A Deep Learning-Based Hybrid Model of Global Terrestrial Evaporation

Akash Koppa (akash.koppa@ugent.be)
Ghent University https://orcid.org/0000-0001-5671-0878
Dominik Rains
Ghent University
Petra Hulsman
Ghent University
Diego Miralles
Ghent University https://orcid.org/0000-0001-6186-5751

Article

Keywords: terrestrial evaporation, climate, deep learning

Posted Date: September 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-827869/v1

Version of Record: A version of this preprint was published at Nature Communications on April 8th, 2022. See the published version at https://doi.org/10.1038/s41467-022-29543-7.

A Deep Learning–Based Hybrid Model of Global Terrestrial Evaporation

Akash Koppa[1,*], Dominik Rains[1], Petra Hulsman[1], and Diego G. Miralles[1]

[1]Hydro-Climate Extremes Lab (H-CEL), Ghent University, Ghent, Belgium

Abstract

Terrestrial evaporation (E) is a key climatic variable that depends on a plethora of environmental factors. The constraints that modulate the evaporation from plant leaves (or transpiration, E_t) are particularly complex, yet often assumed to interact linearly in global models due to our limited knowledge based on local experimental studies. Here, we combine *in situ* and satellite observations with deep learning to model transpiration stress (S_t), i.e. the reduction of E_t from its theoretical maximum. Then, we embed the new S_t formulation within a process-based model of E to yield a global hybrid E model. In this hybrid, the S_t formulation is bidirectionally coupled to the the host model at daily timescales. Comparisons against *in situ* data and satellite-based proxies demonstrate an enhanced ability to estimate S_t and E globally. Therefore, the proposed approach provides a framework to improve the estimation of E in Earth System Models and our understanding of this crucial climatic variable.

Main

E is a key element of the global water cycle: approximately two-thirds of the precipitation over land is evaporated back to the atmosphere[1]. Due to its influence on water vapor and cloud feedbacks, E plays a crucial role in global warming, and its projected increase is expected to intensify the global hydrological cycle [2]. Changes in E will not only have far-reaching consequences on water availability and climate[3,4], but can also severely affect the occurrence of hydroclimatic extremes[5] and the ability of ecosystems and river basins to recover from them[6–8]. Moreover, E is an important indicator of vegetation stress, thus it is widely used for estimating drought conditions[9] and their implications for water management, ecosystem health, and agricultural production[10]. Its reliable representation in hydrological and climate models is therefore crucial, and so is its accurate global monitoring from space. However, E cannot be derived directly from satellite imagery, and current retrieval algorithms also rely on process-based formulations.

Several approaches exist to estimate E at large scales based on process-based models. Some simulate E as a residual of the energy balance or derive it empirically using vegetation, temperature and radiation information. This approach is primarily employed by high-resolution satellite retrieval algorithms, especially in agricultural areas, owing to minimal input data requirements[11]. Other models employ a flux-based approach to derive E using process-based methods such as the Monin-Obukhov similarity theory to calculate the gradients of specific humidity between the atmosphere and land (vegetated or non-vegetated) surface, and explicitly model the surface resistance to the diffusion of water vapor. Such an approach is prevalent in climate models [12]. Finally, a third and a common approach in hydrological models[13], as well as satellite retrieval algorithms [14–16], is the calculation of the theoretical maximum, or potential evaporation (E_p), for given land cover and meteorological conditions. Subsequently, actual E is calculated by reducing E_p by a certain factor (S), which is designed to account for the 'evaporative stress' experienced by the vegetated (or non-vegetated) surface. Despite this wide range of approaches, significant uncertainties exists in the current global estimates of E, and that applies to both climate models[17] and satellite-based algorithms[18].

In this study, we focus on stress-based models of E, the most common means to derive global evaporation from satellite data[19]. In such models, uncertainty arises from the formulations of E_p and S (and particularly S_t). While several process-based formulations of E_p exist[20,21], they differ in their estimates substantially, and even the mere definition of E_p as a concept remains elusive[22]. Nevertheless, the chosen E_p function forms the least empirical part of the stress-based E models, and while parameters within E_p formulations can be better constrained with more data[23], the opportunities to improve stress-based models via modifications to E_p remain limited[24]. Therefore, we focus on the other source of uncertainty: the S formulation. Here, the major uncertainty arises from the lack of understanding of the response of vegetation to environmental stressors, particularly at the spatial resolution at which global models operate. The E_t stress factor (i.e. S_t) should encapsulate multiple interacting hydroclimatic variables that affect different aspects of plant physiology and structure in a highly non-linear manner. However, stress formulations used in existing global models are simple, not capturing all the influences and interactions between stressors. This occurs because they are based on a limited number of experimental studies whose extrapolation to global scale is hindered by their local nature[25–27]. The complexity of the interactions among these stressors, and the fact that they involve physiological processes that are unobserved, calls for machine learning techniques as a suitable solution to this long-standing challenge.

Machine learning methods have become popular in Earth sciences in recent years, enabling the discrete classification of important geo-spatial variables which are hard to map, such as clouds[28], soils[29], and forest cover[30]; but also estimation of dynamic variables, such as carbon fluxes[31], precipitation[32], or river discharge[33]. In fact, machine learning models trained on *in situ* measurements of E and other hydro-meteorological covariates, have already been used to estimate global E in recent years[31]. However, pure machine learning–based approaches have several disadvantages in realistically modeling earth system processes. Machine learning models do not implicitly obey the physical limits which constrain earth system processes at different scales such as the closure of water and energy balances, unless they are externally imposed. Further, the black–box nature of machine learning hinders the interpretability of such models, an important requirement if the importance of individual covariates need to be realistically represented, and if such models are to be used improving process understanding. However, advances in the growing field of explainable artificial intelligence have shown promise in mitigating this issue[34].

An emerging research direction, and the approach adopted in this study, is to combine process-based and machine learning models in a symbiotic manner. Here machine learning, and specifically deep learning, is used to directly model the earth system process of concern, with the hypothesis being that deep learning methods can learn the functional relationship between covariates (stress drivers in this study) and the target process or phenomenon (evaporative stress). For example, deep learning methods have proven to be very effective in learning sub-grid processes such as convection in coarse resolution climate models by emulating computationally expensive high resolution models[35,36]. Further, such formulations can be embedded within process-based models to create 'hybrid' models which retain the advantages of process-based models, i.e. physical consistency and interpretability, and machine learning models, i.e. more realistic data-driven formulation of processes that are insufficiently understood[37]. Several proof-of-concept implementations have demonstrated the advantages of hybrid modeling in climate sciences with machine learning sub-models employed for representing different processes[35,38] or for improved model parameter discovery[39]. For modeling E in particular, attempts have been made to physically constrain pure machine learning models to improve the accuracy of E estimates[23]. However, an important research question is whether hybrid models capable of operating at a global scale with machine learning used for completely replacing a specific process formulation, as opposed to more accurate parameter discovery or emulating high resolution models, can be developed.

Here, we exploit recent progresses in satellite-based remote sensing, deep learning, and an unprecedented number of *in situ* observation stations spread across the globe to develop a novel formulation of S_t from the ground-up without any prior assumptions. Further, we implement the new formulation of S_t, and execute it online, in a process-based model of global evaporation which provides physical constraints to the deep learning-based stress formulation. In doing so, we develop a hybrid model capable of simulating E daily at the global scale. We develop the hybrid model in such a way that the new deep learning-based formulations of S_t is tightly coupled to the process-based model. A comprehensive validation of the model is carried out with *in situ* observations. Further, the improvements, or lack thereof, compared to the process-based model is evaluated.

Results

Hybrid Model Architecture. A hybrid model at the highest level of abstraction is made up of two components: a process-based host model and machine learning-based sub-models/formulations/parameterizations embedded in the host for representing certain processes[37]. For the process-based model, we choose the Global Land Evaporation Amsterdam model (GLEAM)[14,40]. GLEAM simulates E as a summation of its constituents: E_t, bare-soil evaporation (E_b), and interception loss (E_i). E_t and E_b are estimated for every grid cell of the global model using a Priestley Taylor-based approach and their respective stress factors (S_t and S_b). Interception is based on the Gash analytical model[41] (Figure 1). The model contains a multi-layer soil water balance model in which satellite-based surface soil moisture data is assimilated. S_b is a function of soil moisture availability (see Methods), while S_t accounts for the stress experienced by vegetation due to the shortage of plant available water (PAW) and the phenological state (represented by vegetation optical depth, VOD). However, in reality, several additional stressors are responsible for limiting E_t, with the exact responses to these stressors being species-dependent and difficult to model. The hypothesis here is that, by combining sufficient reliable data of the stressors using deep learning, functional relationships among the different stressors and S_t can be uncovered.

Deep learning models are developed for tall and short vegetation separately (see Methods for the details of the target variable and covariates used in the deep learning model). We consider four other stress drivers in addition to PAW and VOD that are known to regulate stomatal conductance and hence S_t: (a) vapor pressure deficit (VPD), an indicator of atmospheric dryness[42], (b) air temperature (T_a), to include the effects of sub-optimal temperature and heat stress[43], (c) incoming shortwave radiation (SW_i), to incorporate the influence of light limitation [44], and (d) atmospheric carbon dioxide (CO_2) concentration, which exhibits a first order control on stomatal opening [45]. We note that the potential effect of phosphorous and nitrogen limitations on S_t[46] is not considered in this study due to the lack of dynamic global data.

Finally, the hybrid model of global E is created by coupling the deep learning-based model of S_t to the GLEAM process-based model. At every (daily) time step, and at every (0.25 degree) grid cell of the global model, the soil water balance module of GLEAM translates precipitation (P) into PAW. Then, PAW, VOD, T_a, VPD, SW_i ,CO_2 are input to the (trained) deep learning model (see Methods). The deep learning model is run in predictive mode to generate S_t. S_t is then used to constrain E_p and thus compute E by the process-based host model. Finally, E is used to update the soil moisture (and PAW) before the next time step (Figure 1).

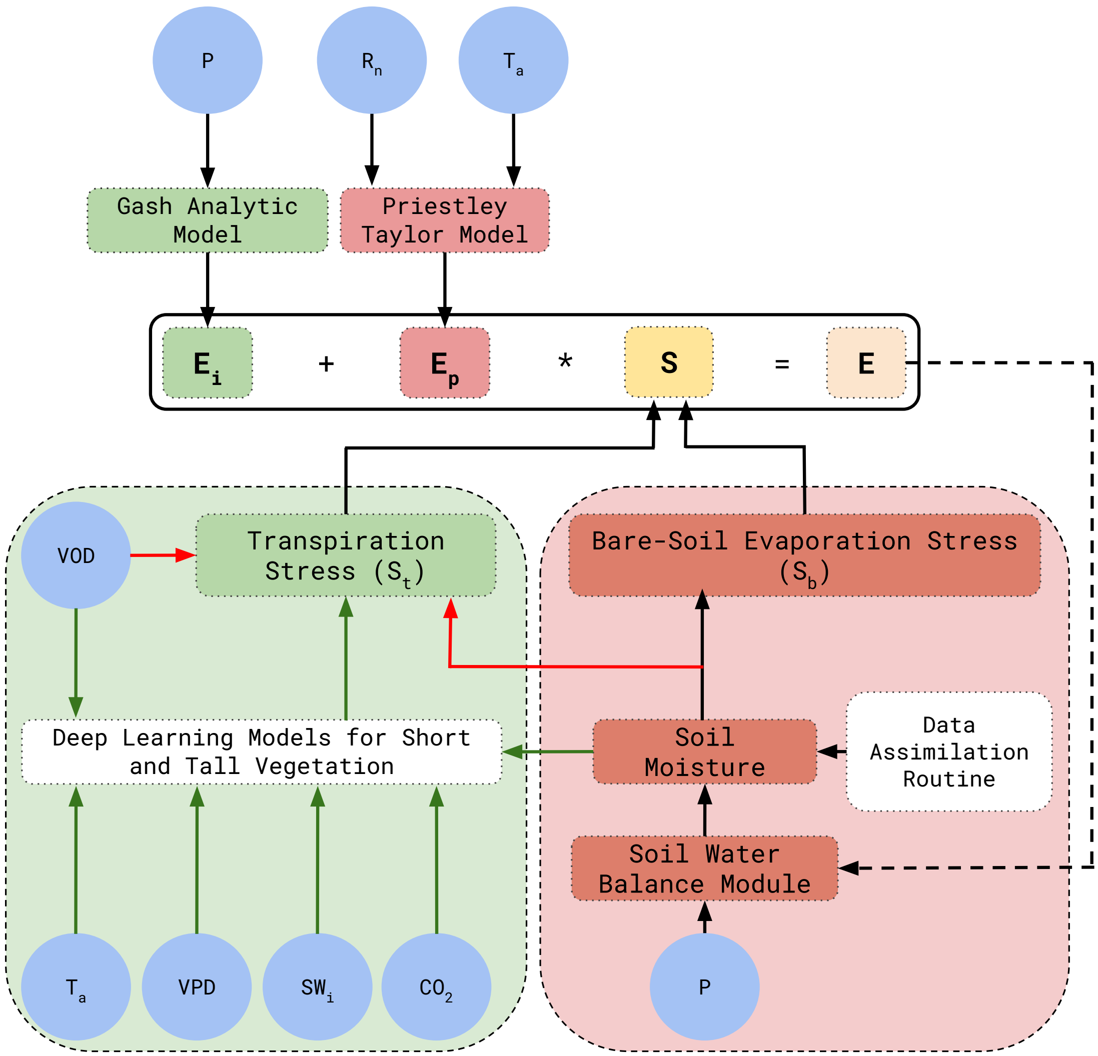

Figure 1: Schematic of the hybrid terrestrial evaporation model. E_i is interception, E_p is potential evaporation, S is the evaporative stress factor, and E is actual evaporation. The red arrows indicate modeling steps which are exclusive to the processed-based model, the green arrows are steps which have been added in the hybrid, and the black arrows are steps common to both the models.

Validation with *in situ* measurements. S_t and E estimates from the hybrid model are validated at 368 *in situ* monitoring stations (see Figure 10 in Supplementary Information) sourced from several flux tower databases (refer to the Methods section for the calculation of S_t from flux tower data). The hybrid model performance is compared to that of the fully process-based model using violin plots and spatial maps illustrating the Kling-Gupta Efficiency (KGE), a metric which combines correlation, variability bias, and mean bias (see Methods). KGE values theoretically range from $-\infty$ to 1.0, with values greater -0.41 implying that the model is a better predictor than the mean seasonal cycle[47].

The violin plots (Figure 2a) show the distribution of KGE values calculated for the 231 stations in short-vegetation ecosystems, and the 137 stations in tall-vegetation ecosystems. We see that both the process-based model and the hybrid model perform well in estimating S_t for short vegetation ecosystems (including Croplands, Shrub and Grasslands, and Wetlands; see Table 3 in Supplementary Information for station-wise land cover classification). For most stations ($>$ 75%), KGE values from the process-based model exceed -0.41. However, the deep learning-based model of S_t improves these results considerably (the median KGE value is positive, unlike that of the process-based model). This improvement is even more pronounced for tall vegetation (consisting of Broadleaf, Needleleaf, and Mixed forests; see Table 3 in Supplementary Information) – see Figure 2a. The observed improvement in KGE is attributable to improvements in the bias and variability components of KGE rather than correlation – refer to Figure 1 in Supplementary Information for violin plots of correlation and root mean square error (RMSE). In fact, the average correlations of the process-based model estimates of S_t are similar or even marginally better than for the hybrid model. However, the RMSE of the hybrid model estimates of S_t are substantially lower, particularly for the tall vegetation class.

Next, we check whether the improvement in the estimation of S_t in the hybrid model is propagated to the simulation of E. From Figure 2b, it is evident that the improvements in S_t are not linearly translated to E estimates. This can be attributed to the fact that the vast majority of the flux towers are in energy-limited regions, where E dynamics are influenced more by E_p than by S_t. Overall, both models exhibit high, and similar, KGE values (median value of approximately 0.5) for short vegetation. For tall vegetation, the hybrid model outperforms the process-based model. In terms of correlation and RMSE, we see that both models are performing similarly: the process-based model exhibits marginally higher correlations, while the RMSE of the hybrid model is lower for both vegetation classes (see Figure 1 in Supplementary Information).

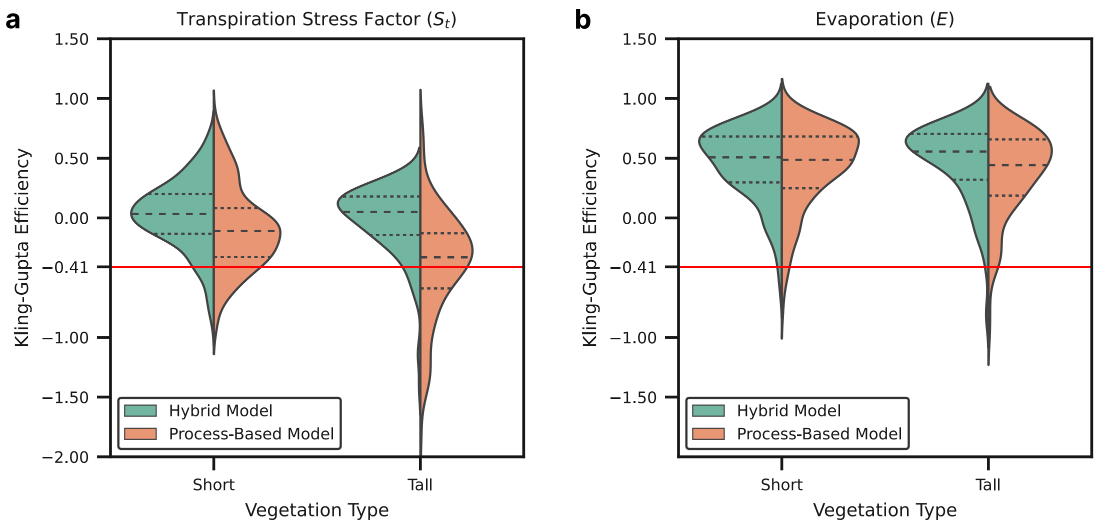

Figure 2: **a** and **b** Violin plots showing the distribution of the Kling-Gupta Efficiency (KGE) metric for transpiration stress factor (S_t) and evaporation (E), respectively, calculated for all flux tower sites. The KGE distribution for the hybrid and process-based models are classified according to short and tall vegetation types. The dashed lines represent the median (large dashes) and the interquartile range (small dashes). The red line represents a KGE value of -0.41, above which a model prediction or simulation is considered better than the mean seasonal cycle.

To understand the difference between the hybrid and process-based models better, we compare the spatial distribution of difference in KGE values for S_t and E estimates from the two models for different geographical zones (Figure 3, also see Figure 2 and Figure 3 in Supplementary Information for absolute values of KGE for S_t and E). In North America (NA), which has the largest number of flux towers, the hybrid model outperforms the process-based model, especially in the humid eastern and north-eastern parts. In comparison, both models tend to inaccurately simulate S_t in the arid south-west region. In Europe (EU), the hybrid model performs better than the process-based model across the majority of the flux tower stations, including the stations which are in the relatively arid south. However, in Asia (AS) and rest of the world (RW), the performance of the hybrid model is very similar to the process-based model. One reason could be that the AS and RW regions have a very sparse distribution, and thus flux towers in those ecosystems may have distinct biophysical characteristics from the majority of sites in the training database. Further, we compare the spatial maps of correlation and RMSE (see Figure 4–Figure 7 in Supplementary Information) to understand the source of the disparity in KGE values. In terms of correlation, the two models perform very similar to each other across the different regions. Therefore, the major source of improvement in the hybrid model can be traced to the better estimation of the variability, and to a smaller extent, the bias, seen in the observation, a fact supported by the violin plots (Figure 2a). Further, we notice that the discrepancy in the S_t estimates between the two models, does not translate to E estimation in energy limited regions (Figure 3), which are poorly represented in the training data.

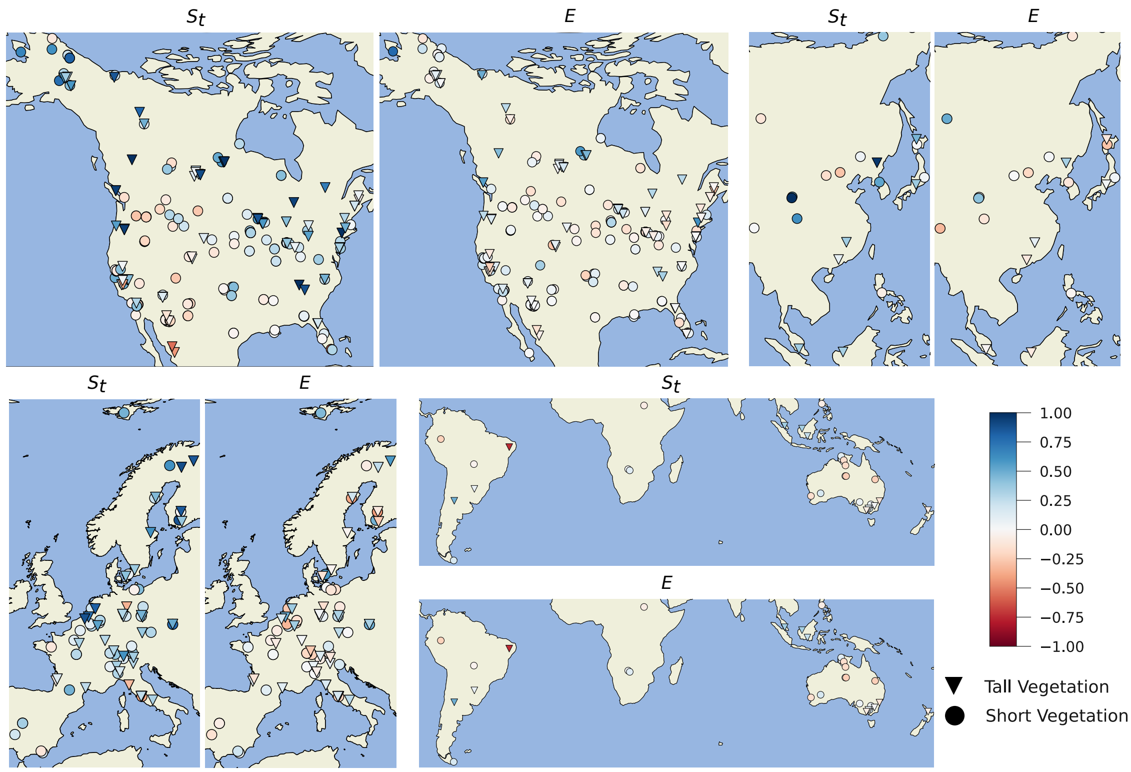

Figure 3: Maps showing the difference in Kling-Gupta Efficiency (KGE) metric between the hybrid model and process-based model for transpiration stress factor (S_t) and evaporation (E) calculated using observations at flux tower sites in different geographical zones: North America (NA), Asia (AS), Europe (EU), Rest of the World (RW). Blue (red) tones indicate improvement (degradation) in the hybrid model compared to the process-based counterpart.

Comparison with global datasets. In contrast to point-scale measurements in flux towers, which have a small footprint, the GLEAM model generates spatially and temporally continuous estimates of S_t and E over the entire continental surface. Therefore, it is important to validate the hybrid model against independent global estimates of both S_t and E. We validate S_t and E_t by comparing their seasonal aggregates with other global datasets in Figure 4 and Figure 5 respectively. To further investigate the realism of the global S_t and E estimates, the temporal dynamics are investigated in Figure 6 by displaying correlation maps based on monthly time series.

Due to the absence of direct observations of S_t at those scales, we choose a satellite-retrieved proxy that has been shown to represent the evaporative stress experienced by vegetation reasonably well: the ratio of solar-induced chlorophyll fluorescence to photosynthetically-active radiation (SIF/PAR)[48] (see Methods). We note here that the scale and range of SIF/PAR values is different from that of S_t, but that the spatial gradients and temporal dynamics are expected to be comparable. In June-July-August (JJA), summer season in the Northern Hemisphere, we see that the spatial patterns of S_t in the hybrid model are similar to those in the process-based model (Figure 4a and c).However, the hybrid model captures better the higher vegetation stress that is suggested by the low values of SIF/PAR (Figure 4e). For December-January-February (DJF), the picture is similar (Figure 4b,d,f).On the other hand, we see a possible underestimation of S_t (too much stress) by the hybrid model in the rainforests of Congo, Amazonia and Eastern Asia, both in JJA (Figure 4a) as well as DJF (Figure 4a). This points again to the importance of sufficient data availability for deep-learning methods. Figures 6a,c show the temporal correspondence between S_t and SIF/PAR for the hybrid and process-based models, respectively, while Figure 6e shows the difference between the two previous maps. We see that the hybrid model exhibits positive correlation with SIF/PAR over a majority of the grid cells with parts of Amazonia, Congo, and South East Asia (Figure 6a) being the exception. It shows better correlation with SIF/PAR compared to the process-based model's S_t formulation (Figure 6b) in eastern China and in the norther latitudes (Figure 6c). In contrast, the process-based model shows higher correlation in large parts of eastern North America, Europe, and Australia. In addition, the hybrid model shows a marked improvement in the spatial correlation with SIF/PAR (0.66 compared to 0.59 for the JJA season and 0.42 compared to 0.34 for the DJF season).

We also compare the E estimates from the hybrid and process-based models with a pure machine learning-based E dataset (FLUXCOM) which is trained on a subset of the global flux towers used in this study[31]. We see that in both seasons, JJA and DJF, the spatial patterns of E from our hybrid and process-based models are similar to that of FLUXCOM (Figure 5). Regions of divergence are seen in the north eastern parts of South America, and southern and eastern Africa where the FLUXCOM E estimates are higher than that of the hybrid and process-based models, especially in the JJA season. As far as the correlation maps (Figure 6b, d) are concerned, the hybrid model estimates of E are highly correlated with FLUXCOM. A major region of divergence that stands out in both the hybrid and process-based models is Amazonia. This maybe due to the fact that very few stations are available in the region for model training, and therefore FLUXCOM estimates may also be more uncertain in Amazonia. The difference between the hybrid and process-based model correlation is nominal (Figure 6f).

The hybrid model also shows mild improvements in the spatial correlation metric (0.84 compared to 0.81 for JJA and 0.95 compared to 0.94 for DJF).

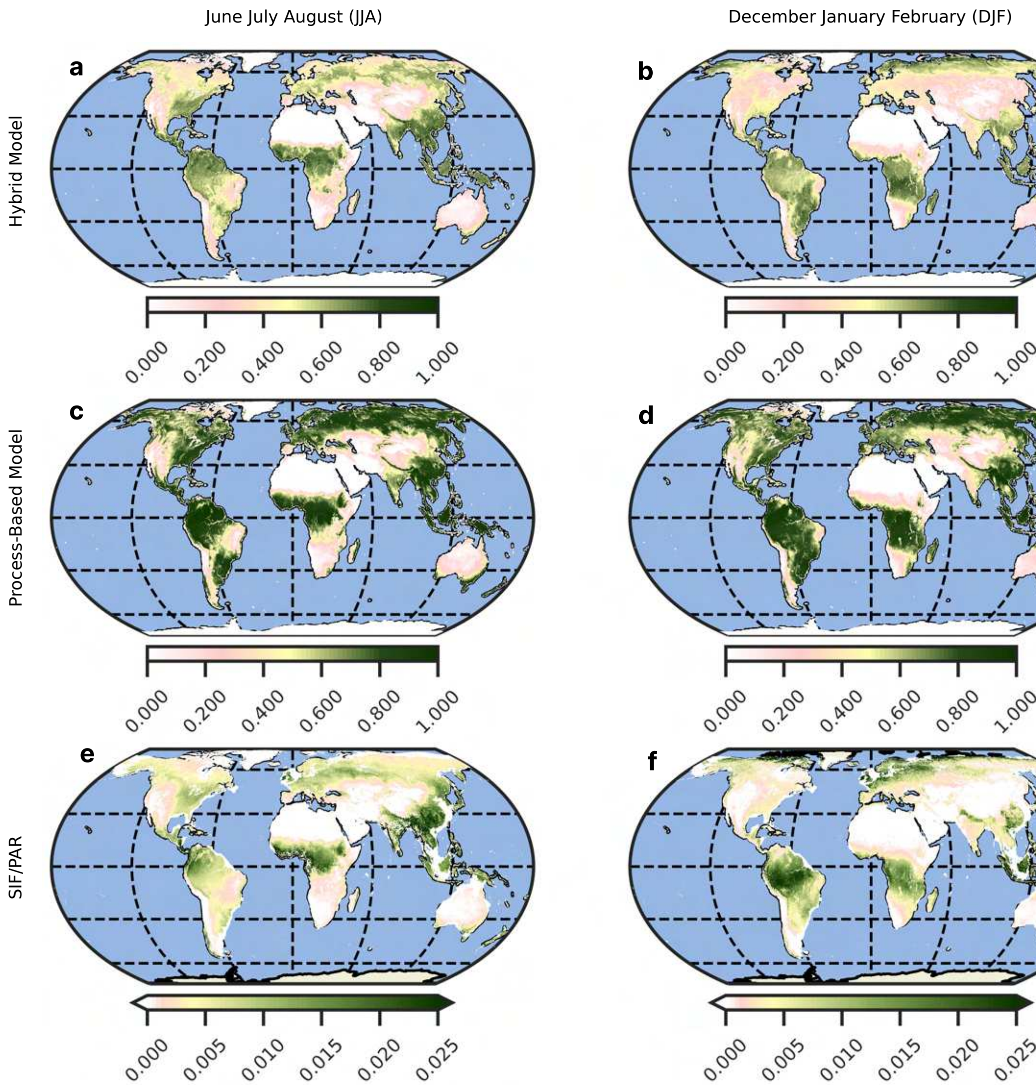

Figure 4: Comparison of the seasonal aggregates of the transpiration stress factor from the processed-based and hybrid models compared with the ratio of solar-induced chlorophyll fluorescence and photosynthetically-active radiation (SIF/PAR) for June-July-August (JJA) (**a**, **c**, and **e**) and December-January-February (DJF) (**b**, **d**, and **f**) seasons. Note: The units of SIF is $mWm^2/sr/nm$ and PAR is represented in W/m^2.

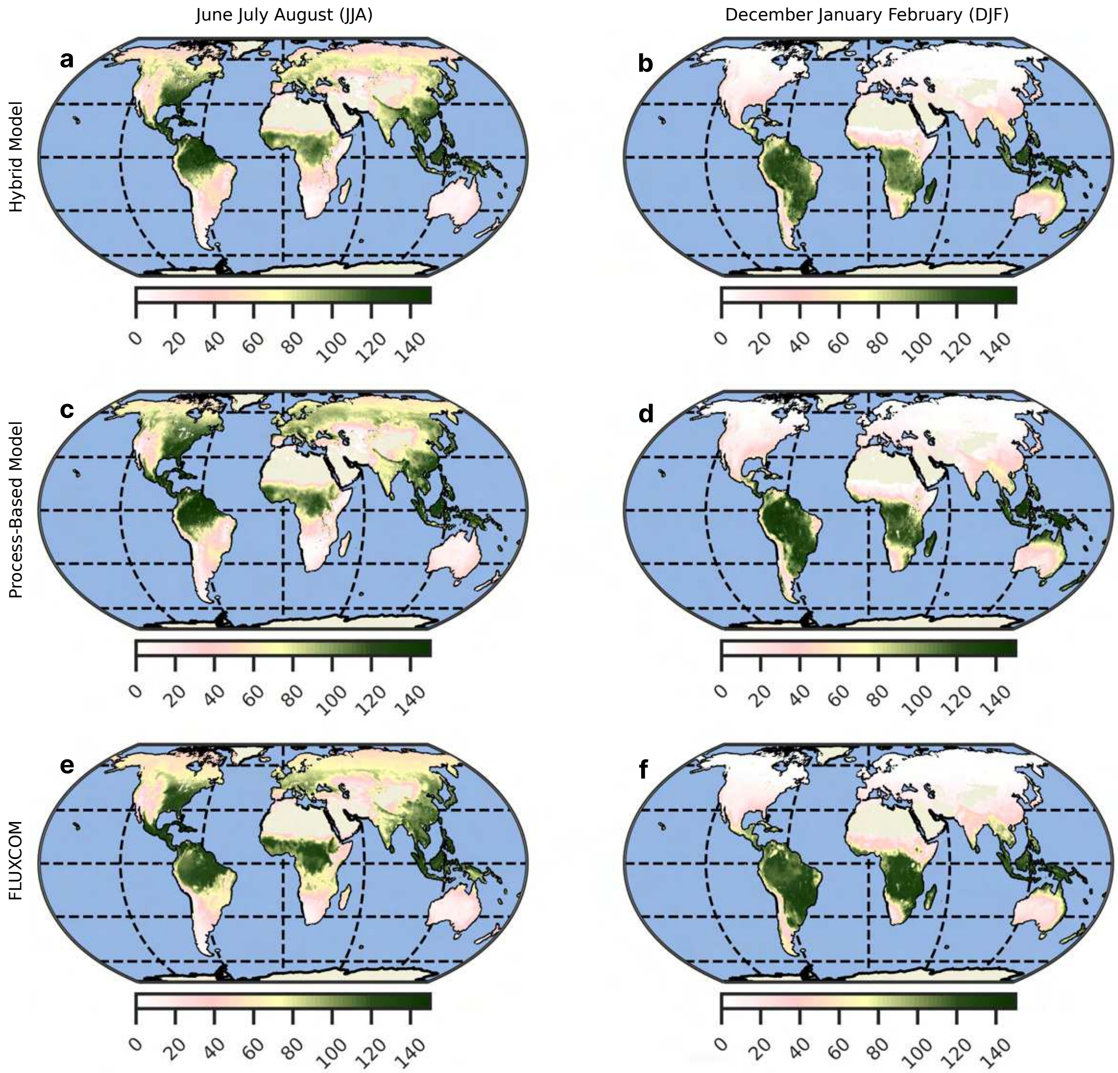

Figure 5: Comparison of the seasonal aggregates of evaporation (E) from the processed-based and hybrid models compared with a purely machine learning-based model trained directly on evaporation from FLUXNET sites as the target variable (FLUXCOM[31]) for JJA (**a**, **c**, and **e**) and DJF (**b**, d, and **f**) seasons. Note: The units of E is $mm/month$.

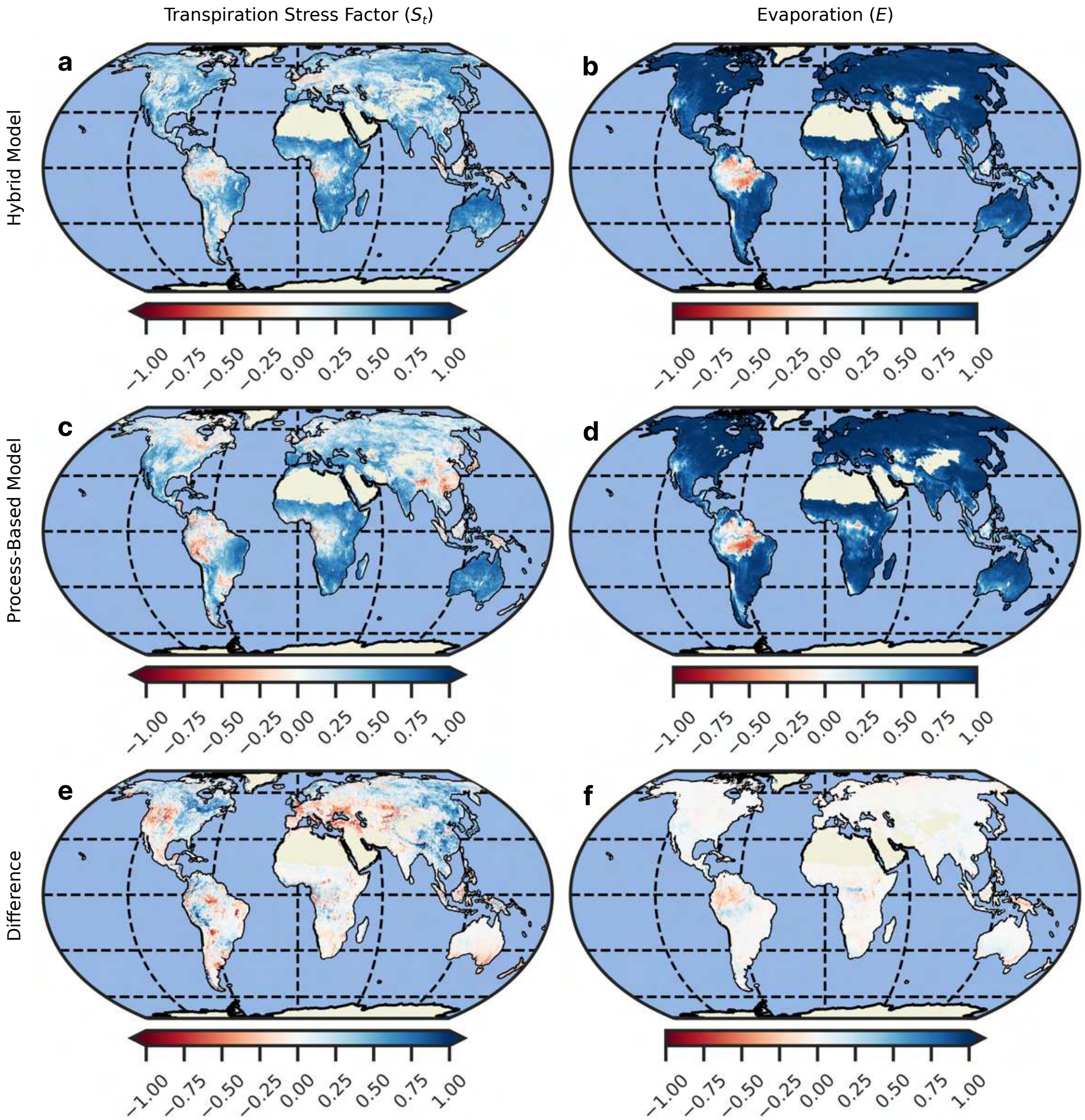

Figure 6: **a** and **c**, Comparison of correlation maps for transpiration stress factor (S_t) between processed-based and hybrid models with observational SIF/PAR. **b** and **d**, Comparison of correlation maps for evaporation (E) between processed-based and hybrid models with machine learning-based estimates (FLUXCOM). **e**, Difference between **a** and **c**. **f**, Difference between **b** and **d**.

Discussion

The growing complexity of large-scale earth system and climate models requires increasingly high computational resources. More importantly, new processes are frequently represented based on limited experimental understanding and are thus uncertain in their application at larger scales. Hybrid modelling approaches have the potential to reduce the ill-effects of over-parameterization, reduce computing time, and even improve accuracy in process representation[49]. Here, we focus on one of the main unknowns in the global water cycle and a key variable in climate models: terrestrial evaporation (E). We develop and apply a global-scale hybrid model of E, in which a deep learning-based formulation of vegetation stress is embedded within a process-based model at daily timescales. We show that the deep learning model, designed with *a priori* assumptions based on expert knowledge, is overall more accurate than the traditional process-based counterpart at capturing the non-linear interacting processes that yield transpirational stress (S_t). Specifically, the biggest improvements in S_t are seen in northern latitudes, likely due to the consideration of incoming radiation (a key driver of stomatal conductance). On the contrary, the deep learning-based S_t tends to overestimate the stress in tropical rainforests, primarily in the DJF season. This highlights a limitation of any deep learning model, in which sufficient availability of training data is crucial: the majority of the flux towers used for training are located in NA and EU. This is especially relevant for modeling Earth system processes such as S_t, which exhibit large regional (and local) variability and for which the ability of any data-driven formulation to generalize over the entire globe will by default be imperfect. Further, the estimates of E from the hybrid model accurately capture the temporal dynamics and the spatial patterns of E seen in both the *in situ* network of flux tower observations and a (purely) machine learning-based dataset (FLUXCOM). From a computational perspective, the model was developed in Tensorflow, a popular Python library for deep learning, which scales across a wide range of hardware, operating systems, and programming languages. Therefore, the transpiration stress model is agnostic of the host model, and hence can be embedded in different global scale earth system models.

Methods

Stress formulation in the process-based model. In GLEAM, the total evaporative stress (S) is composed of S_t and S_b. S_t is defined as

$$S_t = \sqrt{\frac{VOD}{VOD_{max}}}\left(1 - \left(\frac{w_c - w_w}{w_c - w_{wp}}\right)^2\right) \tag{1}$$

where VOD_{max} is the maximum (99th percentile) VOD, w_c is critical soil moisture, w_w is the soil moisture content of the wettest soil layer, w_{wp} is wilting point. S_t is calculated separately for tall and short vegetation.

S_b is defined as

$$S_b = 1 - \left(\frac{w_c - w_1}{w_c - w_r}\right) \tag{2}$$

where w_1 is the surface soil moisture (first layer in the soil water balance module of GLEAM) and w_r is the residual soil moisture content. The values of w_{wp}, w_c and w_r are taken from... Full detail can be found in Martens et al...

Development of the deep learning-based stress formulation. The first step consists of defining the target variable, and the appropriate predictors or covariates. Here, the target variable is the tower-scale S_t, calculated as

$$S_t = \frac{E_t}{E_{pt}} \tag{3}$$

where, E_t is actual transpiration and E_{pt} is potential transpiration.

To estimate E_t in Equation 3, we use daily *in situ* measurements of E, assembled from a total of 557 flux towers. These towers were compiled from FLUXNET[50] (https://fluxnet.org/data/fluxnet2015-dataset/), FLUXNET-CH4 (https://fluxnet.org/data/fluxnet-ch4-community-product/), AmeriFlux (https://ameriflux.lbl.gov/), European Eddy Fluxes Database Cluster (http://www.europe-fluxdata.eu/), and the Integrated Carbon Observation System (ICOS) (https://www.icos-cp.eu/). After the removal of inconsistent values, we end up with 368 stations, out of which 231 stations (approximately 173,000 data points) are classified as having dominantly short vegetation and 137 stations (approximately 103,000 data points) are classified as tall vegetation (refer to Table 3 in Supplementary Information for site-specific details and for the mapping of flux tower land cover class to tall and short vegetation). To separate E_t from E at the flux stations, we use empirical functions relating the ratio of E_t to E to the leaf area index (LAI) for different vegetation classes[51] (see Section 2 in Supplementary Information). We remove rainy days from the flux tower datasets to minimize the impact of interception loss on the measurements of E and sensor errors during rain.

Next, we obtain from GLEAM daily values of E_{pt} for each station. GLEAM uses a Priestley-Taylor formulation to calculate E_{pt} which has been shown to be generally accurate at ecosystem scales [22]. To account for the scale mismatch between grid-scale estimates of GLEAM and point-scale measurements at the flux tower sites, we scale the E_{pt} values with E_t values using days following rain days as follows:

$$E_{pt}^{scaled} = (\frac{E_{pt}^{raw} - E_{pt,mean}^{raw}}{E_{pt,sd}^{raw}}) * E_{t,sd}^{flux} + E_{t,mean}^{flux} \quad (4)$$

where E_{pt}^{raw} is the raw GLEAM E_{pt} for the specific flux tower site, $E_{pt,mean}^{raw}$ is the mean of the raw GLEAM E_{pt} estimates for the specific flux tower site, $E_{pt,sd}^{raw}$ is the standard deviation of the raw GLEAM E_{pt} for the specifc flux tower site, $E_{t,mean}^{flux}$ is the mean of the observed E_t at the specific flux tower, and $E_{t,mean}^{flux}$ is the standard deviation of the observed E_t at the flux tower.Inherent in this bias-correction approach is the assumption that vegetation transpire at their potential on days after rainfall.

The covariates used for modeling S_t are the absolute values and seasonal anomalies of the following variables: a) PAW, b) VPD, c) T_a, d) SW_i e) VOD, f) CO_2. PAW is commonly defined[52] as

$$PAW = \frac{w_w - w_{wp}}{w_c - w_{wp}} \tag{5}$$

The absolute values and anomalies of PAW for the flux tower sites are derived from GLEAM[40](see section 3 in Supplementary Information for input data used in GLEAM). VPD is derived from relative humidity and T_a data sourced from Atmospheric Infrared Sounder (AIRS) aboard the Aqua satellite mission[53]. SW_i is derived from the Clouds and the Earth's Radiant Energy System (CERES) satellite mission[54]. VOD is derived from the Vegetation Optical Depth Climate Archive (VODCA) dataset[55]. CO_2 data is sourced from the Copernicus Atmopsh-eric Monitoring Service Global Inversion of Greenhouse Gas Fluxes and Concentrations project (https://ads.atmosphere.copernicus.eu). Finally, within the GLEAM model's soil water balance model, Equation 5 is solved for short and tall vegetation separately for every grid cell and aggregated based on the fraction of tall and short vegetation in every grid cell. For tall (or short) vegetation flux tower sites, PAW weighted by the corresponding tall (or short) vegetation fraction is extracted. In GLEAM, for tall vegetation, w_w calculated based on three soil layers, and for short vegetation w_w is based on two soil layers.

Deep learning model architecture and training. Designing an optimal deep learning model involves optimizing a number of model-related variables (hyper-parameters) such as the number of layers, number of neurons in each layer, the activation functions in each layer, the rate of dropout to prevent over-fitting, the optimal learning rate, and a loss or objective function along with an appropriate validation metric for evaluating the progress of model training. Here, we design the model architecture, optimize the hyper-parameters, and train the deep learning model using TensorFlow version 2.4[56]. To optimize the hyper-parameters, we employ an automated optimization library available in TensorFlow; specifically, a Bayesian optimization procedure with maximization of the Kling Gupta Efficiency (KGE)[57] as both the training objective and validation metric. In training the objective fucntion is implemented as minimization of $1 - KGE$. KGE is selected as it combines correlation, variability bias, and mean bias into a single metric. KGE is defined as

$$KGE = 1 = \sqrt{(r-1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2 + (\frac{\mu_{sim}}{\mu_{obs}} - 1)^2} \tag{6}$$

where r is linear correlation between simulated and observed values, σ_{sim} and σ_{obs} are standard deviation of simulations and observations, and μ_{sim} and μ_{obs} are mean values of simulations and observations.

First, the Bayesian hyper-parameter optimization was carried out for short vegetation data (231 sites). The most optimal deep learning architecture was found after approximately 1000 iterations of the Bayesian optimization procedure. The resulting deep learning architecture was manually tuned. The final model was then trained for short vegetation S_t with a training:validation data split of 85:15, a batch size of 100, a learning rate of 0.000142, and a maximum epoch size of 1000. The training was automatically stopped when the validation objective function start degrading (while the training objective function keeps improving), a sign that the model is overfitting (Figure 9 in Supplementary Information shows the evolution of the objective during the training process). The same model architecture and training setup was used for training the model for tall vegetation S_t (137 sites). As the model performed satisfactorily with some minor changes, the time consuming hyper-parameter optimization procedure was not performed separately for the tall vegetation dataset (see Figure 8 in Supplementary Information for the final deep learning models).

Calculation of SIF/PAR. SIF data is sourced from the contiguous Orbiting Carbon Observatory-2 (OCO-2) dataset, which is available at 0.05^o resolution and 16-day time step[58]. This dataset uses machine learning to gap-fill SIF data to produce a spatially continuous data from the OCO-2 satellite, which has a smaller footprint and infrequent overpass times. The data was spatially aggregated to 0.25^o and temporally aggregated to monthly timescales for calculating the correlation maps (Figure 6) and to seasonal time scales for Figure 4. PAR data is from the CERES mission[54]. PAR data is available at 1.0^o resolution at hourly to monthly resolution. Here, the monthly PAR data was used to normalize SIF data.

Data availability

The outputs of the hybrid model are available at https://doi.org/10.5281/zenodo.5220753.

Code availability

The deep learning-based stress formulations for tall and short vegetation and all the codes required for reproducing the results in this study are available at https://doi.org/10.5281/zenodo.5220753.

References

1. Good, S. P., Noone, D. & Bowen, G. Hydrologic connectivity constrains partitioning of global terrestrial water fluxes **349**, 175–177 (2015). URL `https://science.sciencemag.org/content/349/6244/175`.

2. Masson-Delmotte, V. *et al.* Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel of climate change. *Global warming of 1.5 C. An IPCC Special Report* (2021).

3. Milly, P. C. D., Dunne, K. A. & Vecchia, A. V. Global pattern of trends in streamflow and water availability in a changing climate. *Nature* **438**, 347–350 (2005). URL `https://doi.org/10.1038/nature04312`.

4. Konapala, G., Mishra, A. K., Wada, Y. & Mann, M. E. Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. *Nature Communications* **11**, 3044 (2020). URL `https://doi.org/10.1038/s41467-020-16757-w`.

5. Miralles, D. G., Gentine, P., Seneviratne, S. I. & Teuling, A. J. Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. *Annals of the New York Academy of Sciences* **1436**, 19–35. URL `https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1111/nyas.13912`.

6. Schwalm, C. R. *et al.* Global patterns of drought recovery. *Nature* **548**, 202–205 (2017). URL `https://doi.org/10.1038/nature23021`.

7. Sippel, S. *et al.* Drought, heat, and the carbon cycle: a review. *Current Climate Change Reports* **4**, 266–286 (2018). URL `https://doi.org/10.1007/s40641-018-0103-4`.

8. Peterson, T. J., Saft, M., Peel, M. C. & John, A. Watersheds may not recover from drought **372**, 745–749 (2021). URL `https://science.sciencemag.org/content/372/6543/745`.

9. Vicente-Serrano, S. M., Beguería, S. & López-Moreno, J. I. A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate* **23**, 1696 – 1718 (2010). URL `https://journals.ametsoc.org/view/journals/clim/23/7/2009jcli2909.1.xml`.

10. Anderson, M. C. *et al.* The evaporative stress index as an indicator of agricultural drought in brazil: An assessment based on crop yield impacts. *Remote Sensing of Environment* **174**, 82–99 (2016). URL https://www.sciencedirect.com/science/article/pii/S0034425715302212.

11. Kalma, J. D., McVicar, T. R. & McCabe, M. F. Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surveys in Geophysics* **29**, 421–469 (2008). URL https://doi.org/10.1007/s10712-008-9037-z.

12. Lawrence, D. M. *et al.* The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems* **11**, 4245–4287. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001583.

13. Niu, G.-Y. *et al.* The community noah land surface model with multiparameterization options (noah-mp): 1. model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres* **116**. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD015139.

14. Miralles, D. G. *et al.* Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences* **15**, 453–469 (2011). URL https://hess.copernicus.org/articles/15/453/2011/.

15. Fisher, J. B., Tu, K. P. & Baldocchi, D. D. Global estimates of the land–atmosphere water flux based on monthly avhrr and islscp-ii data, validated at 16 fluxnet sites. *Remote Sensing of Environment* **112**, 901–919 (2008). URL https://www.sciencedirect.com/science/article/pii/S0034425707003938.

16. Mu, Q., Zhao, M. & Running, S. W. Improvements to a modis global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment* **115**, 1781–1800 (2011). URL https://www.sciencedirect.com/science/article/pii/S0034425711000691.

17. Mueller, B. & Seneviratne, S. I. Systematic land climate and evapotranspiration biases in cmip5 simulations. *Geophysical Research Letters* **41**, 128–134. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013GL058055.

18. Koppa, A., Alam, S., Miralles, D. G. & Gebremichael, M. Budyko-based long-term water and energy balance closure in global watersheds from earth observations. *Water Resources Research* **57**, e2020WR028658. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020WR028658.

19. Fisher, J. B. *et al.* The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. *Water Resources Research* **53**, 2618–

2626. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016WR020175.

20. Penman, H. L. & Keen, B. A. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **193**, 120–145 (1948). URL https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1948.0037.

21. Priestley, C. H. B. & Taylor, R. J. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review* **100**, 81 – 92 (1972). URL https://journals.ametsoc.org/view/journals/mwre/100/2/1520-0493_1972_100_0081_otaosh_2_3_co_2.xml.

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

23. Zhao, W. L. *et al.* Physics-constrained machine learning of evapotranspiration. *Geophysical Research Letters* **46**, 14496–14507. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL085291.

24. Miralles, D. G. *et al.* The wacmos-et project – part 2: Evaluation of global terrestrial evaporation data sets. *Hydrology and Earth System Sciences* **20**, 823–842 (2016). URL https://hess.copernicus.org/articles/20/823/2016/.

25. Verhoef, A. & Egea, G. Modeling plant transpiration under limited soil water: Comparison of different plant and soil hydraulic parameterizations and preliminary implications for their use in land surface models. *Agricultural and Forest Meteorology* **191**, 22–32 (2014). URL https://www.sciencedirect.com/science/article/pii/S0168192314000483.

26. Powell, T. L. *et al.* Confronting model predictions of carbon fluxes with measurements of amazon forests subjected to experimental drought. *New Phytologist* **200**, 350–365. URL https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.12390.

27. Wu, X. *et al.* Parameterization of the water stress reduction function based on soil–plant water relations. *Irrigation Science* **39**, 101–122 (2021). URL https://doi.org/10.1007/s00271-020-00689-w.

28. Zhang, J., Liu, P., Zhang, F. & Song, Q. Cloudnet: Ground-based cloud classification with deep convolutional neural network. *Geophysical Research Letters* **45**, 8665–8672. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL077787.

29. Hengl, T. *et al.* Soilgrids250m: Global gridded soil information based on machine learning. *PLOS ONE* **12**, 1–40 (2017). URL https://doi.org/10.1371/journal.pone.0169748.

30. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change **342**, 850–853 (2013). URL https://science.sciencemag.org/content/342/6160/850.

31. Jung, M. *et al.* The fluxcom ensemble of global land-atmosphere energy fluxes. *Scientific Data* **6**, 74 (2019). URL https://doi.org/10.1038/s41597-019-0076-8.

32. McGovern, A. *et al.* Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bulletin of the American Meteorological Society* **98**, 2073 – 2090 (2017). URL https://journals.ametsoc.org/view/journals/bams/98/10/bams-d-16-0123.1.xml.

33. Kratzert, F. *et al.* Toward improved predictions in ungauged basins: Exploiting the power of machine learning. *Water Resources Research* **55**, 11344–11354. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019WR026065.

34. Arrieta, A. B. *et al.* Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *CoRR* **abs/1910.10045** (2019). URL http://arxiv.org/abs/1910.10045. 1910.10045.

35. Rasp, S., Pritchard, M. S. & Gentine, P. Deep learning to represent subgrid processes in climate models **115**, 9684–9689 (2018). URL https://www.pnas.org/content/115/39/9684.

36. Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G. & Yacalis, G. Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters* **45**, 5742–5751. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078202.

37. Reichstein, M. *et al.* Deep learning and process understanding for data-driven earth system science. *Nature* **566**, 195–204 (2019). URL https://doi.org/10.1038/s41586-019-0912-1.

38. de Bézenac, E., Pajot, A. & Gallinari, P. Deep learning for physical processes: incorporating prior scientific knowledge. *Journal of Statistical Mechanics: Theory and Experiment* **2019**, 124009 (2019). URL https://doi.org/10.1088/1742-5468/ab3195.

39. Kraft, B., Jung, M., Körner, M. & Reichstein, M. Hybrid modeling: Fusion of a deep learning approach and a physics-based model for global hydrological modeling. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLIII-B2-2020**, 1537–1544 (2020). URL https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLIII-B2-2020/1537/2020/.

40. Martens, B. *et al.* Gleam v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development* **10**, 1903–1925 (2017). URL https://gmd.copernicus.org/articles/10/1903/2017/.

41. Gash, J. H. C. An analytical model of rainfall interception by forests. *Quarterly Journal of the Royal Meteorological Society* **105**, 43–55. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49710544304.

42. Grossiord, C. *et al.* Plant responses to rising vapor pressure deficit. *New Phytologist* **226**, 1550–1566. URL https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.16485.

43. Urban, J., Ingwers, M., McGuire, M. A. & Teskey, R. O. Stomatal conductance increases with rising temperature. *Plant Signaling & Behavior* **12**, e1356534 (2017). URL https://doi.org/10.1080/15592324.2017.1356534. PMID: 28786730, https://doi.org/10.1080/15592324.2017.1356534.

44. Matthews, J. S. A., Vialet-Chabrand, S. & Lawson, T. Role of blue and red light in stomatal dynamic behaviour. *Journal of Experimental Botany* **71**, 2253–2269 (2019). URL https://doi.org/10.1093/jxb/erz563.

45. Xu, Z., Jiang, Y., Jia, B. & Zhou, G. Elevated-co2 response of stomata and its dependence on environmental factors. *Frontiers in Plant Science* **7**, 657 (2016). URL https://www.frontiersin.org/article/10.3389/fpls.2016.00657.

46. Peng, Y., Bloomfield, K. J., Cernusak, L. A., Domingues, T. F. & Colin Prentice, I. Global climate and nutrient controls of photosynthetic capacity. *Communications Biology* **4**, 462 (2021). URL https://doi.org/10.1038/s42003-021-01985-7.

47. Knoben, W. J. M., Freer, J. E. & Woods, R. A. Technical note: Inherent benchmark or not? comparing nash–sutcliffe and kling–gupta efficiency

scores. *Hydrology and Earth System Sciences* **23**, 4323–4331 (2019). URL https://hess.copernicus.org/articles/23/4323/2019/.

48. Pagán, B. R., Maes, W. H., Gentine, P., Martens, B. & Miralles, D. G. Exploring the potential of satellite solar-induced fluorescence to constrain global transpiration estimates. *Remote Sensing* **11** (2019). URL https://www.mdpi.com/2072-4292/11/4/413.

49. Bauer, P. *et al.* The digital revolution of earth-system science. *Nature Computational Science* **1**, 104–113 (2021). URL https://doi.org/10.1038/s43588-021-00023-0.

50. Pastorello, G. *et al.* The fluxnet2015 dataset and the oneflux processing pipeline for eddy covariance data. *Scientific Data* **7**, 225 (2020). URL https://doi.org/10.1038/s41597-020-0534-3.

51. Wei, Z. *et al.* Revisiting the contribution of transpiration to global terrestrial evapotranspiration. *Geophysical Research Letters* **44**, 2792–2801 (2017). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL072235.

52. Granier, A. & Loustau, D. Measuring and modelling the transpiration of a maritime pine canopy from sap-flow data. *Agricultural and Forest Meteorology* **71**, 61–81 (1994). URL https://www.sciencedirect.com/science/article/pii/0168192394901007.

53. Aumann, H. *et al.* Airs/amsu/hsb on the aqua mission: design, science objectives, data products, and processing systems. *IEEE Transactions on Geoscience and Remote Sensing* **41**, 253–264 (2003).

54. Wielicki, B. A. *et al.* Clouds and the earth's radiant energy system (ceres): An earth observing system experiment. *Bulletin of the American Meteorological Society* **77**, 853–868 (1996).

55. Moesinger, L. *et al.* The global long-term microwave vegetation optical depth climate archive (vodca). *Earth System Science Data* **12**, 177–196 (2020). URL `https://essd.copernicus.org/articles/12/177/2020/`.

56. Abadi, M. *et al.* TensorFlow: Large-scale machine learning on heterogeneous systems (2015). URL `https://www.tensorflow.org/`. Software available from tensorflow.org.

57. Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology* **377**, 80–91 (2009). URL `https://www.sciencedirect.com/science/article/pii/S0022169409004843`.

58. Yu, L., Wen, J., Chang, C. Y., Frankenberg, C. & Sun, Y. High-resolution global contiguous sif of oco-2. *Geophysical Research Letters* **46**, 1449–1458. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL081109`.

Acknowledgements

The authors acknowledge support from the European Union Horizon 2020 Programme (DOWN2EARTH, 869550) and the European Research Council (DRY-2-DRY, 715254). This research was possible thanks to the various flux tower networks for making the data publicly available . The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation, Flanders (FWO), and the Flemish Government.

Author Contributions

AK and DM conceived and designed the study. AK developed the deep learning-based stress formulation, implemented it in GLEAM, and conducted all the model runs. DR led the Python implementation of the process-based model (GLEAM). PH contributed to the collection of flux tower data. AK and DM analyzed the results and wrote the manuscript with inputs from DR and PH.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.pdf