https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Hydrology and Earth System Sciences
Discussions

Catchments do not strictly follow Budyko curves over multiple decades but deviations are minor and predictable

Muhammad Ibrahim[1], Miriam Coenders-Gerrits[1], Ruud van der Ent[1], Markus Hrachowitz [1]

[1]Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands

Correspondence to: Muhammad Ibrahim (m.ibrahim@tudelft.nl)

Abstract. Quantification of precipitation partitioning into evaporation and runoff is crucial for predicting future water availability. Within the widely used Budyko Framework, which relates the long-term aridity index to the long-term evaporative index, curvilinear relationships between these indices (i.e., parametric Budyko curves) allow for the quantification of precipitation partitioning under prevailing climatic conditions. A movement along a Budyko curve with changes in the climatic conditions has been used as a predictor for catchment behaviour under change. However, various studies have reported deviations around these curves, which raises questions about the usefulness of the method for future predictions. To investigate whether parametric Budyko curves still have predictive power, we quantified the global, regional, and local evolution of deviations of catchments from their parametric Budyko curves over multiple subsequent 20-year periods throughout the last century, based on historical long-term water balance data from over 2000 river catchments worldwide. This process resulted in up to four 20-year distributions of annual deviations from the long-term mean parametric curve for each catchment. To use these distributions of deviations to predict future deviations, the temporal stability of these four distributions of deviations was evaluated between subsequent periods of time. On average, it was found that the majority of 62 % of study catchments did not significantly deviate from their expected parametric Budyko curves. From the remaining 38 % of catchments that deviated from their expected curves, the long-term magnitude of median deviations remains minor, with 70 % of catchments falling within the range of ±0.025 of the expected evaporative index. Furthermore, a significant majority of catchments, constituting around the same percentage, were found to have stable distributions of deviations across multiple time periods, making them well-suited to statistically predict future deviations with high predictive power. These findings suggest that while trajectories of change in catchments do not strictly follow the expected long-term mean parametric Budyko curves, the deviations are minor and quantifiable. Consequently, taking into account these deviations, the parametric formulations of the Budyko Framework remain a valuable tool for predicting future evaporation and runoff under changing climatic conditions, within quantifiable margins of error.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

1 Introduction

Climate change is likely to have a profound impact on future global water resources (Jaramillo et al., 2018; Xing et al., 2018) by causing major shifts in the water balance of river basins world-wide (Serpa et al., 2015; Hattermann et al., 2017). Robust quantitative estimates of future water resources are therefore required to develop policies and to design engineering interventions that will allow the mitigation of the potentially adverse effects of these shifts on water supply (Destouni et al., 2012).

From the early 20th century onwards, multiple authors have suggested analytical, functionally similar non-parametric, curvilinear relationships that describe the long-term average partitioning of precipitation into runoff and evaporative fluxes in terrestrial hydrological systems (Schreiber, 1904; Oldekop, 1911; Budyko, 1948). In spite of differences in their detailed mathematical formulation (Arora, 2002; Andréassian et al., 2016), all these relationships allow to map the long-term mean fraction of precipitation P that is evaporated, i.e., the evaporative index $I_E = E_A/P$, onto the long-term mean ratio of energy input, expressed as potential evaporation E_P, over precipitation, referred to as aridity index $I_A = E_P/P$. Many studies have demonstrated that empirical evaporative indices I_E of river catchments world-wide indeed scatter rather narrowly around these non-parametric Budyko curves (Turc, 1954; Budyko, 1961; Choudhury, 1999; Zhang et al., 2001; Donohue et al., 2007; Berghuijs et al., 2014; van der Velde et al., 2014; Andréassian et al., 2016; Jaramillo et al., 2018; Reaver et al., 2022). To better account for the scatter, several parametric reformulations of the non-parametric Budyko curves have been proposed (Turc, 1954; Mezentsev, 1955; Tixeront, 1964; Fu, 1981). These one-parameter formulations were shown to be functionally almost equivalent to each other (Yang et al., 2008). Their parameter, hereafter referred to as ω, defines catchment specific parametric Budyko curves that locate each catchment on a uniquely defined position in the space spanned by I_A and I_E i.e., the Budyko Framework. The ω parameter is widely interpreted to encapsulate all combined properties of a catchment that may influence the storage and release of water other than I_A (Milly, 1994; Donohue et al., 2012; Shao et al., 2012).

The fact that the long-term water balance exhibits such a relatively consistent behaviour described across a wide spectrum of hydro-climatically and physio-graphically distinct environments has led to the hypothesis that the general shape of Budyko curves emerges for natural systems in a co-evolution of climate, soil water storage and vegetation properties (Milly, 1994; Porporato et al., 2004; Donohue et al., 2012; Gentine et al., 2012; Troch et al., 2013). Consequently, it may plausibly be assumed that once equilibrium is reached after a change in I_A, the water partitioning in a catchment converges towards a new but predictable stable state (here: I_E), by following its catchment specific parametric Budyko curve defined by ω. By extension, such a space–time symmetry under a changing climate may then allow estimates of future I_E, and thus of E_A and Q, based on changes in I_A, inferred from future projections of P and E_P (Roderick and Farquhar, 2011; Wang et al., 2016; Liu et al., 2020; Bouaziz et al., 2022).

However, parametric Budyko curves and their ω parameters were originally not developed from physical reasoning but rather from a largely process-agnostic, mathematical perspective with the aim to statistically describe observed data.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

They, therefore, do not have a clearly defined physical meaning and the interaction of actual processes, that control ω in specific environments, is poorly understood. Consequently, mechanistic evidence that supports the space–time symmetry hypothesis remains erratic. This poses a serious obstacle for the formulation of a general mechanistic description to quantitatively and mechanistically link ω of parametric Budyko curves (and thus I_E) to catchment properties other than I_A (Xu et al., 2013; Padrón et al., 2017). This further entails that estimates of ω and the associated I_E for ungauged catchments or future climate conditions may be subject to major uncertainties and should therefore be interpreted from a probabilistic perspective (Greve et al., 2015).

Recently, it was also argued that catchments should not be necessarily expected to follow their long-term average, catchment specific parametric Budyko curves when subject to climatic perturbations, expressed as changes in I_A (Berghuijs and Woods, 2016; Jaramillo et al., 2018; Jaramillo et al., 2022; Reaver et al., 2022). Such deviations ($\varepsilon_{IE\omega}$) from the expected parametric Budyko curve, were previously referred to as residual or landscape-driven (Destouni et al., 2012; van der Velde et al., 2014). Where $\varepsilon_{IE\omega}$ is defined as the difference between the observed evaporative index ($I_{E,o}$) and the predicted evaporative index (I_E) derived from the expected parametric Budyko curve. As a consequence, Reaver et al. (2022) have warned that parametric Budyko curves may have no predictive power at all. This may be a too pessimistic perspective. First, the average magnitudes of $\varepsilon_{IE\omega}$ so far reported in the studies remain rather low. Second, there is increasing evidence that estimates in water yield are much less sensitive to fluctuations in ω (and thus $\varepsilon_{IE\omega}$) than to changes in precipitation, in particular for humid environments (Roderick and Farquhar, 2011; Berghuijs et al., 2017). In addition, the assumption of steady conditions might not be applicable (Mianabadi et al., 2020) and the presence of uncertainties in the modelling process are inevitable (Westerberg et al., 2011; Nearing et al., 2016). In other words, some level of deviations from the expected parametric Budyko curves realistically needs to be expected as different time periods will never be characterized by exactly the same environmental conditions and as well the mechanistic processes that control ω are not well understood.

Although part of several previous analyses (Destouni et al., 2012; van der Velde et al., 2014; Berghuijs and Woods, 2016; Jaramillo et al., 2022; Reaver et al., 2022), to our knowledge, there has been no systematic, in-depth analysis of the distributions of $\varepsilon_{IE\omega}$ or their evolution over multiple time periods at global, regional, and local scales explicitly reported in the literature. Jaramillo and Destouni (2014) and Jaramillo et al. (2018) provide estimates of average $\varepsilon_{IE\omega}$ for several regions but limited their analyses to two independent time periods. In contrast, Reaver et al. (2022) quantified $\varepsilon_{IE\omega}$ over multiple time periods but explicitly reported only estimates of $\varepsilon_{IE\omega}$ maxima for individual catchments, thus describing merely the most extreme situations.

Our research question is whether the distributions and average magnitudes of $\varepsilon_{IE\omega}$ remain stable and thus probabilistically predictable over time under changing environmental conditions in space and time. A positive answer to this question would imply that parametric Budyko curves can indeed be, at least over time scales of several decades, considered useful for predicting future I_E under changing conditions within quantifiable margins of error. Based on historical long-term water balance data from > 2000 river catchments worldwide, we therefore here quantify the distributions of deviations of catchments from parametric Budyko curves, i.e., $\varepsilon_{IE\omega}$, at global, regional, and local scales between multiple 20-year periods

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Hydrology and Earth System Sciences
Discussions

throughout the 20th century. Specifically, we test the hypothesis that the distributions of $\varepsilon_{IE\omega}$ are too wide and temporally too unstable so that I_E from parametric Budyko curves needs to be considered practically unpredictable with the available data.

2 Datasets and methods

2.1 Meteorological and hydrological data

Daily precipitation P [mm d^{-1}] as well as maximum and minimum temperature T [°C] data at a spatial resolution of 0.5° x 0.5° were obtained from the Global Soil Wetness Project Phase-3 (GSWP-3); (Dirmeyer et al., 2006) and spatially averaged for each study catchment over the time period 1901 – 2015.

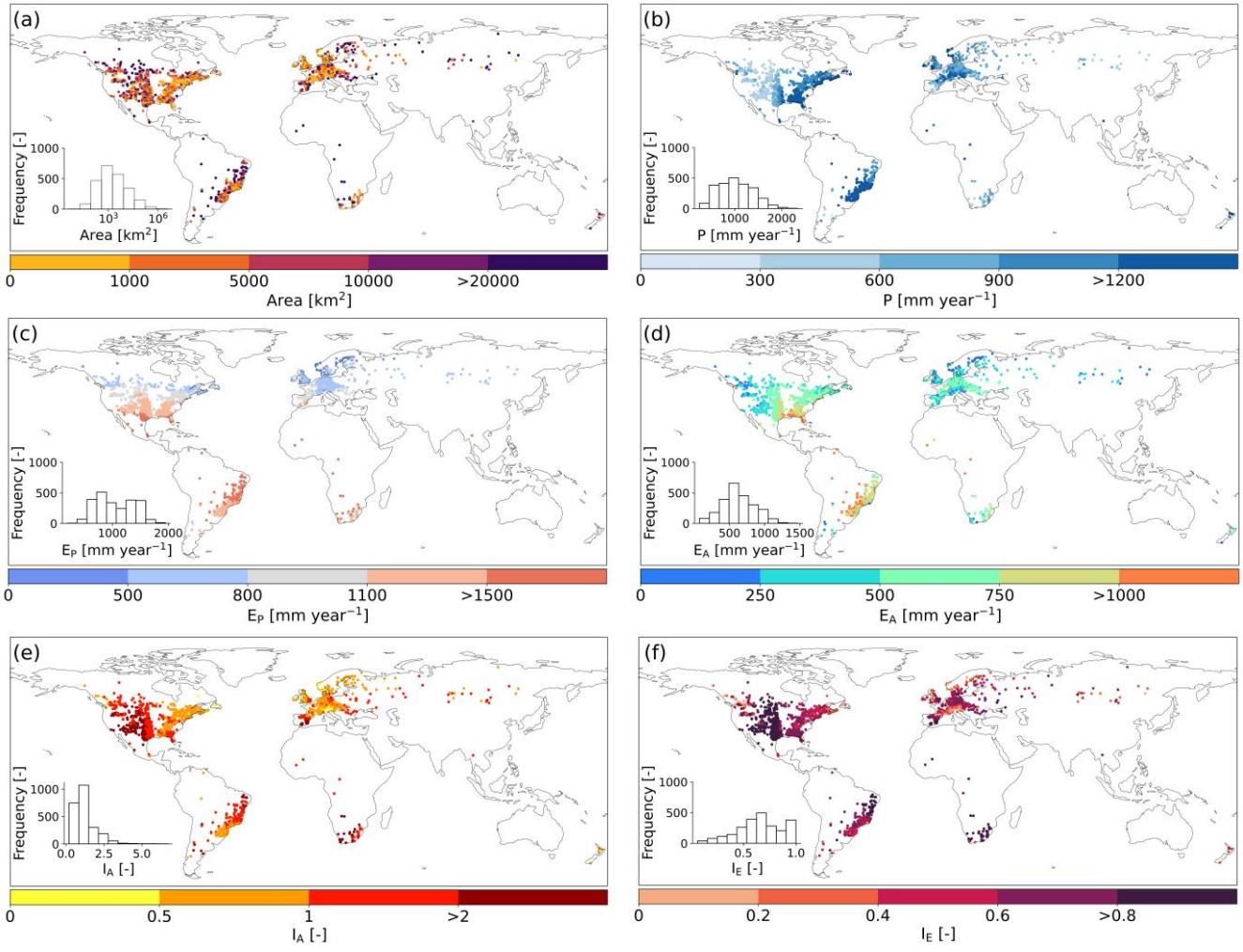


Figure 1: Spatial distribution of 2387 studied catchments along with topographic characteristics and long-term mean (1901-2015) climatic indices: a) Catchment area, b) Precipitation *P*, c) Potential evaporation E_P, d) Actual evaporation $E_A = P-Q$, e) Aridity Index I_A , f) Evaporative Index I_E

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Hydrology and Earth System Sciences
Discussions

Potential evaporation E_P [mm d^{-1}] was estimated based on the method proposed by Hargreaves and Samani (1982):

$$\lambda_E E_P = \alpha R_a (T_a + 17.8)\sqrt{(T_{\max} - T_{\min})} \quad (1)$$

Where λ_E is the latent heat of vaporization with a value of 2.45 [MJ kg^{-1}], constant $\alpha \sim 0.0023$ to convert MJ kg^{-1} to mm/day, R_a is the top atmosphere radiation [MJ m^{-2} day^{-1}] and T_a, $T_{\max}$ and $T_{\min}$ are the daily average, maximum and minimum temperatures [°C], respectively. R_a is estimated by using the method proposed by Duffie and Beckman (1980).

In this study, we obtained annual river flow data from the Global Streamflow Indices and Metadata (GSIM) archive (Do et al., 2018; Gudmundsson et al., 2018) which consists of in situ streamflow observations data for over 30000 gauging stations worldwide. We selected stations with runoff data spanning at least 50 years in the 1901–2015 period, excluding those with a data quality flag marked as 'Caution'. After filtering, we retained 2387 river catchments with data series ranging from 50 years to 100 years (median: 78 years). These catchments vary in size, from 4 to 3,475,000 km^2 (median ~1564 km^2; Fig. 1a). The catchments represent diverse hydro-climatic conditions (Fig. 1b-f), as indicated by the long-term average aridity indices (I_A) that range from 0.19 to 6.66 (median: 0.97; Fig. 1e) and evaporative indices (I_E) that range from 0.06 to 0.99 (median: 0.65; Fig. 1f).

2.2 Methods

The subsequent experiment to estimate for each of the 2387 study catchments the deviations $\varepsilon_{IE\omega}$ from its expected evaporative indices I_E over multiple subsequent time periods is based on the parametric Tixeront-Fu reformulation of the Budyko hypotheses (Tixeront, 1964; Fu, 1981), as illustrated in Fig. 2:

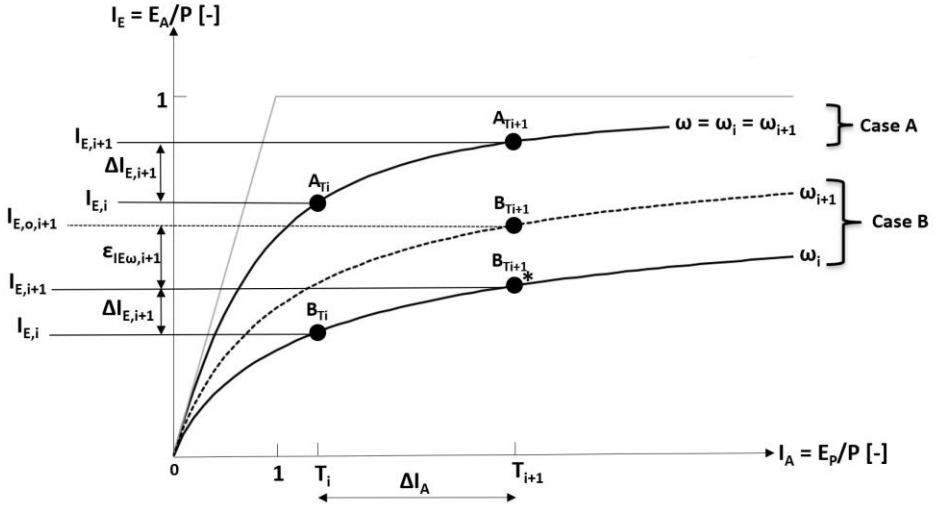


Figure 2: A schematic representation of a catchment movement in Budyko space between two long-term time periods T_i and T_{i+1}. Case A: Catchment A moves along the same Budyko curve from the first period T_i to the next period T_{i+1} i.e., $\omega_i = \omega_{i+1}$. Case B: Catchment B has deviated from its expected parametric Budyko curve i.e., $\omega_i \neq \omega_{i+1}$

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Hydrology and Earth System Sciences
Discussions

This movement in Budyko space is governed by the following equation:

$$I_E = \frac{E_A}{P} = 1 + \frac{E_P}{P} - \left[1 + \left(\frac{E_P}{P}\right)^{\omega}\right]^{\frac{1}{\omega}} \quad (2)$$

where ω is a catchment-specific parameter estimated from long-term averages of observed P, E_P and $E_A = P - Q$, assuming negligible change in storage dS/dt.

Equation (2) suggests that with a given ω, hydro-climatic shifts between two periods T_i and T_{i+1}, expressed as changes in aridity index $\Delta I_A = \Delta(E_P/P)$ will lead to predictable changes $\Delta I_{E,i+1}$ (Case A). In other words, catchments will follow their specific curves in period T_{i+1}, defined by parameter $\omega = \omega_i = \omega_{i+1}$ to an expected new $I_{E,i+1}$, which is expressed as:

$$I_{E,i+1} = I_{E,i} + \Delta I_{E,i+1} \quad \text{(Case A in Fig. 2)} \quad (3)$$

However, in reality, as described above, ω is often not constant over time (Case B). Catchments therefore do not strictly follow their $I_{E,i}$ curve defined by ω_i (from T_i) in a subsequent period T_{i+1}. For period T_{i+1}, this therefore leads to additional deviation $\varepsilon_{IE\omega,i+1}$ which is described as:

$$\varepsilon_{IE\omega,i+1} = I_{E,o,i+1} - I_{E,i+1} \neq 0 \quad (4)$$

representing the difference between the actually observed $I_{E,o,i+1}$ from the expected $I_{E,i+1}$. Thus, for period T_{i+1}, the observed $I_{E,o,i+1}$ depends on the combination of the predicted change and these deviations, i.e.,

$$I_{E,o,i+1} = I_{E,i} + \Delta I_{E,i+1} + \varepsilon_{IE\omega,i+1} \quad \text{(Case B in Fig. 2)} \quad (5)$$

Here, we have sub-divided the available data records of each catchment into up to five individual 20-year periods T_i over the last century (Table 1). We assume that these 20 years periods are long enough to satisfy dS/dt~0. This is supported by Han et al. (2020), who showed that in >80 % of catchments world-wide dS/dt< 5 % for 20 year periods.

Table 1: Symbols used in this study to present 20-year periods, changes between subsequent 20-year periods and distributions of deviations

Time period	Symbols		
	20-year periods	Change between subsequent 20-year periods	Distribution of deviations
1901-1920	T_1	$\Delta_{1\text{-}2}$	$\varepsilon_{IE\Delta1}$
1921-1940	T_2	$\Delta_{2\text{-}3}$	$\varepsilon_{IE\Delta2}$
1941-1960	T_3	$\Delta_{3\text{-}4}$	$\varepsilon_{IE\Delta3}$
1961-1980	T_4	$\Delta_{4\text{-}5}$	$\varepsilon_{IE\Delta4}$
1981-2000	T_5		

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

The experiment to estimate deviations $\varepsilon_{IE\omega}$ between the five individual periods T_1–T_5 for the study catchments was then carried out in a systematic sequence of 5 specific steps as illustrated in Fig. 3 and described in the following:

Step 1: Estimation of catchment-specific $I_{E,i}$ curves and the distribution of annual $I_{E,o}$ around it for each period T_i

For each catchment and each individual 20-year time period T_i, the catchment-specific parametric Budyko curve $I_{E,i}$ defined by parameter ω_i is obtained by fitting Eq.(2) to the mean annual positions of each catchment in the Budyko space, as computed from the observed water balance data. The decision to obtain the ω_i for each 20-year period by fitting Eq.(2) to the set of $n = 20$ corresponding observed annual $I_{E,o}$ values instead of directly to their 20-year averages was a deliberate choice. The fluctuations of the $n = 20$ annual $I_{E,o}$ values explicitly represent annual storage changes dS/dt between individual years. This subsequently allowed to treat the observed annual $I_{E,o}$ probabilistically as distributions around their expected values for that period T_i as defined by $I_{E,i}$ curve (Fig. 3a).

Step 2: Distributions of annual deviations $\varepsilon_{IE\Delta j}$ from expected $I_{E,i+1}$ between subsequent time periods

For each catchment we then used ω_i from each time period T_i to compute the expected $I_{E,i+1}$ for the subsequent period T_{i+1} (i.e. point B_{Ti+1}* in Fig. 2). This then allowed to estimate the individual deviations of the 20 annual observed $I_{E,o}$ values from the expected $I_{E,i+1}$ curve. For each pair of time periods T_i–T_{i+1} (i.e. T_1–T_2, T_2–T_3, etc., hereafter referred to as $\Delta_{1\text{-}2}$, $\Delta_{2\text{-}3}$, etc.) this resulted in an individual distribution of annual deviations $\varepsilon_{IE\Delta j}$ around a 20-year average in each catchment (Fig. 3b). Note, that catchments with data for all five time periods T_1–T_5, have the maximum of $j = 4$ distributions $\varepsilon_{IE\Delta j}$. In contrast, catchments with data for only two periods, e.g., T_2 and T_3, feature only $j = 1$ distribution of between-period deviations $\varepsilon_{IE\Delta j}$.

Non-parametric Wilcoxon Signed Rank Tests were then used to test for each distribution $\varepsilon_{IE\Delta j}$ the null hypothesis that the median deviation is not significantly different from zero. The lower the p-value, the higher the probability that the median deviation of $\varepsilon_{IE\Delta j}$ of observed $I_{E,o,i+1}$ from expected $I_{E,i+1}$ is higher than zero for the comparison of $\varepsilon_{IE\Delta j}$ between periods T_i and T_{i+1}.

Step 3: Fit parametric distributions to the empirical distributions of annual deviations $\varepsilon_{IE\Delta j}$

For each catchment we have then fitted Skew Normal Distributions to each of the $j = 1$–4 empirical distributions of deviations $\varepsilon_{IE\Delta j}$ (Fig. 3c). The probability density function (PDF) of the skew normal distribution is given by:

$$f(x) = \frac{2}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi(\alpha)\left(\frac{x-\xi}{\omega}\right) \quad (6)$$

Where ϕ is the standard normal PDF, Φ is the standard normal cumulative distribution function, ω is a scale parameter, ξ location parameter and α is a shape parameter.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

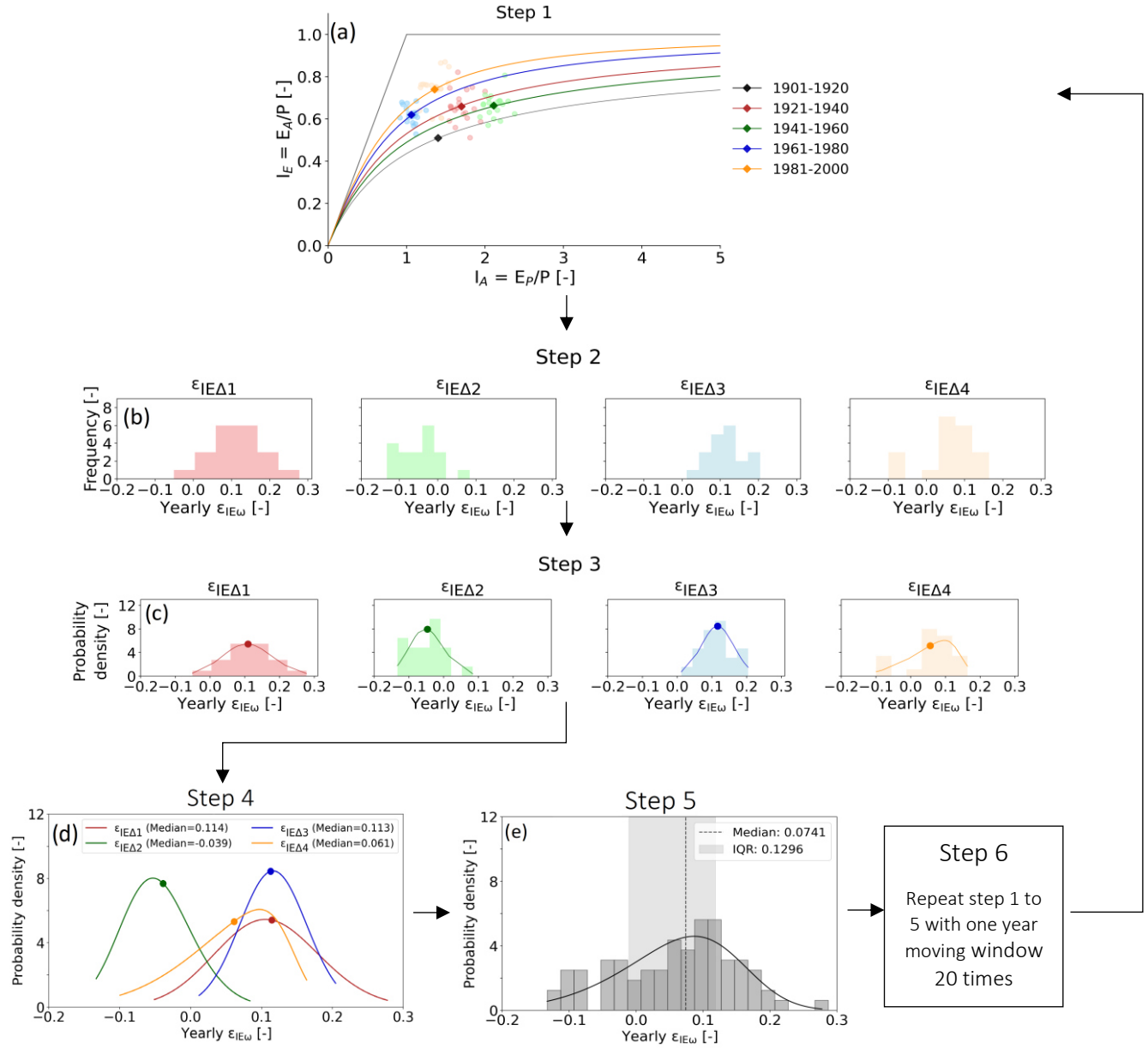


Figure 3: Flow chart of methodology. Step 1: Estimation of catchment-specific $I_{E,i}$ curves and the distribution of annual $I_{E,o}$ around it for each period T_i. Step 2: Distributions of annual deviations $\varepsilon_{IE\Delta j}$ from expected $I_{E,i+1}$ between subsequent time periods Step 3: Fit parametric distributions to the empirical distributions of annual deviations $\varepsilon_{IE\Delta j}$. Step 4: Evaluate temporal stability of the distributions $\varepsilon_{IE\Delta j}$ in subsequent pairs of time periods Step 5: Aggregated long-term marginal distribution of annual deviations $\varepsilon_{IE\omega}$ from expected I_E for each catchment. Step 6: Evaluation of the sensitivity of the marginal distributions of annual deviations $\varepsilon_{IE\omega}$ to the choice of 20-year averaging window (the above generated distributions of $\varepsilon_{IE\omega}$ are for illustration purpose and not based on real data)

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

The mean of the distribution is computed as:

$$\overline{f(x)} = \xi + \left(\sqrt{2/\pi}\right)\frac{\omega\alpha}{\sqrt{1+\alpha^2}} \quad (7)$$

and the standard deviation is represented by:

$$\sigma = \omega\sqrt{\left(1 - \frac{2\alpha^2}{(1+\alpha^2)\pi}\right)} \quad (8)$$

Step 4: Evaluate temporal stability of the distributions $\varepsilon_{IE\Delta j}$ in subsequent pairs of time periods

For distributions of past deviations to be used to estimate deviations $\varepsilon_{IE\omega}$ under projected hydro-climatic future conditions, it is necessary to upfront evaluate whether it is plausible to assume that they retain sufficient explanatory power under future conditions or if there is evidence against that. This was here done by analysing how stable the individual distributions in a catchment are over time.

To do so, we have followed three steps and for each catchment the up to j = 4 distributions of deviations $\varepsilon_{IE\Delta j}$ from expected $I_{E,i+1}$ between subsequent time periods were compared and analysed for their changes over time (Fig. 3, Sub-steps 4.1-4.3).

Sub-step 4.1

At first, non-parametric Kolmogorov-Smirnov Tests with a significance level of 5 % were used on consecutive pairs of distributions, i.e. $\varepsilon_{IE\Delta j}$ and $\varepsilon_{IE\Delta j+1}$ to test the null hypothesis that the distributions are not significantly different from each other. In case the null hypothesis is not rejected ($p > 0.05$), hereafter referred to with symbol "o", we consider the catchment is stable over time and in such a case the past distributions may be directly used to estimate $\varepsilon_{IE\omega}$ under future conditions with some confidence.

Sub step 4.2

If for a catchment significant differences between consecutive pairs of distributions ($p \leq 0.05$) were found, it was further analysed whether the differences can be considered arbitrarily variable or whether there is indicative evidence for the potential presence of alternating fluctuations or systematic shifts over time. Thus, in a second step, we have checked if the median of $\varepsilon_{IE\Delta j}$ systematically decreased ("–") or increased ("+") over time. If the difference between three or more of the j distribution medians were characterized by the same sign, i.e., "–" or "+", this may be evidence for a systematic and thus non-variable shift in the median of $\varepsilon_{IE\Delta j}$ over time. In that case, past distributions $\varepsilon_{IE\Delta j}$ need to be assumed to have limited predictive power for estimating future $\varepsilon_{IE\omega}$.

Sub step 4.3

In the alternative case, when less than three distributions showed the same sign, we have in a third step analysed, whether $\varepsilon_{IE\Delta j}$ for Δ_j is influenced by the magnitude of $I_{E,i}$ and that e.g. after a 20-year period with a low $I_{E,i}$, further future

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

decreases and thus negative $\varepsilon_{IE\Delta j}$ are unlikely and $I_{E,i+1}$ will, more probably, swing back to higher values and thus positive $\varepsilon_{IE\Delta j}$. Similar to above, if the median $\varepsilon_{IE\omega}$ systematically decreased ("–") or increased ("+") with $I_{E,i}$ for three or more of the pairs of time periods *j*, this may be evidence for a systematic and thus non-variable shift in the median $\varepsilon_{IE\omega}$ over time, indicating limited predictive power.

Following the above, each catchment was classified into one of four qualitative categories of temporal stability of $\varepsilon_{IE\omega}$ (Table 2). Note, that the use of formal quantitative statistical test was here hindered by the small sample size of a maximum of four pairs of time periods and thus omitted. The temporal stability was ranked as "Stable", if between more than half of the *j* distributions in a catchment no significant differences in median $\varepsilon_{IE\omega}$ was found, e.g. "o o o o" or "– o o o". A catchment was ranked as "Variable", if it showed an alternating sequence and thus *no* systematic shift of median $\varepsilon_{IE\omega}$ over time e.g. "+ – + –" or "– + + – " *and no* relation between $I_{E,i}$ and median $\varepsilon_{IE\omega}$. In contrast, if a catchment was characterized by an alternating sequence and a dependency between $I_{E,i}$ and median $\varepsilon_{IE\omega}$, it was tagged as "Alternating". If, finally, between three or more of the j consecutive distributions in a catchment the median $\varepsilon_{IE\omega}$ was found to increase or decrease, e.g., "– + + + " or " – – – –", this may indicate the presence of a systematic shift over time and the temporal stability of deviations from expected I_E was tagged as "Shift".

Table 2: Decision criteria to classify the time stability of the j distributions $\varepsilon_{IE\Delta j}$ for each catchment into one of the four qualitative categories "Stable", "Variable", "Alternating", "Shift" and the associated predictive power of the marginal distribution of $\varepsilon_{IE\omega}$ of a catchment, aggregating all j distributions of that catchment

Tag	Kolmogorov-Smirnov Test	Systematic shift of median	Relation between $I_{E,i}$ and $\varepsilon_{IE\omega}$	Examples	Predictive power	No. of catchments
Stable	$p > 0.05$	No	No	"o o o o" or "– o o o"	High	1651
Variable	$p \leq 0.05$	No	No	"+ – + –" or "– + + – "	Moderate	455
Alternating	$p \leq 0.05$	No	Yes	"+ – + –" or "– + + – "	Low	179
Shift	$p \leq 0.05$	yes	-	" – – – –" or "– + + + "	Low	102

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

Step 5: Aggregated long-term marginal distribution of annual deviations $\varepsilon_{IE\omega}$ from expected I_E for each catchment

In this step the up to j = 4 distributions $\varepsilon_{IE\Delta j}$ were aggregated into one marginal distribution of $\varepsilon_{IE\omega}$ for each catchment (Fig. 3e). This distribution reflects the historical range of fluctuations in annual $\varepsilon_{IE\omega}$ based on all available information for each catchment. Consequently, the median $\varepsilon_{IE\omega}$ of the distribution in each catchment represents a measure of uncertainty around expected future I_E based on current estimates of ω for each catchment, thereby making I_E statistically predictable.

To account for the potential effect of systematic shifts in distributions $\varepsilon_{IE\Delta j}$ (Step 4) on the predictive power of the associated marginal distribution of deviations $\varepsilon_{IE\omega}$, we have tagged the marginal distribution of each catchment with a qualitative robustness flag as defined in Step 4. "Stable" distributions are characterized by the highest predictive power, distributions with "Variable" fluctuations can be expected to have moderate predictive power, while distributions tagged as "Alternating" or "Shift" do in the absence of more detailed data have rather low predictive power (Table 2).

Step 6: Evaluate the sensitivity of the marginal distributions of annual deviations $\varepsilon_{IE\omega}$ to the choice of 20-year averaging window

To further quantify the sensitivity of the above aggregated, i.e. marginal distributions to the choice of the individual 20-year averaging time periods, we have, in a last step, repeated the above Steps 1–5 twenty times to test all possible sequences of 20-year periods. More specifically, in a moving window analysis we first shifted each time period T_1–T_5 by one year, i.e. 1902–1921 (T_1), 1922–1941 (T_2), etc. and repeated the above Steps 1–5. Subsequently we shifted T_1–T_5 by another year to 1903–1922 (T_1), 1923–1942 (T_2), etc. and again repeated Steps 1–5. This was done twenty times until all years of the first period, i.e. 1901–1920, were the starting years of T_2.

3 Results

3.1 Changes in hydro-climatic variables and movement in Budyko space (Step 1)

Throughout the 100-year study period and across all study catchments, considerable hydro-climatic variability was observed, with some variables exhibiting trend-like behaviour over time and others more cyclic behaviour. Overall, mean annual precipitation over the individual 20-year periods systematically increased by ~ 18.4 mm century^{-1}, on average (Fig. 4a), with 57 % of the catchments showing an increase between T_1 and T_2 ($\Delta_{1\text{-}2}$) and 83 % for $\Delta_{4\text{-}5}$. In contrast, mean annual temperatures (Fig. 4b) and the associated potential evaporation (Fig. 4c) were characterized by a more fluctuating pattern. Combined this led to slightly more arid conditions in the first half of the 20th century, followed by considerable reduction of aridity index I_A and thus to a shift towards somewhat more humid conditions towards the end of the century (Fig. 4e), in which 78 % and 75 % of the catchments showed decreases in I_A for $\Delta_{3\text{-}4}$ and $\Delta_{4\text{-}5}$, respectively. The changes in I_A were accompanied by related changes in potential evaporation E_P and precipitation (Fig. 4c,a). Furthermore, we observed various

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Hydrology and Earth System Sciences
Discussions

hydroclimatic conditions over time periods along with indications that there are deviations $\varepsilon_{IE\omega}$ from the expected I_E, as elaborated in detail in Fig. S1.

It is worth mentioning here that the sample sizes vary between individual 20-year periods of comparison due to the length of the data availability. Therefore, to distinguish whether the climatic variability in Fig. 4 is associated to the hydroclimatic variables or due to the change in sample size, the same plot for the catchments that are present in all periods of comparison (n = 142) is provided in Fig. S2. It is found that the overall pattern of temporal variability in that sub-sample largely reflects that of the full sample.

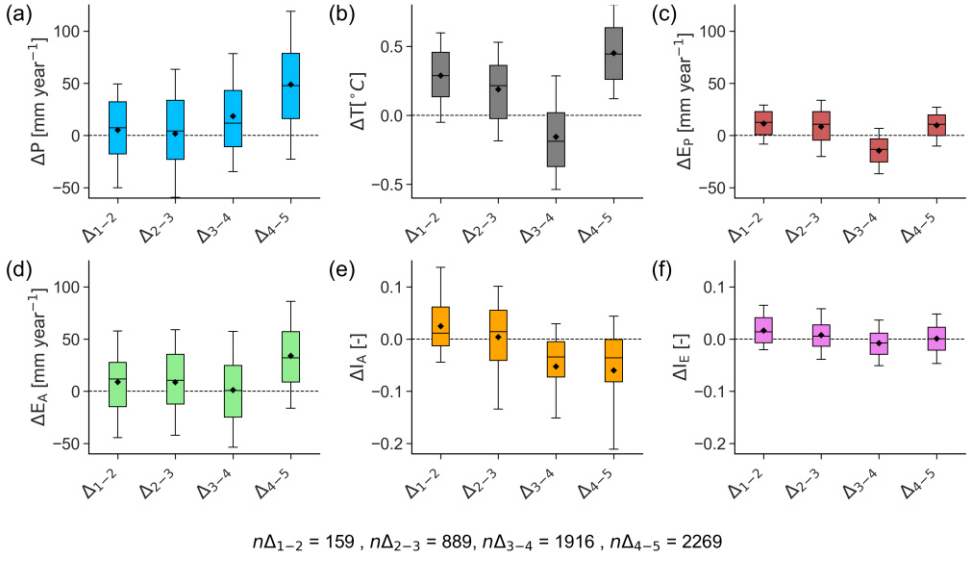


Figure 4: Mean 20-year changes in hydro-climatic variables for the studied catchments between two consecutive periods. a) Precipitation *P*, b) Temperature *T*, c) Potential evaporation E_P, d) Actual evaporation E_A, e) Aridity index I_A, and f) Evaporative index I_E. The boxes represent the 25th to 75th quantiles, while whiskers extend to the 10th and 90th quantiles. Diamonds denote the arithmetic mean. 'n' indicates the number of catchments included in the comparison between two respective periods. Outliers have been removed for improved visualization

3.2 Distributions of annual deviations $\varepsilon_{IE\Delta j}$ from parametric Budyko curve (Steps 2 & 3)

The indicative evidence for presence of deviations $\varepsilon_{IE\omega} \neq 0$ (Fig. S1) in at least some catchments is further supported by a more detailed analysis of the distributions of annual deviations $\varepsilon_{IE\Delta j}$ between the pairs of subsequent 20-year periods for each of the 2387 study catchments. The results of the Wilcoxon Signed Rank Tests indeed suggest that it is likely that the median $\varepsilon_{IE\omega} \neq 0$ for a significant proportion of catchments. For example, at a 95 % confidence level (i.e., $p \leq 0.05$), 34–42 % of the distributions can be considered to feature deviations with a median $\varepsilon_{IE\omega} \neq 0$ (Fig. 5a). However, this conversely also entails that for a majority of 58–66 % of the distributions there is less evidence (i.e., $p > 0.05$) that the median deviations are different form zero. Overall, this is consistent with results from previous studies and shapes a picture in which catchments do not strictly and necessarily follow their expected parametric I_E curves, but that the deviations thereof remain close to zero or very limited for many catchments.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

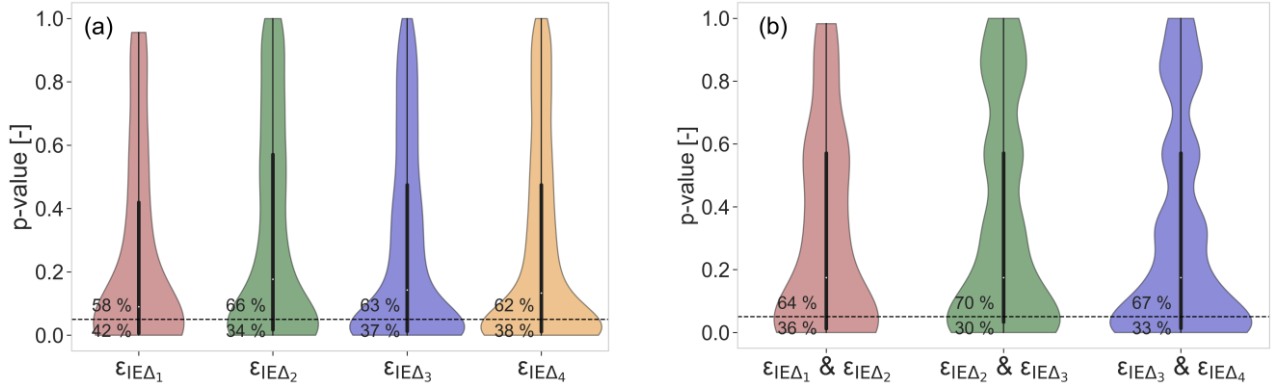


Figure 5: Distribution of p-values from statistical tests a) Wilcoxon Signed Rank Test performed to know whether the long-term median $\epsilon_{IE\omega}$ of the individual 20-year distribution $\varepsilon_{IE\Delta j}$ is significantly differs from zero or not, b) Kolmogorov-Smirnov Test performed on two consecutive distributions ($\varepsilon_{IE\Delta j}$ and $\varepsilon_{IE\Delta j+1}$) to know whether the two distributions of $\epsilon_{IE\omega}$ are significantly different to each other or not. Each violin plot displays the distribution of p-values, with the central black box representing the interquartile range (25th to 75th quantiles) and whiskers extending to the smallest and largest data points within 1.5 times the interquartile range. The dotted black line represents the significance level of 0.05. The percentages above and below this line indicate the proportion of data points with p-values above or below the significance level of 0.05, respectively

A characteristic selected example for the latter is the sequence of the four distributions of annual deviations in the Chemung River at Chemung (New York; 6455 km^2; ID US_0000832) across the four pairs of subsequent 20-year periods over the 20th century (Fig. 6a-c). The annual deviations $\varepsilon_{IE\omega}$ with an interquartile range of IQR ~ 0.062 (Fig. 6c) concentrate quite narrowly around the medians. The medians themselves range between merely $\varepsilon_{IE\omega}$ = -0.006–0.012 (Fig. 6b). The associated p-values from the Wilcoxon Signed Rank Test (p= 0.27–0.84) further suggest that there is only limited evidence that the median deviations $\varepsilon_{IE\omega}$ of distribution $\varepsilon_{IE\Delta j}$ are different from zero. In spite of a somewhat wider spread with an IQR ~ 0.094 (Fig. 6f), a similar pattern with consistently low median $\varepsilon_{IE\omega}$= -0.017–0.018 (Fig. 6e) (p = 0.09–0.47) was observed in the second selected example, the Lee River (Ireland; 1019 km^2; ID GB_0000078).

In contrast, more variable patterns were found for other catchments (Fig. 6g-o). For example, in the Sava River at Radece (Slovenia; 6004 km^2; ID SI_0000007) the four distributions of the annual deviations all display a wider spread, with IQR ~ 0.113, indicating a higher degree of storage fluctuation between individual years (Fig. 6i). In addition, the medians do considerably deviate from zero, as indicated by median $\varepsilon_{IE\omega}$ ranging between -0.023 and 0.118 (Fig. 6h).

The set of 20-year average I_E values and the associated parameters of the fitted parametric distributions of deviations for each of the time periods in the individual study catchments are provided in the Supplementary data downloadable from Zenodo (10.5281/zenodo.10925965).

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

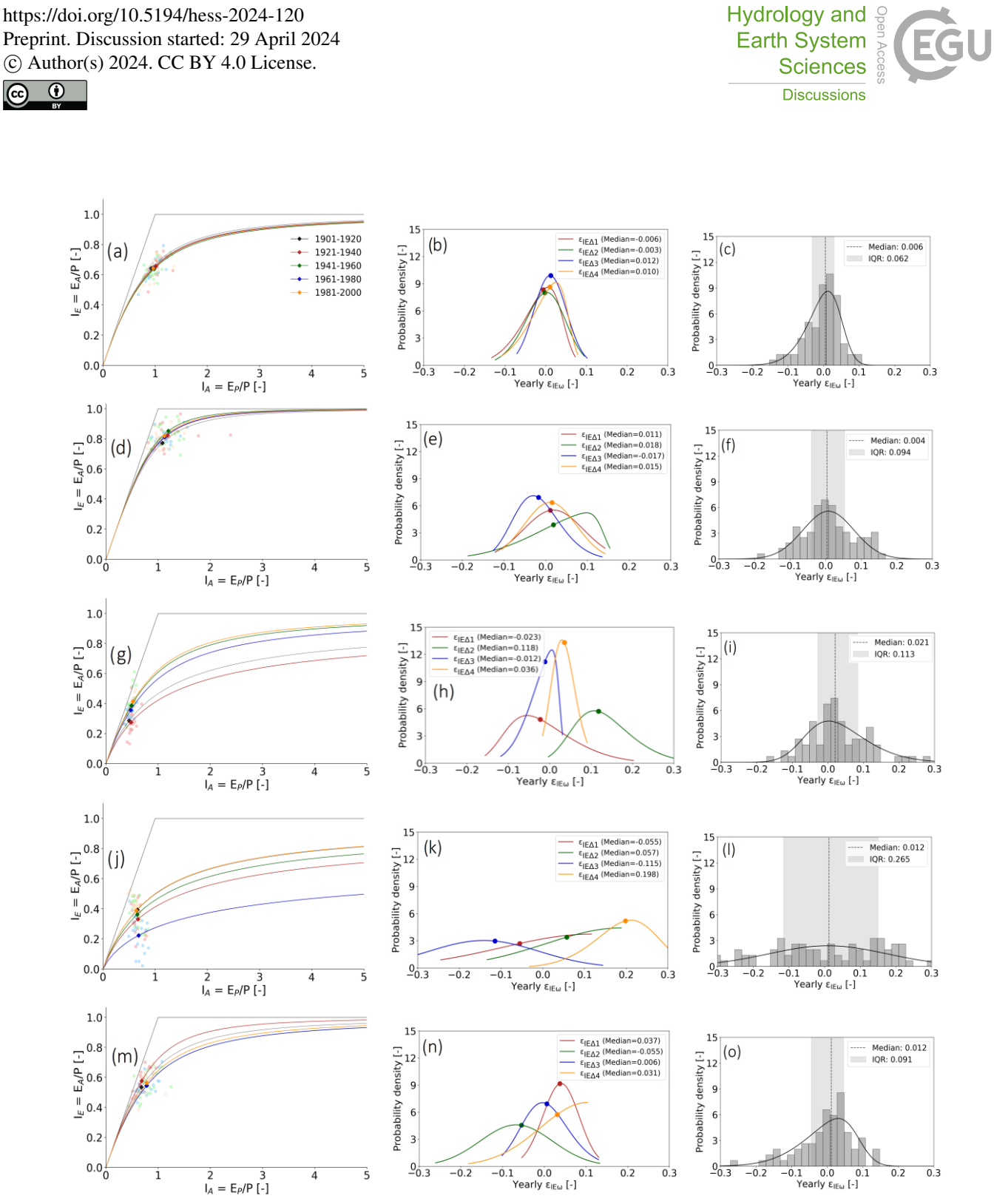


Figure 6: Mean annual position of catchments (light colour dots) in Budyko space along with long-term mean (dark colour dots) and predictive parametric Budyko curves (left column). Individual distribution of deviations ($\varepsilon_{IE\Delta1}$, $\varepsilon_{IE\Delta2}$, $\varepsilon_{IE\Delta3}$ and $\varepsilon_{IE\Delta4}$) with long term median deviation $\varepsilon_{IE\omega}$ values (middle column) and long-term marginal distribution of annual deviations along with long-term median values of $\varepsilon_{IE\omega}$ and IQR of $\varepsilon_{IE\omega}$ values (right column) for five example catchment: Chemung River (a-c), Lee River (d-f), Sava River (g-i), Kaituna River (j-l) and Zschopau River (m-o)

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

3.3 Temporal stability of the distributions $\varepsilon_{IE\Delta j}$ (Step 4)

Based on the Kolmogorov-Smirnov Tests, it was found that overall 68 % of the distributions $\varepsilon_{IE\Delta j}$ between consecutive pairs of time periods are not significantly different from each other ($p > 0.05$; Fig. 5b). Following the criteria defined in Sect. 2.2, this resulted in 1651 catchments classified as "Stable" (Table 2; Fig. 7a). For these catchments, together corresponding to ~ 70 % of all 2387 study catchments, their respective marginal distribution of $\varepsilon_{IE\omega}$ can thus be plausibly considered to have rather high predictive power. Example cases are the Chemung and Lee rivers (Fig. 6a-f), which are characterized by sequences "o o o o" and "o o o o", respectively.

Similarly, 455 additional catchment (19 % of all study catchments), whose distributions exhibited fluctuations over time (Kolmogorov-Smirnov Test $p \leq 0.05$), but which featured only limited evidence for both, the presence of systematic shifts over time as well as for a dependency between $I_{E,i}$ and $\varepsilon_{IE\omega}$, were tagged as "Variable". An example of such a case is shown in Fig. 6g-i for the Sava River at Radece (Slovenia; 6004 km^2; ID SI_0000007). It can be seen that the distributions $\varepsilon_{IE\Delta j}$ between the four pairs of time periods vary considerable with medians ranging from -0.023 and 0.118. However, the fluctuations appear to occur alternatively (" - + - +") and do suggest neither the presence of a systematic shift over time (Fig. S3b) nor a dependency between $I_{E,i}$ and median $\varepsilon_{IE\omega}$ (Fig. S3a). Despite these fluctuations, the marginal distribution of $\varepsilon_{IE\omega}$ aggregated from the 4 individual distributions $\varepsilon_{IE\Delta j}$ (Fig. 6i) can, although wider than for catchments tagged as "Stable", thus be assumed to be an at least moderately robust representation of $\varepsilon_{IE\omega}$.

In contrast, 7 % of the catchments were tagged as "Alternating" and a dependency between $I_{E,i}$ and $\varepsilon_{IE\omega}$ could not be ruled out. A characteristic example for this type of catchments is the Kaituna catchment (New Zealand; 706 km^2, ID NZ_0000003) in Fig. 6j-l. This catchment features major fluctuations with median $\varepsilon_{IE\omega}$ between -0.115 and 0.198. In addition, although no systematic evolution of median $\varepsilon_{IE\omega}$ over time was evident (Fig. S3d), the data suggest the potential presence of a dependency on $I_{E,i}$ as shown in Fig. S3c.

The remaining 102 catchments (4 %) were tagged as "Shift", as they exhibit a rather consistent shift of median $\varepsilon_{IE\omega}$ over time. This can be seen for a selected example in Fig. 6m-o. The median $\varepsilon_{IE\omega}$ in this catchment of the Zschopau River (Germany; 1544 km2; ID DE_0000027) does not only significantly vary between -0.055 and 0.037 but it does so rather systematically into one dominant direction after $\varepsilon_{IE\Delta 1}$ ("+ - ++"; Fig. 6n).

As can be seen in Fig. 7a, the time stability of the study catchments is geographically rather homogenously distributed. Catchments tagged as "Stable" and "Variable" can be found globally, while also no regional concentrations of catchments tagged as "Alternating" and "Shift" could be identified.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

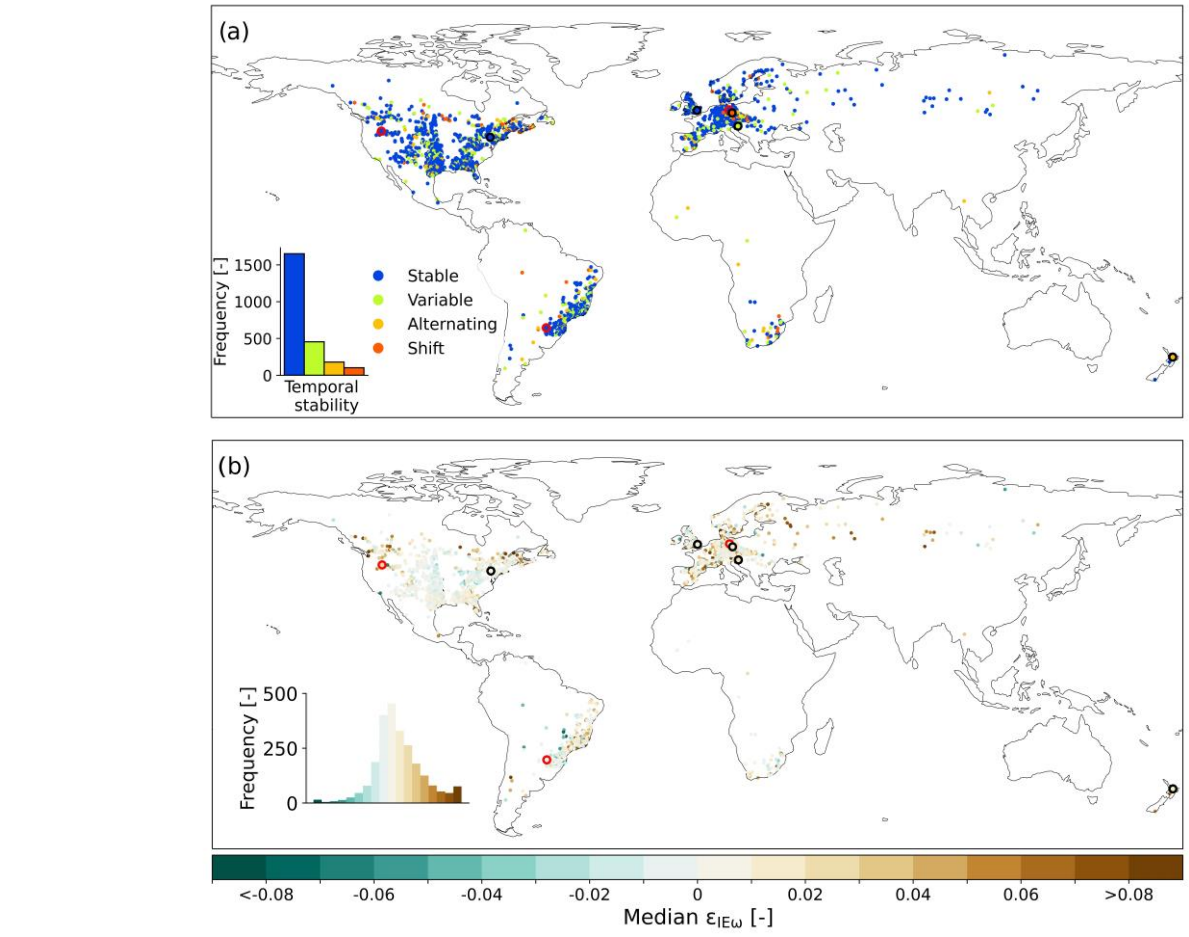


Figure 7: a) Temporal stability and b) long-term median $\varepsilon_{IE\omega}$ values map of aggregated long-term marginal distributions for the study catchments. Catchments highlighted with a black border represent the 5 selected examples from Fig. 6, while those outlined in red denote three additional selected example catchments shown in the supplement (Fig. S4)

3.4 Aggregated long-term marginal distribution of annual deviations $\varepsilon_{IE\omega}$ (Step 5)

By aggregating its j individual distributions, a long-term marginal distribution of $\varepsilon_{IE\omega}$ for each catchment was build. For a large majority of catchments, the long-term median $\varepsilon_{IE\omega}$ remains very close to zero. More specifically, ~50 % of all study catchments are characterized by a median deviation $\varepsilon_{IE\omega}$ that does not exceed ± 0.015 and ~ 70 % by a median within the interval ± 0.025 (Fig. 8a). Depending on the time stability of the *j* individual distributions in a catchment, the spread of annual deviations around these medians showed more variable pattern. Overall, for > 50 % of the study catchments, the IQR of annual deviations remained below 0.08 and for ~ 70 % below 0.10 (Fig. 8b). While catchments with "Stable" distributions exhibit in general a rather narrow spread with an average IQR ~ 0.08, catchments with distributions tagged as "Variable" feature a bit wider spread with average IQR ~ 0.10, while still centring closely around zero. This can also clearly be seen by

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Hydrology and Earth System Sciences
Discussions

the selected examples in Fig. 6. The medians of the marginal distributions of the Chemung and Lee rivers, both tagged as "Stable", are ~0.006 and ~0.004 respectively, with narrow IQRs of 0.062 and 0.094 (Fig. 6c,f). In contrast, while also featuring a marginal distribution with a median deviation $\varepsilon_{IE\omega}$~ 0.021, the Sava River catchment (Fig. 6i), tagged as "Variable", is characterized by a considerably wider scatter of the annual deviations around the median, as evident by the higher IQR of ~0.113. Three additional illustrative examples of well-known river basins are presented in Fig. S4

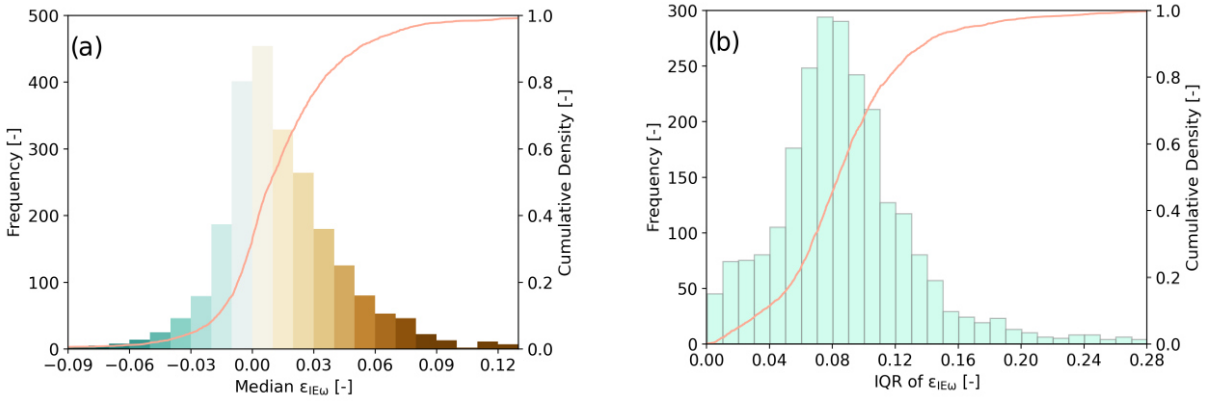


Figure 8: Visualization of long-term a) Median $\epsilon_{IE\omega}$ and b) Interquartile Range (IQR) of $\epsilon_{IE\omega}$ for aggregated long-term marginal distribution of $\epsilon_{IE\omega}$ across all catchments (2387). The light orange solid line represents the Cumulative Distribution Function (CDF)

Overall, it can be observed that median deviations $\varepsilon_{IE\omega}$ close to zero are dominant globally, with no obvious spatial clustering of more pronounced deviations (Fig. 7b). However, it can also be seen that there is some geographic grouping in the direction, i.e. the sign, of the median $\varepsilon_{IE\omega}$. While for many catchments in the central US and southern Brazil median deviations are negative, i.e. $\varepsilon_{IE\omega} < 0$, the rest of the study catchments globally are dominated by $\varepsilon_{IE\omega} > 0$.

For a more regional evaluation we have further aggregated the marginal distributions of $\varepsilon_{IE\omega}$ of the individual catchments into regional marginal distributions, stratified according to the long-term mean aridity index I_A and varied latitude bands (Fig. 9a). The general pattern found across most regions with available data are broadly consistent. 16 out of 20 regions are characterized by median deviations $\varepsilon_{IE\omega}$ that do not exceed ± 0.02. Similarly, no consistent directional pattern in the magnitude of regional median $\varepsilon_{IE\omega}$ could be identified either (Fig. 9b). For higher latitude regions beyond ±30°, the minor fluctuations in median $\varepsilon_{IE\omega}$ bear no evidence for a relationship with I_A. On the other hand, the data suggest that the spread around the regional medians consistently decreases with increasing I_A across all latitude bands except 50° N–90° N band as shown by the sequence of IQR in Fig. 9c. This indicates that catchments in more humid regions across the study domain are subject to more pronounced annual water storage fluctuations.

The parameters of the fitted parametric regional and catchment-specific marginal distributions together with the associated predictive robustness flags, as defined by the time stability of the j individual distributions for each catchment are

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Hydrology and Earth System Sciences
Discussions

provided in Supplementary data (10.5281/zenodo.10925965) and can be used, depending on their robustness flag, to estimate $I_{E,t} = I_{E,i} + \varepsilon_{IE\omega}$ for a catchment under future hydro-climatic conditions.

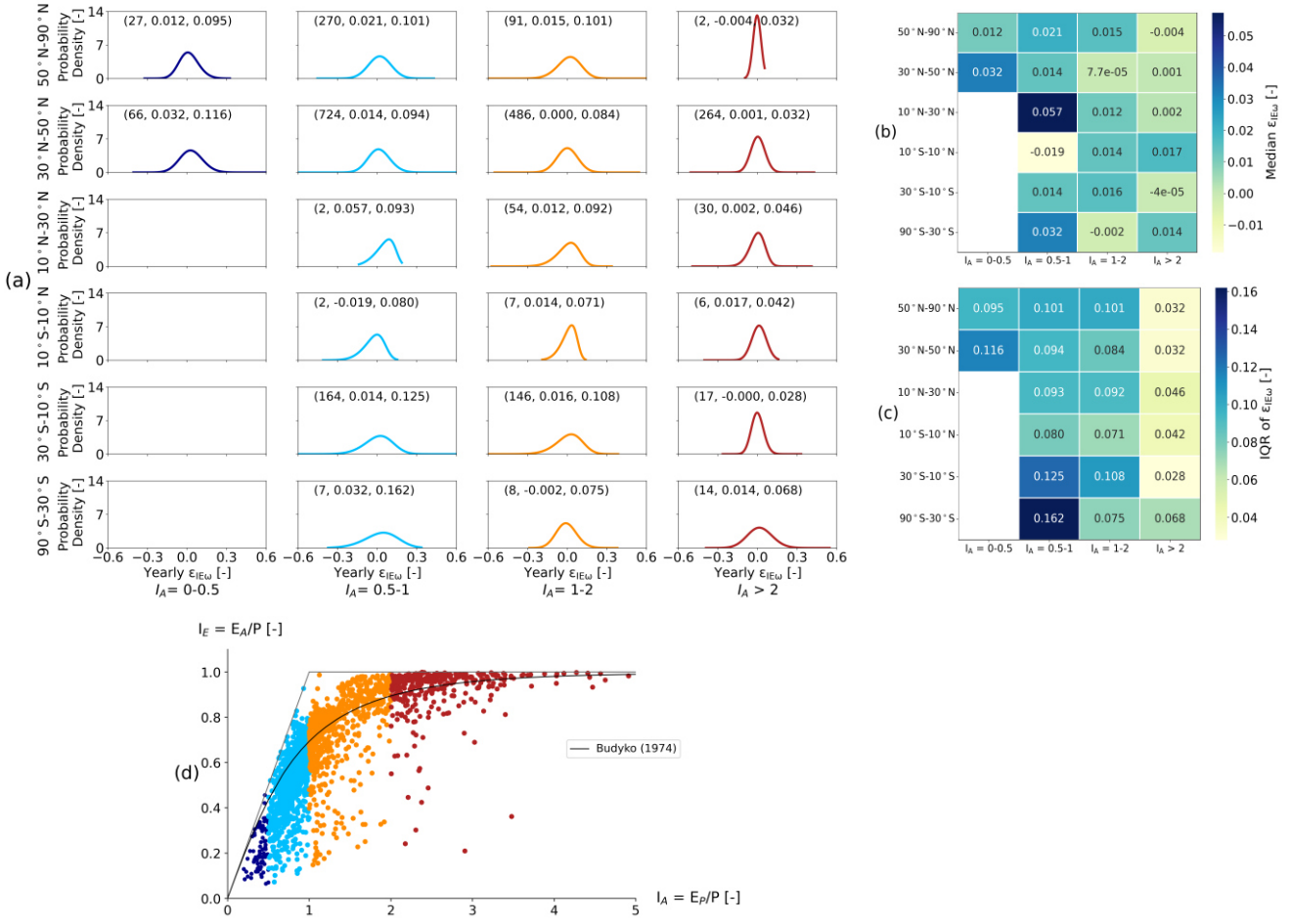
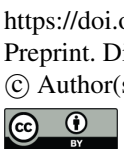


Figure 9: a) Regional marginal distributions of $\epsilon_{IE\omega}$ for defined latitude and I_A bins. The three numerical values in small brackets at the top of each figure presents number of catchments in that category, long-term median $\epsilon_{IE\omega}$ and IQR value of $\epsilon_{IE\omega}$ respectively, b) and (c) presents variation of median and IQR values of $\epsilon_{IE\omega}$ for the regional marginal distribution of $\epsilon_{IE\omega}$, d) Long-term position (1901-2015) of catchments in Budyko space. The colour of the dots corresponds to the regional marginal distributions of $\epsilon_{IE\omega}$ for the corresponding I_A bin

3.5 Sensitivity of marginal distributions of deviations $\varepsilon_{IE\omega}$ to the choice of 20-year averaging window (Step 6)

The moving window analysis to quantify the sensitivity of the marginal distributions of $\varepsilon_{IE\omega}$ resulted in 20 individual, yet not uncorrelated, marginal distributions for each study catchment. Through this approach, we observed that the aggregated marginal distributions of $\varepsilon_{IE\omega}$, may indeed be subject to differences. The magnitudes of the fluctuations vary between catchments, but remain in general rather minor. The Chemung and Lee rivers are two examples for a very low

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024



Hydrology and Earth System Sciences
Discussions

560 sensitivity of the marginal distributions of $\varepsilon_{IE\omega}$ to the choice of time periods. The differences of the medians of the two most extreme marginal distributions does not exceed ~ 0.008 for the Chemung (Fig. 10a) and 90 % of the medians of the 20 marginal distributions fall within an interval of merely 0.01. In addition, the distributions maintain comparable shapes. A similar behaviour was observed for the Lee River (Fig. 10b), with the medians of the two most extreme distributions differing only by ~ 0.014. In this case, 75 % of the medians are observed within an interval of 0.01.

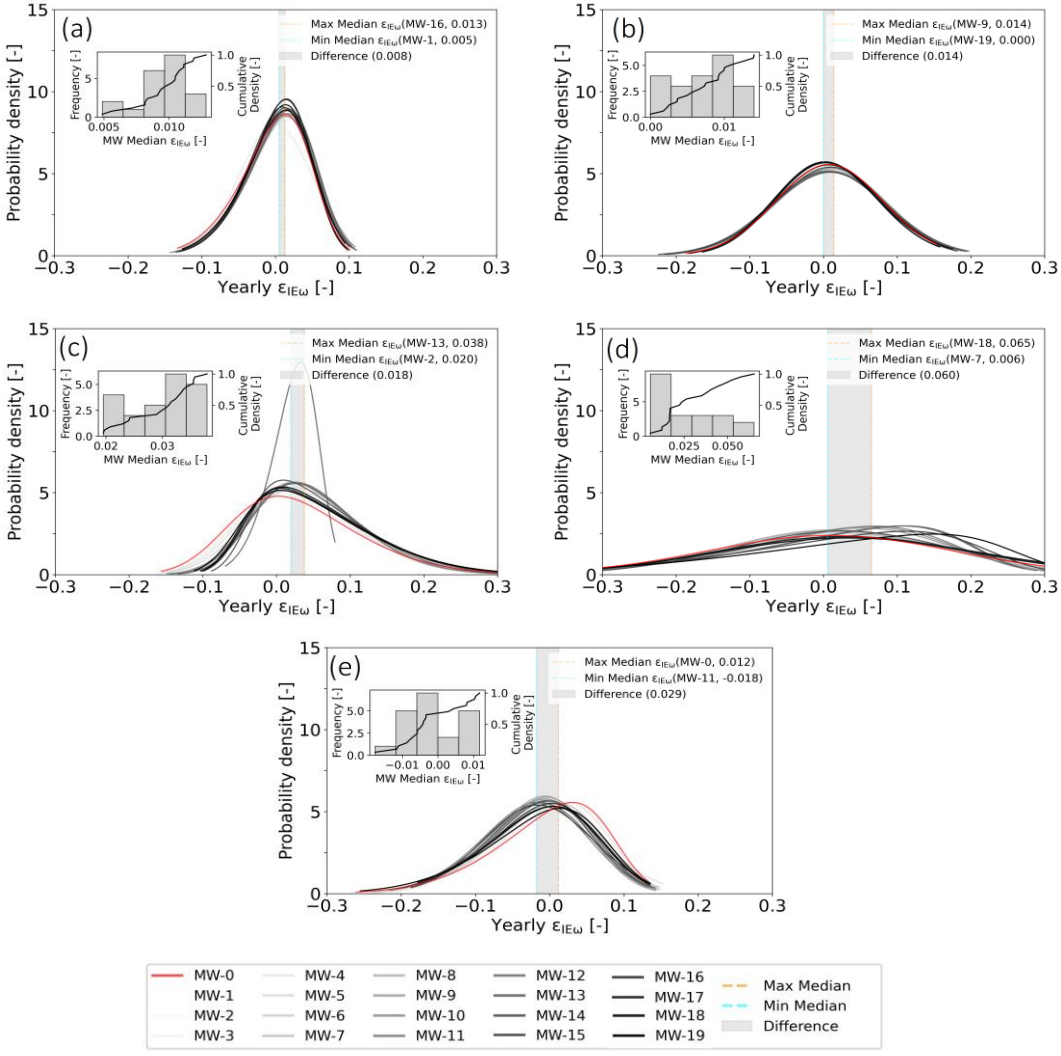
565

570

575

580

585



590 **Figure 10: Aggregated marginal distributions of $\varepsilon_{IE\omega}$ for 20 moving window time periods for five example catchments: a) Chemung River, b) Lee River, c) Sava River, d) Kaituna River, and e) Zschopau River. The red line represents the original marginal distribution of $\epsilon_{IE\omega}$. The orange and aqua-coloured dotted lines depict the maximum and minimum median $\epsilon_{IE\omega}$ values corresponding to their respective moving window time periods. The grey shaded area visually portrays the difference between the extreme maximum and minimum median $\epsilon_{IE\omega}$ values across the moving window time periods**

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

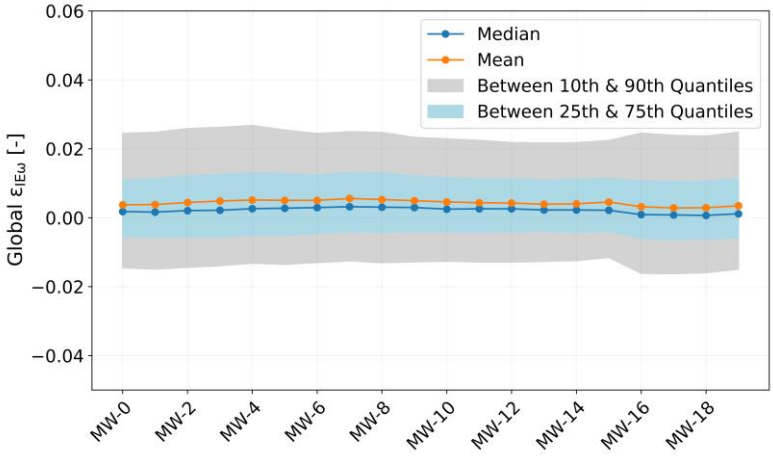
Hydrology and Earth System Sciences
Discussions

595

However, in case of catchments, that are tagged as Variable, Alternating or Shift, the difference in medians of the two extreme marginal distributions is observed to be increased. For the Sava River catchment (Fig. 10c), tagged as Variable, the difference between the two extreme marginal distributions is ~ 0.018 with 75 % of the medians within an interval of 0.032. For Kaituna River (Fig. 10d), tagged as Alternating, the difference between the medians of the two extreme marginal
600 distributions is quite large with a value of 0.060. For 15 out of the 20 marginal distributions, the medians are found to fall within the range of 0.046. A similar pattern for Zschopau River (Fig. 10e), tagged as Shift, is observed with a median difference for two extreme two marginal distributions to be 0.029.

Overall, it was found that for 78 % of the study catchments 15 out of 20 time windows, i.e. 75 %, feature median deviations $\varepsilon_{IE\omega}$ within an interval of ±0.035. This further suggests that, although some sensitivity to the choice of time period
605 can occur, the magnitude of the fluctuations remains rather minor for a large majority of catchments.

Similarly, the distribution of the median $\varepsilon_{IE\omega}$ of all study catchments, i.e. Fig. 7b, remains rather stable when evaluated over the 20 subsequent moving windows, as shown in Fig. 11, with neither the medians nor the spread of the distributions experiencing marked variations. Although lumping the medians of all catchments into one distribution may conceal fluctuations between moving windows of individual catchments, it nevertheless allows the observation that there is
610 no systematic larger scale effect.



615

620

Figure 11: Variation of global long-term median, mean, 10th & 90th quantiles and 25th & 75th quantiles of $\epsilon_{IE\omega}$ for all of the catchments with respect to each moving window

4 Discussion

Our analysis revealed that most study catchments underwent continuous multi-decadal hydro-climatic fluctuations
625 throughout the 20th century (Fig. 4 & Fig. S1). Unlike previous studies comparing only two time periods (Jaramillo and

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

Destouni, 2014), here the higher temporal resolution into with up to five 20-year periods, showed that these fluctuations were not one-directional, with the first half of the century trending towards higher aridity and the latter half towards increased humidity, suggesting cyclic behaviour over longer time scales.

In alignment with previous studies (Berghuijs and Woods, 2016; Jaramillo et al., 2018; Jaramillo et al., 2022; Reaver et al., 2022), our analysis suggests that following disturbances and thus changes in I_A, catchments do not necessarily and strictly follow their specific parametric Budyko curves as defined by parameter ω. In our analysis we found that the general magnitudes of the median deviations $\varepsilon_{IE\omega}$ across all study catchments throughout the 20th century are very minor with median $\varepsilon_{IE\omega} \leq \pm 0.015$ for 50 % and $\leq \pm 0.025$ for 70 % of the catchments. This corresponds well with the results of Jaramillo and Destouni (2014) and Jaramillo et al. (2018), who estimated over two multi-decade periods absolute mean deviations from the expected I_E of $\varepsilon_{IE\omega} \sim 0.01–0.02$, for different regions in the world, based on an analysis of several hundred catchments.

Based on annual water balance data of ~400 catchments in the United States, Berghuijs and Woods (2016) reported an average difference of around 28 % between the spatial and the temporal sensitivity of I_E to changes in I_A. However, a back-of-the-envelope calculation assuming an average $\omega = 2.6$ (Greve et al., 2015) suggests that even with a pronounced shift in aridity of $\Delta I_A = 0.2$ (Jaramillo and Destouni, 2014) such a 28 % difference in sensitivity leads to only minor absolute deviations from the expected I_E with $\varepsilon_{IE\omega} \sim 0.01–0.04$ (4–8 %) for regions with the most common $I_A = 0.5–2.5$, which broadly corresponds with the results of our study. In contrast, (Reaver et al., 2022), using data from ~700 CAMELS US and UK catchments, provided a detailed and exhaustive analysis of possible temporal trajectories through the Budyko space over several decades. They report the mean of all study catchments' *maximum* relative deviations of the actually observed, empirical values $I_{E,o}$ from the predicted values of I_E by catchment-specific curves with $\varepsilon_{IE\omega,max} = I_{E,o,max} - I_{E,max} \sim 26$ %. However, that mean value of all catchment *maxima* is strongly biased by a few rather extreme outliers in their analysis and the vast majority of their study catchments (>650 out of 728) exhibits much lower errors, with a median maximum deviation of $\varepsilon_{IE\omega,max} \sim 9$ % (see Figure. 3 in Reaver et al. (2022). It may thus prove more informative to interpret the results of Reaver et al. (2022) based on the mean instead of the maximum deviations as these average conditions do almost certainly occur more frequently. Doing so, it is plausible to assume that the deviations $\varepsilon_{IE\omega}$ will be considerably lower than the maximum $\varepsilon_{IE\omega,max} \sim 9$ % and potentially closer to the range of 4–8 % estimated above and thus overall consistent with the results of our analysis.

However, we also note that these minor deviations may have different practical implications in different climates. For example, in a humid catchments with $I_A = 0.5$ (e.g. mean annual $P = 2000$ mm year^{-1}, $E_P = 1000$ mm year^{-1} and Q = 1120 mm year^{-1}), a deviation of $\varepsilon_{IE\omega} = 0.02$ results in $\Delta Q \sim 40$ mm year^{-1}, equivalent to merely 3 % of water yield, which hardly affects water supply. In contrast, the practical effects are more pronounced in arid environments. In a typical catchment with $I_A = 2$ (e.g., $P = 500$ mm year^{-1}, $E_P = 1000$ mm year^{-1}, $Q = 60$ mm year^{-1}), the same deviation will lead to a $\Delta Q \sim 10$ mm year^{-1}, equivalent to ~15 % of the available water yield, and thus have considerable higher relevance for water resources

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

planning. For such environments, a robust quantification of expected deviations may thus prove beneficial for future estimates of water resources availability.

Despite some spatial clustering, the deviations $\varepsilon_{IE\omega}$ from the expected parametric Budyko curves do not exhibit any clear and unambiguous relationships with several climatic variables (Fig. S5). The detailed processes and reasons underlying the deviations thus remain so far unknown and may be assumed to be manifold and to vary depending on the characteristics of specific sites. In any case it is plausible to assume that the reasons are a combination of factors, including amongst others changes in precipitation volumes, seasonality and phase, changes in atmospheric water demand, changes in land cover, human interventions, such as reservoir operation or irrigation, but also violations of the assumption that $dS/dt \sim 0$ over the 20-year time periods (Han et al., 2020) and other uncertainties in the available data (Beven, 2016). Note, that a detailed exploration of this issue is beyond the scope of this paper.

To our knowledge, this is the first study to quantify the evolution of median $\varepsilon_{IE\omega}$ over multiple time periods. This allowed to build distributions to predict future $\varepsilon_{IE\omega}$ based on historic data, together with an indicative robustness flag, describing their temporal stability and thus their suitability to predict $\varepsilon_{IE\omega}$ under future hydro-climatic conditions. It was found that, globally, median $\varepsilon_{IE\omega}$ does not only remain minor but, perhaps even more importantly, also rather stable over time. For ~70 % of the study catchments the annual distributions of $\varepsilon_{IE\omega}$, and thus also their 20-year medians, were classified as “Stable”. In other words, the available data suggest that over multiple 20-year periods in the past century the samples of annual deviations originate from the same distribution. This further allows some confidence to plausibly assume that $\varepsilon_{IE\omega}$ and the associated I_E under projected future hydro-climatic conditions can, at least for several decades, be robustly predicted based on these distributions.

Further 19 % of catchments were classified as “Variable” as their distributions of annual deviations for the individual 20-year periods exhibit some variability. Despite this, there is no indicative evidence to link this variability to alternating fluctuations or systematic, one-directional shifts and thus to quantifiable deterministic processes. In this case, the fluctuations can be assumed to be arbitrarily variable, allowing the aggregation of a marginal distribution that reflects all available past knowledge. Although the uncertainty of that distribution may often exceed that of 'Stable' catchments, resulting in somewhat lower predictive power (Montanari and Koutsoyiannis, 2014), it is reasonable to assume that $\varepsilon_{IE\omega}$ remains predictable. The fitted parametric marginal distributions of catchments tagged as “Stable” and “Variable” can be directly used to sample distributions of future annual $\varepsilon_{IE\omega}$ and to estimate the average $\varepsilon_{IE\omega}$ for that future period from the expected future I_E based on ω of the past 20-year period.

For catchments tagged as “Alternating” or “Shift”, the above assumption may be too optimistic. Although the sample size characterizing the evolution of $\varepsilon_{IE\omega}$ over the study period is with a maximum of $j = 4$ pairs of 20-year periods very small and thus no meaningful formal statistical tests could be executed, the data do not rule out the possibility that $\varepsilon_{IE\omega}$ in these catchments is characterized by alternating or shift-like behaviours. In other words, $\varepsilon_{IE\omega}$ may not be sampled from different distributions that change arbitrarily over time, but from distributions that (here) depend either on I_E of the preceding

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

time period or systematically increase or decrease over time. In these cases, the aggregated marginal distribution may produce spurious predictions of $\varepsilon_{IE\omega}$.

For catchments tagged as "Alternating", the user may instead want to consider to construct and use a conditional distribution in the form of $\varepsilon_{IE|\omega i}$, i.e. a distribution of $\varepsilon_{IE\omega}$ given the position $I_{E,i}$, for more reliable estimates. However, note that the limited data available for a maximum of four pairs of time periods, poses a practical complication to construct a meaningful distribution $\varepsilon_{IE|\omega i}$, which is necessary to infer $\varepsilon_{IE|\omega i}$. Alternatively, the user can decide to base predictions only on basis of the $\varepsilon_{IE\omega}$ distribution of the last available time period to avoid the use of the marginal distribution (Koutsoyiannis and Montanari, 2015). For predictions in catchments with suspected presence of a systematic shift, tagged as "Shift", users may choose to extrapolate the fitted distribution parameters of the individual pairs of periods to account for their shifts over time. However, here the reliability of this will depend on the strength of the individual relationship over the past and needs to be evaluated on a case-to-case basis. However, both cases are likely to lead to rather unreliable future estimates.

It was further found that the choice of a specific 20-year window can indeed lead to fluctuations in the distributions of $\varepsilon_{IE\omega}$. However, the magnitude of these fluctuations remains rather limited for the vast majority of catchments. To avoid misinterpretations, we have therefore added the IQR of median $\varepsilon_{IE\omega}$ from the 20 individual moving windows as additional robustness flag for each catchment in Supplementary data downloadable from Zenodo repository (10.5281/zenodo.10925965). Lower IQR then indicate lower sensitivity to the choice of the 20-year window and thus a higher robustness of the marginal distribution for predictions of $\varepsilon_{IE\omega}$ under future conditions.

A complete list of the parameters and robustness flags of the individual 20-year distributions as well as of the local aggregated marginal distributions for each of the 2387 study catchments, but also of the regional distributions as stratified by latitude and I_A are provided in the Supplementary data (10.5281/zenodo.10925965). These distributions of annual $\varepsilon_{IE\omega}$ can be directly used to predict the median $\varepsilon_{IE\omega}$ under future conditions locally in these catchments or regionally by sampling over 20 projected future years.

5 Conclusions

Based on up to 100 years of hydro-climatic and streamflow data for 2387 river catchments world-wide we have here tested whether catchments follow their specific parametric Budyko curves as defined by parameter ω over multiple 20-year periods throughout the 20th century.

We have found that:

(1) 62 % of the catchments do not significantly deviate from their expected parametric Budyko curves. However, this also entails that a fraction of 38 % does indeed deviate.

(2) The overall magnitude of deviations is minor. For ~70 % of the catchments the median deviations do not exceed $\varepsilon_{IE\omega} = \pm 0.025$, which is equivalent to ~ 1–4 %, depending on I_E.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024


Hydrology and Earth System Sciences
Discussions

(3) For 89 % of the study catchments, $\varepsilon_{IE\omega}$ can be considered highly or at least moderately well predictable based on historical data, as distributions of $\varepsilon_{IE\omega}$ in the past were shown to be stable over multiple time periods or characterized by variable fluctuations.

The above implies that while catchments indeed may not strictly follow their parametric Budyko curves, as defined by parameter ω, the deviations remain in general minor and predictable. The latter is of particular importance for catchments in water-limited regions, where already small deviations can considerably affect available water supply and where robust predictions of these deviations are instrumental for effective future water resources planning and management.

Code and data availability. Daily precipitation and temperature data were acquired via the GSWP-3 dataset, accessible at https://data.isimip.org/10.48364/ISIMIP.886955 (Lange, 2020). GSIM discharge data was obtained from https://doi.pangaea.de/10.1594/PANGAEA.887477 (Do et al., 2018) and https://doi.pangaea.de/10.1594/PANGAEA.887470 (Gudmundsson et al., 2018).

Author contributions. MI conceptualized the study, conducted formal analysis, and prepared the manuscript with input from MCG, RE and MH.

Competing interests. MCG and MH are members of the Editorial Board of HESS. The authors have no other competing interests to declare.

Financial support. This research has been supported by Higher Education Commission of Pakistan (HEC) and TU Delft.

References

Andréassian, V., Coron, L., Lerat, J. and Le Moine, N. 2016. Climate elasticity of streamflow revisited – an elasticity index based on long-term hydrometeorological records. Hydrology and Earth System Sciences 20(11), 4503-4524, https://doi.org/10.5194/hess-20-4503-2016.

Arora, V.K. 2002. The use of the aridity index to assess climate change effect on annual runoff. Journal of hydrology 265(1-4), 164-177.

Berghuijs, W.R., Larsen, J.R., van Emmerik, T.H.M. and Woods, R.A. 2017. A Global Assessment of Runoff Sensitivity to Changes in Precipitation, Potential Evaporation, and Other Factors. Water Resources Research 53(10), 8475-8486, https://doi.org/10.1002/2017wr021593.

Berghuijs, W.R. and Woods, R.A. 2016. Correspondence: Space-time asymmetry undermines water yield assessment. Nat Commun 7, 11603, https://doi.org/10.1038/ncomms11603.

Berghuijs, W.R., Woods, R.A. and Hrachowitz, M. 2014. A precipitation shift from snow towards rain leads to a decrease in streamflow. Nature Climate Change 4(7), 583-586, https://doi.org/10.1038/nclimate2246.

Beven, K. 2016. Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. Hydrological Sciences Journal 61(9), 1652-1665, https://doi.org/10.1080/02626667.2015.1031761.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Bouaziz, L.J.E., Aalbers, E.E., Weerts, A.H., Hegnauer, M., Buiteveld, H., Lammersen, R., Stam, J., Sprokkereef, E., Savenije, H.H.G. and Hrachowitz, M. 2022. Ecosystem adaptation to climate change: the sensitivity of hydrological predictions to time-dynamic model parameters. Hydrology and Earth System Sciences 26(5), 1295-1318, https://doi.org/10.5194/hess-26-1295-2022.

Budyko 1948. Isparenie v estestvennykh usloviyakh [Evaporation under natural conditions]. Leningrad: Gidrometeoizdat.

Budyko, M.I. 1961. The heat balance of the earth's surface. Soviet Geography 2(4), 3-13.

Choudhury, B. 1999. Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model. Journal of Hydrology 216(1-2), 99-110.

Destouni, G., Jaramillo, F. and Prieto, C. 2012. Hydroclimatic shifts driven by human water use for food and energy production. Nature Climate Change 3(3), 213-217, https://doi.org/10.1038/nclimate1719.

Dirmeyer, P.A., Gao, X., Zhao, M., Guo, Z., Oki, T. and Hanasaki, N. 2006. GSWP-2: Multimodel analysis and implications for our perception of the land surface. Bulletin of the American Meteorological Society 87(10), 1381-1398.

Do, H.X., Gudmundsson, L., Leonard, M. and Westra, S. 2018. The Global Streamflow Indices and Metadata Archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata. Earth System Science Data 10(2), 765-785, https://doi.org/10.5194/essd-10-765-2018.

Donohue, R., Roderick, M. and McVicar, T.R. 2007. On the importance of including vegetation dynamics in Budyko's hydrological model. Hydrology and Earth System Sciences 11(2), 983-995.

Donohue, R.J., Roderick, M.L. and McVicar, T.R. 2012. Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model. Journal of Hydrology 436-437, 35-50, https://doi.org/10.1016/j.jhydrol.2012.02.033.

Duffie, J.A. and Beckman, W.A. (1980) Solar engineering of thermal processes, Wiley New York.

Fu, B. 1981. On the calculation of the evaporation from land surface. Sci. Atmos. Sin 5(1), 23-31.

Gentine, P., D'Odorico, P., Lintner, B.R., Sivandran, G. and Salvucci, G. 2012. Interdependence of climate, soil, and vegetation as constrained by the Budyko curve. Geophysical Research Letters 39(19), https://doi.org/10.1029/2012gl053492.

Greve, P., Gudmundsson, L., Orlowsky, B. and Seneviratne, S.I. 2015. Introducing a probabilistic Budyko framework. Geophysical Research Letters 42(7), 2261-2269, https://doi.org/10.1002/2015gl063449.

Gudmundsson, L., Do, H.X., Leonard, M. and Westra, S. 2018. The Global Streamflow Indices and Metadata Archive (GSIM) – Part 2: Quality control, time-series indices and homogeneity assessment. Earth System Science Data 10(2), 787-804, https://doi.org/10.5194/essd-10-787-2018.

Han, Z., Long, D., Huang, Q., Li, X., Zhao, F. and Wang, J. 2020. Improving Reservoir Outflow Estimation for Ungauged Basins Using Satellite Observations and a Hydrological Model. Water Resources Research 56(9), https://doi.org/10.1029/2020wr027590.

Hargreaves, G.H. and Samani, Z.A. 1982. Estimating potential evapotranspiration. Journal of the irrigation and Drainage Division 108(3), 225-230.

Hattermann, F.F., Krysanova, V., Gosling, S.N., Dankers, R., Daggupati, P., Donnelly, C., Flörke, M., Huang, S., Motovilov, Y., Buda, S., Yang, T., Müller, C., Leng, G., Tang, Q., Portmann, F.T., Hagemann, S., Gerten, D., Wada, Y., Masaki, Y., Alemayehu, T., Satoh, Y. and Samaniego, L. 2017. Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins. Climatic Change 141(3), 561-576, https://doi.org/10.1007/s10584-016-1829-4.

Jaramillo, F., Cory, N., Arheimer, B., Laudon, H., van der Velde, Y., Hasper, T.B., Teutschbein, C. and Uddling, J. 2018. Dominant effect of increasing forest biomass on evapotranspiration: interpretations of movement in Budyko space. Hydrology and Earth System Sciences 22(1), 567-580, https://doi.org/10.5194/hess-22-567-2018.

Jaramillo, F. and 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.

Jaramillo, F., Piemontese, L., Berghuijs, W.R., Wang-Erlandsson, L., Greve, P. and Wang, Z. 2022. Fewer Basins Will Follow Their Budyko Curves Under Global Warming and Fossil-Fueled Development. Water Resour Res 58(8), e2021WR031825, https://doi.org/10.1029/2021WR031825.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Koutsoyiannis, D. and Montanari, A. 2015. Negligent killing of scientific concepts: the stationarity case. Hydrological Sciences Journal 60(7-8), 1174-1183, https://doi.org/10.1080/02626667.2014.959959.

Lange, S., & Büchner, M. 2020. ISIMIP2a atmospheric climate input data. https://doi.org/https://doi.org/10.48364/ISIMIP.886955.

Liu, H., Wang, Z., Ji, G. and Yue, Y. 2020. Quantifying the Impacts of Climate Change and Human Activities on Runoff in the Lancang River Basin Based on the Budyko Hypothesis. Water 12(12), https://doi.org/10.3390/w12123501.

Mezentsev, V. 1955. More on the calculation of average total evaporation. Meteorologiya i Gidrologiya 5, 24.

Mianabadi, A., Davary, K., Pourreza-Bilondi, M. and Coenders-Gerrits, A.M.J. 2020. Budyko framework; towards non-steady state conditions. Journal of Hydrology 588, https://doi.org/10.1016/j.jhydrol.2020.125089.

Milly, P. 1994. Climate, soil water storage, and the average annual water balance. Water Resources Research 30(7), 2143-2156.

Montanari, A. and Koutsoyiannis, D. 2014. Modeling and mitigating natural hazards: Stationarity is immortal! Water Resources Research 50(12), 9748-9756, https://doi.org/10.1002/2014wr016092.

Nearing, G.S., Tian, Y., Gupta, H.V., Clark, M.P., Harrison, K.W. and Weijs, S.V. 2016. A philosophical basis for hydrological uncertainty. Hydrological Sciences Journal 61(9), 1666-1678, https://doi.org/10.1080/02626667.2016.1183009.

Oldekop, E. (1911) Collection of the Works of Students of the Meteorological Observatory, p. 209, University of Tartu-Jurjew-Dorpat Tartu, Estonia.

Padrón, R.S., Gudmundsson, L., Greve, P. and Seneviratne, S.I. 2017. Large-Scale Controls of the Surface Water Balance Over Land: Insights From a Systematic Review and Meta-Analysis. Water Resources Research 53(11), 9659-9678, https://doi.org/10.1002/2017wr021215.

Porporato, A., Daly, E. and Rodriguez-Iturbe, I. 2004. Soil water balance and ecosystem response to climate change. The American Naturalist 164(5), 625-632.

Reaver, N.G.F., Kaplan, D.A., Klammler, H. and Jawitz, J.W. 2022. Theoretical and empirical evidence against the Budyko catchment trajectory conjecture. Hydrology and Earth System Sciences 26(5), 1507-1525, https://doi.org/10.5194/hess-26-1507-2022.

Roderick, M.L. and Farquhar, G.D. 2011. A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties. Water Resources Research 47(12), https://doi.org/10.1029/2010wr009826.

Schreiber, P. 1904. Über die Beziehungen zwischen dem Niederschlag und der Wasserführung der Flüsse in Mitteleuropa. Z. Meteorol 21(10), 441-452.

Serpa, D., Nunes, J., Santos, J., Sampaio, E., Jacinto, R., Veiga, S., Lima, J., Moreira, M., Corte-Real, J. and Keizer, J. 2015. Impacts of climate and land use changes on the hydrological and erosion processes of two contrasting Mediterranean catchments. Science of the Total Environment 538, 64-77.

Shao, Q., Traylen, A. and Zhang, L. 2012. Nonparametric method for estimating the effects of climatic and catchment characteristics on mean annual evapotranspiration. Water Resources Research 48(3), https://doi.org/10.1029/2010wr009610.

Tixeront, J. 1964. Prévision des apports des cours d'eau. Publication de l'Association internationale d'hydrologie scientifique (63), 118-126.

Troch, P.A., Carrillo, G., Sivapalan, M., Wagener, T. and Sawicz, K. 2013. Climate-vegetation-soil interactions and long-term hydrologic partitioning: signatures of catchment co-evolution. Hydrology and Earth System Sciences 17(6), 2209-2217, https://doi.org/10.5194/hess-17-2209-2013.

Turc, L. 1954. Le bilan d'eau des sols: Relation entre la precipitations, l'evaporation et l'ecoulement. Annales Agronomiques, Série A 5, 491.

van der Velde, Y., Vercauteren, N., Jaramillo, F., Dekker, S.C., Destouni, G. and Lyon, S.W. 2014. Exploring hydroclimatic change disparity via the Budyko framework. Hydrological Processes 28(13), 4110-4118, https://doi.org/10.1002/hyp.9949.

Wang, C., Wang, S., Fu, B. and Zhang, L. 2016. Advances in hydrological modelling with the Budyko framework. Progress in Physical Geography: Earth and Environment 40(3), 409-430, https://doi.org/10.1177/0309133315620997.

https://doi.org/10.5194/hess-2024-120
Preprint. Discussion started: 29 April 2024

Westerberg, I.K., Guerrero, J.L., Younger, P.M., Beven, K.J., Seibert, J., Halldin, S., Freer, J.E. and Xu, C.Y. 2011. Calibration of hydrological models using flow-duration curves. Hydrology and Earth System Sciences 15(7), 2205-2227, https://doi.org/10.5194/hess-15-2205-2011.

Xing, W., Wang, W., Shao, Q., Yong, B., Liu, C., Feng, X. and Dong, Q. 2018. Estimating monthly evapotranspiration by assimilating remotely sensed water storage data into the extended Budyko framework across different climatic regions. Journal of Hydrology 567, 684-695, https://doi.org/10.1016/j.jhydrol.2018.10.014.

Xu, X., Liu, W., Scanlon, B.R., Zhang, L. and Pan, M. 2013. Local and global factors controlling water-energy balances within the Budyko framework. Geophysical Research Letters 40(23), 6123-6129, https://doi.org/10.1002/2013gl058324.

Yang, H., Yang, D., Lei, Z. and Sun, F. 2008. New analytical derivation of the mean annual water-energy balance equation. Water Resources Research 44(3), https://doi.org/10.1029/2007wr006135.

Zhang, L., Dawes, W.R. and Walker, G.R. 2001. Response of mean annual evapotranspiration to vegetation changes at catchment scale. Water Resources Research 37(3), 701-708, https://doi.org/10.1029/2000wr900325.