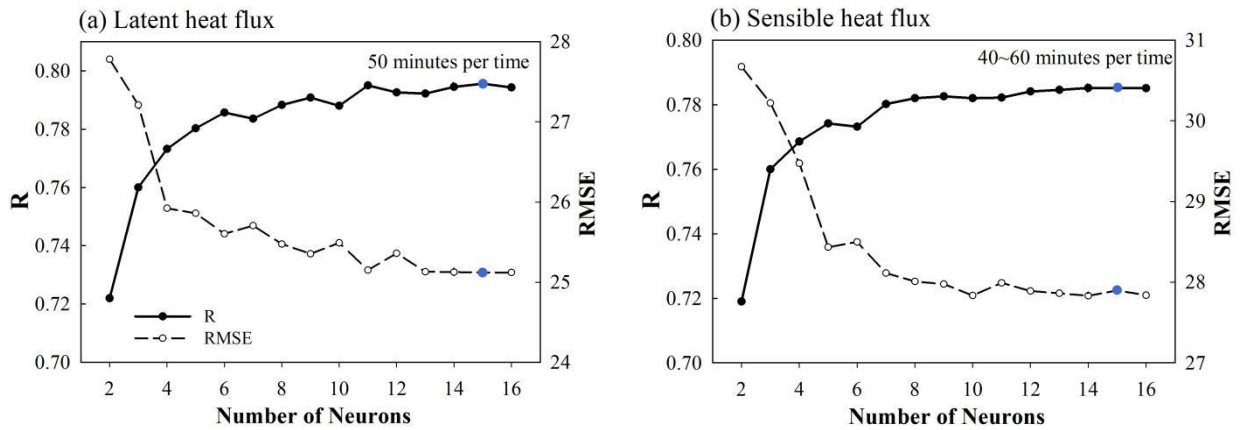
Ren Wang et al.

Correspondence to:

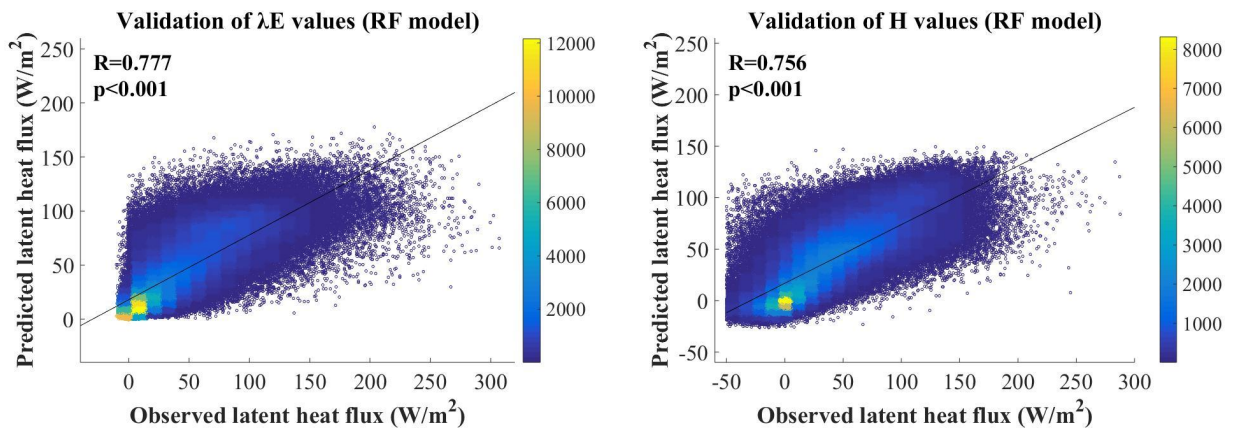
Ren Wang (wangr67@mail2.sysu.edu.cn) or Pierre Gentine (pg2328@columbia.edu)



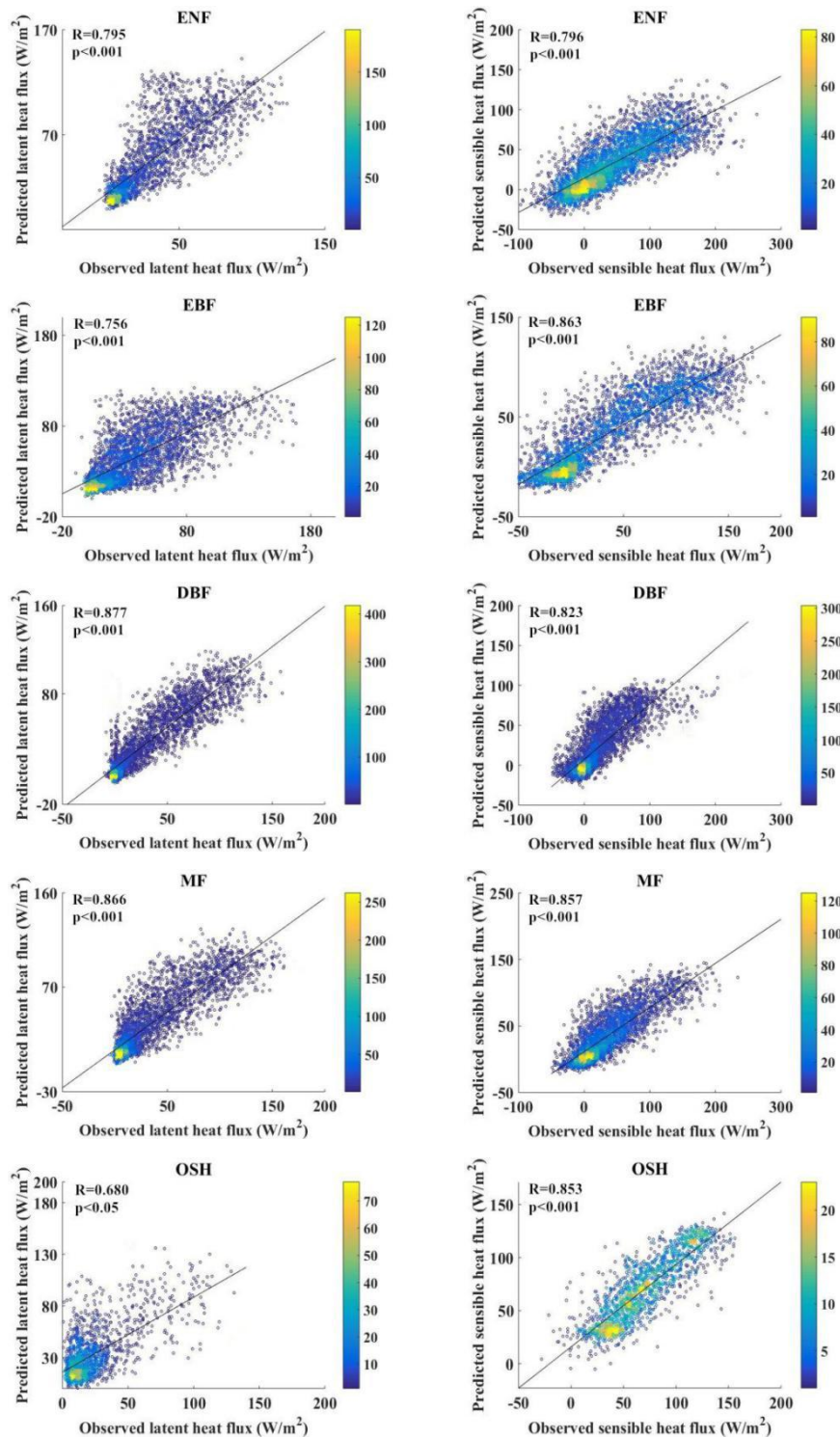
Supplementary Figure S1. Spatial distribution of (a) the flux towers in the FLUXNET2015 and (b) the weather stations used in this study. The plant function types of the flux towers include Croplands (CRO), Deciduous Needleleaf Forests (DNF), Evergreen Needleleaf Forest (ENF), Evergreen Broadleaf Forest (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).



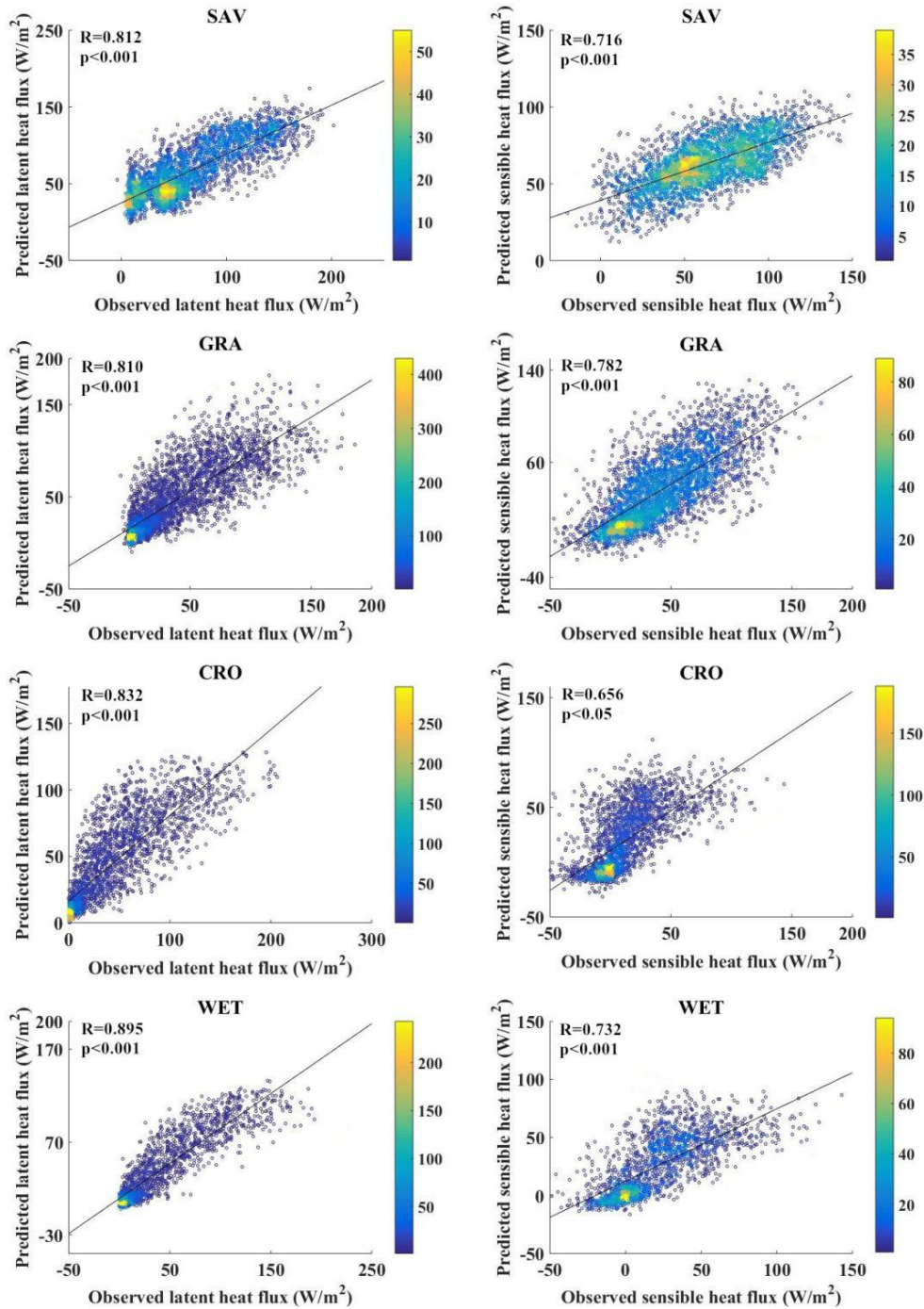
Supplementary Figure S2. The performance of the ANN model using 2 hidden layers and different number of neurons.



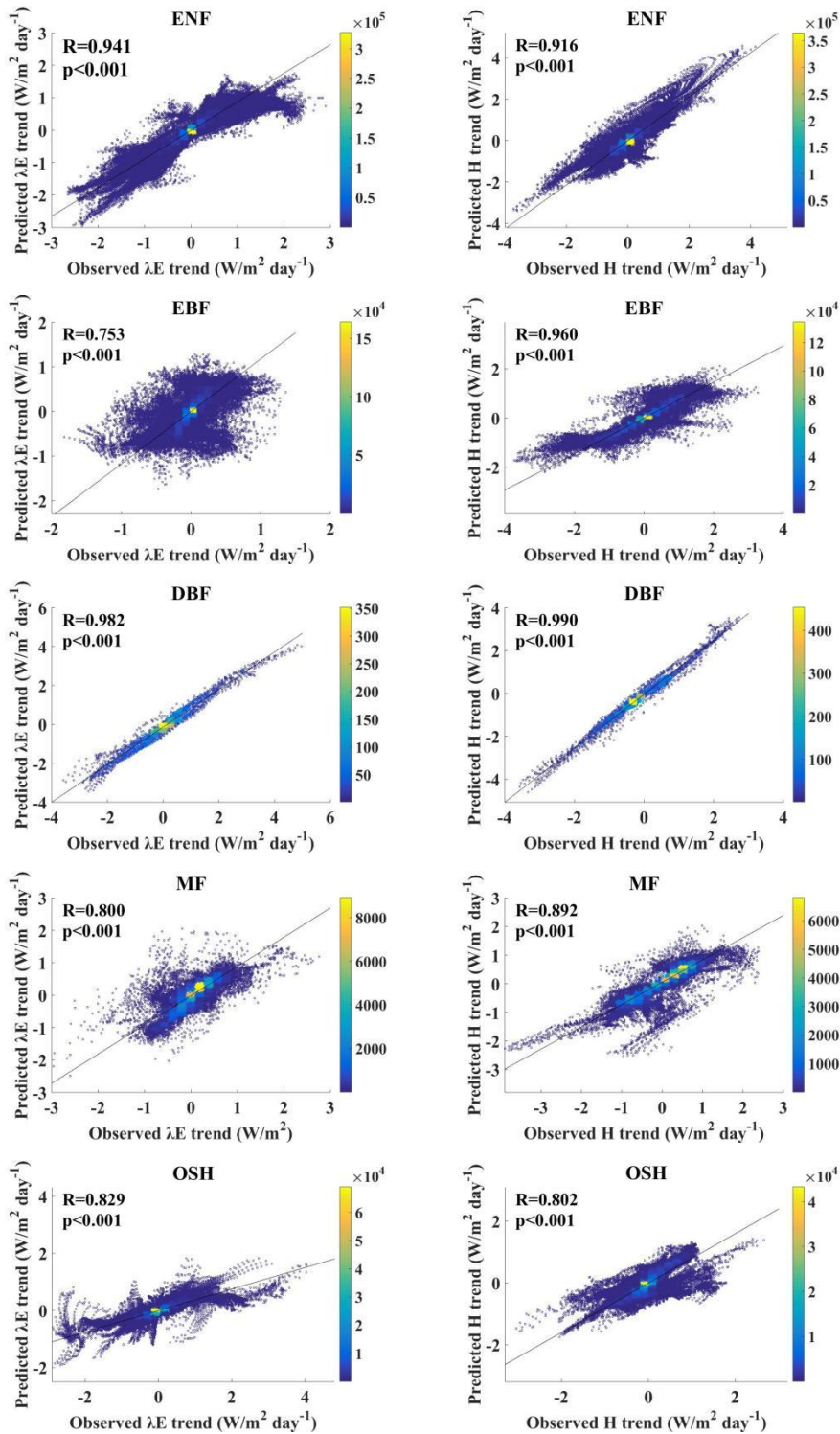
Supplementary Figure S3. The performance of a well-trained random forest (RF) model for predicting daily λE and H based on the same Fluxnet2015 data as the ANN model. The RF model is trained using 500 trees. The test samples are randomly selected from the training data using a ratio of 10%.



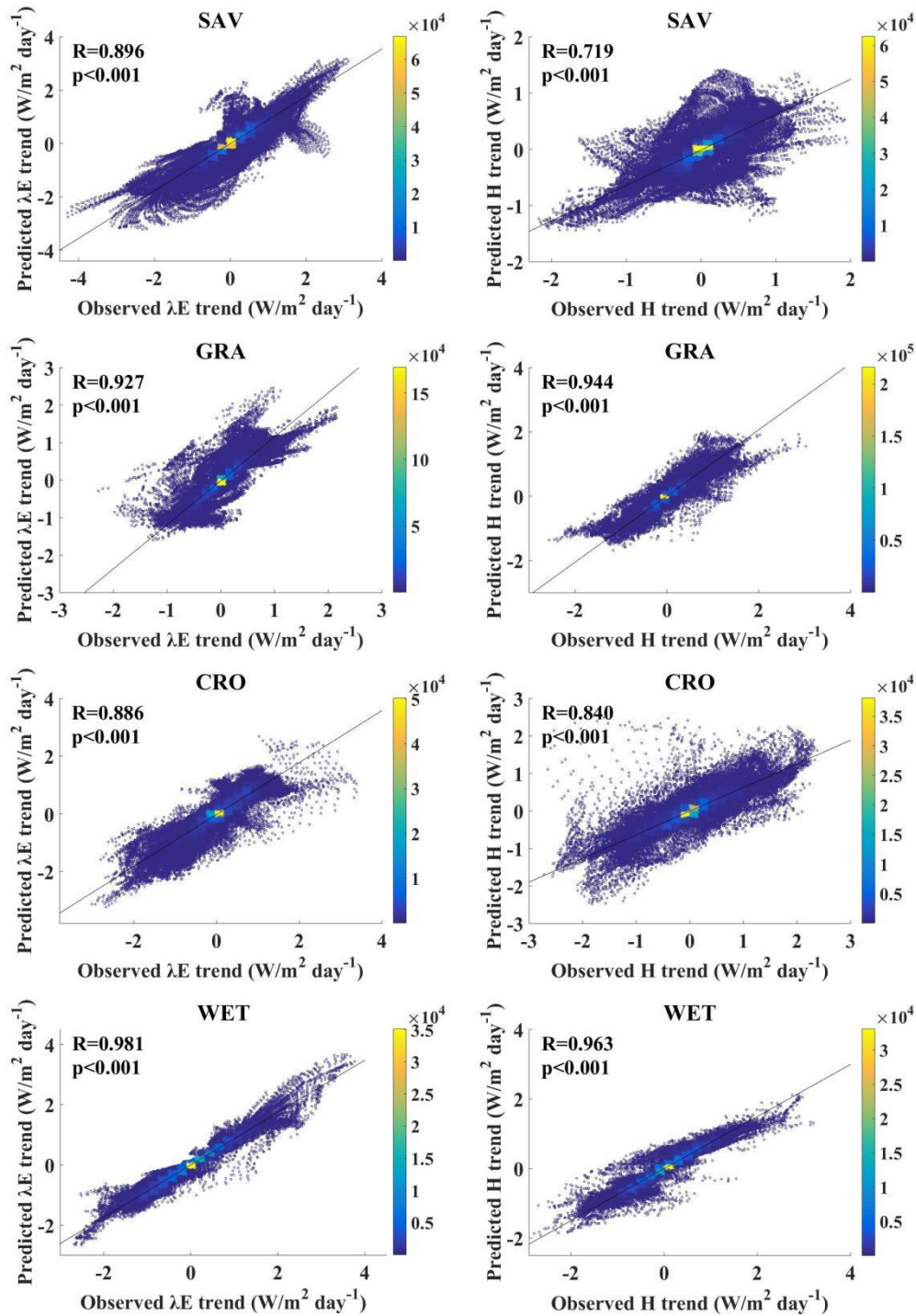
Supplementary Figure S4. Density scatter plot for the cross-validation in terms of values for evergreen needleleaf forest (ENF), evergreen broadleaf forest (EBF), deciduous broadleaf forest (DBF), mixed forest (MF), and open shrublands (OSH). Samples of the validation set are randomly composed of five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



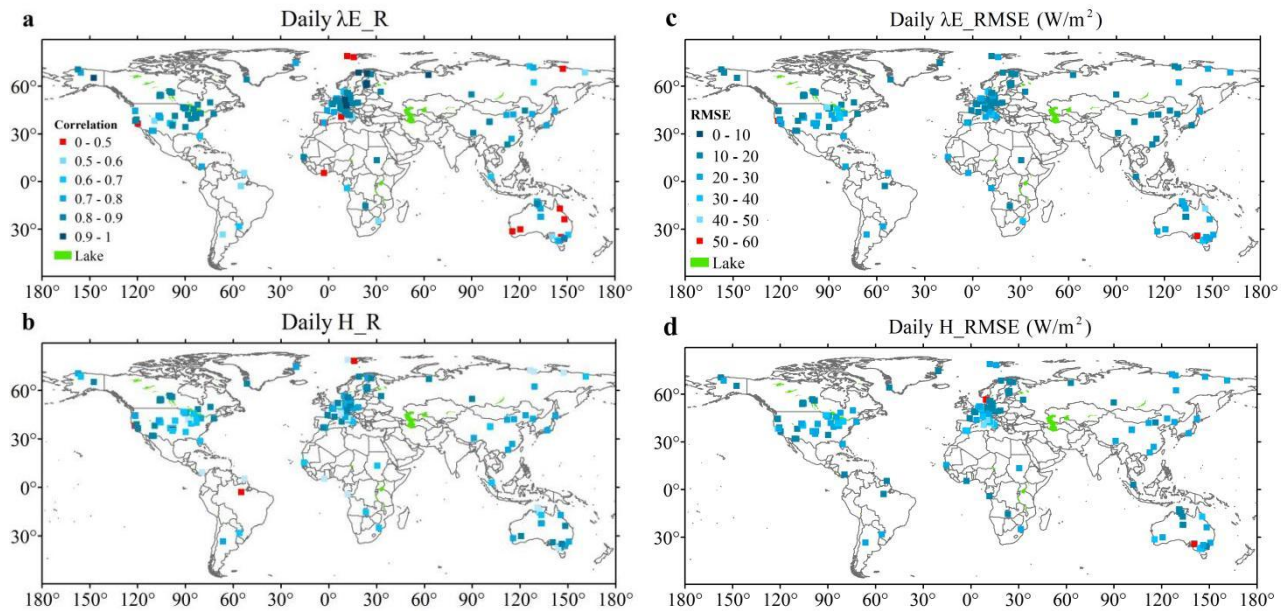
Supplementary Figure S5. Density scatter plot for the cross-validation in terms of values for savannas (SAV), grasslands(GRA), croplands(CRO), and wetlands(WET). Samples of the validation set are randomly composed of five flux towers from one plant function type.



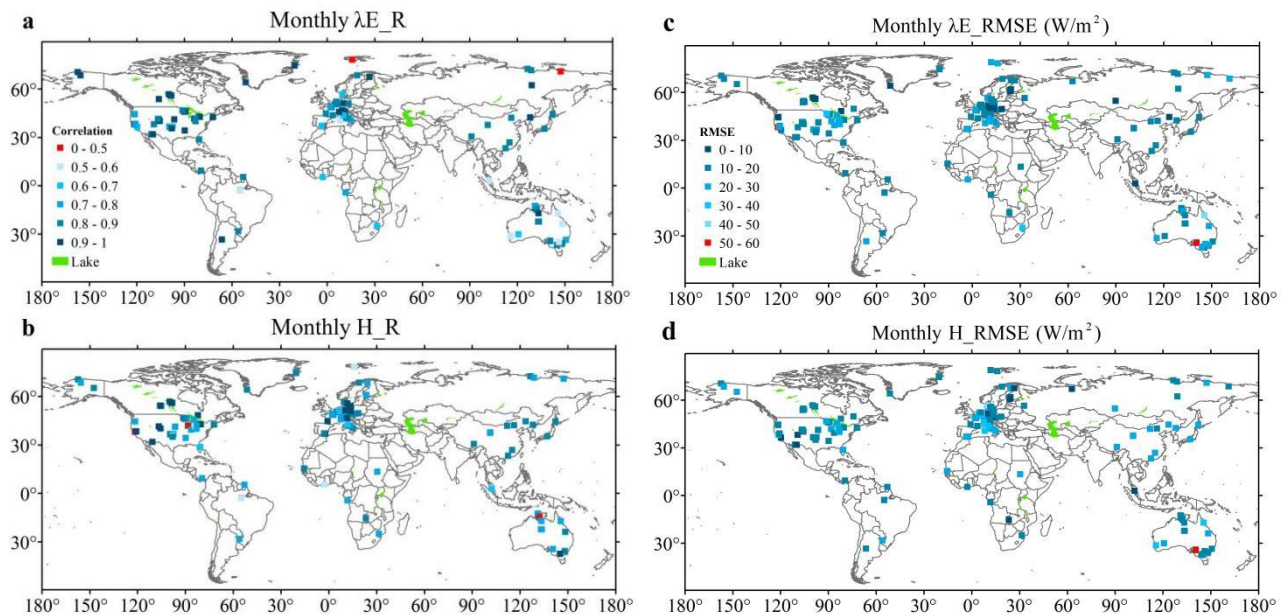
Supplementary Figure S6. Density scatter plot for the cross-validation in terms of trends for different samples from ENF, EBF, DBF, MF, and OSH, respectively. The validation set is randomly composed of one flux tower from one plant function type, and the trends are estimated for all time periods.



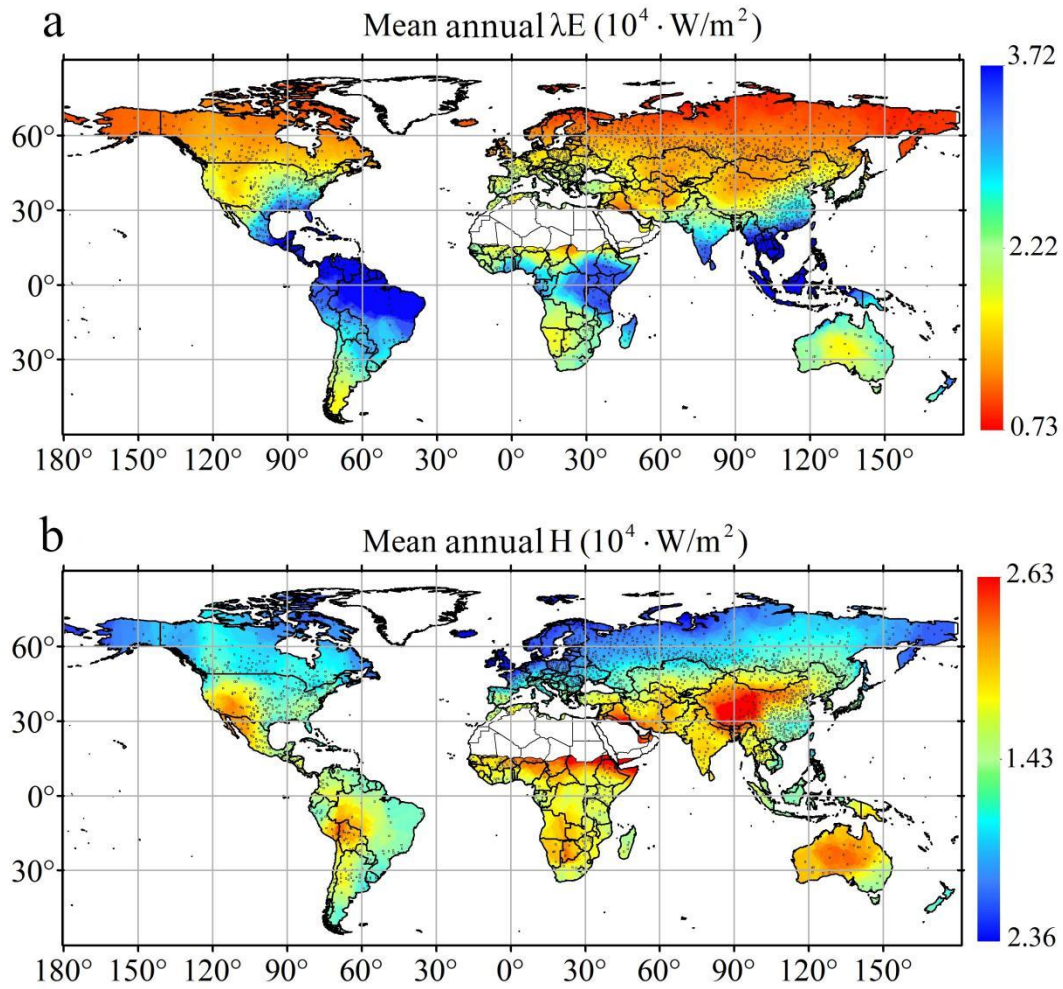
Supplementary Figure S7. Density scatter plot for the cross-validation in terms of trends for different samples from SAV, GRA, CRO, and WET, respectively. The validation set is randomly composed of one flux tower from one plant functional type, and the trends are estimated for all time periods.



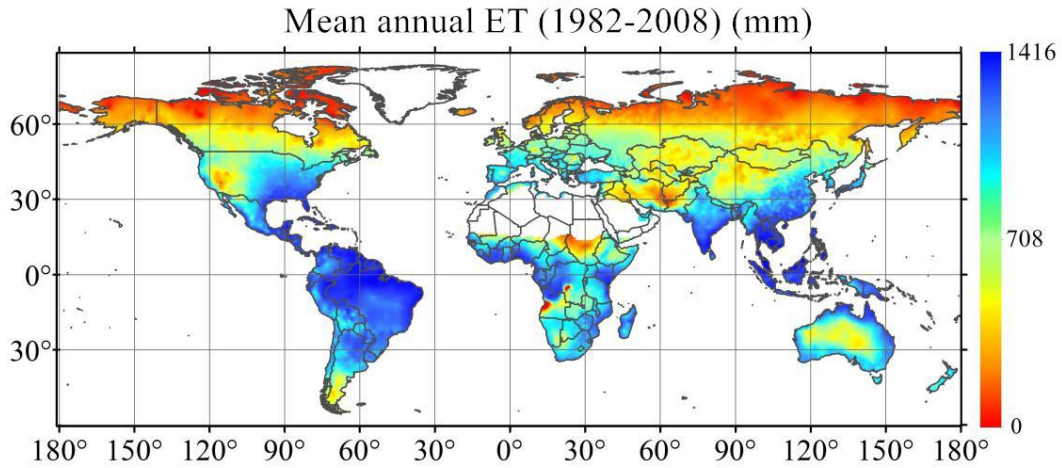
Supplementary Figure S8. Spatial distribution of correlation coefficient (R) and root mean squared error (RMSE) between the predicted latent heat flux (λE) (sensible heat flux (H)) and observed λE (H) at the daily scale for all flux towers.



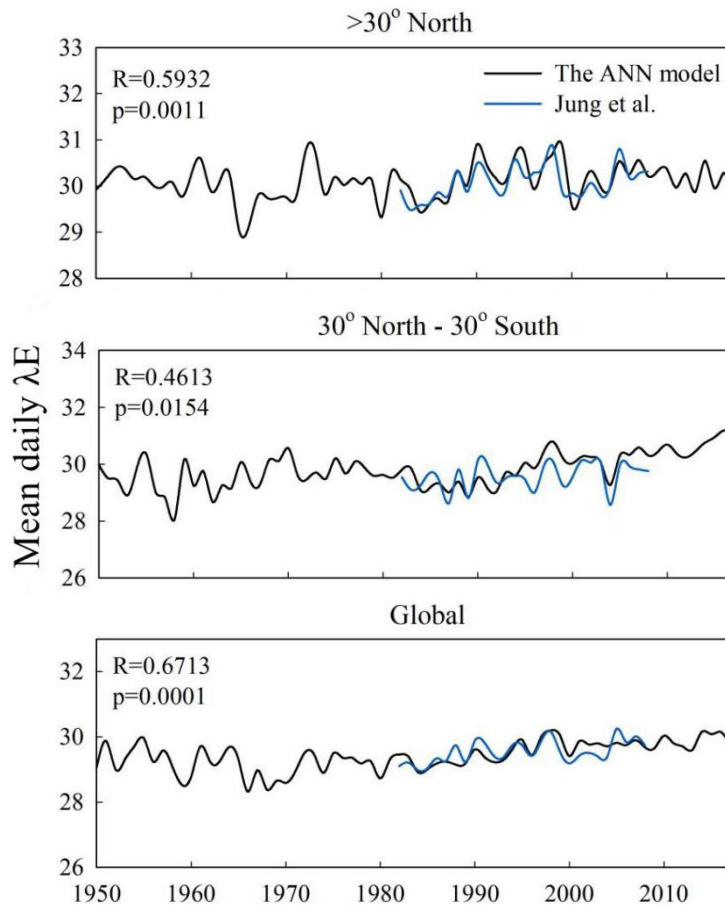
Supplementary Figure S9. Spatial distribution of R and RMSE between the predicted λE (H) and observed λE (H) at the monthly scale for all flux towers.



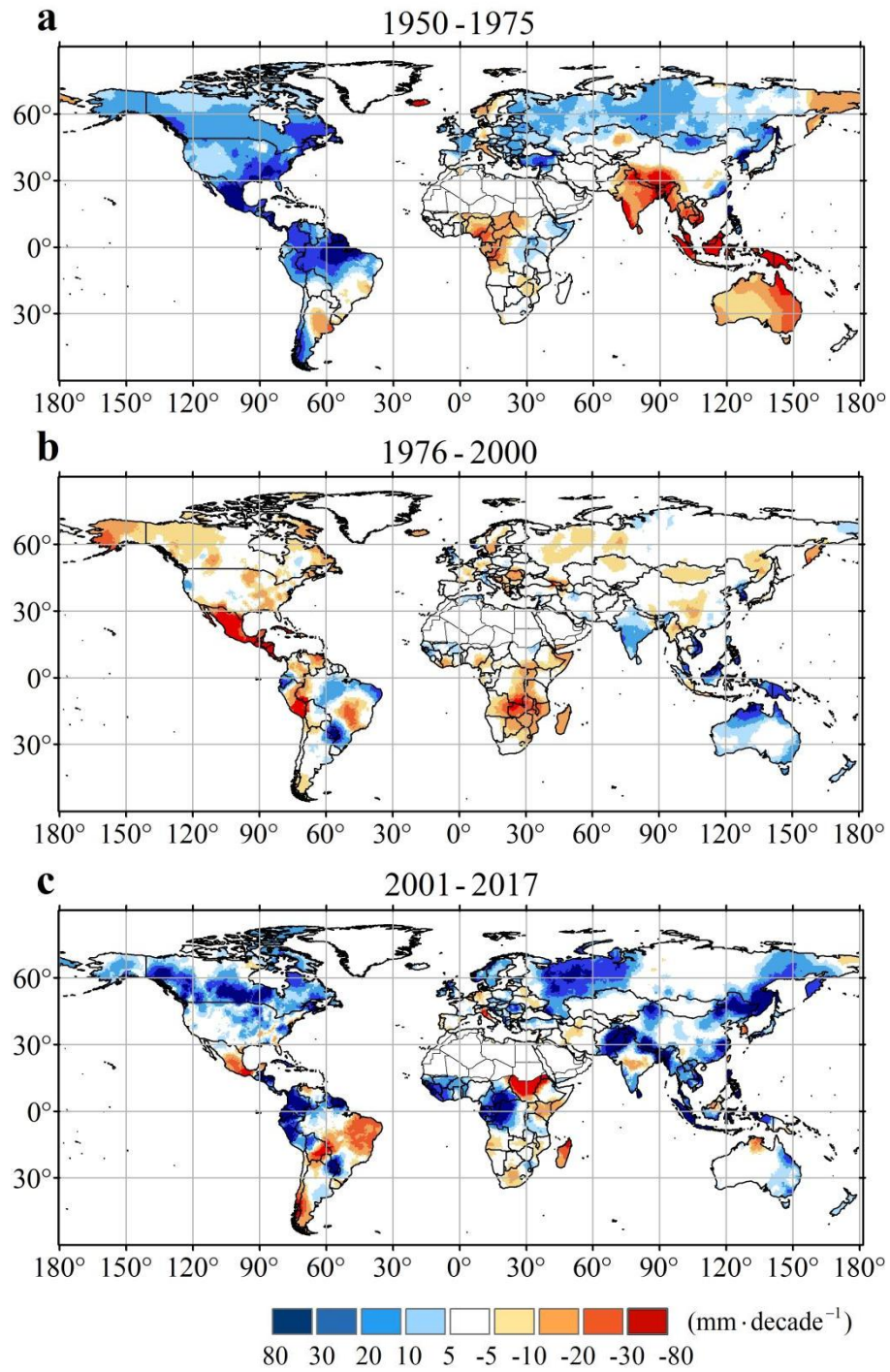
Supplementary Figure S10. Spatial distribution of the mean annual latent heat and sensible heat fluxes.



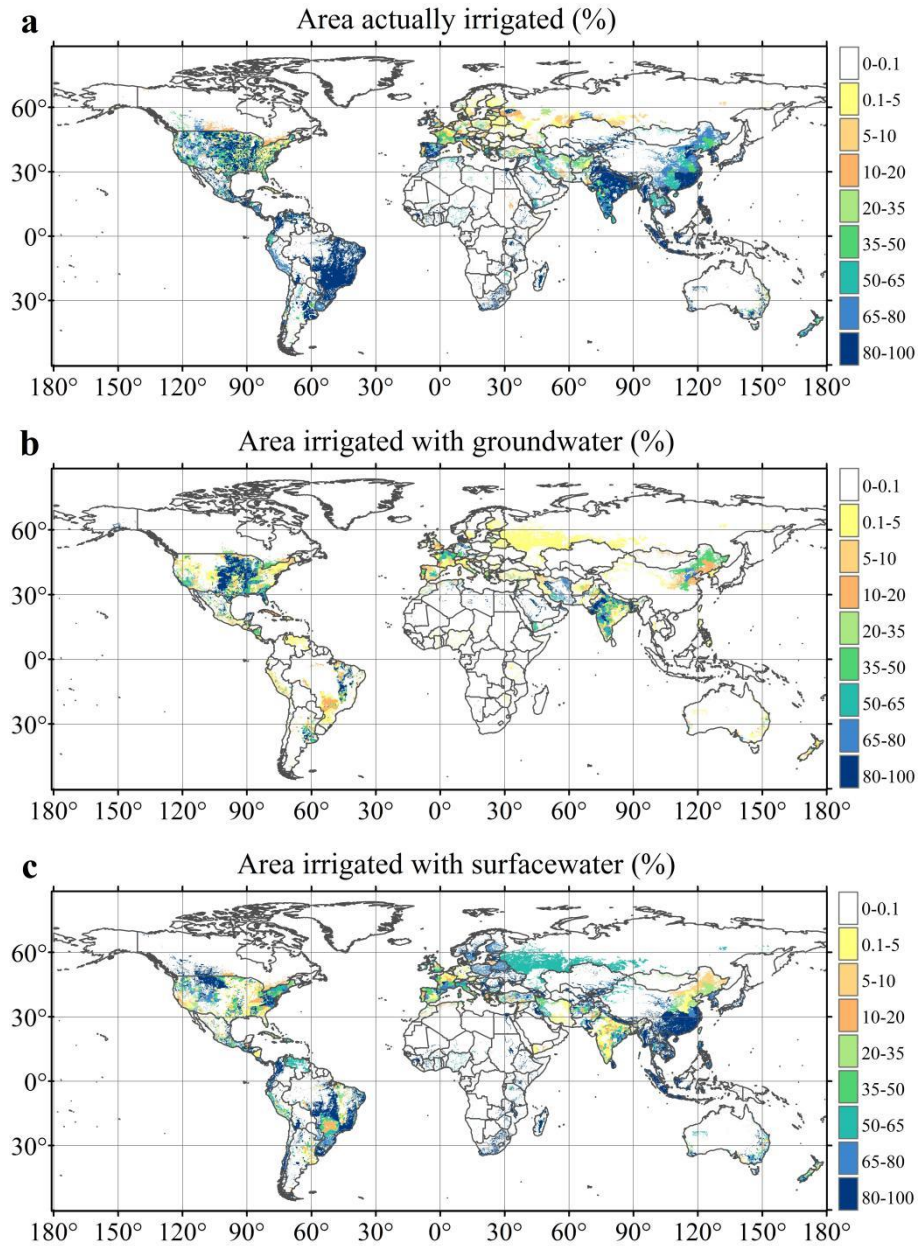
Supplementary Figure S11. Mean annual evapotranspiration (ET) during 1982–2008.



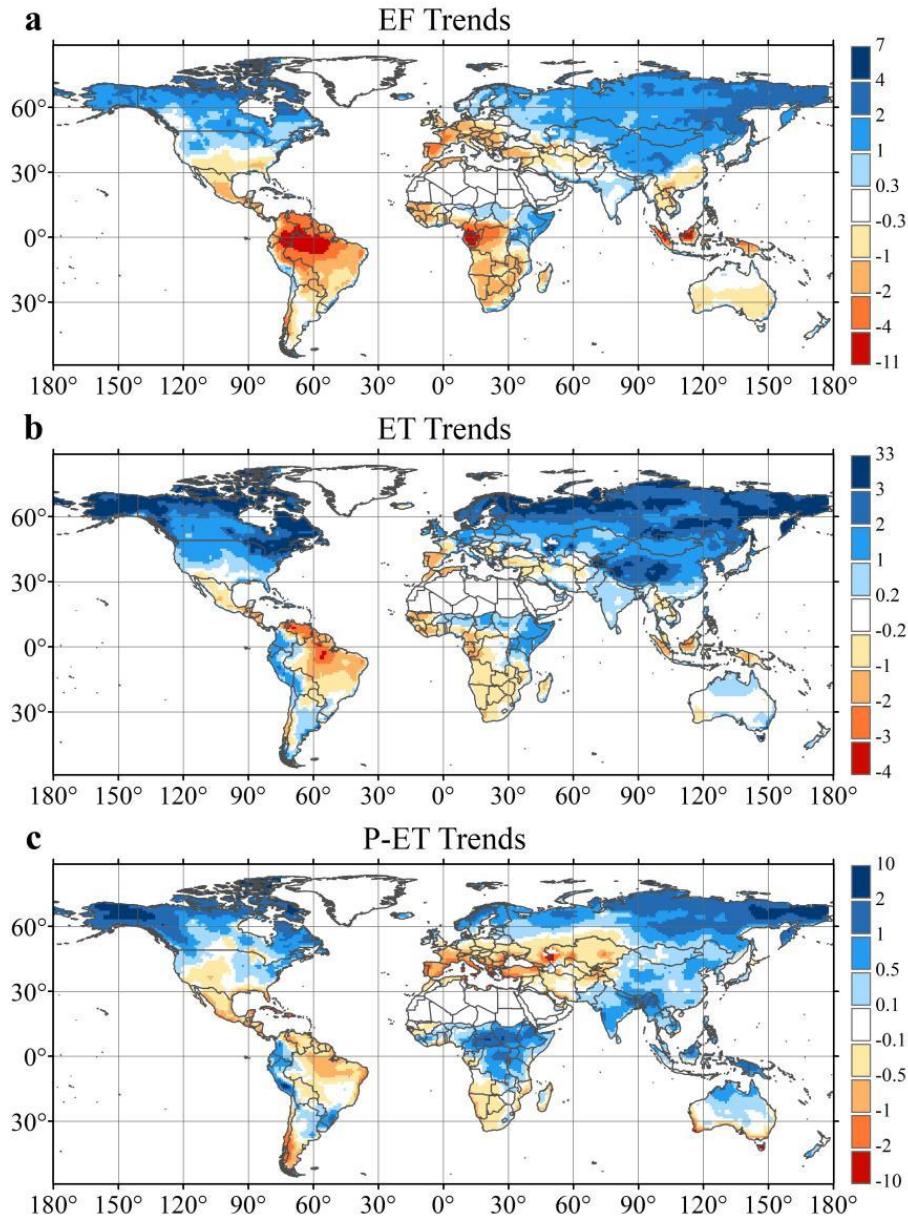
Supplementary Figure S12. Regionally temporal changes of the λE estimated by the ANN model in this study and the λE estimated by the Model Tree Ensemble (MTE) over the regions of $>30^\circ$ North, 30° North– 30° South, and global land, respectively.



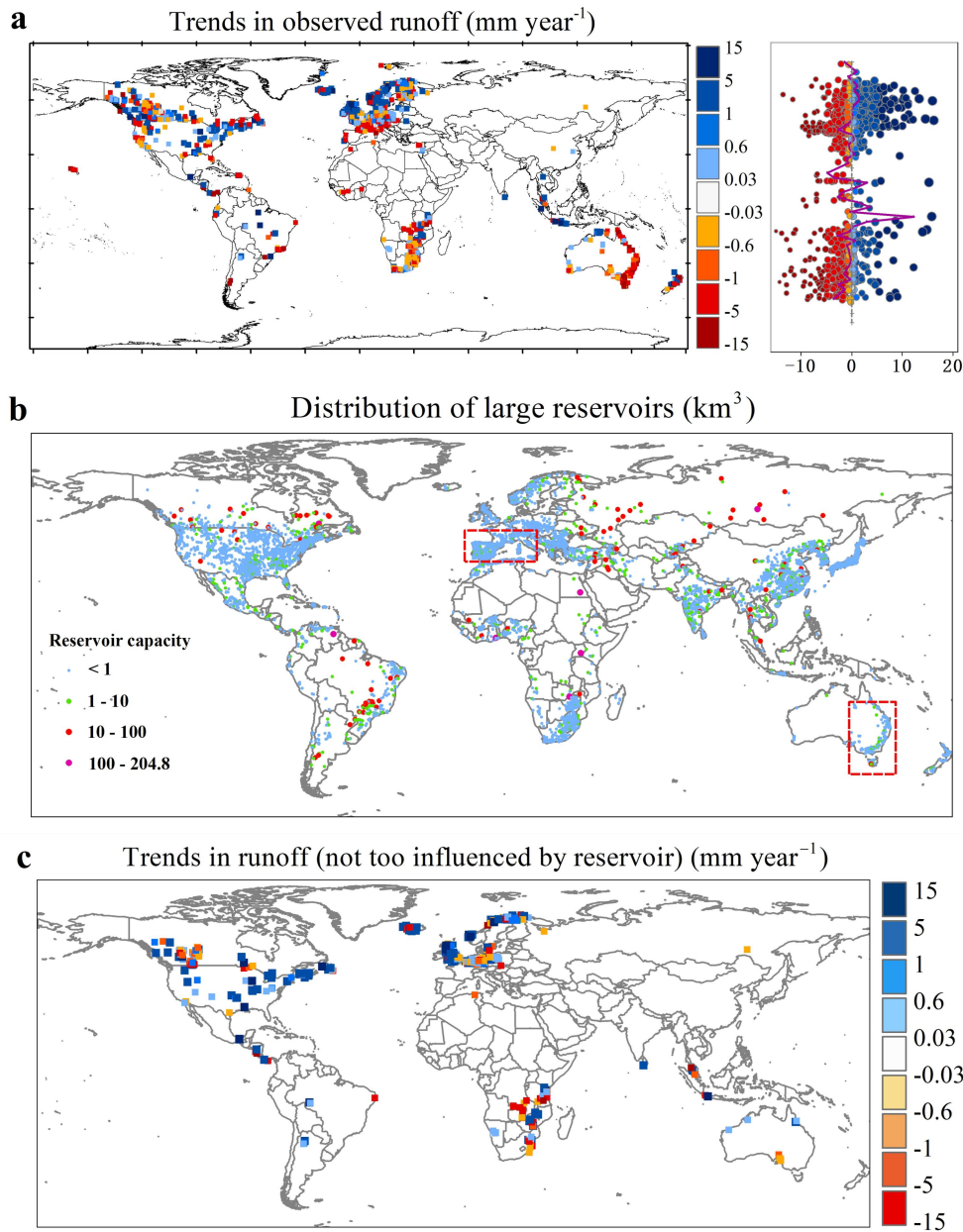
Supplementary Figure S13. Trends in annual precipitation during different historical periods.



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^2 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 $< 1 \text{ km}^3$ and (2) 5 latitude/longitude degree away from the reservoirs with a capacity of $> 1 \text{ km}^3$.

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 top-of-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 (λE) 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, λE 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.

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

Combination of different variables	λE		H	
	R	RMSE (W m ⁻²)	R	RMSE (W m ⁻²)
{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; WS; P}	0.77	27.12	0.74	34.34

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

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
λE	W/m ²	The daily integrated dataset	Output variable
H	W/m ²	The daily integrated dataset	Output variable

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