Hydrol. Earth Syst. Sci., 29, 1703–1723, 2025
https://doi.org/10.5194/hess-29-1703-2025

Hydrology and Earth System Sciences
Open Access
EGU

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

Muhammad Ibrahim, Miriam Coenders-Gerrits, Ruud van der Ent, and Markus Hrachowitz

Department of Water Management, Faculty of Civil Engineering and Geosciences,
Delft University of Technology, Delft, the Netherlands

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

Received: 22 April 2024 – Discussion started: 29 April 2024
Revised: 9 January 2025 – Accepted: 18 January 2025 – Published: 26 March 2025

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 common assumption is that movement along a specific Budyko curve with changes in the aridity index over time can be used as a predictor for catchment responses to changing climatic conditions. 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 (62 %) of study catchments did not significantly deviate from their expected parametric Budyko curves. Out of 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. When these median deviations were expressed as relative changes in discharge, catchments in arid regions showed higher susceptibility to larger discharge shifts compared to those in humid regions. Furthermore, a significant majority of catchments, constituting around the same percentage, was 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.

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 worldwide (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 the water supply (Destouni et al., 2013).

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 for mapping 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 the aridity index $I_A = E_P/P$. Many studies have demonstrated that the empirical evaporative index I_E of river catchments worldwide indeed scatters 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 across a wide spectrum of hydroclimatically and physiographically 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 for estimates of future I_E and thus of E_A and Q based on changes in I_A, which are 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. 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, 2022; Reaver et al., 2022). Such deviations ($\varepsilon_{IE\omega}$) from the expected parametric Budyko curve are here defined as the dimensionless absolute difference between the observed evaporative index ($I_{E,o}$) and the predicted evaporative index (I_E) derived from the expected parametric Budyko curve. They were previously referred to as residual or landscape-driven, indicating that many factors other than I_A, such as human-induced changes in water and land use (e.g. afforestation, deforestation, irrigation, reservoir construction) also play a role (Donohue et al., 2007; Wang and Hejazi, 2011; Sterling et al., 2012; Destouni et al., 2013; van der Velde et al., 2014; Jaramillo and Destouni, 2015; Levi et al., 2015; Nijzink et al., 2016; Daly et al., 2019; Gan et al., 2021; Hrachowitz et al., 2021). As a consequence, Reaver et al. (2022) have warned that parametric Budyko curves may have no predictive power at all. This may be too pessimistic of a perspective. First, the average magnitudes of $\varepsilon_{IE\omega}$ so far reported in studies remain rather low (e.g. Tempel et al., 2024; Wang et al., 2024). 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). Yet, as the assumption of steady conditions might not be applicable (Mianabadi et al., 2020), the presence of uncertainties in the modelling process is inevitable (Westerberg et al., 2011; Nearing et al., 2016). In other words, some level of deviation from the parametric Budyko curves is to be expected as different time periods will never be characterized by exactly the same environmental conditions. However, the mechanistic processes that control these deviations, and thus ω, are not well understood.

Although part of several previous analyses (Destouni et al., 2013; 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 its evolution over multiple time periods at global, regional, and local scales explicitly reported in the literature. Jaramillo and Destouni (2014), Jaramillo et al. (2018) and Tempel et al. (2024) provided estimates of average $\varepsilon_{IE\omega}$ for several regions but limited their analyses to two independent time periods, while Wang et al. (2024) analysed distributions of $\varepsilon_{IE\omega}$ over multiple decades in one single river

basin. In contrast, Reaver et al. (2022) quantified $\varepsilon_{\mathrm{IE}\omega}$ over multiple time periods but explicitly reported only estimates of $\varepsilon_{\mathrm{IE}\omega}$ maxima for individual catchments, thus describing merely the most extreme situations.

Our research question is whether the distributions and average magnitudes of $\varepsilon_{\mathrm{IE}\omega}$ 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 timescales 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 here quantify the distributions of deviations of catchments from parametric Budyko curves, i.e. $\varepsilon_{\mathrm{IE}\omega}$, at global, regional, and local scales between multiple 20-year periods throughout the 20th century. Specifically, we test the hypothesis that the distributions of $\varepsilon_{\mathrm{IE}\omega}$ are too wide and temporally too unstable, so 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 [$\mathrm{mm\,d^{-1}}$] as well as maximum and minimum temperature T [°C] data at a spatial resolution of $0.5° \times 0.5°$ 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 of 1901–2015.

Potential evaporation, E_{P} [$\mathrm{mm\,d^{-1}}$], was estimated based on the method proposed by Hargreaves and Samani (1982):

$$E_{\mathrm{P}} = \alpha R_{\mathrm{a}}(T_{\mathrm{a}} + 17.8)\sqrt{(T_{\max} - T_{\min})}, \quad (1)$$

where $\alpha \sim 0.0023$ is a constant used to convert $\mathrm{MJ\,m^{-2}\,d^{-1}}$ to $\mathrm{mm\,d^{-1}}$; R_{a} is the extraterrestrial radiation at the top of the atmosphere [$\mathrm{MJ\,m^{-2}\,d^{-1}}$]; and T_{a}, $T_{\max}$, and $T_{\min}$ are the daily average, maximum and minimum temperatures [°C], respectively. R_{a} is estimated 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 Archive (GSIM; Do et al., 2018a; Gudmundsson et al., 2018a), which consists of in situ streamflow observations data for over 30 000 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 labelled Caution. After filtering, we retained 2387 river catchments with data series ranging from 50 to 100 years (with a median of 78 years). These catchments vary in size, from 4 to 3 475 000 $\mathrm{km^2}$ (with a median of $\sim 1564\,\mathrm{km^2}$; Fig. 1a). The catchments represent diverse hydroclimatic conditions (Fig. 1b–f), as indicated by the long-term average aridity index (I_{A}) that ranges from 0.19 to 6.66 (with a median of 0.97; Fig. 1e) and evaporative index (I_{E}) that ranges from 0.06 to 0.99 (with a median of 0.65; Fig. 1f).

2.2 Methods

The subsequent experiment to estimate for each of the 2387 study catchments the deviation $\varepsilon_{\mathrm{IE}\omega}$ values from its expected evaporative index 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.

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

$$I_{\mathrm{E}} = \frac{E_{\mathrm{A}}}{P} = 1 + \frac{E_{\mathrm{P}}}{P} - \left[1 + \left(\frac{E_{\mathrm{P}}}{P}\right)^{\omega}\right]^{\frac{1}{\omega}}, \quad (2)$$

where ω is a catchment-specific parameter estimated from long-term averages of observed P, E_{P}, and $E_{\mathrm{A}} = P - Q$, assuming negligible change in storage $\mathrm{d}S/\mathrm{d}t$.

Equation (2) suggests that with a given ω, hydroclimatic shifts between two periods, T_i and T_{i+1}, expressed as changes in aridity index $\Delta I_{\mathrm{A}} = \Delta(E_{\mathrm{P}}/P)$, will lead to predictable change $\Delta I_{\mathrm{E},i+1}$ (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 $I_{\mathrm{E},i+1}$, which is expressed as

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

However, in reality, as described above, ω is often not constant over time (Case B). Catchments therefore do not strictly follow their $I_{\mathrm{E},i}$ 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 $\varepsilon_{\mathrm{IE}\omega,i+1}$, which is described as

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

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

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

Here, we have sub-divided the available data records of each catchment into up to five individual 20-year periods over the last century, denoted as T_i (Table 1), where T_i represents the ith 20-year period. This 20-year period was chosen deliberately to balance the need for a sufficiently long period to minimize the impact of storage changes while preserving the temporal sequence in the data that allowed us to place each catchment into a specific temporal stability category (as described in Step 4). We assume that 20-year periods are long enough to satisfy $\mathrm{d}S/\mathrm{d}t \approx 0$, which is supported by

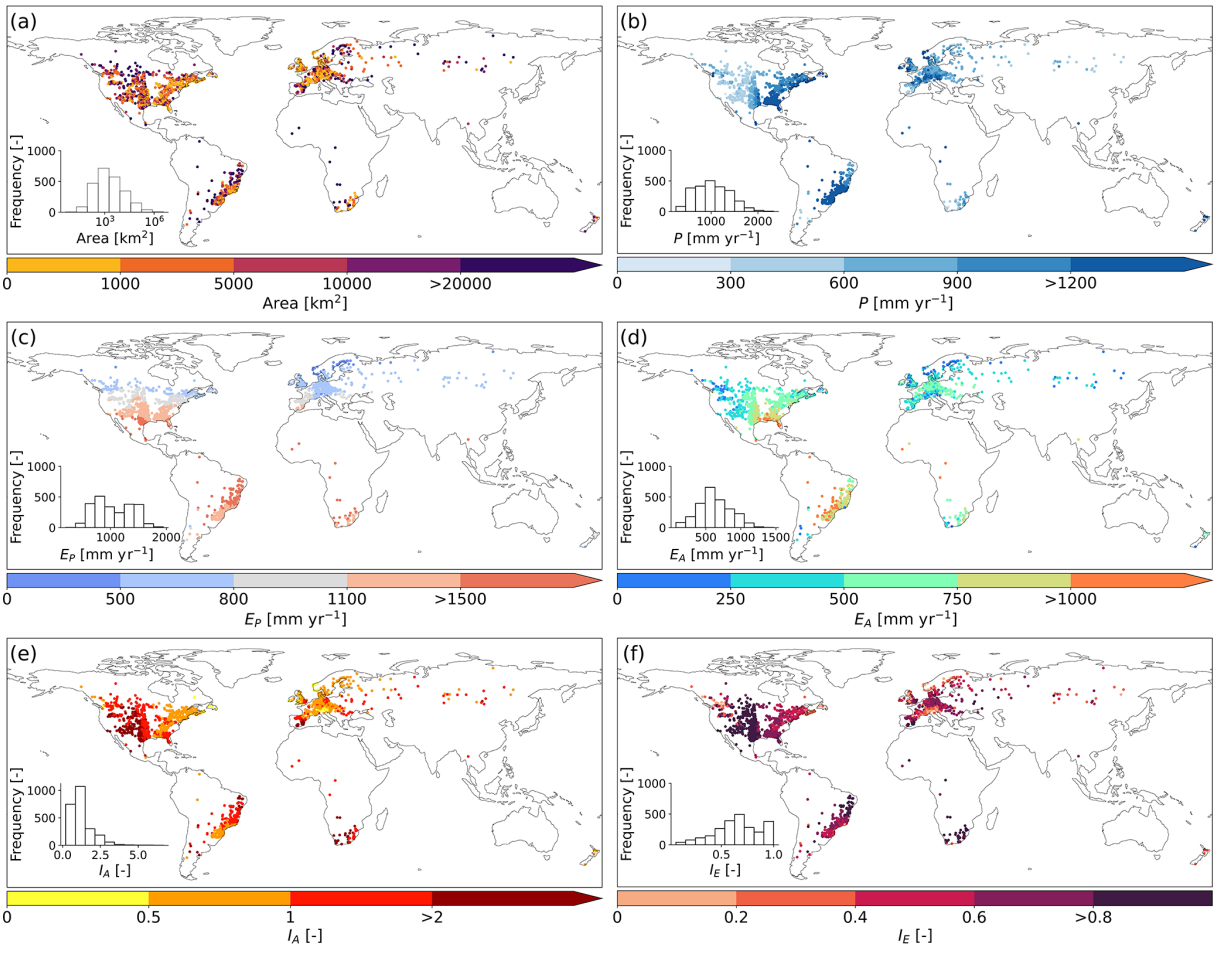

Figure 1. Spatial distribution of 2387 studied catchments along with topographic characteristics and long-term mean (1901–2015) climatic indices: **(a)** catchment area; **(b)** precipitation, P; **(c)** potential evaporation, E_{P}; **(d)** actual evaporation, $E_{\mathrm{A}} = P - Q$; **(e)** aridity index, I_{A}; and **(f)** evaporative index, I_{E}.

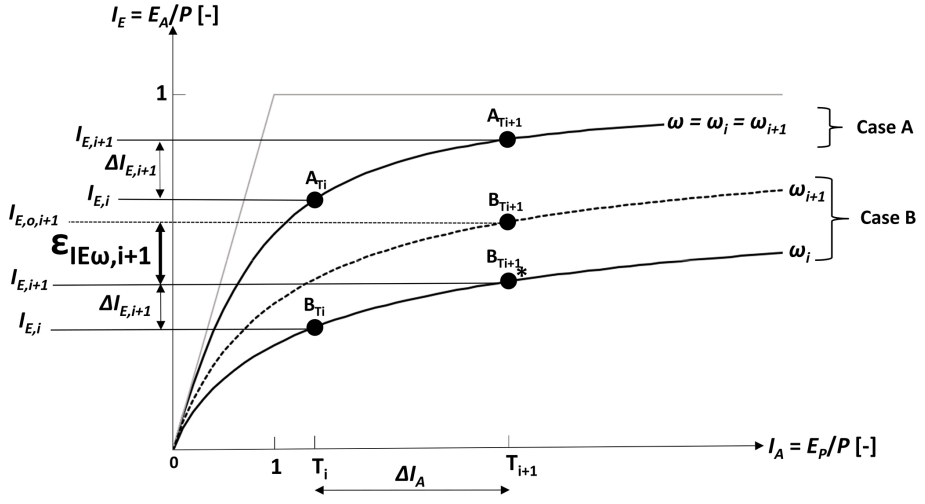

Figure 2. A schematic representation of a catchment movement in Budyko space between two long-term time periods, T_i and T_{i+1}. Case A: catchment A moves along the same Budyko curve from the first period T_i to the next period T_{i+1} (i.e. $\omega_i = \omega_{i+1}$). Case B: catchment B has deviated from its expected parametric Budyko curve (i.e. $\omega_i \neq \omega_{i+1}$), resulting in deviation $\varepsilon_{\mathrm{IE}\omega,i+1}$ (Eq. 4).

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

Time period	Symbols		
	20-year periods	Change between subsequent 20-year periods	Distributions of deviations
1901–1920	T_1	Δ_{1-2}	$\varepsilon_{\mathrm{IE}\Delta 1}$
1921–1940	T_2	Δ_{2-3}	$\varepsilon_{\mathrm{IE}\Delta 2}$
1941–1960	T_3	Δ_{3-4}	$\varepsilon_{\mathrm{IE}\Delta 3}$
1961–1980	T_4	Δ_{4-5}	$\varepsilon_{\mathrm{IE}\Delta 4}$
1981–2000	T_5		

Han et al. (2020), who demonstrated that in more than 80 % of catchments worldwide, $\mathrm{d}S/\mathrm{d}t$ is less than 5 % over 20-year periods. Using longer periods, such as 30 years, as used in previous studies (e.g. Destouni et al., 2013), would have smoothed out potential shifts and limited the ability to detect systematic changes. In addition, 20-year periods align with planning horizons in many water resources management decisions.

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

Step 1 – estimation of catchment-specific $I_{\mathrm{E},i}$ curves and the distribution of annual $I_{\mathrm{E,o}}$ 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 $I_{\mathrm{E},i}$ defined by parameter ω_i is obtained by fitting Eq. (2) to the set of 20 annual values 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 $I_{\mathrm{E,o}}$ values instead of directly to their 20-year averages was a deliberate choice. The fluctuations in the $n = 20$ annual $I_{\mathrm{E,o}}$ values explicitly represent annual storage changes ($\mathrm{d}S/\mathrm{d}t$) between individual years. This subsequently allowed us to treat the observed annual $I_{\mathrm{E,o}}$ probabilistically as distributions around their expected values for that period T_i as defined by $I_{\mathrm{E},i}$ curve (Fig. 3a).

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

For each catchment, we then used ω_i from each time period T_i to compute the expected $I_{\mathrm{E},i+1}$ for the subsequent period T_{i+1} (i.e. point $\mathrm{B_{Ti+1}}$* in Fig. 2). This then allowed us to estimate the individual deviations of the 20 annual observed $I_{\mathrm{E,o}}$ values from the expected $I_{\mathrm{E},i+1}$ 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 Δ_{1-2}, Δ_{2-3}, etc.), this resulted in an individual distribution of annual deviation $\varepsilon_{\mathrm{IE}\omega}$ values around a 20-year average in each catchment (Fig. 3b). This approach using a temporally changing (dynamic) baseline was chosen as it is more sensitive for capturing trends and shifts in hydrological behaviour of catchments over time than a fixed baseline. For completeness, we also performed the same analysis using a fixed baseline (i.e. using the earliest available period as a fixed baseline) and provide the results thereof in the Supplement.

Note that catchments with data for all five time periods T_1–T_5 have the maximum of $j = 4$ distributions $\varepsilon_{\mathrm{IE}\Delta j}$. In contrast, catchments with data for only two periods, e.g. T_2 and T_3, feature only $j = 1$ distribution of between-period deviation $\varepsilon_{\mathrm{IE}\omega}$ values.

Non-parametric Wilcoxon signed rank tests were then used to test for each distribution $\varepsilon_{\mathrm{IE}\Delta j}$ 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 $\varepsilon_{\mathrm{IE}\Delta j}$ of observed $I_{\mathrm{E,o},i+1}$ from expected $I_{\mathrm{E},i+1}$ is higher than zero for the comparison of $\varepsilon_{\mathrm{IE}\Delta j}$ between periods T_i and T_{i+1}.

Step 3 – fitting parametric distributions to the empirical distributions of annual deviations $\varepsilon_{\mathrm{IE}\Delta j}$.

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

$$f(x) = \frac{2}{\lambda}\phi\left(\frac{x-\xi}{\lambda}\right)\Phi(\alpha)\left(\frac{x-\xi}{\lambda}\right), \tag{6}$$

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

The mean of the distribution is computed as

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

and the standard deviation is represented by

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

Step 4 – evaluating temporal stability of the distributions $\varepsilon_{\mathrm{IE}\Delta j}$ in subsequent pairs of time periods.

For distributions of past deviations to be used to estimate deviation $\varepsilon_{\mathrm{IE}\omega}$ values under projected hydroclimatic future conditions, it is necessary to evaluate upfront whether it is plausible to assume that they retain sufficient explanatory power

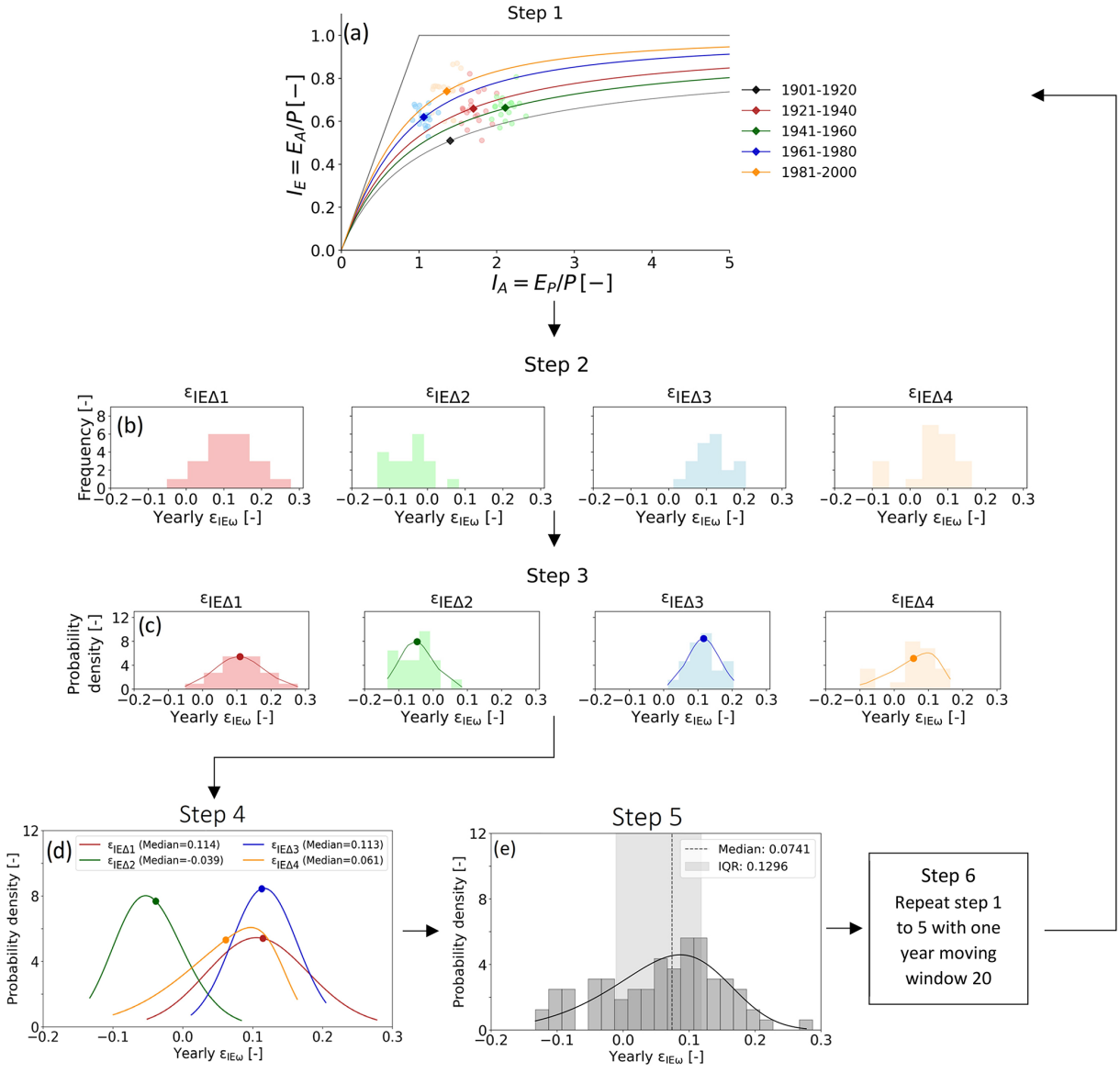

Figure 3. Flow chart of methodology. Step 1: estimation of catchment-specific $I_{E,i}$ curves and the distribution of annual $I_{E,o}$ around it for each period T_i. Step 2: distributions of annual deviations $\varepsilon_{IE\Delta j}$ from expected $I_{E,i+1}$ between subsequent time periods. Step 3: fitting parametric distributions to the empirical distributions of annual deviations $\varepsilon_{IE\Delta j}$. Step 4: evaluating temporal stability of the distributions $\varepsilon_{IE\Delta j}$ in subsequent pairs of time periods. Step 5: aggregated long-term marginal distribution of annual deviation $\varepsilon_{IE\omega}$ values from expected I_E for each catchment. Step 6: evaluation of the sensitivity of the marginal distributions of annual deviation $\varepsilon_{IE\omega}$ values to the choice of 20-year averaging window. Note that the generated distributions of $\varepsilon_{IE\omega}$ are illustrative examples that are not based on real data.

under future conditions or if there is evidence against that. This was done here by analysing how stable the individual distributions in a catchment are over time. To do so, 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). We followed three sub-steps.

– *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 the case the null hypothesis is not rejected ($p > 0.05$), hereafter referred to with the symbol ∘, 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 fluctuations or systematic shifts over time. Thus, in a second step, we checked if the median of $\varepsilon_{\mathrm{IE}\Delta j}$ 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 of a systematic and thus non-variable shift in the median of $\varepsilon_{\mathrm{IE}\Delta j}$ over time. In that case, past distributions $\varepsilon_{\mathrm{IE}\Delta j}$ need to be assumed to have limited predictive power for estimating future $\varepsilon_{\mathrm{IE}\omega}$.

– *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 $\varepsilon_{\mathrm{IE}\Delta j}$ for Δ_j is influenced by the magnitude of $I_{\mathrm{E},i}$ and that, for example, after a 20-year period with a low $I_{\mathrm{E},i}$, further future decreases and thus negative $\varepsilon_{\mathrm{IE}\Delta j}$ are unlikely, and $I_{\mathrm{E},i+1}$ will, more probably, swing back to higher values and thus positive $\varepsilon_{\mathrm{IE}\Delta j}$. Similar to above, if the median $\varepsilon_{\mathrm{IE}\omega}$ systematically decreased (–) or increased (+) with $I_{\mathrm{E},i}$ for three or more of the pairs of values of time period j, this may be evidence of a systematic and thus non-variable shift in the median $\varepsilon_{\mathrm{IE}\omega}$ over time, indicating limited predictive power.

Following the above, each catchment was classified into one of four qualitative categories of temporal stability of $\varepsilon_{\mathrm{IE}\omega}$ (Table 2). Note that the use of a 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 $\varepsilon_{\mathrm{IE}\omega}$ values were found, e.g. ∘ ∘ ∘ ∘ or – ∘ ∘ ∘. A catchment was ranked as variable if it showed an alternating sequence and thus no systematic shift in median $\varepsilon_{\mathrm{IE}\omega}$ values over time, e.g. + – + – or – + + – *and* no relation between $I_{\mathrm{E},i}$ and median $\varepsilon_{\mathrm{IE}\omega}$. In contrast, if a catchment was characterized by an alternating sequence and a dependency between $I_{\mathrm{E},i}$ and median $\varepsilon_{\mathrm{IE}\omega}$, it was tagged as alternating. If, finally, between three or more of the j consecutive distributions in a catchment the median $\varepsilon_{\mathrm{IE}\omega}$ values were 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.

Step 5 – aggregated long-term marginal distribution of annual deviation $\varepsilon_{\mathrm{IE}\omega}$ values from expected I_{E} for each catchment.

In this step, the up to $j = 4$ distributions $\varepsilon_{\mathrm{IE}\Delta j}$ were aggregated into one marginal distribution of $\varepsilon_{\mathrm{IE}\omega}$ values for each catchment (Fig. 3e). This distribution reflects the historical range of fluctuations in annual $\varepsilon_{\mathrm{IE}\omega}$ values based on all available information for each catchment. Consequently, the median $\varepsilon_{\mathrm{IE}\omega}$ 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 $\varepsilon_{\mathrm{IE}\Delta j}$ (Step 4) on the predictive power of the associated marginal distribution of deviation $\varepsilon_{\mathrm{IE}\omega}$ values, we 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 a rather low predictive power (Table 2).

Step 6 – evaluating the sensitivity of the marginal distributions of annual deviation $\varepsilon_{\mathrm{IE}\omega}$ values 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, 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 1 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 20 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 hydroclimatic variables and movement in Budyko space (Step 1)

Throughout the 100-year study period and across all study catchments, considerable hydroclimatic 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 per century, on average (Fig. 4a), with 57 % of the catchments showing an increase between T_1 and T_2 (Δ_{1-2}) and 83 % for Δ_{4-5}. In contrast, mean annual temperatures (Fig. 4b) and the associated potential evaporation (Fig. 4c) were characterized by a more fluctuating pattern. These combined factors led to slightly

Table 2. Decision criteria to classify the time stability of the j distributions $\varepsilon_{\mathrm{IE}\Delta j}$ for each catchment into one of the four qualitative categories stable, variable, alternating, shift and the associated predictive power of the marginal distribution of $\varepsilon_{\mathrm{IE}\omega}$ values of a catchment, aggregating all j distributions of that catchment.

Tag	Kolmogorov–Smirnov test	Systematic shift in median	Relation between $I_{\mathrm{E},i}$ and $\varepsilon_{\mathrm{IE}\omega}$	Examples	Predictive power	No. of catchments
Stable	$p > 0.05$	No	No	∘ ∘ ∘ ∘ or – ∘ ∘ ∘	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

more arid conditions in the first half of the 20th century, followed by a considerable reduction in aridity index I_{A}, and thus to a shift towards somewhat more humid conditions towards the end of the century across all of the temporal stability categories (Fig. 4e), in which, on average, 78 % and 75 % of the catchments showed decreases in I_{A} for Δ_{3-4} and Δ_{4-5}, respectively. The changes in I_{A} were accompanied by related changes in potential evaporation (E_{P}) and precipitation (Fig. 4c and a). The overall movement of catchments in the Budyko space due to hydroclimatic changes is illustrated in Fig. S1 in the Supplement (Jaramillo and Destouni, 2014). If these movements were driven only by changes in I_{A}, catchments would be expected to move within the directional range of $45 < \alpha < 90°$ or $225 < \alpha < 270°$ (Jaramillo et al., 2022). However, observed movement of catchments is also found in other directions, indicating deviation ($\varepsilon_{\mathrm{IE}\omega} \neq 0$) 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 with the hydroclimatic variables or is present 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 in the Supplement. It is found that the overall pattern of temporal variability in that sub-sample largely reflects that of the full sample.

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

The indicative evidence for the presence of deviation $\varepsilon_{\mathrm{IE}\omega} \neq 0$ in at least some catchments is further supported by a more detailed analysis of the distributions of annual deviations $\varepsilon_{\mathrm{IE}\Delta j}$ 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 $\varepsilon_{\mathrm{IE}\omega} \neq 0$ 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 $\varepsilon_{\mathrm{IE}\omega} \neq 0$ (Fig. 5a). Conversely, this also entails that for a majority of 58 %–66 % of the distributions there is less evidence (i.e. $p > 0.05$) that the median $\varepsilon_{\mathrm{IE}\omega}$ is different from zero. Note that minor $\varepsilon_{\mathrm{IE}\omega}$ values were observed in most catchments. Although these $\varepsilon_{\mathrm{IE}\omega}$ values were not classified as significant based on the Wilcoxon signed rank test used here, it may be too naive to assume that the deviation $\varepsilon_{\mathrm{IE}\omega}$ values are strictly zero, as also demonstrated by Reaver et al. (2022). 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.

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 deviation $\varepsilon_{\mathrm{IE}\omega}$ values with an interquartile range of IQR $\sim$ 0.062 (Fig. 6c) concentrate quite narrowly around the medians. The medians themselves range between merely $\varepsilon_{\mathrm{IE}\omega} = -0.006$ and 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 deviation $\varepsilon_{\mathrm{IE}\omega}$ values of distributions $\varepsilon_{\mathrm{IE}\Delta j}$ are different from zero. In spite of a somewhat wider spread with an IQR $\sim$ 0.094 (Fig. 6f), a similar pattern with consistently low median, $\varepsilon_{\mathrm{IE}\omega} = -0.017$–$0.018$ (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 varied patterns were found for other catchments (Fig. 6g–o). For example, in the Sava River at Radeče (Slovenia; 6004 km^2; ID SI_0000007), all four distributions of the annual deviations display a wider spread, with IQR $\sim$ 0.113, indicating a higher degree of storage fluctuation between individual years (Fig. 6i). This variability may largely be attributed to hydropower developments and the associated changes in hydropower production levels (Levi et al., 2015), which disrupt natural flow regimes by increasing runoff during high demand and altering seasonal flow patterns (Renofalt et al., 2010; Lee et al., 2023). In addition, the medians do considerably deviate from zero, as indicated by median $\varepsilon_{\mathrm{IE}\omega}$ values ranging between -0.023 and 0.118 (Fig. 6h).

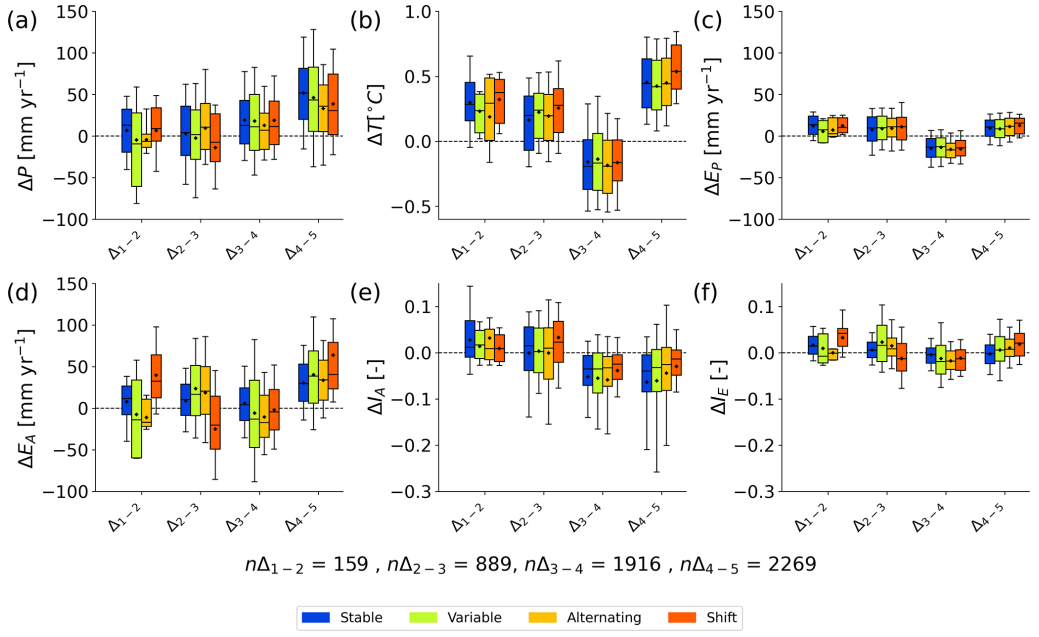

Figure 4. Temporal stability category-wise mean 20-year changes in hydroclimatic variables for the studied catchments between two consecutive periods. **(a)** Precipitation, P; **(b)** temperature, T; **(c)** potential evaporation, E_P; **(d)** actual evaporation, E_A; **(e)** aridity index, I_A; and **(f)** evaporative index I_E. The boxes represent the 25th to 75th percentiles, while whiskers extend to the 10th and 90th percentiles. Diamonds denote the arithmetic mean, and outliers are not shown.

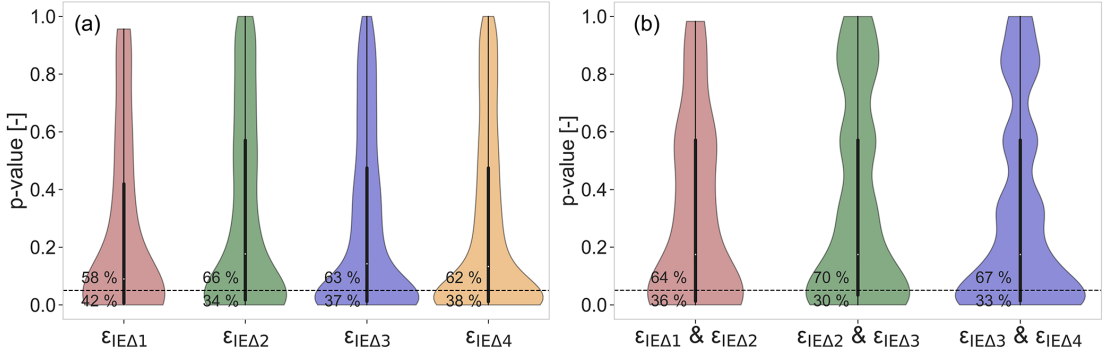

Figure 5. Distribution of p values from statistical tests. **(a)** Wilcoxon signed rank test performed to test whether the long-term median $\varepsilon_{\mathrm{IE}\omega}$ value of the individual 20-year distributions $\varepsilon_{\mathrm{IE}\Delta j}$ is significantly different from zero. The percentage above the significance line (0.05) shows data points where median $\varepsilon_{\mathrm{IE}\omega}$ values are not significantly different from zero, while the percentage below indicates those that are significantly different **(b)** Kolmogorov–Smirnov test performed on the distributions of two consecutive time periods ($\varepsilon_{\mathrm{IE}\Delta j}$ and $\varepsilon_{\mathrm{IE}\Delta j+1}$) to test whether the two distributions of deviations are significantly different from each other. Here, the percentage above the significance line presents data points where the distributions of two consecutive time periods are not significantly different from each other, while the percentage below indicates significant differences. Each violin plot displays the distribution of p values, with the central black box representing the interquartile range (25th to 75th percentiles) and whiskers extending to the smallest and largest data points within 1.5 times the interquartile range. The dashed black line represents the significance level of 0.05.

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 (https://doi.org/10.5281/zenodo.14060926, Ibrahim et al., 2024).

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

Based on the Kolmogorov–Smirnov tests, it was found that overall 68 % of the distributions $\varepsilon_{\mathrm{IE}\Delta j}$ between consecutive pairs of time periods are not significantly different from each other ($p > 0.05$; Fig. 5b). Following the criteria defined in

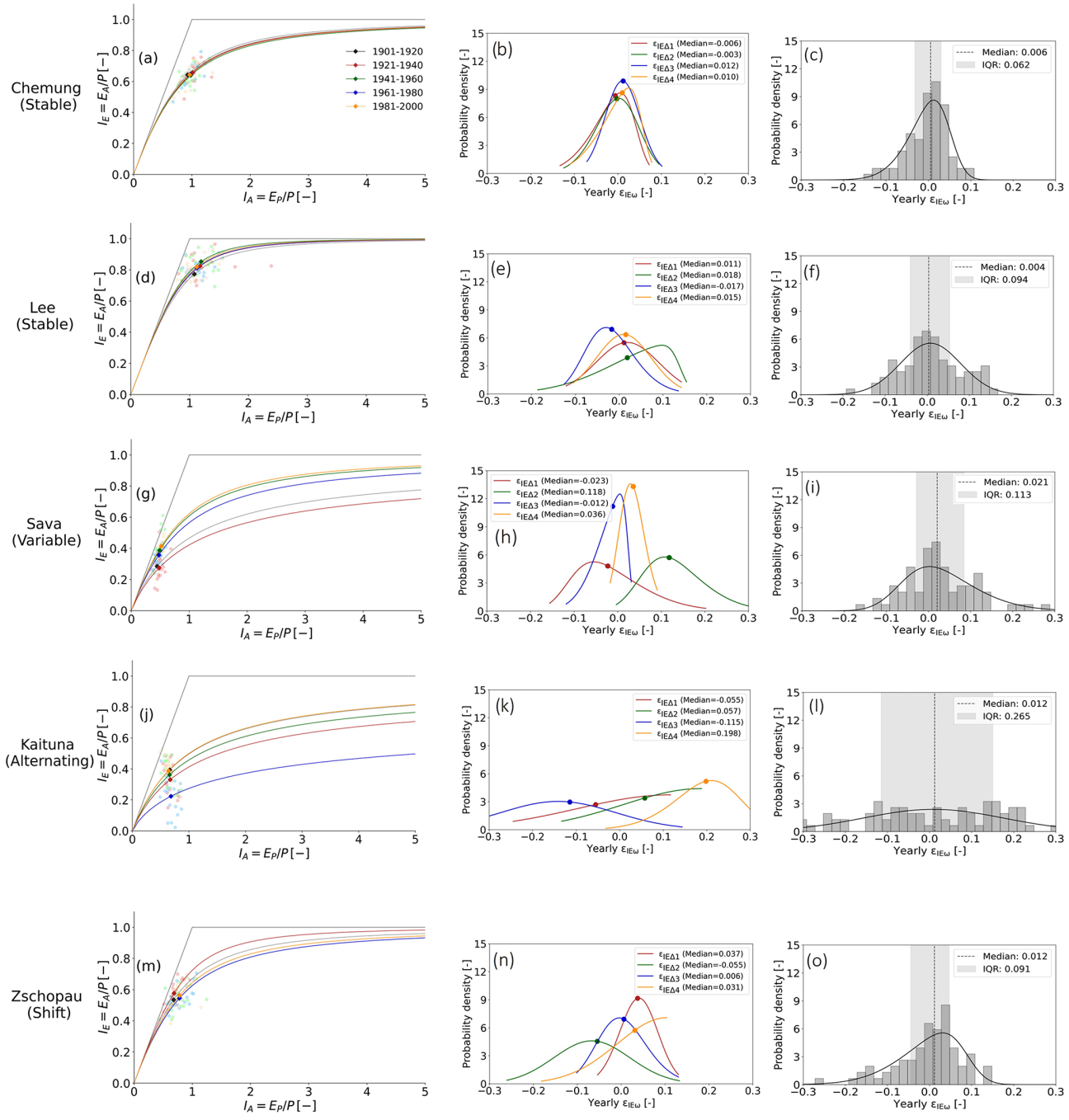

Figure 6. Mean annual position of catchments (faint dots) in Budyko space along with long-term mean (darker dots) and expected parametric Budyko curves (left column). Individual distribution of deviations ($\varepsilon_{IE\Delta1}$, $\varepsilon_{IE\Delta2}$, $\varepsilon_{IE\Delta3}$, and $\varepsilon_{IE\Delta4}$) 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 catchments: Chemung River **(a–c)**, Lee River **(d–f)**, Sava River **(g–i)**, Kaituna River **(j–l)**, and Zschopau River **(m–o)**.

Sect. 2.2, this resulted in 1651 catchments classified as stable (Table 2; Fig. 7a). For these catchments, together corresponding to $\sim 70\,\%$ 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 ∘ ∘ ∘ ∘ and ∘ ∘ ∘ ∘, respectively.

Similarly, 455 additional catchments (19 % of all study catchments), whose distributions exhibited fluctuations over time (Kolmogorov–Smirnov test $p \leq 0.05$), but which fea-

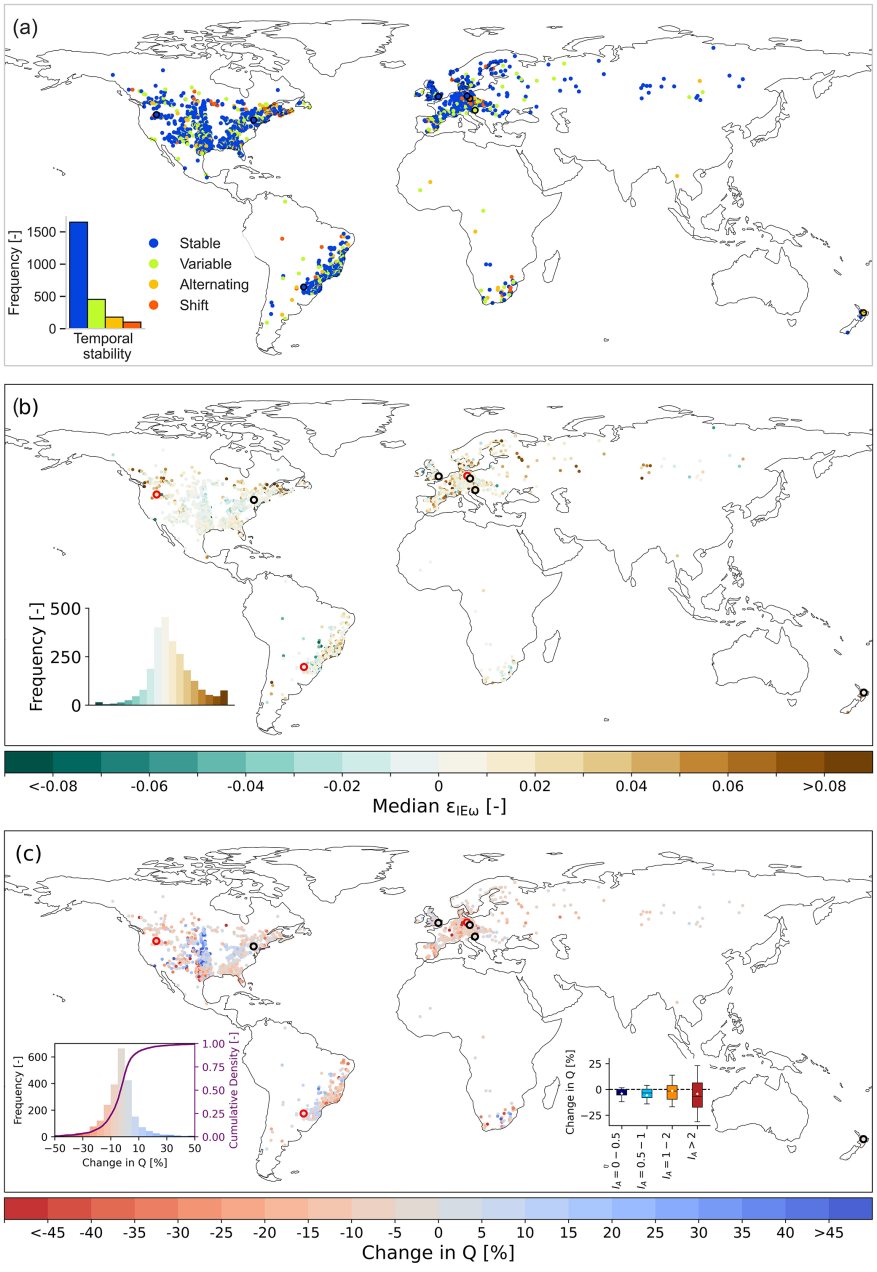

Figure 7. **(a)** Temporal stability and **(b)** long-term median $\varepsilon_{IE\omega}$ value maps of aggregated long-term marginal distributions for the study catchments and **(c)** change in Q as a result of long-term median $\varepsilon_{IE\omega}$ values. Histogram and cumulative density of change in Q and change in Q across different I_A bins are presented as two insets. Change in Q reflects the change due to median deviation $\varepsilon_{IE\omega}$ values from the expected parametric Budyko curve only (i.e. excluding any change resulting from a change in aridity and its associated movement along the expected curve). Catchments highlighted with a black border represent the five selected examples from Fig. 6, while those outlined in red denote three additional selected example catchments shown in the Supplement (Fig. S5). The boxes represent the 25th to 75th percentiles, while whiskers extend to the 10th and 90th percentiles. Diamonds denote the arithmetic mean, and outliers are not shown.

tured only limited evidence for both the presence of systematic shifts over time and 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 Radeče (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 considerably, with medians ranging from −0.023 to 0.118. However, the fluctuations appear to occur alternatively (− + − +) and suggest neither the presence of a systematic shift over time (Fig. S4b in the Supplement) nor a dependency between $I_{E,i}$ and median $\varepsilon_{IE\omega}$ values (Fig. S4a). Despite these fluctuations, the marginal distribution of $\varepsilon_{IE\omega}$ aggregated from

the four individual distributions $\varepsilon_{\mathrm{IE}\Delta j}$ (Fig. 6i) can, although wider than for catchments tagged as stable, thus be assumed to be an at least moderately robust representation of $\varepsilon_{\mathrm{IE}\omega}$.

In contrast, 7 % of the catchments were tagged as alternating, and a dependency between $I_{\mathrm{E},i}$ and $\varepsilon_{\mathrm{IE}\omega}$ 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 $\varepsilon_{\mathrm{IE}\omega}$ values between -0.115 and 0.198. In addition, although no systematic evolution of median $\varepsilon_{\mathrm{IE}\omega}$ over time was evident (Fig. S4d), the data suggest the potential presence of a dependency on $I_{\mathrm{E},i}$, as shown in Fig. S4c. The pronounced alternating behaviour of the $\varepsilon_{\mathrm{IE}\omega}$, fluctuations between -0.115 and 0.198, could not be readily explained by factors such as land use changes as estimated from the Hilda+ gridded land cover product (Winkler et al., 2021), seasonality index (SI) of liquid precipitation input (i.e. rainfall + snowmelt), and Pardé coefficients or median rainfall intensity (Fig. S3a, c, e, and f in the Supplement). The SI was calculated using the formula proposed by Gao et al. (2014). A higher SI value indicates that most of the precipitation falls within a few months, while a lower value reflects more evenly distributed precipitation throughout the year. This suggests that other additional drivers, or a combination of drivers, influence this catchment's alternating behaviour.

The remaining 102 catchments (4 %) were tagged as shift as they exhibit a rather consistent shift in median $\varepsilon_{\mathrm{IE}\omega}$ over time. This can be seen for a selected example in Fig. 6m–o. The median $\varepsilon_{\mathrm{IE}\omega}$ in this catchment of the Zschopau River (Germany; 1544 km^2; 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 $\varepsilon_{\mathrm{IE}\Delta 1}$ ($+-++$; Fig. 6n). This shift aligns with a gradual decrease in the 20-year seasonality index (SI) (Fig. S3e) of liquid precipitation input (i.e. rainfall + snowmelt). In the Zschopau River catchment, this decrease in SI towards the end of the century signifies a shift towards a more evenly distributed precipitation pattern. These changes coincide with an increase in forest cover towards the end of century, as estimated from Hilda+ data (Fig. S3b). Additionally, Renner et al. (2014) and Renner and Hauffe (2024) reported a gradual recovery of forests in the Zschopau catchment during this period, which may further contribute to the observed shift.

As can be seen in Fig. 7a, the time stability of the studied 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.

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

By aggregating its j individual distributions, a long-term marginal distribution of $\varepsilon_{\mathrm{IE}\omega}$ for each catchment was built. For a large majority of catchments, the long-term median $\varepsilon_{\mathrm{IE}\omega}$ value remains very close to zero. More specifically, ~ 50 % of all study catchments are characterized by a median deviation $\varepsilon_{\mathrm{IE}\omega}$ value 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 a 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 of a wider spread with an average IQR ~ 0.10 while still centring closely around zero. This can also clearly be seen by 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 and f). In contrast, while also featuring a marginal distribution with a median deviation $\varepsilon_{\mathrm{IE}\omega} \sim 0.021$, 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 in the higher IQR of ~ 0.113. Three additional illustrative examples of well-known river basins are presented in Fig. S5 in the Supplement. In contrast, the analysis, which uses the earliest available period as a fixed baseline, shows an increase in the number of stable catchments along with a slightly higher median $\varepsilon_{\mathrm{IE}\omega}$ values. Further details are provided in the Supplement (Fig. S6a–f).

Overall, it can be observed that median deviation $\varepsilon_{\mathrm{IE}\omega}$ values 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 $\varepsilon_{\mathrm{IE}\omega}$. While for many catchments in the central US and southern Brazil median deviations are negative, i.e. $\varepsilon_{\mathrm{IE}\omega} < 0$, the rest of the study catchments globally are dominated by $\varepsilon_{\mathrm{IE}\omega} > 0$. Overall, median deviations from the expected parametric Budyko curve resulted in regionally distinct relative changes in Q across the studied catchments, with around ~ 68 % of the catchments exhibiting changes (ΔQ) of less than ± 10 % (Fig. 7c). However, catchments in some regions, notably in the central US and southern Africa, can be characterized by ΔQ exceeding ± 25 %. Overall, the results indicate that catchments in more arid regions ($I_{\mathrm{A}} > 2$) are particularly susceptible to relative changes in discharge as compared to more humid regions (inset Fig. 7c).

For a more regional evaluation, the yearly $\varepsilon_{\mathrm{IE}\omega}$ values for individual catchments were aggregated into regional marginal distributions of $\varepsilon_{\mathrm{IE}\omega}$ stratified according to the long-term mean aridity index (I_{A}) and varied latitude bands (Fig. 9a). These regional distributions capture the variability in yearly $\varepsilon_{\mathrm{IE}\omega}$ across regions, with the median $\varepsilon_{\mathrm{IE}\omega}$ serving as a robust measure of central tendency. The general pattern

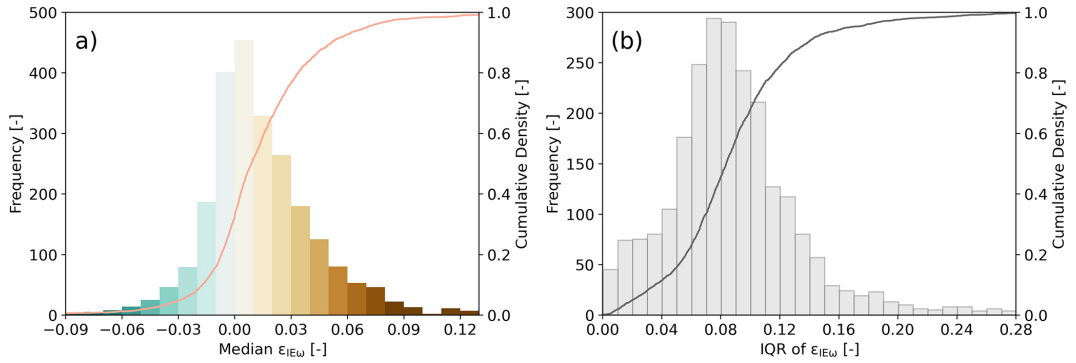

Figure 8. Visualization of long-term **(a)** median $\varepsilon_{IE\omega}$ and **(b)** interquartile range (IQR) of $\varepsilon_{IE\omega}$ for aggregated long-term marginal distribution of $\varepsilon_{IE\omega}$ across all catchments (2387) along with the corresponding cumulative distribution function (CDF). The varying colour palette in Fig. 8a aligns with the palette used in Fig. 7b to maintain consistency. In Fig. 8b, a uniform colour is used since IQR values are all positive.

found across most regions with available data is broadly consistent. Overall, 16 out of 20 regions are characterized by median deviation $\varepsilon_{IE\omega}$ values 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 $\pm 30°$, 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 the 50–90° 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 provided in Ibrahim et al. (2024, https://doi.org/10.5281/zenodo.14060926) 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 hydroclimatic conditions.

3.5 Sensitivity of marginal distributions of deviation $\varepsilon_{IE\omega}$ values 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 sensitivity of the marginal distributions of $\varepsilon_{IE\omega}$ to the choice of time periods. The differences between 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. 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.

However, in the 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 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 marginal distributions to be 0.029.

Overall, it was found that for 78 % of the study catchments in 15 out of 20 time windows, i.e. 75 %, feature median deviation $\varepsilon_{IE\omega}$ values within an interval of ± 0.035. This further suggests that, although some sensitivity to the choice of time period 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, e.g. 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 for the observation that there is no systematic larger-scale effect.

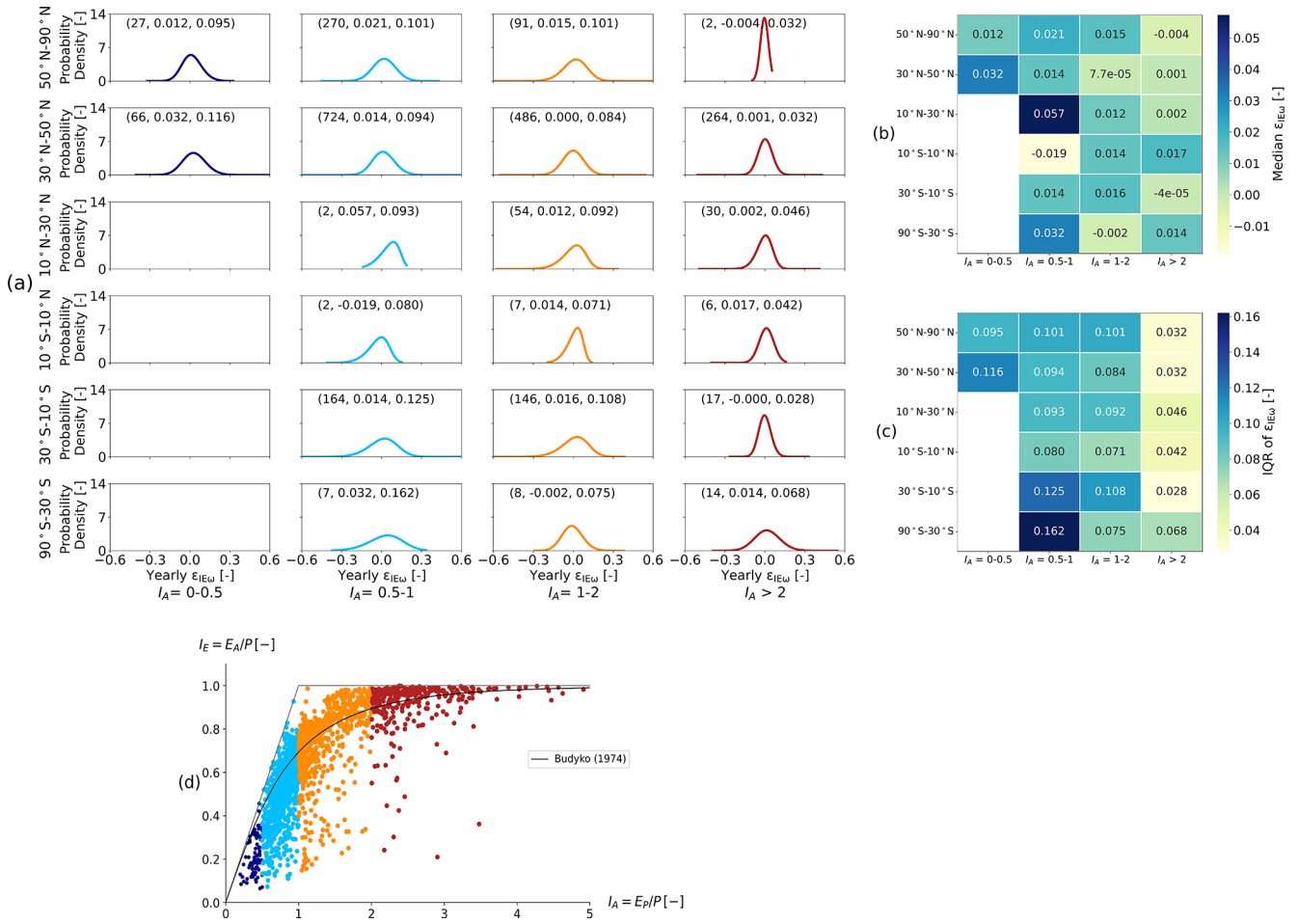

Figure 9. (a) Regional marginal distributions of $\varepsilon_{\mathrm{IE}\omega}$ for defined latitude and I_{A} bins. The three numerical values in small brackets at the top of each panel presents number of catchments in that category, long-term median $\varepsilon_{\mathrm{IE}\omega}$, and IQR value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

4 Discussion

Our analysis revealed that most study catchments underwent continuous multi-decadal hydroclimatic fluctuations throughout the 20th century (Figs. 4 and S1). Notably, these fluctuations were largely consistent across the different temporal stability categories. Unlike previous studies comparing only two time periods (Jaramillo and Destouni, 2014), here the higher temporal resolution, 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 rather than trends over longer timescales.

In alignment with previous studies (Berghuijs and Woods, 2016; Jaramillo et al., 2018, 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 deviation $\varepsilon_{\mathrm{IE}\omega}$ values across all study catchments throughout the 20th century are very minor, with median $\varepsilon_{\mathrm{IE}\omega} \leq \pm 0.015$ 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 absolute mean deviations from the expected I_{E} of $\varepsilon_{\mathrm{IE}\omega} \sim 0.01$–$0.02$ over two multi-decade periods, 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 spa-

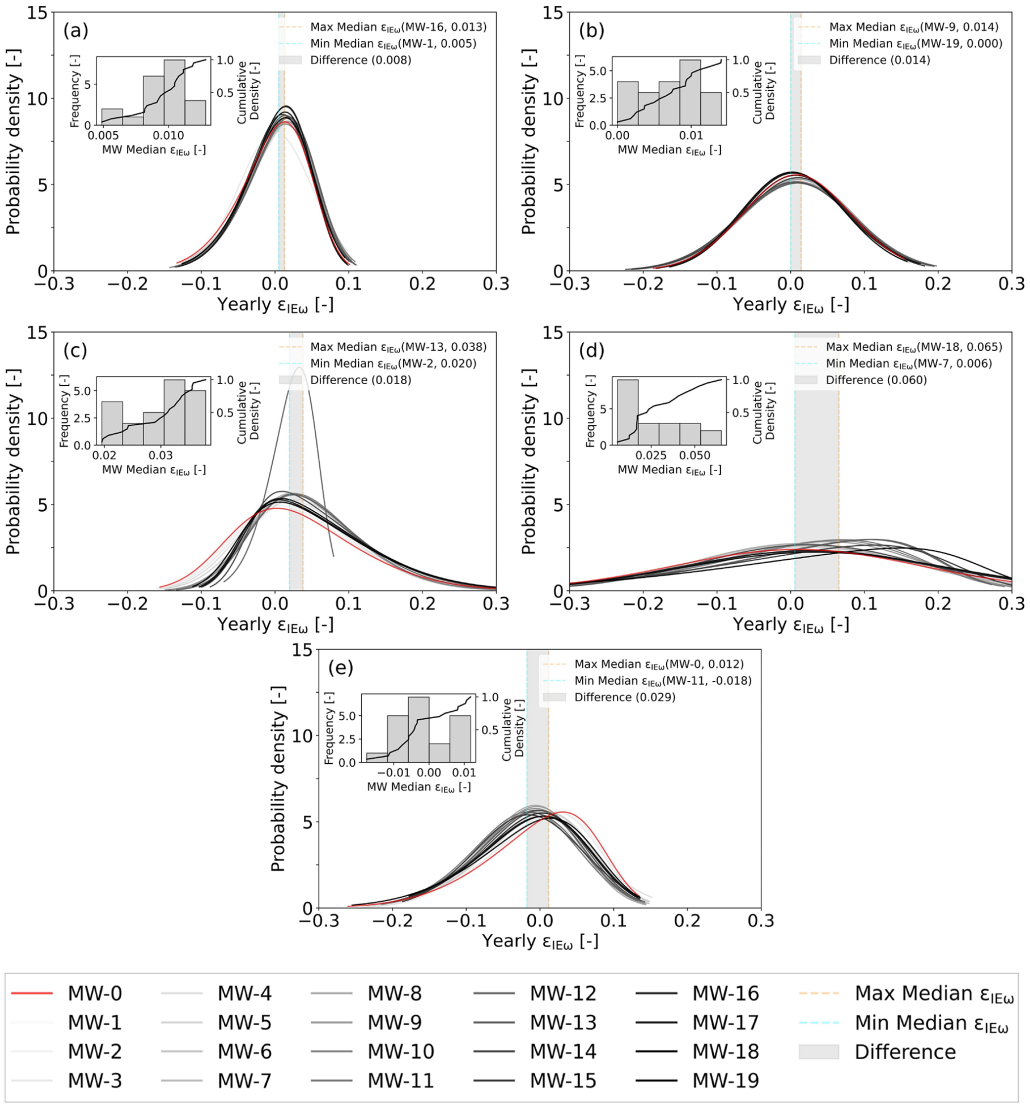

Figure 10. Aggregated marginal distributions of $\varepsilon_{\mathrm{IE}\omega}$ for 20 moving-window time periods for five example catchments: **(a)** Chemung River, **(b)** Lee River, **(c)** Sava River, **(d)** Kaituna River, and **(e)** Zschopau River. The red line represents the original marginal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

tial and 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 $\Delta I_{\mathrm{A}} = 0.2$ (Jaramillo and Destouni, 2014), such a difference (28 %) in sensitivity leads to only minor absolute deviations from the expected I_{E} with $\varepsilon_{\mathrm{IE}\omega} \sim 0.01$–$0.04$ (4 %–8 %) for regions with the most common $I_{\mathrm{A}} = 0.5$–2.5, which broadly corresponds to the results of our study. In contrast, Reaver et al. (2022), using data from ~ 700 catchments from the CAMELS-US (Newman et al., 2015) and the CAMELS-UK datasets (Addor et al., 2017; Coxon et al., 2020), 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 of $I_{\mathrm{E,o}}$ from the predicted values of I_{E} by catchment-specific curves, with $\varepsilon_{\mathrm{IE}\omega,\mathrm{max}} = I_{\mathrm{E,o,max}} - I_{\mathrm{E,max}} \sim 26\,\%$. 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 $\varepsilon_{\mathrm{IE}\omega,\mathrm{max}} \sim 9\,\%$ (see Fig. 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 fre-

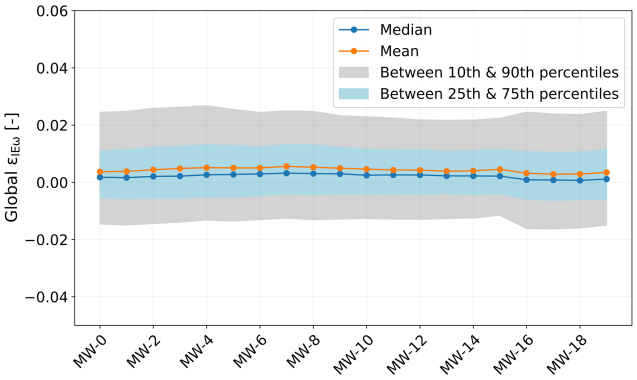

Figure 11. Variation in global long-term median, mean, 10th and 90th percentiles, and 25th and 75th percentiles of $\varepsilon_{IE\omega}$ for all of the catchments with respect to each moving window.

quently. By doing so, it is plausible to assume that the deviation $\varepsilon_{IE\omega}$ values 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 (Fig. 7c). For example, in a humid catchment with $I_A = 0.5$ (e.g. mean annual $P = 2000\,\mathrm{mm\,yr^{-1}}$, $E_P = 1000\,\mathrm{mm\,yr^{-1}}$ and $Q = 1120\,\mathrm{mm\,yr^{-1}}$), a deviation of $\varepsilon_{IE\omega} = 0.02$ results in $\Delta Q \sim 40\,\mathrm{mm\,yr^{-1}}$, 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. $P = 500\,\mathrm{mm\,yr^{-1}}$, $E_P = 1000\,\mathrm{mm\,yr^{-1}}$, $Q = 60\,\mathrm{mm\,yr^{-1}}$), the same deviation will lead to a $\Delta Q \sim 10\,\mathrm{mm\,yr^{-1}}$, equivalent to $\sim 15\,\%$ of the available water yield, and thus have considerable higher relevance for water resources 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 deviation $\varepsilon_{IE\omega}$ values from the expected parametric Budyko curves do not exhibit any clear and unambiguous relationships with several climatic variables (Fig. S7 in the Supplement). 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; and also violations of the assumption that $\mathrm{d}S/\mathrm{d}t \sim 0$ 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 us 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 hydroclimatic 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 $\sim 70\,\%$ 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 for some confidence to plausibly assume that $\varepsilon_{IE\omega}$ and the associated I_E under projected future hydroclimatic conditions can, at least for several decades, be robustly predicted based on these distributions. However, it is important to note that the 20-year time periods used in this study, while effective for medium-term projections, may limit the ability to make long-term climate projections.

Further 19 % of catchments were classified as variable because 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 for 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) either depend on I_E of the preceding 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 constructing and using 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 conditional 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 (Montanari and Koutsoyiannis, 2014). For predictions in catchments with a 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 as both categories are likely to lead to rather unreliable future estimates.

In any case, empirical models like the Budyko framework have inherent weakness in dealing with previously unseen changes in the underlying distribution of a specific variable (here $\varepsilon_{IE\omega}$). Our classification of catchments into stable, variable, alternating, and shift categories aims to capture varying levels of sensitivity to changes in underlying distributions. Catchments classified as alternating or shift are more likely to have experienced large changes in the underlying distributions and may thus remain sensitive to future changes, making empirical models less robust for predictions in these cases. Conversely, stable and variable catchments exhibited much less sensitivity to past climatic variability. In the absence of statistical evidence for changing distributions, it is reasonable to assume that they remain relatively insensitive to change in the near future, allowing empirical models to provide plausible predictions.

It is important to note that approximately 89 % of the study catchments are either stable or variable, with only a small minority ($\sim$ 11 %) exhibiting alternating or shift behaviour. This predominance of stable or variable catchments supports the broad applicability of the Budyko framework for predictive purposes. Although there is no clear spatial pattern, the regional distributions of $\varepsilon_{IE\omega}$ remain, with medians of $\sim$ 0–0.02 (Fig. 9a), broadly consistent with the global distribution (Fig. 8a) but also with each other across most spatial and climatic classes. This indeed suggests that the overall pattern is rather homogenous and regional effects remain limited, making probabilistic predictions feasible in the absence of a deterministic description (Montanari and Koutsoyiannis, 2014). Thus, the presented distributions (Figs. 8a and 9a) are in the absence of further information useful to quantitatively estimate the uncertainty for any specific catchment based on past information in a probabilistic way. However, caution is advised for out-of-sample catchments, where the assumption of stationarity may lead to less reliable predictions as the framework cannot take into account systematic shifts or alternating behaviour.

Despite the challenges associated with catchments classified as alternating and shift, the Budyko framework remains useful for identifying human-driven changes to the water cycle. Although many catchments showed only minor deviations, these deviations are key for recognizing drivers of change. Categorizing catchments into stable, variable, alternating, and shift can guide targeted future research. For example, catchments in the alternating and shift categories my either have been, in the past, subject to more substantial human interference than those in the other categories or be more sensitive to human-induced changes. Further investigations into the drivers of these deviations may strengthen our understanding of how human-induced changes influence catchment responses differently in different environments.

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 an additional robustness flag for each catchment in Supplementary data downloadable from a Zenodo repository (https://doi.org/10.5281/zenodo.14060926, Ibrahim et al., 2024). Lower IQR then indicates 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 with associated changes in Q for each of the 2387 study catchments, but also of the regional distributions as stratified by latitude and I_A, is provided in Ibrahim et al. (2024, https://doi.org/10.5281/zenodo.14060926). 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 hydroclimatic and streamflow data for 2387 river catchments worldwide, we test here whether catchments follow their specific parametric Budyko curves as defined by parameter ω over multiple 20-year periods throughout the 20th century.

The main findings of our analysis are as follows:

1. Overall, 62 % of the catchments do not significantly deviate from their expected parametric Budyko curves, although minor deviations were still observed. However, this also entails that a fraction of 38 % does indeed deviate.

2. The overall magnitude of deviations is minor. For $\sim$ 70 % of the catchments, the median deviations do not exceed $\varepsilon_{IE\omega} = \pm 0.025$, which is equivalent to $\sim$ 1 %–4 %, depending on I_E. These median $\varepsilon_{IE\omega}$ values, when expressed as relative changes in Q, result in less than a $\pm$10 % change in discharge for most catchments.

3. For 89 % of the study catchments, $\varepsilon_{\mathrm{IE}\omega}$ can be considered highly or at least moderately well predictable based on historical data as distributions of $\varepsilon_{\mathrm{IE}\omega}$ in the past were shown to be stable over multiple time periods or characterized by variable fluctuations. The framework works well for most catchments; however, for out-of-sample catchments showing systematic shifts or alternating behaviour, additional analysis may be required.

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.

Data availability. Daily precipitation and temperature data were acquired via the GSWP-3 dataset, accessible at https://doi.org/10.48364/ISIMIP.886955 (Lange and Büchner, 2020). GSIM discharge data were obtained from https://doi.org/10.1594/PANGAEA.887477 (Do et al., 2018a, b) and https://doi.org/10.1594/PANGAEA.887470 (Gudmundsson et al., 2018a, b). Supplementary data containing topographic and climatic characteristics, parameters for fitted 20-year parametric distributions of deviations, robustness flags indicating the temporal stability of aggregated marginal distributions, and associated changes in Q across 2387 study catchments are available at https://doi.org/10.5281/zenodo.14060926 (Ibrahim et al., 2024). Regional parameters for fitted parametric marginal distributions are also provided.

Supplement. The supplement related to this article is available online at https://doi.org/10.5194/hess-29-1703-2025-supplement.

Author contributions. MI, MH, MCG, and RvdE conceptualized the study. MI conducted the formal analysis and prepared the paper in discussion and with input from MCG, RvdE, and MH.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Hydrology and Earth System Sciences*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. The authors greatly appreciate the constructive and detailed comments by two anonymous reviewers which significantly improved the manuscript. We also thank Muhammad Fraz Ismail and Fransje Van Oorschot for their support and guidance in resolving coding challenges during this work.

Financial support. This research has been supported by the Higher Education Commission of Pakistan (grant no. Ref:1(2)/HRD/OSS-III/2021/HEC/19607).

Review statement. This paper was edited by Nunzio Romano and reviewed by two anonymous referees.

References

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrol. Earth Syst. Sci., 21, 5293–5313, https://doi.org/10.5194/hess-21-5293-2017, 2017.

Andréassian, V., Mander, Ü., and Pae, T.: The Budyko hypothesis before Budyko: The hydrological legacy of Evald Oldekop, J. Hydrol., 535, 386–391, https://doi.org/10.1016/j.jhydrol.2016.02.002, 2016.

Arora, V. K.: The use of the aridity index to assess climate change effect on annual runoff, J. Hydrol., 265, 164–177, https://doi.org/10.1016/S0022-1694(02)00101-4, 2002.

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

Berghuijs, W. R., Woods, R. A., and Hrachowitz, M.: A precipitation shift from snow towards rain leads to a decrease in streamflow, Nat. Clim. Change, 4, 583–586, https://doi.org/10.1038/nclimate2246, 2014.

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

Beven, K.: Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication, Hydrolog. Sci. J., 61, 1652–1665, https://doi.org/10.1080/02626667.2015.1031761, 2016.

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.: Ecosystem adaptation to climate change: the sensitivity of hydrological predictions to time-dynamic model parameters, Hydrol. Earth Syst. Sci., 26, 1295–1318, https://doi.org/10.5194/hess-26-1295-2022, 2022.

Budyko, M. I.: Evaporation under natural conditions, Gidrometeorizdat, Leningrad, English translation by IPST, Jerusalem, 1948.

Budyko, M. I.: The heat balance of the earth's surface, Sov. Geogr., 2, 3–13, 1961.

Choudhury, B.: Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model, J. Hydrol., 216, 99–110, https://doi.org/10.1016/S0022-1694(98)00293-5, 1999.

Coxon, G., Addor, N., Bloomfield, J. P., Freer, J., Fry, M., Hannaford, J., Howden, N. J. K., Lane, R., Lewis, M., Robinson, E. L., Wagener, T., and Woods, R.: CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain, Earth Syst. Sci. Data, 12, 2459–2483, https://doi.org/10.5194/essd-12-2459-2020, 2020.

Daly, E., Calabrese, S., Yin, J., and Porporato, A.: Hydrological Spaces of Long-Term Catchment Water Balance, Water Resour. Res., 55, 10747–10764, https://doi.org/10.1029/2019wr025952, 2019.

Destouni, G., Jaramillo, F., and Prieto, C.: Hydroclimatic shifts driven by human water use for food and energy production, Nat. Clim. Change, 3, 213–217, https://doi.org/10.1038/nclimate1719, 2013.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel analysis and implications for our perception of the land surface, B. Am. Meteorol. Soc., 87, 1381–1398, https://doi.org/10.1175/BAMS-87-10-1381, 2006.

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

Do, H. X., Gudmundsson, L., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive – Part 1: Station catalog and Catchment boundary, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.887477 (last access: 17 March 2025), 2018b.

Donohue, R. J., Roderick, M. L., and McVicar, T. R.: On the importance of including vegetation dynamics in Budyko's hydrological model, Hydrol. Earth Syst. Sci., 11, 983–995, https://doi.org/10.5194/hess-11-983-2007, 2007.

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

Duffie, J. A. and Beckman, W. A.: Solar engineering of thermal processes, Wiley New York,, 762 pp., ISBN 978-0471050667, 1980.

Fu, B.: On the calculation of the evaporation from land surface, Sci. Atmos. Sin., 5, 23–31, 1981.

Gan, G., Liu, Y., and Sun, G.: Understanding interactions among climate, water, and vegetation with the Budyko framework, Earth-Sci. Rev., 212, 103451, https://doi.org/10.1016/j.earscirev.2020.103451, 2021.

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., and Savenije, H. H. G.: Climate controls how ecosystems size the root zone storage capacity at catchment scale, Geophys. Res. Lett., 41, 7916–7923, https://doi.org/10.1002/2014gl061668, 2014.

Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G., and Salvucci, G.: Interdependence of climate, soil, and vegetation as constrained by the Budyko curve, Geophys. Res. Lett., 39, L19404, https://doi.org/10.1029/2012gl053492, 2012.

Greve, P., Gudmundsson, L., Orlowsky, B., and Seneviratne, S. I.: Introducing a probabilistic Budyko framework, Geophys. Res. Lett., 42, 2261–2269, https://doi.org/10.1002/2015gl063449, 2015.

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

Gudmundsson, L., Do, H. X., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive (GSIM) – Part 2: Time Series Indices and Homogeneity Assessment, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.887470 (last access: 17 March 2025), 2018b.

Han, J., Yang, Y., Roderick, M. L., McVicar, T. R., Yang, D., Zhang, S., and Beck, H. E.: Assessing the Steady-State Assumption in Water Balance Calculation Across Global Catchments, Water Resour. Res., 56, e2020WR027392, https://doi.org/10.1029/2020wr027392, 2020.

Hargreaves, G. H. and Samani, Z. A.: Estimating potential evapotranspiration, J. Irr. Drain. Div.-ASCE, 108, 225–230, 1982.

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.: Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins, Climatic Change, 141, 561–576, https://doi.org/10.1007/s10584-016-1829-4, 2017.

Hrachowitz, M., Stockinger, M., Coenders-Gerrits, M., van der Ent, R., Bogena, H., Lücke, A., and Stumpp, C.: Reduction of vegetation-accessible water storage capacity after deforestation affects catchment travel time distributions and increases young water fractions in a headwater catchment, Hydrol. Earth Syst. Sci., 25, 4887–4915, https://doi.org/10.5194/hess-25-4887-2021, 2021.

Ibrahim, M., Coenders-Gerrits, M., van der Ent, R., and Hrachowitz, M.: Catchments do not strictly follow Budyko curves over multiple decades but deviations are minor and predictable, Zenodo [data set], https://doi.org/10.5281/zenodo.14060926 (last access: 17 March 2025), 2024.

Jaramillo, F. and Destouni, G.: Developing water change spectra and distinguishing change drivers worldwide, Geophys. Res. Lett., 41, 8377–8386, https://doi.org/10.1002/2014gl061848, 2014.

Jaramillo, F. and Destouni, G.: Local flow regulation and irrigation raise global human water consumption and footprint, Science, 350, 1248–1251, https://doi.org/10.1126/science.aad1010, 2015.

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

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

Lange, S. and Büchner, M.: ISIMIP2a atmospheric climate input data, ISIMIP Repository [data set], https://doi.org/10.48364/ISIMIP.886955, 2020.

Lee, T.-Y., Chiu, C.-C., Chen, C.-J., Lin, C.-Y., and Shiah, F.-K.: Assessing future availability of water resources in Taiwan

based on the Budyko framework, Ecol. Indic., 146, 109808, https://doi.org/10.1016/j.ecolind.2022.109808, 2023.

Levi, L., Jaramillo, F., Andricevic, R., and Destouni, G.: Hydroclimatic changes and drivers in the Sava River Catchment and comparison with Swedish catchments, Ambio, 44, 624–634, https://doi.org/10.1007/s13280-015-0641-0, 2015.

Liu, H., Wang, Z., Ji, G., and Yue, Y.: Quantifying the Impacts of Climate Change and Human Activities on Runoff in the Lancang River Basin Based on the Budyko Hypothesis, Water, 12, 3501, https://doi.org/10.3390/w12123501, 2020.

Mezentsev, V.: More on the calculation of average total evaporation, Meteorologiya i Gidrologiya, 5, 24–26, 1955.

Mianabadi, A., Davary, K., Pourreza-Bilondi, M., and Coenders-Gerrits, A. M. J.: Budyko framework; towards non-steady state conditions, J. Hydrol., 588, 125089, https://doi.org/10.1016/j.jhydrol.2020.125089, 2020.

Milly, P.: Climate, soil water storage, and the average annual water balance, Water Resour. Res., 30, 2143–2156, https://doi.org/10.1029/94WR00586, 1994.

Montanari, A. and Koutsoyiannis, D.: Modeling and mitigating natural hazards: Stationarity is immortal!, Water Resour. Res., 50, 9748–9756, https://doi.org/10.1002/2014wr016092, 2014.

Nearing, G. S., Tian, Y., Gupta, H. V., Clark, M. P., Harrison, K. W., and Weijs, S. V.: A philosophical basis for hydrological uncertainty, Hydrolog. Sci. J., 61, 1666–1678, https://doi.org/10.1080/02626667.2016.1183009, 2016.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance, Hydrol. Earth Syst. Sci., 19, 209–223, https://doi.org/10.5194/hess-19-209-2015, 2015.

Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., Han, D., Wagener, T., McGuire, K., Savenije, H., and Hrachowitz, M.: The evolution of root-zone moisture capacities after deforestation: a step towards hydrological predictions under change?, Hydrol. Earth Syst. Sci., 20, 4775–4799, https://doi.org/10.5194/hess-20-4775-2016, 2016.

Oldekop, E.: Collection of the Works of Students of the Meteorological Observatory, University of TartuJurjew-Dorpat Tartu, Estonia, p. 209, 1911.

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

Porporato, A., Daly, E., and Rodriguez-Iturbe, I.: Soil water balance and ecosystem response to climate change, Am. Nat., 164, 625–632, https://doi.org/10.1086/424970, 2004.

Reaver, N. G. F., Kaplan, D. A., Klammler, H., and Jawitz, J. W.: Theoretical and empirical evidence against the Budyko catchment trajectory conjecture, Hydrol. Earth Syst. Sci., 26, 1507–1525, https://doi.org/10.5194/hess-26-1507-2022, 2022.

Renner, M. and Hauffe, C.: Impacts of climate and land surface change on catchment evapotranspiration and runoff from 1951 to 2020 in Saxony, Germany, Hydrol. Earth Syst. Sci., 28, 2849–2869, https://doi.org/10.5194/hess-28-2849-2024, 2024.

Renner, M., Brust, K., Schwärzel, K., Volk, M., and Bernhofer, C.: Separating the effects of changes in land cover and climate: a hydro-meteorological analysis of the past 60 yr in Saxony, Germany, Hydrol. Earth Syst. Sci., 18, 389–405, https://doi.org/10.5194/hess-18-389-2014, 2014.

Renofalt, B. M., Jansson, R., and Nilsson, C.: Effects of hydropower generation and opportunities for environmental flow management in Swedish riverine ecosystems, Freshwater Biol., 55, 49–67, https://doi.org/10.1111/j.1365-2427.2009.02241.x, 2010.

Roderick, M. L. and Farquhar, G. D.: A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties, Water Resour. Res., 47, W00G07, https://doi.org/10.1029/2010wr009826, 2011.

Schreiber, P.: Über die Beziehungen zwischen dem Niederschlag und der Wasserführung der Flüsse in Mitteleuropa, Z. Meteorol., 21, 441–452, 1904.

Serpa, D., Nunes, J., Santos, J., Sampaio, E., Jacinto, R., Veiga, S., Lima, J., Moreira, M., Corte-Real, J., and Keizer, J.: Impacts of climate and land use changes on the hydrological and erosion processes of two contrasting Mediterranean catchments, Sci. Total Environ., 538, 64–77, https://doi.org/10.1016/j.scitotenv.2015.08.033, 2015.

Shao, Q., Traylen, A., and Zhang, L.: Nonparametric method for estimating the effects of climatic and catchment characteristics on mean annual evapotranspiration, Water Resour. Res., 48, W03517, https://doi.org/10.1029/2010wr009610, 2012.

Sterling, S. M., Ducharne, A., and Polcher, J.: The impact of global land-cover change on the terrestrial water cycle, Nat. Clim. Change, 3, 385–390, https://doi.org/10.1038/nclimate1690, 2012.

Tempel, N., Bouaziz, L., Taormina, R., van Noppen, E., Stam, J., Sprokkereef, E., and Hrachowitz, M.: Catchment response to climatic variability: implications for root zone storage and streamflow predictions, Hydrol. Earth Syst. Sci., 28, 4577–4597, https://doi.org/10.5194/hess-28-4577-2024, 2024.

Tixeront, J.: Prévision des apports des cours d'eau, Publication de l'Association internationale d'hydrologie scientifique, 63, 118–126, http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=19271864 (last access: 17 March 2024), 1964.

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

Turc, L.: Le bilan d'eau des sols: Relation entre la precipitations, l'evaporation et l'ecoulement, Annales Agronomiques, Série A, 5, 491–596, 1954.

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

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

Wang, D. and Hejazi, M.: Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow

in the contiguous United States, Water Resour. Res., 47, W00J12, https://doi.org/10.1029/2010wr010283, 2011.

Wang, S., Hrachowitz, M., and Schoups, G.: Multi-decadal fluctuations in root zone storage capacity through vegetation adaptation to hydro-climatic variability have minor effects on the hydrological response in the Neckar River basin, Germany, Hydrol. Earth Syst. Sci., 28, 4011–4033, https://doi.org/10.5194/hess-28-4011-2024, 2024.

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

Winkler, K., Fuchs, R., Rounsevell, M., and Herold, M.: Global land use changes are four times greater than previously estimated, Nat. Commun., 12, 2501, https://doi.org/10.1038/s41467-021-22702-2, 2021.

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

Xu, X., Liu, W., Scanlon, B. R., Zhang, L., and Pan, M.: Local and global factors controlling water-energy balances within the Budyko framework, Geophys. Res. Lett., 40, 6123–6129, https://doi.org/10.1002/2013gl058324, 2013.

Yang, H., Yang, D., Lei, Z., and Sun, F.: New analytical derivation of the mean annual water-energy balance equation, Water Resour. Res., 44, W03410, https://doi.org/10.1029/2007wr006135, 2008.

Zhang, L., Dawes, W. R., and Walker, G. R.: Response of mean annual evapotranspiration to vegetation changes at catchment scale, Water Resour. Res., 37, 701–708, https://doi.org/10.1029/2000wr900325, 2001.