Water Resources Research

RESEARCH ARTICLE
10.1029/2022WR033447

Which Potential Evapotranspiration Formula to Use in Hydrological Modeling World-Wide?

R. Pimentel[1,2,3], B. Arheimer[3], L. Crochemore[3,4], J. C. M. Andersson[3], I. G. Pechlivanidis[3], and D. Gustafsson[3]

[1]Fluvial Dynamics and Hydrology, Andalusian Institute for Earth System Research, University of Cordoba, Córdoba, Spain, [2]Department of Agronomy, Unit of Excellence María de Maeztu (DAUCO), University of Córdoba, Córdoba, Spain, [3]Hydrology Research Unit, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden, [4]IRD, CNRS, INRAE, Grenoble INP, IGE, Université Grenoble Alpes, Grenoble, France

Key Points:

- We propose to apply different formulas for potential evapotranspiration in hydrological modeling globally based on climate regions
- Higher complexity in potential evapotranspiration formula only improves predictions in temperate and polar regions
- Choosing potential evapotranspiration formula strongly affects catchment modeling accuracy of evapotranspiration and streamflow world-wide

Correspondence to:
R. Pimentel,
rpimentel@uco.es

Citation:
Pimentel, R., Arheimer, B., Crochemore, L., Andersson, J. C. M., Pechlivanidis, I. G., & Gustafsson, D. (2023). Which potential evapotranspiration formula to use in hydrological modeling world-wide? *Water Resources Research*, *59*, e2022WR033447. https://doi.org/10.1029/2022WR033447

Received 11 AUG 2022
Accepted 20 APR 2023

Author Contributions:
Conceptualization: R. Pimentel, B. Arheimer
Data curation: R. Pimentel, L. Crochemore
Formal analysis: R. Pimentel
Investigation: R. Pimentel, B. Arheimer, L. Crochemore, J. C. M. Andersson, I. G. Pechlivanidis, D. Gustafsson
Methodology: R. Pimentel, B. Arheimer, L. Crochemore
Resources: B. Arheimer
Software: R. Pimentel, L. Crochemore, J. C. M. Andersson
Supervision: B. Arheimer
Validation: R. Pimentel
Visualization: R. Pimentel, B. Arheimer, L. Crochemore

Abstract Although many potential evapotranspiration (PET) formulas are available, there is still a lack of knowledge on when and where to use them in catchment modeling world-wide. Here we experimented with three different formulas in a global hydrological model (the World-wide HYPE), using 15 years of observations from 5,338 streamflow gauges and global evapotranspiration from Earth-observations (MOD16). We tested model performance in a multi-process approach to select the best formula for catchments covering the global landmass. From comparing the results with land-cover, climate classification, water-energy limitations, we found that climate is the main driver behind the spatial patterns in model performance. Hargreaves was the best PET formula in 50% of the catchments, most of them located in the Amazonas, central Europe, and Oceania; Jensen-Haise was better for catchments in northern latitudes (36%). Finally, Priestley-Taylor was the best formula for India and latitudes above 65° North. The selection of a PET formula seems to be more critical in tropical regions close to the equator, where the differences in performance are above 50%. This is also where PET is highest. We found a strong connection between the five main Köppen regions and the PET formulas, further supported by landcover analysis. Hence, the PET formulas differed in their capacity to provide useful input to the water balance modeling, with complex formulas only giving improved predictions in temperate and polar regions; however, for the rest of the globe simpler formulas were better. We thus recommend to apply different PET formulas based on climatic regions world-wide.

1. Introduction

The role of evapotranspiration (ET) for meteorology, agriculture, and hydrology has been identified as key during centuries (Miralles et al., 2016). Our current understanding of this process arises two hundred years ago (e.g., Dalton, 1802; Horton, 1919; Penman, 1948; Tanner et al., 1967). However, special emphasis due to the role of ET in mass and energy exchanges between land and atmosphere has been done in last decades, following the recognition of evaporation and transpiration in climate impact assessments (e.g., Dolman & De Jeu, 2010; Lemaitre-Basset et al., 2022; Wang & Dickinson, 2012) and hydrological process understanding (Savenije, 2004).

Potential ET (PET) is a concept for predicting ET, defined as the evaporation in case of an unlimited amount of water available. Many different formulas have been proposed to estimate PET (e.g., Baier & Robertson, 1965; Hargreaves & Samani, 1985; Monteith, 1965; Oudin et al., 2005; Penman, 1948; Priestley & Taylor, 1972; Thornthwaite, 1948), which further condition the actual ET (AET) in hydrological modeling. These equations commonly represent, implicitly or explicitly, the available energy for the phase change of liquid water to vapor, and the efficiency in the moisture transport away from the evaporating surface. Subsequently, these methods can be grouped according to how the available energy and transport efficiency is represented and depending on the meteorological inputs required, including temperature, radiation, humidity, and radiation-aerodynamics (Dingman, 2015). Catchment modeling normally uses lumped representation of physical processes with as few input variables and calibration parameters as possible, such as simplified PET formulas based on temperature and radiation, to avoid too many calibrated parameters and thereby risk of over-calibration/equifinality. However, the available PET formulas can result in different PET estimates due to their different assumptions and inputs requirements (Allen & Breshears, 1998).

The impact of PET formula selection on streamflow estimation is yet not fully understood (e.g., Lemaitre-Basset et al., 2022). While some studies show that modeled streamflow appears to be insensitive to different PET inputs

Writing – original draft: R. Pimentel, B. Arheimer
Writing – review & editing: L. Crochemore, J. C. M. Andersson, I. G. Pechlivanidis, D. Gustafsson

(Andersson, 1992; Andréassian et al., 2004; Oudin et al., 2005), other studies have indicated a dependency between the two (Ellenburg et al., 2018; Jayathilake & Smith, 2020; Samain & Pauwles, 2013; Seong et al., 2018; Yang et al., 2021). Most of these studies were carried out at catchment or national scale or assessed other aspects of choosing a PET formula, such as ET estimates (Miralles et al., 2016), implication on extreme events (Valipour et al., 2017; Zhou et al., 2020) or uncertainty in climate change projections (Lemaitre-Basset et al., 2022). However, to our knowledge, no study evaluated the impact on hydrological predictions across the globe form the choice of PET. Therefore, we still lack guidance on which formula(s) better describes PET at a specific site. Although generalization cannot be made since previous studies lack large sampling, the Priestley-Taylor equation has been presented as the most suitable for global hydrological applications (Weiß & Menzel, 2008). However, other studies indicate that the selection of the PET formula should be subject to the local climatic conditions (Bai et al., 2016). Following this last recommendation, the hydrological model World-wide HYPE (WWH; Arheimer et al., 2020) used different PET formulas in different climate regions, linked to the five major Köppen-Geiger classes (Kottek et al., 2006). The choice of PET formula in a given region was then based on previous experiences from modeling continental hydrology (e.g., India, by Pechlivanidis & Arheimer, 2015; Europe, by Donnelly et al., 2016; West Africa, by Andersson, Arheimer, et al., 2017), but without further quantitative evaluation.

New opportunities for global assessments are accessible today with data from Earth Observation (EO) products. Different global PET and AET products are available (Miralles et al., 2016; Mu et al., 2011) offering larger spatial coverage than what is monitored by streamflow gauges. The monthly MOD16 product (Mu et al., 2011) is one the most used ET product for model evaluation (e.g., Immerzeel & Droogers, 2008; Jiang et al., 2020; Lopez-Lopez et al., 2017; Musuuza et al., 2020; Nijzink et al., 2018; Odusanya et al., 2019). MOD16 uses the Penman-Monteith formula (Monteith, 1965), combining meteorological conditions from local observations and re-analyses with satellite-based surface characteristics such as leaf area index and surface resistance. Penman-Monteith formula explicitly considers the impact of windspeed and land surface characteristics on the moisture transport efficiency and vegetation control, as well as a more physically based representation of the surface energy balance for estimating the available energy for ET.

In the current study, we explain the connections between a PET-formula selection and performance in catchment modeling across the world. To do so, we identify global spatial patterns in the ability of a hydrological PET model to simulate PET, AET and streamflow using different PET formulas, compare similarities and dissimilarities in the PET estimates and, finally recommend the PET formula to be applied based on different catchment and climate characteristics world-wide. We explore the following scientific questions: (a) Which PET formula is best suited to describe either PET, AET or streamflow world-wide? (b) Where is each PET formula best for all three variables (i.e., PET, AET and streamflow)? (c) How much can the catchment model performance be improved depending on the choice of the PET formula? (d) Over which physiographic characteristics does each PET formula perform best? and (e) Can the PET formula performance in catchment modeling be linked to energy or water-limited limitations? To answer these questions, we experimented with three simplified PET models over the globe: Jensen-Haise (Jensen & Haise, 1963), Hargreaves (Hargreaves & Samani, 1985) and Priestley-Taylor (Priestley & Taylor, 1972). The same global hydrological model setup (Arheimer et al., 2020) was applied to assess the accuracy of streamflow and ET simulations against observed streamflow and the MOD16 product, respectively. Catchment physiography and a classification in the Budyko space (Budyko, 1948) were investigated to explain the results from the comparative analysis. Hence, the global hydrological model is applied as a virtual laboratory where we keep all other hydrological predictors constant except for PET to examine its influence on the model performance. The aim is to identify global spatial patterns in the PET model performance, compare similarities and dissimilarities in results and, finally, recommend which PET formula should be applied for different catchments world-wide.

2. Data

2.1. MOD16 Evapotranspiration Product

We used the monthly MODIS global ET products (MOD16, Mu et al., 2007, 2011) as an independent reference source of ET. MOD16 was chosen because, to our understanding is the most widely used remote-sensing-based ET data set. In addition, it is based on the Penman-Monteith model, which is the most common formula when retrieving ET from remote sensing, and it is independent from the PET formulas assessed in this study (see Section 3). This source was used to assess the performance of the catchment model when tuned against the

different PET formulas. MOD16 is one of the most used remote sensing ET products applied to assess modeled ET, due to its high spatial resolution (1×1 km^2), length of historical availability (since 2000) and its global coverage, except for ice-driven catchments, desert areas and water bodies where the ET cannot be derived with the MOD16 algorithm (Mu et al., 2011). This product computes PET and AET using the Penman-Monteith formula (Monteith, 1965; Penman, 1948), which is based on water and energy balances. This PET formula requires net radiation, temperature, humidity, and wind speed as inputs. For MOD16, the inputs are taken from the reanalysis products Global Modeling and Assimilation Office (GMAO). In addition, information about the canopy aerodynamics and surface resistance from the vegetation are required. In this case MOD16 utilizes the band 10 from MOD43C1 (Jin et al., 2003; Salomon et al., 2006; Schaaf et al., 2002); the global land cover MOD12Q1 (Friedl et al., 2002); and the MOD15A2 (Myneni et al., 2002). MOD16 estimates the individual evaporation and transpiration components for soil and vegetation. Then, the final PET provided is the sum of the evaporation from intercepted rain by the canopy before it reaches the soil, the potential plant transpiration, the evaporation from wet soil and the potential ET from the soil components. AET is retrieved from remote sensing data as the sum of evaporation losses from the wet canopy, actual plant transpiration and actual soil evaporation (Mu et al., 2007, 2011).

MOD16 has been evaluated against ground measurements derived from eddy flux towers in several parts of the world with different results in terms of accuracy (Hu et al., 2015). For instance, Mu et al. (2011) showed a mean bias in daily ET of 0.31 mm/day and a mean relative error (RE) of 24% in 46 sites in North America. In South America, Ruhoff et al. (2013) found that the RE in ET was about 19% in two sites over the Rio Grande (Brazil). In Europe, Hu et al. (2015) obtained correlation coefficients ranging within the interval (0.45–0.98) and a mean bias within the interval (−0.9–1.11 mm/day). Trambauer et al. (2014) showed inconsistencies in ET when comparing to MOD16, with extremely changeable RE among the evaluated years (−55%–33%) in two sites in South Africa. In Asia, Kim et al. (2012) found ET correlations ranging from 0.27 to 0.82 and biases ranging from −24 to 2.3 mm/8 days, over 17 sites. In all these comparisons, it is important to highlight that ground measurements have their own uncertainties, with an accuracy in the range (10%–30%) when ET is retrieved from eddy flux towers.

Since the global coverage of eddy flux towers is not homogenous, MOD16 has also been assessed by comparison with other distributed ET estimates. For instance, Velpuri et al. (2013) showed that MOD16 underestimates ET over poorly vegetated surfaces in USA when comparing to GFET (Jung et al., 2009, 2010). For Africa, Trambauer et al. (2014) affirmed that MOD16 underestimated and was poorly correlated with other products in semi-arid, Sahel and Southern Africa in comparison to GLEAM and ERA-Interim. Hu et al. (2015) affirmed that MOD16 was consistent over most of Europe when compared with LSA-SAR MSG ETa. At the global scale Miralles et al. (2016) showed that MOD16 has good correlation (0.89 and 0.93) with ET estimates from GLEAM (Miralles et al., 2011) and PT-JPL ETa (Fisher et al., 2005, 2008), resulting in a root mean squared error of 180 mm/year for both cases. The differences found were mainly caused by a general tendency in MOD16 to underestimate the flux in the tropics and subtropics.

2.2. Streamflow Data

In total 5,338 streamflow gauges providing openly available streamflow time-series were used to identify the optimum parameter sets (i.e., calibration) for each of the PET formulas in step 3 of the world-wide calibration procedure for global catchment modeling (Arheimer et al., 2020). These streamflow gauges were selected among the 11,369 gauges downloaded and quality controlled during the hydrological model setup (Crochemore et al., 2020). Among the selected ones, we chose 318 streamflow gauges for model re-calibration. This calibration set was the same as in the original stepwise model calibration by Arheimer et al. (2020). The geographical coverage and spread of the selected streamflow gauges guarantee a good representation of model performance variability world-wide (Figure 1).

3. Methods

Three simplified PET formulas commonly used in catchment hydrology, namely (a) Jensen-Haise (Jensen & Haise, 1963), (b) Hargreaves (Hargreaves & Samani, 1985), and (c) Priestley Taylor (Priestley & Taylor, 1972) were explored. Here we use the same hydrological model, WWH, keeping all other hydrological features constant except for PET. This was possible thanks to the model setup with a stepwise calibration procedure, in which parameters for specific processes are tuned using only specific representative gauges (Arheimer et al., 2020).

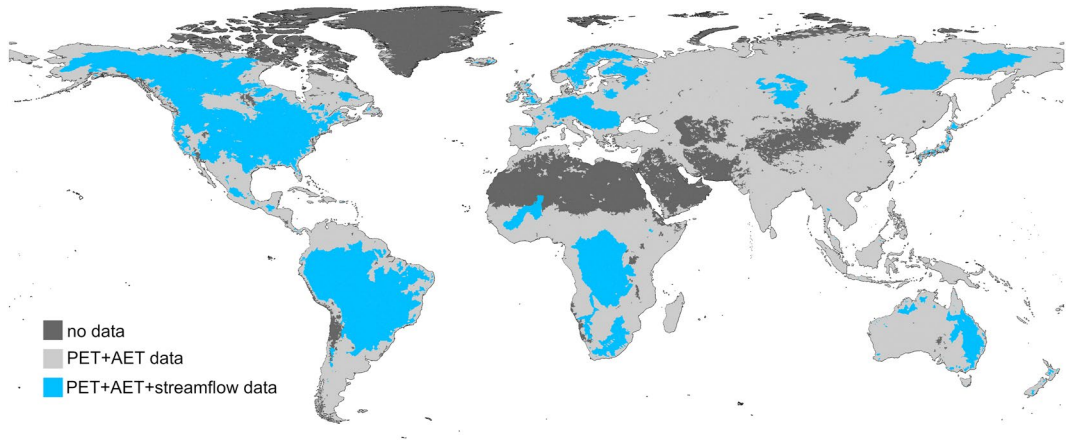

Figure 1. Location of the analyzed catchments: dark gray represents areas without information, light gray represents catchments with information from MODIS (potential evapotranspiration (PET) and actual evapotranspiration (AET)), and light blue represents areas with information from PET, AET and streamflow (upstream areas of the selected streamflow gauge).

Changing the PET formula, however, requires new parameters as the original WWH set-up used the three PET formulas in a predefined distribution world-wide (Arheimer et al., 2020) while in this study we want to apply them all globally. The same 318 catchments used in step 3 of the stepwise calibration of WWH (Arheimer et al., 2020) were used and they represent gauged basin of the different Hydrological Response Units (HRUs) in WWH and are independent of posterior streamflow evaluation. The model performance was assessed individually against PET and AET from the MOD16 products and in-situ streamflow. A multi-process analysis based on these performances was carried out to identify regions where the same formula provided the best result for all three variables. Catchments were then assigned the PET formula resulting in the best (multi-process) model performance. The three groups of catchments assigned to each PET formula were further analyzed with respect to (dis)similarities in: (a) physiographic characteristics in the WWH model setup, and (b) the Budyko space (Budyko, 1948) to separate energy—from water—limited catchments. The above investigation not only aims to improve our understanding of the drivers that lead PET formulas to perform differently depending on the regions (e.g., Jayathilake & Smith, 2020) but also to identify the objectively "best" PET formula to be applied in different catchments world-wide.

3.1. The PET Formulas

The three selected PET formulas are relatively simple algorithms in the sense that they rely only on radiation and daily temperature. The Jensen-Haise formula (Jensen & Haise, 1963) was developed as a simple technique for calculating ET estimates in arid and semiarid irrigated areas in the USA. It requires extraterrestrial radiation, depending only on latitude and the day of the year, and mean daily surface air temperature. We have chosen the modification proposed by Oudin et al. (2005) in which two extra temperature constants were added to scale mean temperature in the original equation making it more tunable (e.g., adapting the minimum value of air temperature for which PET is not zero) when applied to catchments with very different characteristics. Then, PET (mm day^{-1}) is computed as:

$$\begin{aligned} \mathrm{PET} &= kc \cdot \frac{R_e}{\lambda\rho}\frac{(T_a + T_{\mathrm{add}})}{T_{\mathrm{scale}}} \text{ if } T_a + T_{\mathrm{add}} > 0 \\ \mathrm{PET} &= 0 \text{ otherwise} \end{aligned} \tag{1}$$

where R_e (MJ m^{-2} day^{-1}) and T_a (C) are extraterrestrial radiation and mean daily air temperature, respectively; λ (MJ kg^{-1})is the latent heat of vapourization, ρ(kg m^{-3}) is the water density, T_{add} (C) and T_{scale} (C) are the two temperature parameters to scale the mean daily temperature; and k_c is an additional parameter added in the implementation in the HYPE model to provide the ability to adjust the PET for different land-cover types (similar to so-called crop-coefficients commonly used with the Penman-Monteith formula, see further in Allen and Breshears (1998)). Note that in the WWH model, T_{add} and T_{scale} are global parameters meaning that they have the

same values for all catchments, while k_c is a land-cover dependent parameter that can be calibrated independently for each hydrological response unit (HRU).

The second formula used is the Hargreaves equation (Allen & Breshears, 1998; Hargreaves & Samani, 1985), which, in its original form, is based on extraterrestrial radiation, daily mean air temperatures, and the range between daily minimum and maximum air temperature as a proxy for the atmospheric turbidity. To allow for alternative turbidity estimates, using solar radiation or cloudiness, the following modified version of the Hargreaves equation is implemented in the WWH model:

$$\mathrm{PET} = kc \cdot 0.0023 \frac{R_e}{\lambda \rho} \frac{\tau}{k_{rs}} (T_a - 17.8) \tag{2}$$

When solar radiation and cloudiness is missing (as in this study), the atmospheric turbidity τ is calculated following the Hargreaves formula, but limited by an upper and lower constraint following Allen & Breshears (1998).

$$\tau = \max\left(0.25, \min\left(0.75, k_{rs}\sqrt{T_{\max} - T_{\min}}\right)\right) \tag{3}$$

and the modified Hargreaves formula essentially is identical to the original. The parameter k_{rs} is a constant to relate the solar radiation to the extraterrestrial radiation multiplied by the square-root of the daily temperature range. According to Allen & Breshears, 1998 k_{rs} varies approximately between 0.16 at inland locations to 0.19 at coastal locations. The constant 0.0023 is the product of $k_{rs} = 0.16$ and Hargreaves radiation-temperature to evaporation coefficient (0.0135); kc as in the previous equation is a land-cover dependent parameter.

The last PET formula considered is the Priestley-Taylor equation (Priestley & Taylor, 1972). As opposed to the previous ones, this formula was not initially developed for irrigation purposes at local scale but to assess surface heat flux and evaporation at the large scale. This is the most complex of the three formulas chosen since it requires net radiation as an input, which involves estimation of net shortwave and longwave radiation at the land surface adding more parameters such as the surface albedo and input variables such as the vapor pressure. Here, PET (mm day^{-1}) is computed as:

$$\mathrm{PET} = kc \cdot \frac{\alpha_{pt} \Delta R_n}{\lambda \rho (\Delta \gamma)} \tag{4}$$

where α_{pt} is a constant with an empirical value of 1.26 for open water (Priestley & Taylor, 1972) representing the relation between latent and sensible heat fluxes; Δ (kPa C^{-1}) is the slope of the vapor pressure curve derived from the mean daily temperature; R_n (MJ m^{-2} day^{-}1) is the net radiation derived as the net downward shortwave minus the net upwards longwave radiation; and γ(kPa C^{-1})is the psychrometric constant. Here, kc as in the previous equations is a land-cover dependent parameter. Net downward shortwave radiation is calculated as the downward solar radiation estimated using the same extraterrestrial radiation and atmospheric turbidity as described for the Hargreaves equation and multiplied by one minus a land-use dependent albedo parameter. The net outgoing radiation is calculated using Stephan-Boltzmann law as a function of the daily minimum and maximum air temperatures and compensated for the downward atmospheric emissions as a function of the actual vapor pressure and relative solar radiation as described in Allen & Breshears (1998).

3.2. The World-Wide HYPE Model Laboratory

WWH is the application of the HYPE model (Lindström et al., 2010) at the global scale (Arheimer et al., 2020). HYPE is an integrated catchment model characterizing the main water fluxes and pathways at the catchments scale. HYPE has specific routines for water storages, such as snow, glacier, soil moisture, groundwater, lakes, wetlands and floodplains. Water fluxes are calculated, for example, regulation, abstractions, irrigation, streamflow, and ET. Streamflow from individual catchments is routed and aggregated along the river network, where lakes and reservoirs may dampen the flow according to rating curves.

The HYPE model is currently applied in national forecast services of floods and droughts (e.g., Pechlivanidis et al., 2014; Pechlivanidis et al., 2020), to retrieve infrastructure-design values (Andersson, Ali, et al., 2017) and assessments of water quality (e.g., Arheimer et al., 2015; Bartosova et al., 2019), as well as impacts from hydro-morphological alterations (e.g., Arheimer & Lindström, 2015) and climate change (e.g., Arheimer &

Lindström, 2015; Arheimer et al., 2017). The HYPE model was initiated in 2003, applied for entire Sweden in 2008 (Strömqvist et al., 2012) and then for different global regions (Andersson, Arheimer, et al., 2017; Donnelly et al., 2016; Krysanova et al., 2017; Pechlivanidis & Arheimer, 2015) and finally for the entire landmass of the globe except Antarctica (Arheimer et al., 2020).

For the WWH v1.3 configuration (Arheimer et al., 2020), a total of 131,296 catchments were delineated with an average size of about 1,020 km^2. The model is forced by daily precipitation and temperature, which initialize the water balance in each HRU, HRUs being the basic calculation units within the catchment, defined as a combination of land-cover, elevation, and climate. WWH was calibrated against some 5,000 streamflow gauges and evaluated for some 4,000 independent ones (Arheimer et al., 2020). The average monthly Kling-Gupta Efficiency for streamflow in the period 1981–2014 was 0.4 globally (maximum was 0.93, 75th percentile was 0.64, and 25th percentile was 0.03).

For the current study, the WWH hydrological model was used as virtual laboratory for experimenting with the different PET formulas. When testing one formula at a time the rest of the model was kept constant, hence we only changed the parameters conditioning evaporation processes.

Parameters in each PET formula were calibrated individually world-wide in the objective to optimize streamflow in representative gauged basins of the different HRUs, according to the procedure by Arheimer et al. (2020). Automatic calibration was applied for each subset of parameters using the differential evolution Markov chain (DEMC) approach (Ter Braak, 2006) to obtain the optimum parameter set for each PET formula. The identified parameters were then applied wherever relevant in the whole geographical domain.

3.3. Multi-Process Analysis

To account for the equifinality problem and low parameter sensitivity that has been noticed in hydrological modeling (Andersson, 1992; Andréassian et al., 2004; Oudin et al., 2005; Santos et al., 2022), the PET formulas were also evaluated against independent data and EO products. Simulated PET, AET and streamflow from using each one of the three formulas were evaluated, respectively, against: (a) PET from MOD16, (b) AET from MOD16, and (c) independent observed streamflow not used in the calibration procedure. Here, gridded MOD16 PET and AET products (1 × 1 km) were aggregated to WWH sub-basin scale (~1,000 km^2) for the period 2000–2015. This information was aggregated in 115,151 of the 131,296 sub-basins delineated in WWH, since MOD16 was unavailable in some location, for instance in deserts or permanent ice areas (see Section 2.1).

Since we are assessing PET formulas and consequently water volumes in the water balance, we choose the Relative Error (RE) as the performance metric. RE was defined as:

$$\mathrm{RE} = \frac{\sum_{i=1}^{m}(\mathrm{sim}_i - \mathrm{obs}_i)}{\sum_{i=1}^{m}\mathrm{obs}_i} \cdot 100 \tag{5}$$

where i represents each time step in the time series, m the total number of time steps in the time series, and sim and obs the simulated and observed time series, respectively.

For each catchment and analyzed variable (PET, AET, and streamflow), the PET formula that minimizes RE was chosen as the best performing one. Therefore, three different spatial distributions of optimum PET formulas, one per variable, were defined. These three distributions were combined in a multi-process analysis, defining the best PET formula in a catchment only when all three variables identified the same PET formula as the optimum one; otherwise, it was classified as "no match." Since only 38% of the global land surface was covered by streamflow gauges, we made an additional classification using only PET and AET, which both have global coverage (except for large deserts and ice sheets).

3.4. Linking Performance to Catchment Physiography

Following the common practice in catchment hydrology of linking processes with physiographic characteristics for hydrological understanding and predictions (e.g., Donnelly et al., 2016; Hundecha et al., 2016; Kuentz

Table 1
Catchment Physiographic Characteristics (C: Climate, LC: Land Cover) Used to Examine Relations to the Performance of the PET Formulas

Variable name	Description	Range
Pmean (mm)	Annual precipitation in mm yr^{-1}	[0–4,876]
Tmean (C)	Mean annual temperature in degrees	[0.08–30.6]
5 Köppen regions (%)	% of the catchment area within the following Köppen regions: A (Tropical), B (Arid), C (Temperate), D (Continental) and E (Polar)	[0–100]
13 land cover variables (%)	% of the catchment area covered by the following land cover types: Water, Urban, Snow and Ice, Bare, Crop, Mosaic, Tree Broadleaf Evergreen and Flooded areas (TreeBrEvFlood), Tree Broadleaf Deciduous and Mixture (TreeBrdecMix), Trees Needleleaf Evergreen (TreeNeEv), Trees Needleleaf Deciduous (TreeNeDec), Shrub, Grass and Sparse	[0–100]

et al., 2017; Pechlivanidis et al., 2020; Rice et al., 2015; Santos et al., 2022; Sawicz et al., 2014), we computed several catchment physiographic characteristics from the WWH model setup. These were used to examine relations between the best performing PET formula and catchment characteristics to diagnostically guide where each PET formula is suitable. In the original WWH model setup (Arheimer et al., 2020), experience from regional HYPE modeling was extrapolated to global coverage using the World Map of the Köppen-Geiger climate classification (Kottek et al., 2006), but without quantitative evaluation. Hence, this classification was also added to examine potential relations between catchment physiographic characteristics and the performance of the PET formulas (Table 1).

3.5. Performance Attribution Using the Budyko Space

The Budyko space (Budyko, 1948) relates precipitation (P), AET and PET, through the comparison of the evaporative index (AET/P) and the aridity index (PET/P). Although simple, it is an effective way for linking energy and water availability in the water balance at the catchment scale and further determine the drivers controlling the water balance. The Budyko curve was defined empirically based on more than 1,200 medium-size catchments as follows:

$$\frac{\mathrm{AET}}{P} = \left[\frac{\mathrm{PET}}{P}\tanh\left(\frac{P}{\mathrm{PET}}\right)\left(1 - e^{-\left(\frac{\mathrm{PET}}{P}\right)}\right)\right]^{1/2} \quad (6)$$

Here, we used the Budyko space to analyze the selected catchments and assess how the water balance controllers vary depending on the PET formula used.

4. Results and Discussion

4.1. Different PET Formulas in Catchment Modellign Versus Indepent Data of Three Variables

Here, we re-calibrated the PET formulas in WWH and compared results with independent data for the three variables (AET, PET and streamflow) in each catchment world-wide. For all three PET formulas, the performance of WWH in the selected representative catchments with respect to KGE somewhat improves after re-calibration (Table 2), which was expected as some regions were not calibrated before. Nevertheless, equifinality where other model parameters compensate for changes in PET (e.g., Andersson, 1992; Andréassian et al., 2004; Jayathilake & Smith, 2020) probably still exist, even though such comparison are constrained by the stepwise calibration. In many catchments, streamflow may be limited by water availability rather than energy availability. Nevertheless, the parameter values obtained here were within the range of previous HYPE setups (e.g., Andersson, Arheimer, et al., 2017; Donnelly et al., 2016; Lebedeva & Gustafsson, 2021; Pechlivanidis & Arheimer, 2015; Stadnyk et al., 2020), which limits (to a certain extent) the risk of achieving a good model performance for the wrong reasons.

When comparing the new global results against the three independent data sources (i.e., MOD16 PET, MOD16 AET, and streamflow data not involved in the calibration process), we achieved three different world-wide distributions of best PET formula performances (Figure 2).

Table 2
Comparison of the Performance in Streamflow of World-Wide HYPE (WWH) in Terms of the Kling-Gupta Efficiency (KGE) in Selected Catchments for Each Potential Evapotranspiration (PET) Formula Using the Calibration Parameter Values by Arheimer et al. (2020) ("Before," Column 3) and the New Re-Calibration Parameter Values ("After," Column 4)

		Median monthly KGE		
PET formula	WWH parameters	Before	After	Parameter value(s)
Jensen-Haise	10 HRUs: kc2, soilcorr, srrcs	0.26	0.31	[0.90–4.89], [1.00–9.39], [0.02–1.00]
Hargreaves	10 HRUs: kc3, soilcorr, srrcs	0.36	0.37	[0.40–1.77], [0.55–1.87], [0.03–1.00]
Priestley-Taylor	10 HRUs: kc4, alb, soilcorr, srrcs	−0.02	0.17	[0.20–1.90], [0.20–0.80], [0.90–6.02], [0.02–1.00]

Note. Parameter names and values are given in the same order of appearance (column 2 and 5).

All PET formulas showed a general underestimation of PET in the WWH model compared to MOD16 PET product (Table 3, column 2). The Hargreaves formula is, in this case, the best compromise for representing PET globally (Figure 2a) with generally low differences against EO in more than half of the analyzed catchments (53.5%) and a median RE of about 8% from this EO product. Geographically, the Hargreaves formula is best performing best, predominantly in the Americas, Central Europe, and Australia. The longitudinal pattern is more equally spread, due to the fewer catchments best represented by this formula, among the other two PET formulas with a predominance of Jensen-Haise in central Africa and Priestley-Taylor in the Tibetan Plateau. There is a clear pattern in latitudinal bands for these two models, with a clear predominance of Jensen-Haise between 45°N and 65°N and Priestley-Taylor in latitudes above 65°N.

The representativeness of each PET formula changes when assessing them against MOD16 AET, with a more equal distribution between models being found (Figure 2b). The percentage of selected catchments for each PET formula is around 30% (Table 3, column 5). Even though Jensen-Haise exhibits the best median RE overall at about 1%, Hargreaves exhibits the best median RE (in terms of capturing the MOD16 AET) at about −2% simulating AET in the catchments where it performs best, that is, the Amazonas, Central Europe, and Australia. RE was larger for the catchments selected as best in the other two models, with a 7% underestimation for Jensen-Haise and 7% overestimation for Priestley-Taylor. Geographically, Jensen-Haise is the most suitable PET formula to derive AET in latitudes between 45°N and 65°N and Priestley-Taylor the most suitable in India and the eastern coast of South America below the Amazons.

Finally, the assessment against independent streamflow gauges shows a clear predominance of Priestley-Taylor as the best formula to use as a compromise world-wide (Figure 3c), as it minimizes the RE in about 56% of the analyzed catchments (Table 3, column 7). Jensen-Haise is a suitable alternative PET formula in central North America, South Africa and some spots in central Europe, central Asia, and Australia. The most accurate representations of streamflow are based on the Hargreaves formula in central South America and eastern Siberia.

4.2. Optimal PET Formula for All Three Variables in Each Catchment

The overall best PET formula for each catchment was identified based on a multi-process analysis, which selects a PET formula for a catchment if, and only if, it is optimal for all three analyzed variables (Figure 3a). Only 10% of the total area covered by the independent streamflow gauges (3% of the global land area) fulfills this criterion. The Hargreaves formula was selected in about 60% of these catchments with a clear presence in the Amazonas, Africa, and Australia. The Jensen-Haise formula best captured the three assessed variables in about 27% of the catchments but with no clear geographical distribution. The same occurs with the Priestley-Taylor formula, which only satisfies the multi-process analysis in about 14% of the catchments (with no clear global distribution) due to little agreement between the two EO products.

Since the global coverage was limited for independent streamflow gauges (cf. Figure 1), the multi-process analysis was performed again just using MODIS evaporation products (Figure 3b) to increase the representativeness in the analysis. In this case, the global coverage increased from 39% with three variables to 82% of the total number of catchments. Well-defined spatial patterns of the different PET formulas were found; in particular, Hargreaves was the best PET formula in 50% of the catchments, most of them located in the Amazonas, central

10.1029/2022WR033447

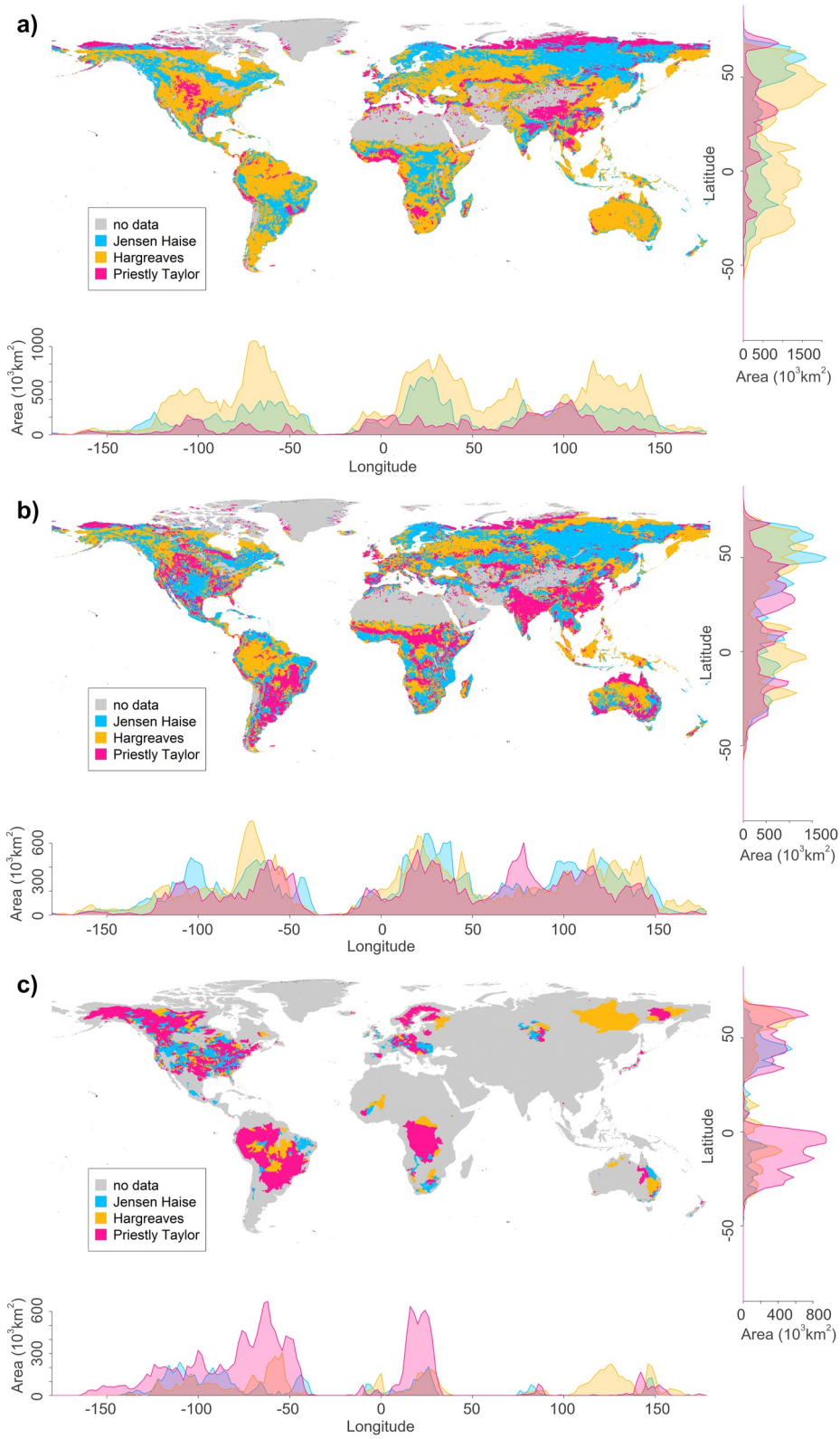

Figure 2. Global distribution of best PET formula according to: (a) MOD16 PET (115,151 catchments), (b) MOD16 AET (115,151 catchments), and (c) streamflow observations (4,367 catchments).

Table 3
Number of Catchments, Area, and Median Relative Error (RE) Between World-Wide HYPE (WWH) and Observations Using the Three Potential Evapotranspiration (PET) Formulas (Rows) for PET, Actual Evapotranspiration (AET) and Streamflow (Columns) Globally (All World) and for Catchments Where Each Formula Was Performing Best (Best Performing Catchments)

	MOD16 PET		MOD16 AET		Streamflow	
PET formula	All world	Best performing catchments	All world	Best performing catchments	All world	Best performing catchments
Jensen-Haise						
No.	115,151	32.1% (36,948)	115,151	36.9% (42,467)	36,931	(19.0%) 7,032
Area (km^2)	110,796,384	32,807,917	110,796,384	39,300 441	35,231,436	6,473,656
Median RE (%)	−17.02	−12.99	1.14	−7.27	−21.74	−2.93
Hargreaves						
No.	115,151	53,5% (61,650)	115,151	35.4% (40,758)	36,931	(25.1%) 9,264
Area (km^2)	110,796,384	62,363,685	110,796,384	39,639 251	35,231,436	9,020,790
Median RE (%)	−10.67	−7.85	−2.54	−2.16	−14.06	−8.39
Priestley-Taylor						
No.	115,151	14.4% (16,553)	115,151	27.7% (31,926)	36,931	(55.8%) 20,635
Area (km^2)	110,796,384	15,624,782	110,796,384	31,812 655	35,231,436	19,736,990
Median RE (%)	−21.74	−17.5	−14.06	6.64	−16.51	−15.87

Europe, and Oceania, while Jensen-Haise was better representing catchments (36%) in northern latitudes. Finally, Priestley-Taylor was identified as the best formula for India, China, Mainland Southeast Asia, and latitudes above 65°N.

4.3. Sensitivity of the Hydrological Model Performance to the Choice of PET Formula

Here, we compare the difference between the best and the worst performance of WWH (in terms of RE against EO and streamflow observations) with the different PET formulas (Figure 4). The spatial patterns of the RE differences were quite similar for the two EO products (Figures 4a and 4b), with the RE difference being slightly larger for PET than for AET. The selection of a PET formula seems to be more critical in tropical regions close to the equator, where the differences are above 50%, than in other parts of the world. In regions located at southern and northern latitudes, the impact of the choice of a PET formula is smaller and so is the PET absolute values. When analyzing the performance improvement in terms of streamflow (Figure 4c), similar results were found, although the representativeness of the analyzed areas is much smaller due to the low global coverage of independent streamflow gauges.

The methodology used here, which combines remote sensing information (EO) and in-situ streamflow data, not only allowed the evaluation of internal model variables such as PET and AET, but also enlarged the representative area for model evaluation. This allowed for a model evaluation beyond streamflow data. However, we also recognize that ET information from EO products is not directly observations from space but rather retrieved by different algorithms, and therefore subject to uncertainties (see Data Section 2.1 above). For instance, Miralles et al. (2016) demonstrated an underestimation on ET retrieval from the MOD16 product in the tropical region, showing about 250 mm/year difference to the ensemble mean of MOD16, GLEAM and PT-JPL models. Hence, this might to some extent explain the large differences in results we found over these areas. However, our results show differences beyond 50% with absolute values of 600 mm/year () and is thus much larger than the MOD16 underestimation assessed by Miralles et al. (2016). Therefore, we can conclude that the catchment model performance is sensitive to the PET formula selection, at least in tropical regions. The differences we found (Figure 4) are lower in other regions world-wide and it masked could thus be hidden by the general uncertainties in ET retrieval from remote sensing data (Hu et al., 2015; Kim et al., 2012; Ruhoff et al., 2013; Trambauer et al., 2014) or from the fact that ET is a less sensitive parameter in catchment modeling at higher latitudes where the ET is lower.

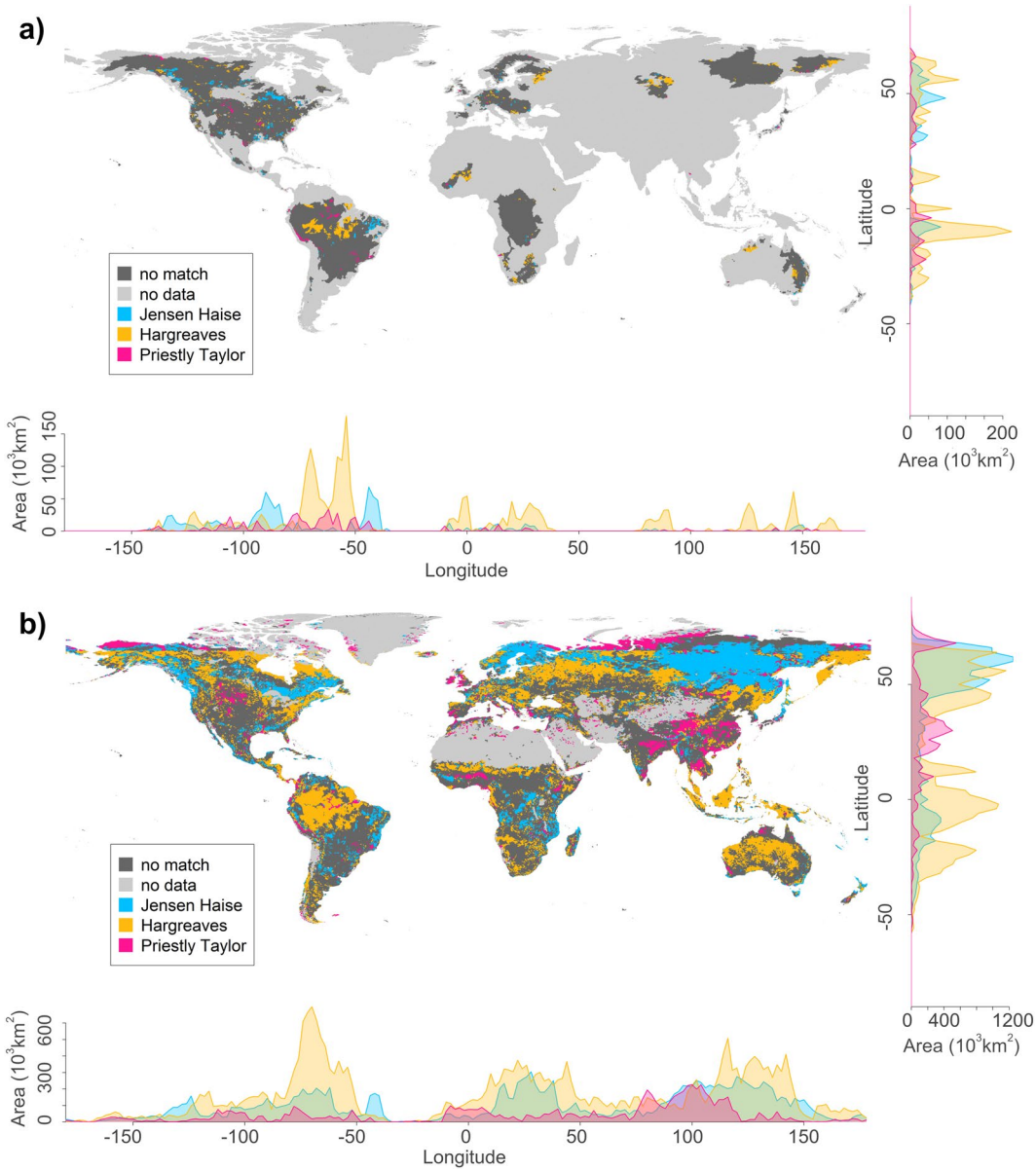

Figure 3. Global distribution of the best potential evapotranspiration (PET) formula according to the multi-process analysis for: (a) a combination of the three variables PET, actual evapotranspiration (AET) and streamflow against MOD16 PET, MOD16 AET and streamflow observations, respectively (4,367 catchments), and (b) a combination of the two variables PET and AET against MOD16 PET and MOD16 AET respectively (115,151 catchments). No match means that no single PET formula was optimal for all three variables in the catchment.

4.4. Linking Physiographic Characteristics to Best Performing PET Formula

Here, we analyzed the relation between the best identified PET formula (Figure 3b) and some physiographic features (Table 1) of each catchment. There is a clear connection between the five main Köppen regions and the PET formulas (Figure 5, top left panel); for example, Priestley-Taylor is the best PET formula in polar areas (Köppen region E). This is directly connected to the fact that Köppen climate regions are defined using monthly temperature range as one of the main factors and the PET formulas analyzed here are highly dependent on temperature.

Tropical (A) and dry (B) climatic regions, which are mostly located between the tropics, are best represented by the Hargreaves formula. The main difference between both climate regions is related to precipitation ranges;

10.1029/2022WR033447

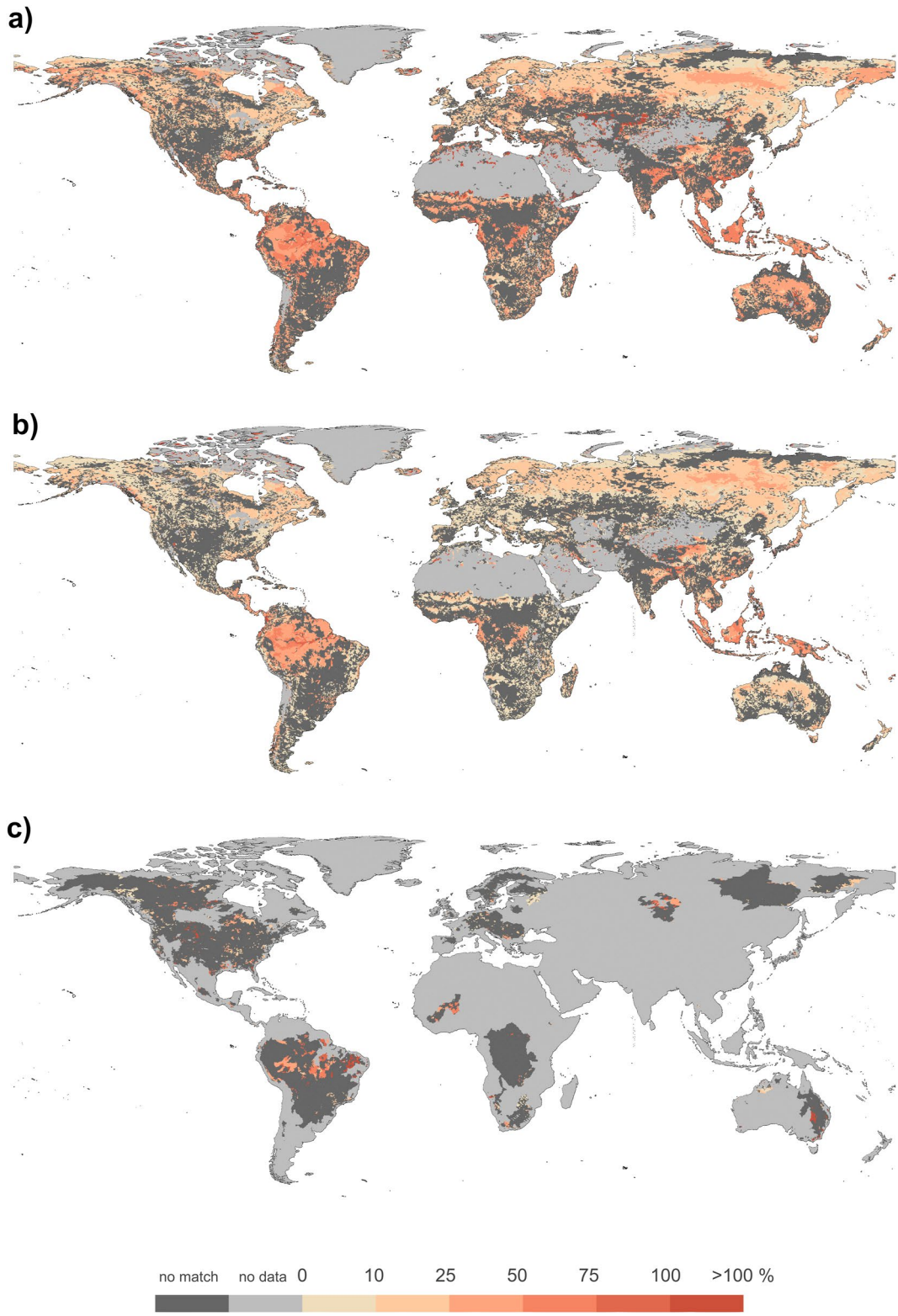

Figure 4. Global distribution of the gain in relative error (RE) from choosing the best over the worst performing potential evapotranspiration (PET) formula according to: (a) MOD16 PET (115,151 catchments), (b) MOD16 actual evapotranspiration (AET) (115,151 catchments), and (c) streamflow observations (4,367 catchments). No match means that the three variables were not optimal for the same PET formula in the catchment.

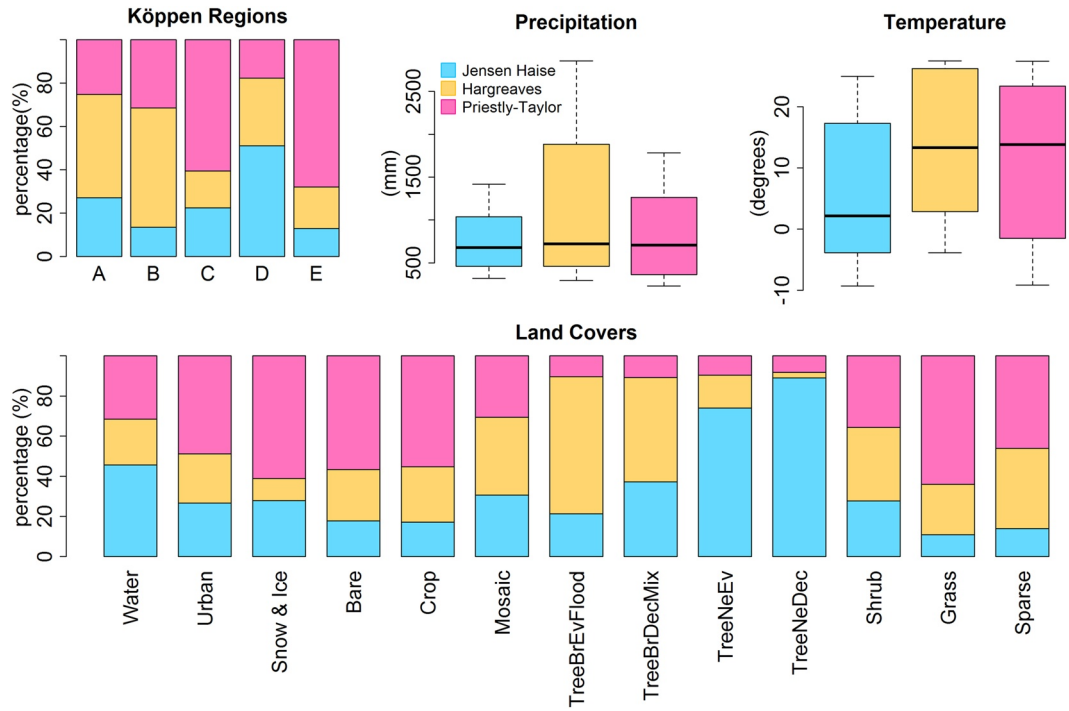

Figure 5. Capacity of the potential evapotranspiration (PET) formulas to capture, PET and actual evapotranspiration (AET) depending on the physiographic characteristics of the catchments (i.e., the 115,151 catchments with MOD16 data). Physiographic characteristics are found in Table 1.

tropical climate (A) has high precipitation, while dry climate (B) has low precipitation. In addition, temperatures in tropical climate are always above 18°C for the coolest month, but in dry climate the monthly coolest temperature is always below this threshold. Both behaviors are reflected in the precipitation and temperature boxplots for the Hargreaves formula (Figure 5, top middle, and top right panels).

For the Hargreaves formula, the precipitation range was larger and the lowest temperature values were higher than for the other formulas. It is to be mentioned that daily temperature range over the tropical (A) and dry (B) climatic regions is large, especially in the dry climate (Zanetti et al., 2019). This might be the main cause for the better performance of this PET formula over this region, since it uses the daily temperature range for its calculations instead of the mean values like the Jensen-Haise formula.

The Jensen-Haise formula best captures the process dynamics in the continental climate (D), which is geographically located in inner continents and usually at latitudes above 40°N. Overall, 51% of the catchments located in this climatic region are better represented by this PET formula. Monthly temperature in this climatic region is always above 10°C in the warmest month and below 0°C for the coldest months (Peel et al., 2007). This temperature range is supported by the monthly temperature values in these catchments (Figure 5, top right panel). Despite this annual variability, the daily temperature range in these regions is smaller than in more southern regions. This fact might make the Jensen-Haise formula (which only uses daily mean temperature) perform appropriately in this climate region.

The Priestley-Taylor formula seems to perform better in temperate (C) and polar (E) climatic regions. These two regions differ strongly, especially regarding temperature. While the warmest monthly temperature in temperate climate is always above 10°C, it is always below this threshold in polar climate. Priestley-Taylor is a so-called combination formula (water and energy approach), based on an approximation of the surface energy balance. This approach is clearly needed in polar regions (E), characterized by the presence of snow and ice. The high albedo of snow, the potential sunshine hours, and the presence of clouds highly condition the radiative balance. In the case of temperate climate (C) regions, the high variability within the region, with many possible combinations of dry summer/winter and hot/cold/warm summer, makes the physical approximation of Priestley-Taylor formula more capable to capture these ET dynamics.

Finally, we found that land cover was also highly conditioned by climatic region, which was recognized already by Köppen, who had a background in plant sciences and was inspired to develop the classification from the global vegetation map that Grisebach published in 1866 (Peel et al., 2007; Wilcock, 1968). Therefore, the connection we found between PET formulas and climate is implicitly reflected in land covers. The Jensen-Haise formula thus better represents needle-leaved trees (both deciduous and evergreen), which are mainly located in northern latitudes linked to continental climate. On the contrary, the Hargreaves formula is more connected to land covers present in tropical regions, such as broadleaved deciduous evergreen forest. Finally, Priestley-Taylor is the PET formula that performs best in snow-, ice-, and bare soil-ruled catchments from the polar region (E), and some grass- and sparse vegetation-driven catchments from temperate climate (C) (Figure 5, bottom panel).

Hence, this analysis shows that there is a quantitative scientific basis for using the Köppen regions when choosing a PET formula. We also note that, in most areas, the expert-based choice of a PET formula applied by Arheimer et al. (2020) corresponded well with the quantitative evaluation employed in this analysis for tropical (A), dry (B), polar (E) regions. However, Arheimer et al. (2020) swapped the best PET formulas between temperate and continental climates (e.g., Priestley-Taylor should be used for temperate climate instead of continental, and likewise, Jensen-Haise should be used for continental (D) instead of temperate (C) climate). This will be considered in the next version of WWH.

The link between climate and PET formula selection may be important in assessing climate-change impacts via catchment modeling, especially if main Köppen regions are to radically change under global warming. However, Beck et al. (2018) provided a Köppen region map for the end of the 21st century, without finding large changes in the distribution of the main five categories. The main expected changes are limited to some categories in the northern hemisphere, for example, regions with continental humid climate (Df) migrating to more continental Mediterranean climate (Ds). The only main climate region with significant changes is central Europe. Therefore, except for central Europe, the choice of PET formula should not be drastically affected by climate change. This is in line with conclusions by Lemaitre-Basset et al. (2022), who affirmsargue that all climate variables that covary with temperature should have the same relationships in climate projections.

4.5. Linking Water- and Energy-Limited Catchments to Best Performing PET Formula

When plotting all analyzed catchments in the Budyko space according to their best PET formula (in terms of the MODIS evaporation products, Figure 3b) their position along the curve between energy-limited (PET/P < 1) and water-limited (PET/P > 1) catchments were almost negligible, with small differences in the percentages, about 50% of catchment, in one category or the other (Figure 6). However, upon closer analysis we observed areas in the Budyko space with a dense clustering of catchments. In general, these highly dense areas are in the energy-limited region (PET/P < 1) for the three PET formulas but with different positions along the Budyko curve. PriestlyPriestley-Taylor seems to be more suitable over the more energy-limited catchments (AET/P < 0.4); Jensen-Haise is better in catchments with AET being around half of the total catchment precipitation (0.4 < AET/P < 0.75); and Hargreaves is better in catchments closer to the water-limited threshold (AET/P > 0.5). When analyzing the water-limited catchments (PET/P > 1) we observed a much higher spread in the Budyko space, especially in the case of the PriestlyPriestley-Taylor formula.

When plotting all analyzed catchments in the Budyko space according to their best PET formula (in terms of the MODIS evaporation products, Figure 3b) their position along the curve between energy-limited (PET/P < 1) and water-limited (PET/P > 1) catchments were almost negligible, with small differences in the percentages, about 50% of catchment, in one category or the other (Figure 6). However, upon closer analysis we observed areas in the Budyko space with a dense clustering of catchments. In general, these highly dense areas are in the energy-limited region (PET/P < 1) for the three PET formulas but with different positions along the Budyko curve. Priestley-Taylor seems to be more suitable over the more energy-limited catchments (AET/P < 0.4); Jensen-Haise is better in catchments with AET being around half of the total catchment precipitation (0.4 < AET/P < 0.75); and Hargreaves is better in catchments closer to the water-limited threshold (AET/P > 0.5). When analyzing the water-limited catchments (PET/P > 1) we observed a much higher spread in the Budyko space, especially in the case of the Priestley-Taylor formula.

These results of catchment distribution within the Budyko space are in line with previous findings by Jaramillo and Destouni (2014) for each continent. Thus, we observed that overall, the Jensen-Haise formula is best others

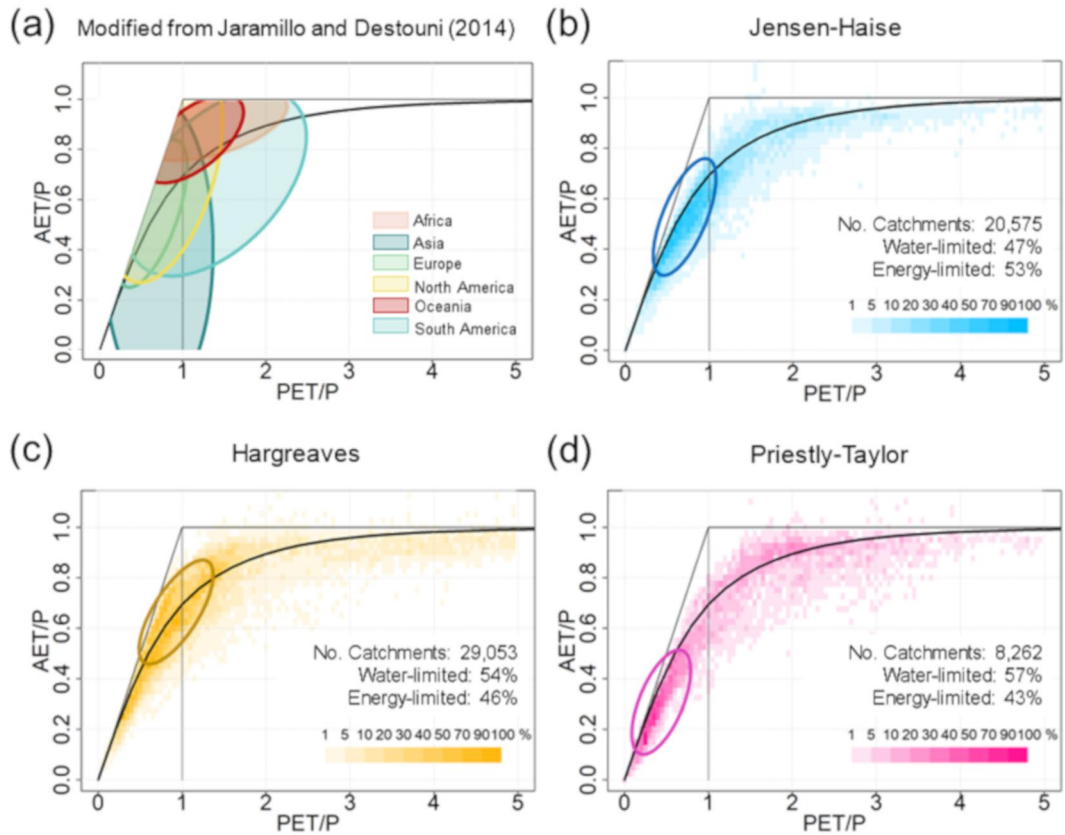

Figure 6. (a) Continental water characteristics in the Budyko space (potential evapotranspiration over precipitation (PET/P) versus actual evapotranspiration over precipitation (AET/P)), modified from Jaramillo and Destouni (2014), (b–d) Distribution of the catchments better represented (according to Figure 3b) by (b) Jensen-Haise, (c) Hargreaves, (d) Priestley-Taylor, within the Budyko space. The intensity of the color indicates the density of the catchments.

representing PET for catchments in Europe, Asia, and North America; the Hargreaves formula in North America, Africa, and Oceania; and the Priestley-Taylor formula in Asia. This is coherent with the above results on the best performing formula for each location and links to the catchment physiography and climate region.

5. Conclusions

Based on this large sample analysis including thousands of catchments over the world, we recommend applying different PET formulas in different climate regions. Our results indicate that higher complexity in the PET formula only gives improved predictions in temperate and polar regions. Based on the Köppen classification, we suggest using the Jensen-Haise formula in continental climate, the Hargreaves formula in tropical and dry climate, and the Priestley-Tailor formula in temperate and polar climate. This recommendation is based on the following findings:

- We have found a clear connection between the best PET formula for catchment modeling and the main climate regions world-wide. The simpler Jensen-Haise formula offers a better model performance in catchments located in the continental climate due to lower daily temperature variations; the Hargreaves formula in the tropical and arid regions where the daily temperature range is clearly affecting the ET rates; and the Priestley-Taylor formula in the temperate and polar areas, which are clearly driven by the radiative budget.
- The relation between the best PET formulas and the predominant land cover reflects that between the best PET formula and the climate type. The main vegetation types are highly conditioned by the climate region where they are located. Hence, the Jensen-Haise formula was found to better perform over needled trees (both deciduous and evergreen), Hargreaves better performs in broadleaved deciduous evergreen forests, and lastly, Priestley-Taylor better performs over snow and ice, and grass and sparse vegetation.

- The analysis using the Budyko space to classify energy or water limitation in catchments and continents further supports the impact of climate on the PET formula performance.
- The selection of the PET formula can strongly affect the model performance in terms of ET (potential and actual) and streamflow variables. In tropical regions, we found a 50% difference in estimated ET and a 75% difference in streamflow between the formulas. It is thus of major importance to apply the most appropriate PET formula in hydrological modeling across the globe.

Data Availability Statement

Regarding data, two different Earth Observation (EO) products were used for this research: (a) MOD16 Potential Evapotranspiration and (b) MOD16 Actual Evapotranspiration (Mu et al., 2011) https://lpdaac.usgs.gov/products/mod16a2v005/ (accessed 2020-04-30). In addition, open streamflow data used as summarized in Crochemore et al. (2020). In relation to software, the global setup of the HYPE model (Arheimer et al., 2020) was used as hydrological model in this study. The HYPE model code and simulations can be download at: https://hypeweb.smhi.se/ (accessed 2022-08-10). Long term mean Potential Evapotranspiration (PET) and Actual Evapotranspiration (EAT) estimates using World-Wide HYPE and different PET-formula are openly available (https://doi.org/10.5281/zenodo.7693381).

Acknowledgments

We would like to thank three anonymous reviewers and the co-editor for useful comments, which helped improving the original manuscript. Furthermore, we thank all the data providers of open data who made their results and observations readily available for re-purposing; without you any global hydrological modelling would not have been possible at all. Especially we would like to express our gratitude to Dai Yamazaki, University of Tokyo, for developing and sharing the global width database for large rivers. WWH was developed at the SMHI hydrological research unit, where much work is done in common, taking advantage of previous work and several projects running in parallel in the group. It was indeed a team effort. We would especially like to acknowledge contributions from our colleagues Kristina Isberg, Jörgen Rosberg, Lotta Pers and Peter Berg, who provide much of the model infrastructure. RP acknowledges the fundings by the Juan de la Cierva-Incorporación program of the Spanish Ministry of Science and Innovation (IJC2018-038093-I). In addition, RP is currently member of DAUCO, Unit of Excellence ref. CEX2019-000968-M, with financial support from the Spanish Ministry of Science and Innovation, the Spanish State Research Agency, through the *Severo Ochoa and María de Maeztu* Program for Centers and Units of Excellence in R&D. Funding for open access charge: Universidad de Córdoba / CBUA.

References

Allen, C. D., & Breshears, D. D. (1998). Drought-induced shift of a forest–woodland ecotone: Rapid landscape response to climate variation. *Proceedings of the National Academy of Sciences*, *95*(25), 14839–14842. https://doi.org/10.1073/pnas.95.25.14839

Andersson, J. C. M., Ali, A., Arheimer, B., Gustafsson, D., & Minoungou, B. (2017). Providing peak river flow statistics and forecasting in the Niger River basin. *Physics and Chemistry of the Earth, Parts A/B/C*, *100*, 3–12. https://doi.org/10.1016/j.pce.2017.02.010

Andersson, J. C. M., Arheimer, B., Traoré, F., Gustafsson, D., & Ali, A. (2017). Process refinements improve a hydrological model concept applied to the Niger River basin. *Hydrological Processes*, *31*(25), 4540–4554. https://doi.org/10.1002/hyp.11376

Andersson, L. (1992). Improvements of runoff models - What way to go? *Nordic Hydrology*, *23*(5), 315–332. https://doi.org/10.2166/nh.1992.0022

Andréassian, V., Perrin, C., & Michel, C. (2004). Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models. *Journal of Hydrology*, *286*(1–4), 19–35. https://doi.org/10.1016/j.jhydrol.2003.09.030

Arheimer, B., Donnelly, C., & Lindström, G. (2017). Regulation of snow-fed rivers affects flow regimes more than climate change. *Nature Communications*, *8*(62), 62. https://doi.org/10.1038/s41467-017-00092-8

Arheimer, B., & Lindström, G. (2015). Climate impact on floods: Changes in high flows in Sweden in the past and the future (1911–2100). *Hydrology and Earth System Sciences*, *19*(2), 771–784. https://doi.org/10.5194/hess-19-771-2015

Arheimer, B., Nilsson, J., & Lindström, G. (2015). Experimenting with coupled hydro-ecological models to explore measure plans and water quality goals in a semi-enclosed Swedish Bay. *Water*, *7*(7), 3906–3924. https://doi.org/10.3390/w7073906

Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J. C. M., Hasan, A., & Pineda, L. (2020). Global catchment modelling using World-Wide HYPE (WWH), open data, and stepwise parameter estimation. *Hydrology and Earth System Sciences*, *24*(2), 535–559. https://doi.org/10.5194/hess-24-535-2020

Bai, P., Liu, X., Yang, T., Li, F., Liang, K., Hu, S., & Liu, C. (2016). Assessment of the influences of different potential evapotranspiration inputs on the performance of monthly hydrological models under different climatic conditions. *Journal of Hydrometeorology*, *17*(8), 2259–2274. https://doi.org/10.1175/JHM-D-15-0202.1

Baier, W., & Robertson, G. W. (1965). Estimation of latent evaporation from simple weather observations. *Canadian Journal of Plant Science*, *45*(3), 276–284. https://doi.org/10.4141/cjps65-051

Bartosova, A., Capell, R., Olesen, J. E., Jabloun, M., Refsgaard, J. C., Donnelly, C., et al. (2019). Future socioeconomic conditions may have a larger impact than climate change on nutrient loads to the Baltic Sea. *Ambio*, *48*(11), 1325–1336. https://doi.org/10.1007/s13280-019-01243-5

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 1–12. https://doi.org/10.1038/sdata.2018.214

Budyko, M. I. (1948). *Evaporation under natural conditions*. Gidrometeorizdat. English translation by IPST.

Crochemore, L., Isberg, K., Pimentel, R., Pineda, L., Hasan, A., & Arheimer, B. (2020). Lessons learnt from checking the quality of openly accessible river flow data worldwide. *Hydrological Sciences Journal*, *65*(5), 699–711. https://doi.org/10.1080/02626667.2019.1659509

Dalton, J. (1802). Experiments and observations to determine whether the quantity of rain and dew is equal to the quantity of water carried off by the rivers and raised by evaporation, with an enquiry into the origin of springs. *Memoirs of the Literary and Philosophical Society of Manchester, V, II* (pp. 346–372).

Dingman, S. L. (2015). *Physical hydrology* (5th ed.). Waveland press.

Dolman, A. J., & De Jeu, R. A. M. (2010). Evaporation in focus. *Nature Geoscience*, *3*(5), 296. https://doi.org/10.1038/ngeo849

Donnelly, C., Andersson, J. C. M., & Arheimer, B. (2016). Using flow signatures and catchment similarities to evaluate a multi-basin model (E-HYPE) across Europe. *Hydrological Sciences Journal*, *61*(2), 255–273. https://doi.org/10.1080/02626667.2015.1027710

Ellenburg, W. L., Cruise, J. F., & Singh, V. P. (2018). The role of evapotranspiration in streamflow modeling – An analysis using entropy. *Journal of Hydrology*, *567*, 290–304. https://doi.org/10.1016/j.jhydrol.2018.09.048

Fisher, J. B., DeBiase, T. A., Qi, Y., Xu, M., & Goldstein, A. H. (2005). Evapotranspiration models compared on a Sierra Nevada forest ecosystem. *Environmental Modelling & Software*, *20*(6), 783–796. https://doi.org/10.1016/j.envsoft.2004.04.009

Fisher, J. B., Tu, K. P., & Baldocchi, D. D. (2008). 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*(3), 901–919. https://doi.org/10.1016/J.RSE.2007.06.025

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *83*(1), 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0

Hargreaves, G. H., & Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, *1*(2), 96–99. https://doi.org/10.13031/2013.26773

Horton, R. E. (1919). Rainfall interception. *Monthly Weather Review*, *47*(9), 603–623. https://doi.org/10.1175/1520-0493(1919)47<603:RI>2.0.CO;2

Hu, G., Jia, L., & Menenti, M. (2015). Comparison of MOD16 and LSA-SAF MSG evapotranspiration products over Europe for 2011. *Remote Sensing of Environment*, *156*, 510–526. https://doi.org/10.1016/J.RSE.2014.10.017

Hundecha, Y., Arheimer, B., Donnelly, C., & Pechlivanidis, I. (2016). A regional parameter estimation scheme for a pan- European multi-basin model. *Journal of Hydrology: Regional Studies*, *6*, 90–111. https://doi.org/10.1016/J.EJRH.2016.04.002

Immerzeel, W. W., & Droogers, P. (2008). Calibration of a distributed hydrological model based on satellite evapotranspiration. *Journal of Hydrology*, *349*(3–4), 411–424. https://doi.org/10.1016/j.jhydrol.2007.11.017

Jaramillo, F., & Destouni, G. (2014). Developing water change spectra and distinguishing change drivers worldwide. *Geophysical Research Letters*, *41*(23), 8377–8386. https://doi.org/10.1002/2014GL061848

Jayathilake, D. I., & Smith, T. (2020). Assessing the impact of PET estimation methods on hydrologic model performance. *Hydrological Research*, *52*(2), 373–388. https://doi.org/10.2166/nh.2020.066J

Jensen, M. E., & Haise, H. R. (1963). Estimating evapotranspiration from solar radiation. *Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division*, *89*(4), 15–41. https://doi.org/10.1002/2014GL061848

Jiang, L., Wu, H., Tao, J., Kimball, J. S., Alfieri, L., & Chen, X. (2020). Satellite-based evapotranspiration in hydrological model calibration. *Remote Sensing*, *12*(3), 428. https://doi.org/10.3390/rs12030428

Jin, Y., Schaaf, C. B., Gao, F., Li, X., Strahler, A. H., Lucht, W., & Liang, S. (2003). Consistency of MODIS surface bidirectional reflectance distribution function and albedo retrievals: 1. Algorithm performance. *Journal of Geophysical Research*, *108*(D5), 4158. https://doi.org/10.1029/2002JD002803

Jung, M., Reichstein, M., & Bondeau, A. (2009). Towards global empirical upscaling of FLUXNET eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, *6*(10), 2001–2013. https://doi.org/10.5194/BG-6-2001-2009

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., et al. (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, *467*(7318), 951–954. https://doi.org/10.1038/nature09396

Kim, H. W., Hwang, K., Mu, Q., Lee, S. O., & Choi, M. (2012). Validation of MODIS 16 global terrestrial evapotranspiration products in various climates and land cover types in Asia. *KSCE Journal of Civil Engineering*, *16*(2), 229–238. https://doi.org/10.1007/s12205-012-0006-1

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*(3), 259–263. https://doi.org/10.1127/0941-2948/2006/0130

Krysanova, V., Vetter, T., Eisner, S., Huang, S., Pechlivanidis, I. G., Strauch, M., et al. (2017). Intercomparison of regional-scale hydrological models in the present and future climate for 12 large river basins worldwide - A synthesis. *Environmental Research Letters*, *12*(10), 105002. https://doi.org/10.1088/1748-9326/aa8359

Kuentz, A., Arheimer, B., Hundecha, Y., & Wagener, T. (2017). Understanding hydrologic variability across Europe through catchment classification. *Hydrology and Earth System Sciences*, *21*(6), 2863–2879. https://doi.org/10.5194/hess-21-2863-2017

Lebedeva, L., & Gustafsson, D. (2021). Streamflow changes of small and large rivers in the Aldan river basin, eastern Siberia. *Water*, *13*, 19. https://doi.org/10.3390/w13192747

Lemaitre-Basset, T., Oudin, L., Thirel, G., & Collet, L. (2022). Unraveling the contribution of potential evaporation formulation to uncertainty under climate change. *Hydrology and Earth System Sciences*, *26*(8), 2147–2159. https://doi.org/10.5194/HESS-26-2147-2022

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., & Arheimer, B. B. (2010). Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology Research*, *41*(3–4), 295–319. https://doi.org/10.2166/nh.2010.007

López López, P., Sutanudjaja, E. H., Schellekens, J., Sterk, G., & Bierkens, M. F. P. (2017). Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. *Hydrology and Earth System Sciences*, *21*(6), 3125–3144. https://doi.org/10.5194/hess-21-3125-2017

Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., & Dolman, A. J. (2011). Magnitude and variability of land evaporation and its components at the global scale. *Hydrology and Earth System Sciences*, *15*(3), 967–981. https://doi.org/10.5194/HESS-15-967-2011

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., McCabe, M. F., et al. (2016). The WACMOS-ET project – Part 2: Evaluation of global terrestrial evaporation data sets. *Hydrology and Earth System Sciences*, *20*(2), 823–842. https://doi.org/10.5194/hess-20-823-2016

Monteith, J. L. (1965). Evaporation and environment. In *Symposia of the society for experimental biology* (Vol. 19, pp. 205–234). Cambridge University Press (CUP).

Mu, Q., Heinsch, F. A., Zhao, M., & Running, S. W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment*, *111*(4), 519–536. https://doi.org/10.1016/j.rse.2007.04.015

Mu, Q., Zhao, M., & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, *115*(8), 1781–1800. https://doi.org/10.1016/j.rse.2011.02.019

Musuuza, J. L., Gustafsson, D., Pimentel, R., Crochemore, L., & Pechlivanidis, I. (2020). Impact of Satellite and in situ data assimilation on hydrological predictions. *Remote Sensing*, *12*(5), 811. https://doi.org/10.3390/rs12050811

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., et al. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, *83*(1), 214–231. https://doi.org/10.1016/S0034-4257(02)00074-3

Nijzink, R. C., Almeida, S., Pechlivanidis, I. G., Capell, R., Gustafssons, D., Arheimer, B., et al. (2018). Constraining conceptual hydrological models with multiple information sources. *Water Resources Research*, *54*(10), 8332–8362. https://doi.org/10.1029/2017WR021895

Odusanya, A. E., Mehdi, B., Schürz, C., Oke, A. O., Awokola, O. S., Awomeso, J. A., et al. (2019). Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data-sparse catchment in southwestern Nigeria. *Hydrology and Earth System Sciences*, *23*(2), 1113–1144. https://doi.org/10.5194/hess-23-1113-2019

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of Hydrology*, *303*(1–4), 290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026

Pechlivanidis, I. G., & Arheimer, B. (2015). Large-scale hydrological modelling by using modified PUB recommendations: The India-HYPE case. *Hydrology and Earth System Sciences*, *19*(11), 4559–4579. https://doi.org/10.5194/hess-19-4559-2015

Pechlivanidis, I. G., Bosshard, T., Spångmyr, H., Lindström, G., Gustafsson, D., & Arheimer, B. (2014). Uncertainty in the Swedish operational hydrological forecasting systems. In *ASCE proceedings: Vulnerability, uncertainty, and risk* (pp. 253–262). https://doi.org/10.1061/9780784413609.026

Pechlivanidis, I. G., Crochemore, L., Rosberg, J., & Bosshard, T. (2020). What are the key drivers controlling the forecasts of seasonal streamflow volumes? *Water Resources Research*, *56*(6), e2019WR026987. https://doi.org/10.1029/2019WR026987

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *193*(1032), 120–145. https://doi.org/10.1098/rspa.1948.0037

Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, *100*(2), 81–92. https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

Rice, J. S., Emanuel, R. E., Vose, J. M., & Nelson, S. A. C. (2015). Continental U.S. streamflow trends from 1940 to 2009 and their relationships with watershed spatial characteristics. *Water Resources Research*, *51*(8), 6262–6275. https://doi.org/10.1002/2014WR016367

Ruhoff, A. L., Paz, A. R., Aragao, L. E. O. C., Mu, Q., Malhi, Y., Collischonn, W., et al. (2013). Assessment of the MODIS global evapotranspiration algorithm using eddy covariance measurements and hydrological modelling in the Rio Grande basin. *Hydrological Sciences Journal*, *58*(8), 1658–1676. https://doi.org/10.1080/02626667.2013.837578

Salomon, J. G., Schaaf, C. B., Strahler, A. H., Feng, G., & Yufang, J. (2006). Validation of the MODIS bidirectional reflectance distribution function and albedo retrievals using combined observations from the aqua and terra platforms. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(6), 1555–1565. https://doi.org/10.1109/TGRS.2006.871564

Samain, B., & Pauwels, V. R. N. (2013). Impact of potential and (scintillometer-based) actual evapotranspiration estimates on the performance of a lumped rainfall–runoff model. *Hydrology and Earth System Sciences*, *17*(11), 4525–4540. https://doi.org/10.5194/hess-17-4525-2013

Santos, L., Andersson, J. C. M., & Arheimer, B. (2022). Evaluation of parameter sensitivity of a rainfall-runoff model over a global catchment set. *Hydrological Sciences Journal*, *67*(3), 342–357. https://doi.org/10.1080/02626667.2022.2035388

Savenije, H. H. G. (2004). The importance of interception and why we should delete the term evapotranspiration from our vocabulary. *Hydrological Processes*, *18*(8), 1507–1511. https://doi.org/10.1002/hyp.5563

Sawicz, K. A., Kelleher, C., Wagener, T., Troch, P., Sivapalan, M., & Carrillo, G. (2014). Characterizing hydrologic change through catchment classification. *Hydrology and Earth System Sciences*, *18*(1), 273–285. https://doi.org/10.5194/hess-18-273-2014

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., et al. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*(1), 135–148. https://doi.org/10.1016/S0034-4257(02)00091-3

Seong, C., Sridhar, V., & Billah, M. M. (2018). Implications of potential evapotranspiration methods for streamflow estimations under changing climatic conditions. *International Journal of Climatology*, *38*(2), 896–914. https://doi.org/10.1002/joc.5218

Stadnyk, T. A., MacDonald, M. K., Tefs, A., Dery, S. J., Koenig, K., Gustafsson, D., et al. (2020). Hydrological modeling of freshwater discharge into Hudson Bay using HYPE. *Elementa-Science of the Anthropocene*, *8*, 43. https://doi.org/10.1525/elementa.439

Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C., & Lindström, G. (2012). Water and nutrient predictions in ungauged basins – Set-up and evaluation of a model at the national scale. *Hydrological Sciences Journal*, *57*(2), 229–247. https://doi.org/10.1080/02626667.2011.637497

Tanner, C. B. (1967). Measurement of evapotranspiration. In R. M. Hagan, H. R. Haise, & T. W. Edminster (Eds.), *Irrigation of agricultural lands* (pp. 534–574). Am. Soc. Agron.

Ter Braak, C. J. F. (2006). A Markov chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter spaces. *Statistical Computing*, *16*(3), 239–249. https://doi.org/10.1007/s11222-006-8769-1

Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical Review*, *38*(1), 55–94. https://doi.org/10.1097/00010694-194807000-00007

Trambauer, P., Dutra, E., Maskey, S., Werner, M., Pappenberger, F., Van Beek, L. P. H., & Uhlenbrook, S. (2014). Comparison of different evaporation estimates over the African continent. *Hydrology and Earth System Sciences*, *18*(1), 193–212. https://doi.org/10.5194/hess-18-193-2014

Valipour, M., Gholami Sefidkouhi, M. A., & Raeini−Sarjaz, M. (2017). Selecting the best model to estimate potential evapotranspiration with respect to climate change and magnitudes of extreme events. *Agricultural Water Management*, *180*, 50–60. https://doi.org/10.1016/J.AGWAT.2016.08.025

Velpuri, N. M., Senay, G. B., Singh, R. K., Bohms, S., & Verdin, J. P. (2013). A comprehensive evaluation of two MODIS evapotranspiration products over the conterminous United States: Using point and gridded FLUXNET and water balance ET. *Remote Sensing of Environment*, *139*, 35–49. https://doi.org/10.1016/J.RSE.2013.07.013

Wang, K., & Dickinson, R. E. (2012). A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Reviews of Geophysics*, *50*(2). https://doi.org/10.1029/2011RG000373

Weiß, M., & Menzel, L. (2008). A global comparison of four potential evapotranspiration equations and their relevance to stream flow modelling in semi-arid environments. *Advances in Geosciences*, *18*, 15–23. https://doi.org/10.5194/adgeo-18-15-2008

Wilcock, A. A. (1968). Koppen after fifty years. *Annals of the Association of American Geographers*, *58*(1), 12–28. https://doi.org/10.1111/j.1467-8306.1968.tb01633.x

Yang, Y., Chen, R., Han, C., & Liu, Z. (2021). Evaluation of 18 models for calculating potential evapotranspiration in different climatic zones of China. *Agricultural Water Management*, *244*, 106545. https://doi.org/10.1016/j.agwat.2020.106545

Zanetti, S. S., Estevez-Dohler, R., Cecílio, R. A., Macedo-Pezzopane, J. E., & Xavier, A. C. (2019). Proposal for the use of daily thermal amplitude for the calibration of the Hargreaves-Samani equation. *Journal of Hydrology*, *571*, 193–201. https://doi.org/10.1016/j.jhydrol.2019.01.049

Zhou, J., Wang, Y., Su, B., Wang, A., Tao, H., Zhai, J., et al. (2020). Choice of potential evapotranspiration formulas influences drought assessment: A case study in China. *Atmospheric Research*, *242*, 104979. https://doi.org/10.1016/j.atmosres.2020.104979