André S. Ballarin[1], Paulo Tarso S. Oliveira[1,2], Bruno K. Marchezepe[1], Raquel Godoi[2], Aline M. Campos[1], Fabrizio S. Campos[1], André Almagro[2], and Antônio A. Meira Neto[3]

[1] Department of Hydraulics and Sanitary Engineering, São Carlos School of Engineering, University of São Paulo, São Carlos, SP, Brazil

[2] Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul, MS, 79070–900, Brazil.

[3]Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, Arizona, USA.

Corresponding author: A.S. Ballarin (andre.ballarin@usp.br)

Email addresses/ORCID:

André S. Ballarin - andre.ballarin@usp.br, ORCID: 0000-0001-6997-8662

Paulo Tarso S. Oliveira - paulotarsoms@gmail.com, ORCID: 0000-0003-2806-0083

Bruno K. Marchezepe - bruno.marchezepe@usp.br, ORCID: 0000-0001-9266-5015

Raquel Godoi - rachgodoi@gmail.com, ORCID: 0000-0003-1869-1850

Aline M. Campos - alinemarcampos96@gmail.com, ORCID: 0000-0002-6074-6256

Fabrizio S. Campos - fabrizio.campos@usp.br, ORCID: 0000-0003-1959-7508

André Almagro - andre.almagro@gmail.com, ORCID: 0000-0002-3822-4865

Antônio A. Meira Neto - antoniomeira@gmail.com, ORCID: 0000-0002-7508-8241

- Key Points:
- We assessed the controls on long-term streamflow components under an open water balance assumption
- Inclusion of inter-catchment groundwater flow significantly improves the performance of the aridity-based long-term streamflow formulations
- Partitioning streamflow into baseflow and quickflow improves the understanding of the water-balance main control mechanisms

Abstract

Understanding how streamflow and its components, baseflow and quickflow, vary spatially according to climate and landscape characteristics is fundamental for dealing with different water-related issues. Analytical formulations have been proposed to investigate their long-term behavior and additional influencing factors, suggesting that they are mainly controlled by the aridity index (ϕ). Nevertheless, these studies assume the catchment as a closed water balance system, neglecting inter-catchment groundwater flow (IGF). This simplification makes the analysis of the long-term streamflow components and their main control mechanisms challenging, given that many catchments cannot be considered as closed hydraulic entities. Here, we assessed the controls of the mean-annual streamflow components and their behavior under an open water balance assumption, using observed data of 731 Brazilian catchments with diverse hydroclimatic conditions. Our results indicate that indeed streamflow components are primarily controlled by ϕ at the mean annual timescale. The consideration of an open water-balance significantly improved the performance of the functional forms to describe streamflow components while also elucidating the assessment of other influencing factors on the streamflow behavior. Land cover, groundwater, climate seasonality and topographic attributes appeared as the main control mechanisms beyond aridity. Overall, our study provides new insights of the main control mechanism of the streamflow behavior at the mean-annual scale, while shedding light on the importance of the open water-balance assumption for model development and water resources management.

1. Introduction

Understanding catchment hydrological responses to rainfall (P), represented by the water-balance components streamflow (Q) and evapotranspiration (ET), and how it varies with climate and landscape properties is crucial for water-related issues and remains as one of the major challenges in hydrological studies (Gnann et al., 2019, Padron et al. 2017, Willian et al., 2012). However, additional insight on water balance partitioning can be gained by looking into the decomposition of streamflow int its components (Sivapalan et al., 2011). The two-step water balance proposed by L'vovich (1979), assumes that Q may be disaggregated into two primary components: direct runoff (Q_D) and baseflow (Q_B). The former represents the fast response of a catchment to a rainfall event, while the latter denotes the slow response, representing the water stored in the catchment system that supplies the total streamflow between rainfall events and during dry periods (Price, 2011; Meira Neto et al., 2020; Zhang and Schilling, 2006). As a quick response, Q_D is usually associated with flood and soil erosion hazards; while Q_B is related to water supply and aquifer recharge issues. Accurate knowledge of these components and how they are controlled by climate, land-use conditions, and catchments' physiographic characteristics is of primary importance to improve water resources management strategies, especially in a context of changing climate and increasing water demand (Miller et al., 2016; Santhi et al., 2008; Wada et al., 2016).

The investigation of this interplay between streamflow components, climate, and catchment properties in the long-term water balance is commonly done through the Budyko framework (Budyko, 1974; Cuthbert et al., 2019; Gnann et al., 2019; Li et al., 2013; Liu and You, 2021; Meira Neto et al., 2020; Wang and

Wu, 2013). This hypothesis considers the partitioning of P into ET and Q to be a function of ϕ (Meira et al., 2020; Wang and Wu, 2013), formalizing this relationship into the Budyko curve, which describes the long-term relationship between the actual and potential evaporative fractions, ET$/P$ and PET$/P$. This framework has been widely used to evaluate the main control mechanisms of water-balance components, as it provides a useful way to compare the hydrologic fluxes of different watersheds for a specific amount of available water and energy (Barnhart et al., 2016). By analyzing how catchments with a specific property deviates from the Budyko curve or how this specific property correlates with the parametric version of the Budyko equation, it can be inferred how and how much this particular property affects the water balance components beyond ϕ (Berghuijs et al., 2020).

Sivapalan et al. (2011), through the mathematical formulation of the two-stage water balance proposed by L'vovich (1979) and extended by Ponce and Shetty (1995), characterized the spatial and temporal variability of the water balance components for over 300 catchments across the United States. Despite their interesting findings of both regional and temporal patterns, the authors did not associate their results with climate or landscape characteristics (Meira Neto et al., 2020). Gnann et al. (2019) utilized the L'vovich (1979) and Ponce-Shetty frameworks (Ponce & Shetty, 1995; Sivapalan et al., 2011), to propose a numerical model for Baseflow Coefficient (BFC = Q_B/P), reporting that the aridity Index (ϕ), which is usually used to express climate conditions by the ratio between energy and water availability (ϕ = PET$/P$), does not satisfactorily explains the geographical distribution of the BFC in humid catchments, but it becomes more important for arid catchments. Meira Neto et al. (2020) also investigated the role of the ϕ on the long-term (mean-annual) Q_D and Q_B, using the L'vovich and Budyko frameworks for 378 catchments within the conterminous United States. In contrast to Gnann et al. (2021), the authors identified ϕ as a suitable predictor of the spatial variability of mean-annual Q_B, suggesting that more studies considering humid catchments are necessary to investigate whether the pattern observed in Gnann et al. (2019) can be seen as the expression of an even higher variability of baseflow fraction over very low ϕ values. Likewise, Cheng et al. (2021) and Yao et al. (2021) proposed analytical formulations of BFC as a function of ϕ considering, additionally, the catchment's storage capacity in their framework. Their results suggest a good performance of their approaches to capture the mean-annual spatial variability of observed BFC across the conterminous United States, Australia and United Kingdom. In both studies, the storage capacity was shown to be an important factor for the prediction of Q_B at humid sites.

The Budyko and L'vovich approaches, however, assume a closed water balance system (Liu et al, 2020), meaning that water gains or losses through inter-catchment groundwater flows (IGF) are considered negligible (Gnann et al., 2019; Kampf et al., 2020). A direct consequence of the closed water-balance assumption is the use of the topographic catchment area (A_{topo}) rather than the effective catchment area (A_{eff}) - an equivalent area proposed by Liu et al.

(2020) that considers inter-catchment groundwater exchanges by evaluating the deviations between recharge and discharge rates - as the basic spatial unit to compute hydrologic fluxes and to define possible mechanisms that may affect the catchment hydrological response (Wagener et al., 2007). Nevertheless, there is strong evidence that catchments are “leaky” and cannot be considered as closed hydrologic entities, as they may lose or gain water through different mechanisms such as IGF, transferring water beyond their topographic boundaries (Fan, 2019; Liu et al., 2021; Muñoz et al., 2016; Schwamback et al., 2022). In fact, given that hydrological connections between neighboring catchments do exist, these cases may be the rule rather than the exception (Le Moine et al., 2007). That is, A_{topo} may be often different from A_{eff} (Fan et al., 2019; Liu et al., 2020; Safeeq et al., 2021).

Schaller and Fan (2009) when evaluating the exporter/importer condition of 1555 catchments across the United States found that catchments might export (import) up to 90% (50%) of their flow to other catchments through IGF. Schwamback et al. (2022) also reported that nearly 32% of a total of 733 Brazilian catchments showed more than 30% of difference between their A_{topo} and A_{eff}, concluding on the need to consider inter-catchment connectivity in hydrologic studies. Safeeq et al. (2021) investigated the water-balance closure assumptions of six river basins (>1000 km^2) and ten headwater catchments (< 5 km^2) in the southern Sierra Nevada. Their findings indicated IGF to be much higher than typical amount of measurement bias for both evaluated spatial scales, arguing for a greater consideration of inter-catchment groundwater fluxes when evaluating hydrological processes, especially for small spatial scales

Despite the importance of the hydrologic connectivity for catchment’s water balance computation and for hydrological implications, they are mostly neglected in hydrological studies (Bouaziz et al., 2018; Fan, 2019; Safeeq et al., 2021). To our knowledge, all the empirical evidence gathered so far at evaluating streamflow components and their main controls mechanisms assumed a closed water-balance assumption, considering that catchments are not connected to their surroundings (Bouazis et al., 2018, Liu et al., 2020). Such simplification may hamper the analysis of the long-term water balance components and hinder the understanding of how different mechanisms control their properties, given that catchments’ leakage violates the initial close water-balance assumption and may consequently hamper any empirical inference (Kampf et al., 2020).

Here, we discuss the effect of the assumption of a closed water balance system on the possible mechanisms that control the catchment’s hydrological responses to climate, following a two-step water balance formulation. We ask the following question: Does the consideration of IGF into the water balance tend to improve our understanding of the mechanisms controlling its components? To do that, we first investigate how ϕ-based formulations of Q_B and Q_D differ when using both A_{topo} and A_{eff} for the water balance closure of 731 Brazilian catchments. Further, we perform a simple assessment of the role of additional climatic and landscape catchment’s attributes as possible controls on long-term streamflow components under both A_{eff} and A_{topo} perspectives. The manuscript is organized as follows: In section 2, we describe the dataset used in the study. Section 3 shows the methods used in the study as well as the long-term water balance frameworks proposed by L'vovich (1979) and Budyko (1974). In sections 4, we present and discuss the results and, in section 5, we highlight the main conclusions of the paper.

1. Data

We used the Catchments Attribute for Brazil dataset (CABra), a large-scale dataset that includes a set of more than 100 attributes, divided into 8 classes (topography, climate, streamflow, groundwater, soil, geology, land-use and land-cover, and hydrologic disturbance) for 735 catchments distributed across the Brazilian territory (Figure 1). CABra includes daily time-series of streamflow, actual and potential evapotranspiration and precipitation for a 30-year period (1980-2010), with up to 10% of missing data. For a detailed description of the methodologies used to obtain the catchments' attributes, see Almagro et al. (2021a). The mean-annual potential evapotranspiration available in the CABra's dataset was computed following the Priestley-Taylor method (Priestley and Taylor (1972). We excluded from our analysis three catchments highly disturbed by human activities, which presented a large discrepancy between streamflow and precipitation (runoff coefficient $(\frac{Q}{P}) > 1$); and one catchment that exhibited inconsistency between precipitation and evapotranspiration data.

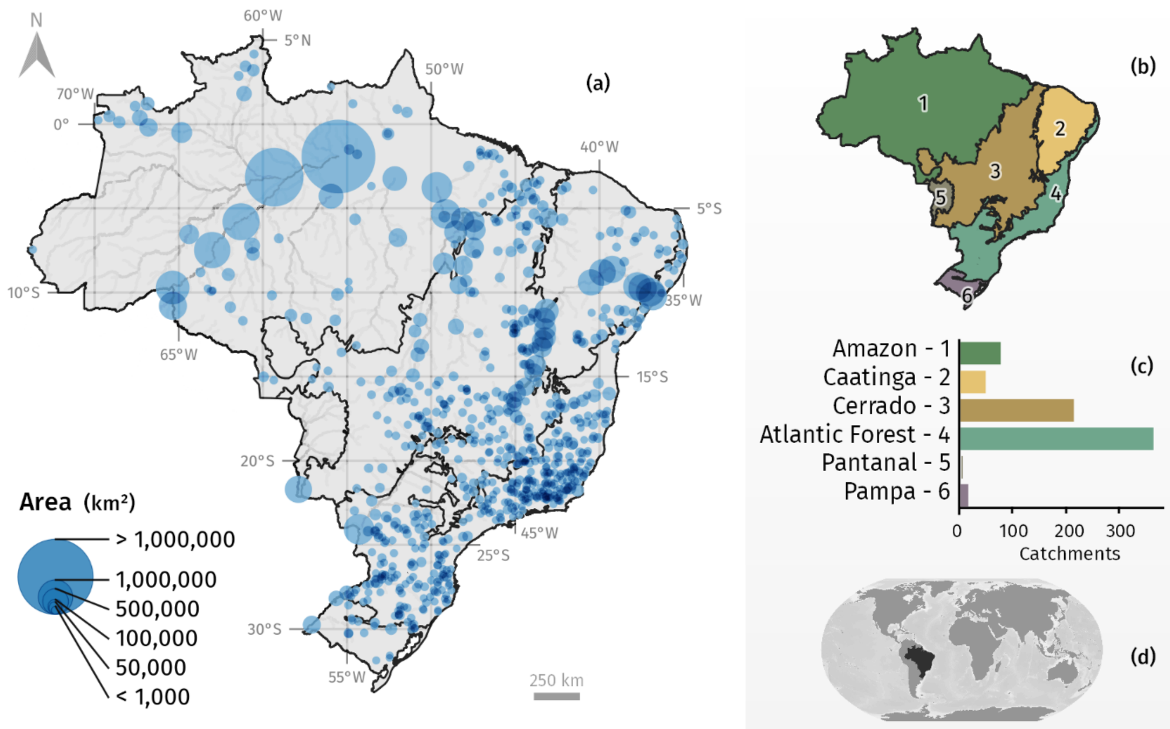

Figure 1 - (a) The 731 catchments of the CABra's dataset evaluated in this study. Dark gray lines indicate rivers. (b) Six main biomes of Brazil with the (c) distribution of evaluated catchments for each biome. (d) Brazil, in dark gray, in the global context.

To partition Q into Q_D and Q_B, we used the one parameter low-pass filter developed by Lyne and Hollick (1979), following previous studies that also evaluate these streamflow components (Gnann et al., 2019; Lucas et al., 2021; Meira Neto et al., 2020; Sivapalan et al., 2011; Trancoso et al., 2017). As well as Gnann et al. (2019), we applied the recursive filter in a 3-times running (forward-back-forward), setting the parameter to 0.925 for all catchments to allow for intercomparison. As pointed out by Meira Neto et al. (2020), there is an inherent uncertainty in this step, as the streamflow components obtained with the digital filter cannot be attributed to specific processes. For this reason, we also tested, for all evaluated catchments, different streamflow partitioning methods: Lyne and Hollick filter (one and three passes), Eckhard filter (Eckhard, 2005), and United Kingdom Institute of Hydrology smoothed minima method - 5 and 90 days (Institute of Hydrology, 1980). With the exception of the Eckhard filter, the partitioning methods showed a good agreement with the 3-time running Lyne-Hollick filter (Figure S1, supplementary material). Therefore, we keep using the 3-time running Lyne-Hollick filter for the subsequent analysis in our study.

1. Methods

(a) Climate-based Formulation for L'vovich Water Balance Components

We used the mathematical formulation proposed by Meira Neto et al. (2020), which combines Budyko (1974) and L'vovich (1979) frameworks to define climate-based functional forms to describe the streamflow components Q_B and Q_D as a function of aridity Index (ϕ). According to the L'vovich formulation, under the assumption of negligible changes in water storage and other water gain or losses processes (Gnann et al., 2019), the annual water balance can be conceptualized as a two-stage partitioning: initially, the precipitation (P) is partitioned into Q_D and an infiltration component, called catchment wetting (W). Later, part of the water that is being stored in the catchment, represented by W, is transformed into Q_B or evapotranspiration (ET). The combination of Q_B and Q_D yields the total streamflow (Q). Thereby, the two-stage water balance partitioning can be mathematically expressed by Equations 1 and 2, respectively.

$P \ = \ Q_D + W$ (1)

$W = Q_B + ET$ (2)

Alternatively, the water balance components may be expressed as:

$P \ = \ Q + ET$ (3)

$Q = Q_B + Q_D$ (4)

The Budyko framework, in turn, hypothesizes that, considering a negligible change in water storage condition over long time scales, the long-term mean partitioning of P into Q and ET (Equation 3) is largely controlled by the climate, expressed in terms of ϕ (Berghuijs et al., 2020; Meira Neto et al., 2020). Despite the complexity and uniqueness of catchments (Beven, 2000), several studies have proved that this general relationship holds true for different catchments around the world (McDonell et al., 2007; Padrón et al., 2017; Wang et al., 2016; Williams et al., 2012). If we normalize Equation 3 by the long-term mean precipitation (P), seeking to synthesize data of different sites and isolate the effects of land cover and climate type (Willians et al., 2012), and reorder it, we obtain:

$\frac{\mathrm{ET}}{P} \ = \ 1 \ - \ \frac{Q}{P} \ = f_E(\phi)$ (5)

where $f_E(\phi)$ represents the functional form that relates ϕ with the evaporative fraction ($\frac{\mathrm{ET}}{P}$). This equation is also known as the Budyko hypothesis. In this framework, the averaged catchment's behavior over many years is constrained by two physical limits: water (or demand) limit and energy (or supply) limit (Padrón et al., 2017). The former indicates that the catchment cannot evaporate more than it receives from precipitation, while the latter implies that the catchment's actual evapotranspiration cannot exceed the potential evapotranspiration (PET) (Bouaziz et al., 2018). If we assume that Q is also largely controlled by ϕ, we can write, from Equation 5, the following expression:

$f_R(\phi) \ = 1 \ - f_E(\phi) \quad (6)$

where $f_R(\phi)$ represents the functional form which relates ϕ with the runoff coefficient ($\frac{Q}{P}$). Since Q can be partitioned into Q_D and Q_B, we can combine Equations 4 and 6 to get:

$f_R(\phi) \ = \frac{\mathrm{Qd}}{P} \ + \ \frac{\mathrm{Qb}}{P} \quad (7)$

$f_R(\phi) \ = f_B(\phi) \ + \ f_D(\phi) \quad (8)$

where $f_B(\phi)$ and $f_D(\phi)$ represent the functional forms that relate the streamflow components, Q_D and Q_B, with ϕ, at a long-term time scale. From Equations 1, 2, 5 and 8, we can rewrite the L'vovich water balance components as a function of $f_B(\phi)$ and $f_D(\phi)$:

$Q_D \ = P.(f_D(\phi)) \quad (9)$

$Q_B \ = P.(f_B(\phi)) \quad (10)$

$W \ = P.(1 \ - \ f_D(\phi)) \quad (11)$

$ET \ = P.(1 \ - f_B(\phi) \ - \ f_D(\phi)) \quad (12)$

1. Functional forms and calibration procedure

Meira Neto et al. (2020) suggested that f_B and f_D may be well represented by a simple exponential decay function, such as following:

$$f_D(\phi) \ = \ \exp\left(-\phi^a \ + \ln\left[\left\{\frac{Q_D}{P}\right\}_{\max}\right]^{\frac{1}{b}}\right)^b \quad (13)$$

$$f_B(\phi) \ = \ \exp\left(-\phi^c \ + \ln\left[\left\{\frac{Q_B}{P}\right\}_{\max}\right]^{\frac{1}{d}}\right)^d \quad (14)$$

where a, b, c, and d are shape parameters, and the terms of the form ln[$*$] are shift coefficients to satisfy the condition when $\phi \to 0$, $Q \ \to 1$, observed by Budyko (1974). This condition means that when P is many times greater than PET ($\phi \to 0$), Q should reach 1. On the other hand, if PET is many times greater than P, Q should reach 0. For a detailed description of these functional forms and the limit conditions, the readers are referred to Meira Neto et al. (2020) and Budyko (1974).

To obtain the best parameters for the functional forms, as well as their uncertainties, we conducted a calibration-validation procedure proposed by Meira Neto et al. (2020). First, we randomly divided the total dataset in two calibration and validation subsets, with sample size equal to half of the total sample size (n = 731 catchments). Then, we fitted the functional forms f_B and f_D to the calibration subset, using the Levenberg-Marquardt nonlinear least square algorithm (Seber & Wild, 2003). We repeated this process 100 times. Finally, we computed the mean and the coefficient of variation of parameters a, b, c, and d, as well as of the determination coefficient (R^2), calculated for the validation subset, using the functional forms fitted to the calibration subset. We repeated these steps varying values (0.1 to 0.9) of $(Q_D/P)_{\max}$. The final value of $(Q_D/P)_{\max}$was set to 0.38, which yielded the best results in terms of R^2.

1. Investigating the Closed-system assumption

To consider the effects of groundwater connectivity, we used the effective catchment area (A_{eff}) approach proposed by Liu et al. (2020). This approach introduces the concept of A_{eff}, an equivalent area that accounts for the inter-catchment groundwater flows (IGF) in the water balance (Schwamback et al., 2022). The idea behind this concept derives from the fact that Q at the catchment outlet represents the response of the A_{eff}, as it is influenced by recharge ($P-\text{ET}$) and inter-catchment groundwater connectivity, whereas $P-\text{ET}$ represents the response of the topographic area (A_{topo}), which is not directly influenced by IGF (Liu et al., 2020). The A_{eff} is defined as:

$A_{\text{eff}} \ = A_{\text{topo}}.\left(\frac{Q}{P-ET}\right)$ (15)

The correction using A_{eff} is based on dividing the long-term mean catchment's streamflow Q [m^3/s] by the A_{eff} [m^2] rather than A_{topo} [m^2], shifting the unit of Q to the same as P and ET [mm]. As Q_B and Q_D derive from Q, they were also corrected. We applied this correction for all 731 evaluated catchments and compared the results with those obtained considering the traditional A_{topo} approach. Following Schwamback et al. (2022), we classified the evaluated catchments as follows: catchments with substantial deviations of the $\frac{A_{\text{eff}}}{A_{\text{topo}}}$ ratio ($\frac{A_{\text{eff}}}{A_{\text{topo}}} \ \geq 1.3$ or $\frac{A_{\text{eff}}}{A_{\text{topo}}} \ \leq 0.7$) were classified as 'Gaining' and 'Losing', respectively. The other catchments were classified as 'Small gain', if $1\frac{\leq A_{\text{eff}}}{A_{\text{topo}}} \ < 1.3$, and 'Small loss', if $0.7\frac{\leq A_{\text{eff}}}{A_{\text{topo}}} < 1$. This categorization considers that a deviation larger than 30% between A_{eff} and A_{topo} may substantially impact hydrological studies.

1. **Identification of other influencing factors on streamflow components**

To investigate the effect of the A_{eff} correction in the identification of influencing factors on streamflow behavior, we computed the Spearman's correlation considering the deviation between estimated and observed streamflow components' ratio $\left(\Delta_{Q'/P} \ = (Q'/P)_{\text{Est}} \ - \left(\frac{Q'}{P}\right)_{\text{Obs}} \right)$, where Q' denotes either Q, Q_B or Q_D, illustrated in Figure 2, and different catchments' attributes (Barnhart et al., 2016; Willians et al., 2012). We used the Spearman's correlation, considering that the relations between deviations and influencing factor may be non-linear. Thus, we assume that the relations must be monotonic (Padrón et al., 2017). It is important to emphasize that in our study positive correlations indicate that the functional form tends to overestimate the observed values (positive deviation) with increasing values of the attribute. In other words, a positive correlation means that increasing values of a certain attribute disfavors the evaluated streamflow component, while negative correlations indicate the opposite.

We selected 12 climatic and physiographic attributes (Table 1) of the CABra's dataset, guided by the idea of fundamental hydrologic landscapes units (FHLU) and their major control mechanisms (Winter, 2001). FHLU are used to describe and compare the hydrological behavior of different catchments and are characterized by the combination of climatic, geological and topographic conditions (Gharari et al., 2011). The selected attributes cover fundamental characteristics of the catchment: land cover, topography, climate, geology, soils, and groundwater (Almagro et al., 2021a; Price, 2011). Additionally, attribute selection was corroborated by previous studies which also evaluated influencing factors/predictors of the streamflow or of other water balance components behavior (Beck et al., 2013, 2015; Liu et al.; 2020; Padrón et al., 2017; Salinas et al., 2013).

Table 1

Climatic and Physiographic Attributes Investigated as Possible Influencing Factors on Streamflow Components.

Attributes' classes	**Attribute**	**Unit**
Climate	Aridity index	-
	Precipitation Seasonality (Woods, 2009)	-
Groundwater	Water table depth (WTD)	m
Geology	Permeability	log (k) m^2
Soil	Fraction Sand	%
	Fraction Silt	%
	Fraction Clay	%
Topography	Mean Slope	%
	Mean Elevation	m
	Catchment Area	km^2
Land Cover	Forest Cover	%
	Crops Cover	%

Note. Attributes were classified according to Almagro et al. (2021a).

1. Results and Discussion

   (a) Predictive performances of -based formulations

We fitted the three functional forms (f_R, f_B and f_D) relating ϕ to observed values of Q/P, Q_B/P, and Q_D/P (Table 2). The low values of coefficient of variation (CV) indicate a robust fit for both scenarios (with and without correction). The differences between final parameter values for the cases with and without correction were somehow expected, given that the correction $\frac{A_{\text{eff}}}{A_{\text{topo}}}$

may significantly alter the values of Q, and consequently, Q_B and Q_D.

Table 2

Results From the Calibration-Validation Procedure

Parameters	No correction	Aeff/Atopo corrected		
	Mean	C.V	Mean	C.V
a	1.1	4.54%	0.71	7.70%
b	1.23	0.77%	1.3	0.65%
c	1.2	3.30%	0.75	2.71%
d	0.74	1.10%	0.8	0.42%

Note. Mean and C.V of the shape parameters (a, b, c, and d) obtained from the fitting procedure for the no correction (Left) and the A_{eff} correction (Right) condition.

Figure 2 shows the functional forms (considering the mean values of the fitted parameters) for the 731 catchments analyzed in this study, considering both no-correction (left panel) and A_{eff} correction (right panel). In general, the curves presented a good fit to the observed data. This result suggests that the functional forms are able to reproduce the long-term streamflow (and its components) relationship with ϕ. Focusing on the left panel, it can be seen that higher scatter is found for catchments with substantial deviation values of the ratio $\frac{A_{\text{eff}}}{A_{\text{topo}}}$ (gaining or losing water condition). In general, the functional forms overestimated the ratios for catchments with losing water condition and underestimated the ratios for catchments with a gaining water condition. It is also worth noting that despite the gaining or losing-water condition or the aridity index, all catchments exhibited low values of $\frac{Q_D}{P}$, which may suggest that for Brazilian catchments Q_B plays a major role in the streamflow water-balance component. In fact, Brazilian catchments presented much higher values of baseflow index (BFI) (Figure S1) in comparison with other studies that adopted the same 3-passes Lyne-Hollick filter for United States and United Kingdom (Gnann et al., 2019; Yao et al., 2019), and New Zealand catchments (Singh et al., 2019).

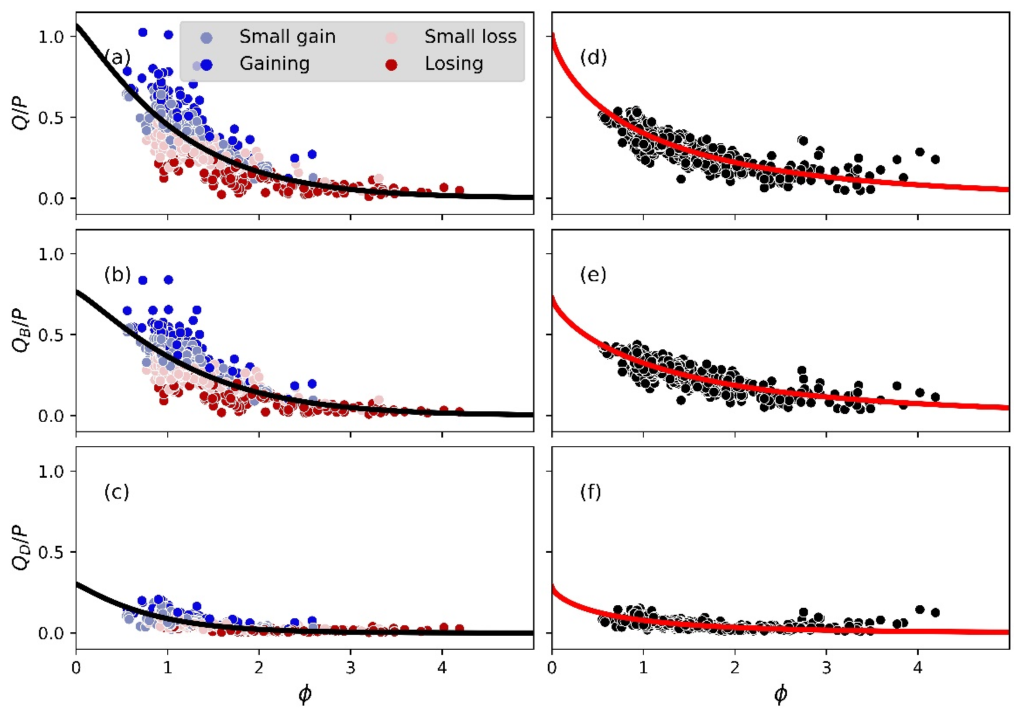

FIGURE 2 - Functional forms and observed values of the streamflow components ratio. Left panel: No correction condition. Catchments classified according to Schwamback et al. (2022). Right panel: A_{eff} corrected condition.

We found a smaller deviation between the observed values and the fitted functions when the effective area correction was applied (Figure 2, right panel). Nevertheless, this is not true for catchments with higher values of the aridity index ($\phi > 2$), which exhibited a worse performance in terms of the RMSE (Figure S2). These arid basins, which are mainly located in the northeastern portion of Brazil, are characterized by non-perennial hydrologic regimes, exhibiting the highest values of streamflow elasticity among Brazilian's catchments (Almagro et al., 2021a). Such atypical hydrological condition may have impacted the performance of the functional forms after the proposed correction. To shed light on this improvement, we plotted the estimated values of the L'vovich water balance components against all the observed values together with the predictive capabilities of the analytical formulations, expressed in terms of the mean value of the coefficient of variation R^2, obtained in the calibration-validation procedure (Figure 3). The results confirm that ϕ is a major control mechanism of the hydrological water balance terms. For the no-correction condition, ϕ was able to explain 0.78, 0.67, 0.79 and 0.98 of the variability of Q_B, Q_D, Q and W, respectively. For the A_{eff} corrected condition, the capability of ϕ to explain these L'vovich water balance components was even greater, exhibiting R^2 mean values of 0.89, 0.68, 0.90 and 0.99, for the same variables, respectively. Similarly to Figure 2, catchments with substantial deviation values of the ratio $\frac{A_{\text{eff}}}{A_{\text{topo}}}$ ratio exhibited a higher deviation from the 1:1 line. Moreover, the A_{eff} correction slightly reduced the C.V. values, indicating an increase in the model's robustness.

The improvement in predictive performance when the A_{eff}-corrected ϕ-based analytical formulations suggests that the inter-catchment connectivity exerts a significant control on explaining the inter-catchment variability of streamflow components. It is worth mentioning that the A_{eff} correction did not significantly improve the estimations of Q_D and W. The former represents the catchment's fast response to precipitation. It is not directly liked to IGF and, therefore, shouldn't be affected by the proposed correction. The latter, on the other hand, is computed as the difference between P and Q_D. Since the proposed correction did not improve the estimations of Q_D, it follows that W should not improve either. Furthermore, even in the no-correction condition, ϕ was able to explain 0.98 of W variability. Thus, there was not much space for improvement.

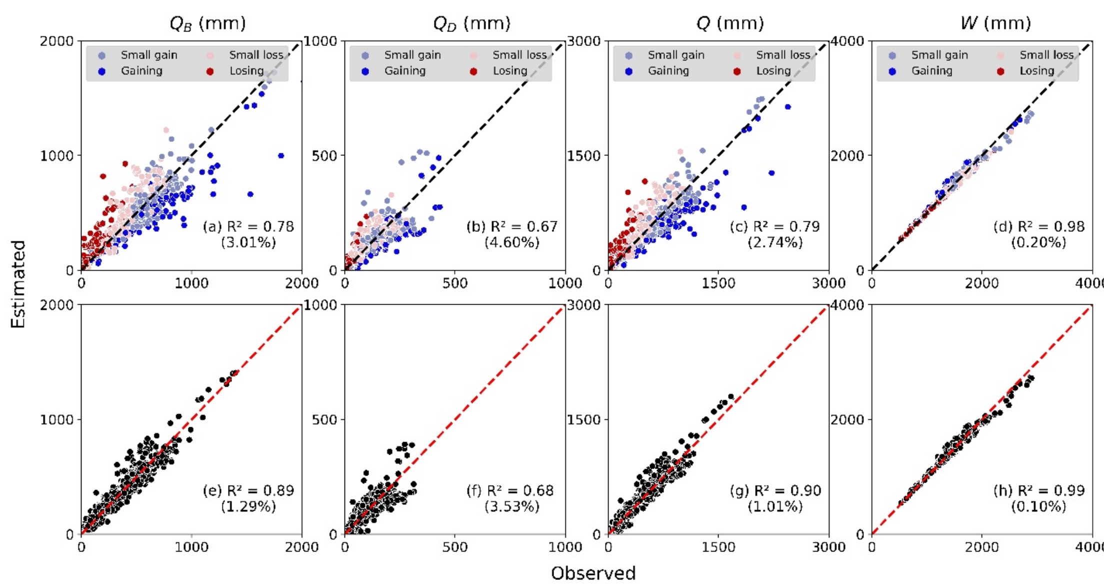

Figure 3 - Observed versus estimated values of the L'vovich water balance components for the no correction (upper panels) and Aeff corrected (lower panels) conditions. Mean (standard deviation) of R^2 is shown in each plot, considering the 100 repetitions of the calibration-validation procedure. An additional 1:1 line is plotted for reference.

Further, we evaluated the performance of the functional forms to represent the relationships of the L'vovich water balance components (Figure 4) at the mean annual scale. This analysis can elucidate the long-term interrelationships between the water balance variables for all considered catchments in a complementary way to the traditional Budyko-based analysis (Sivapalan et al., 2011; Meira Neto et al., 2020). Overall, the functional forms of both conditions were able to reproduce the observed trajectories of the relationships between water-balance components. It is worth noting that the no-correction condition case exhibited, in general, a larger spread of the observed data. Nevertheless, it is evident that the A_{eff} correction improved the modelling, reducing the deviations between observed and estimated values. We focus our analysis on the second partitioning of the L'vovich water balance, as similar conclusions may be inferred for the first partitioning. As can be noted, the functional forms were able to capture the different responses of Q_B and E upon an increase in W. For low wetting values, a quick and almost linear increase in E is observed, while Q_B remains null. As W exceeds a certain threshold, Q_B grows quickly. As observed by Meira Neto et al. (2020), at this threshold, the relationship W x E loses its "linearity", since part of the is being W transformed into Q_B. The scattered pattern observed in Q_B x W relationship (Figure 4d) confirms the role of IGF in the water-balance components. Q_B is strongly influenced by IGF and inter-catchment water exchanges conditions in Brazil are extremely variable (Schwamback et al., 2022). When the influence of IGF is not consider in the computation of streamflow components, a larger amplitude of Q_B values can be observed for a fixed value of W. This scatter pattern is not observed in the first partitioning, which represents the catchment's quick response to rainfall events and consequently is not directly influenced by IGF.

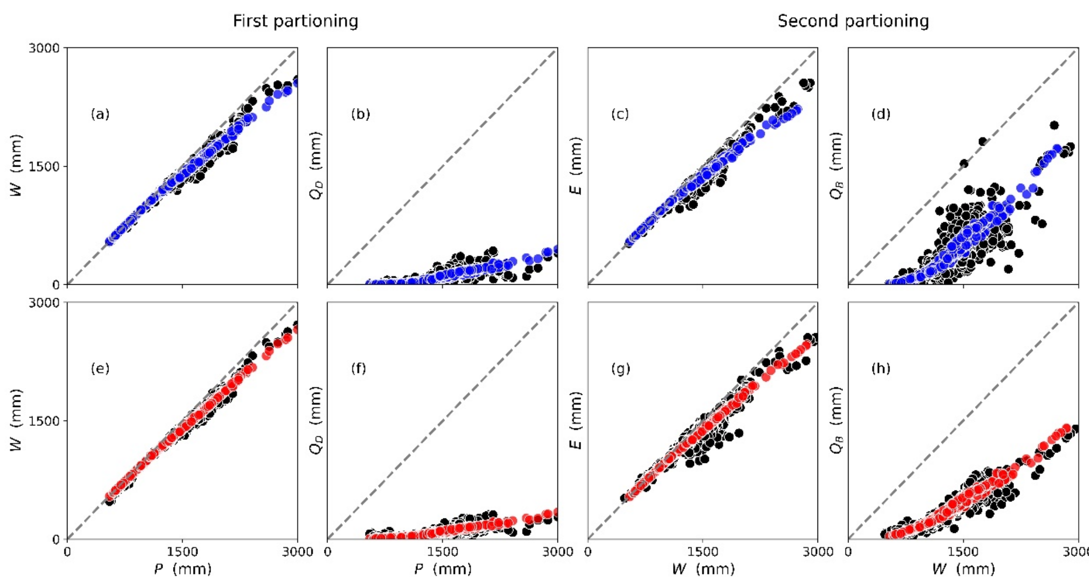

Figure 4 - Observed (black) and estimated (coloured) mean annual values of the L'vovich water balance components, for the no correction (upper panels) and Aeff corrected (lower panels) conditions.

Figure 5 shows the spatial variability of the relative deviation between estimated and observed water balance components. The mean annual catchments' water-balance components are shown in Figure S3. We did not detail the results obtained for W, given that both correction conditions achieved optimal performance to describe it. We note higher correspondence between the deviations' signals of Q_B and Q. This is expected, since the mean baseflow index (BFI) of the CABra's catchments is approximately 0.80, indicating Q_B as the main primary component of the total streamflow. No clear correspondence, however, was found between signals of relative deviations in Q_D, and Q or Q_B. We found higher relative deviations for the Q_D component, given the lower performance of the functional forms to represent it.

For the no-correction condition (Figure 5, upper panel), we observed higher Q_B deviations, in terms of absolute values, for the Amazon and Caatinga biomes (Figure S3), which were pointed out by Schwamback et al. (2022) as two biomes with large discrepancy in terms of the $\frac{A_{\text{eff}}}{A_{\text{topo}}}$ ratio. This can also be seen, with lower magnitude, for the Atlantic Forest, which also exhibited large divergence in the $\frac{A_{\text{eff}}}{A_{\text{topo}}}$ ratio. There is a slight predominance (~60% of the catchments) of positive deviations for the arid Caatinga biome, whose catchments were prone to present, in general, a losing-water condition, suggesting that the functional forms tend to overestimate Q_B for these basins. This pattern is also clear for the northeastern Cerrado, which is border with the Caatinga biome. The opposite situation was observed in Atlantic Forest (~58% of catchments with negative bias), where most catchments (72%) present an effective area larger than the topographic one. In fact, these catchments are highly affected by human hydrologic disturbances, which may hamper an accurate analysis. If we selected only catchments with low anthropic impact (hydrologic disturbance index < 0.1, see Almagro et al. (2021a) for a detailed description of this index), the percentage of catchments with positive (negative) deviations grows up to 70% for Caatinga (Atlantic forest). We did not find a clear pattern for the Pantanal and Pampa biomes, as they present a more heterogeneous condition in terms of the $\frac{A_{\text{eff}}}{A_{\text{topo}}}$ ratio and a low density of monitoring gauges (Almagro et al., 2021a).

The A_{eff} corrected condition (Figure 5) significantly reduced the deviations of Q_B and Q estimates. This improvement cannot be noticed for the other components (Q_D and W), since the A_{eff} correction did not significantly improve their estimations. However, as one can note in Figure S4, the A_{eff} correction decreased the number of their outliers. It is worth noting that the A_{eff} correction was not able to alter significantly the deviations for some catchments, located in the northeastern Brazil, suggesting that other factors than climate (aridity index) and inter-catchment connectivity ($\frac{A_{\text{eff}}}{A_{\text{topo}}}$) may also be controlling their hydraulic behavior. Actually, as mentioned above, this region is largely affected by human hydrologic disturbances (Almagro et al., 2021a; Nascimento and Neto, 2017) due to its complex network of reservoirs that significantly alters the local water budget (Schwamback et al., 2022). Therefore, the proposed correction may not be sufficient to overcome the limitation of the functional forms to characterize streamflow components in these basins.

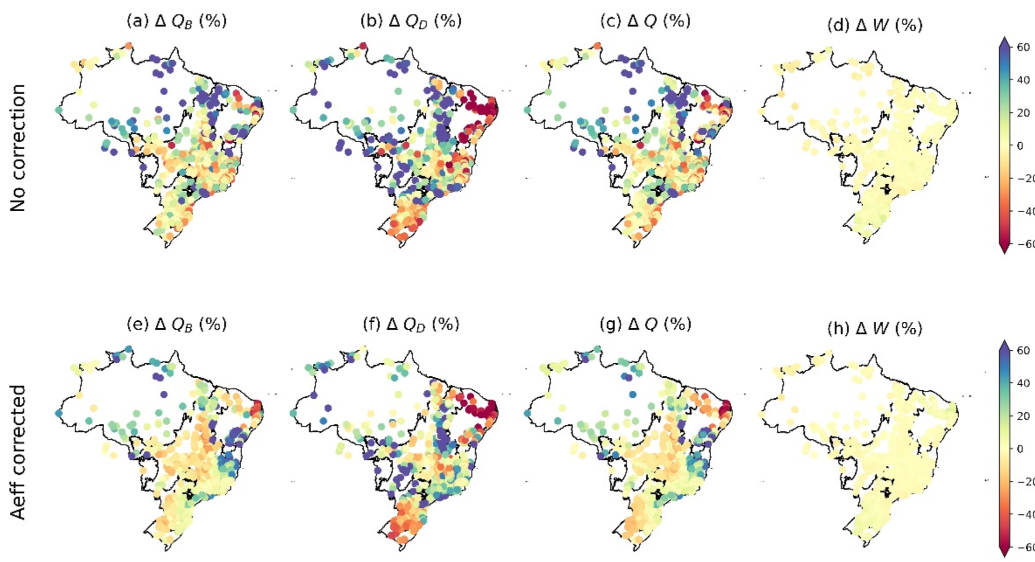

Figure 5 - Spatial distribution (including the 6 main Brazilian biomes) of the relative bias between observed and estimated values of the four L'vovich water balance components for the no correction (upper panels) and Aeff corrected (lower panels) condition.

1. Influencing factors on streamflow components

Figure 6 shows the scatter plot of $\Delta_{Q/P}$ versus the 12 selected attributes. Similar plots considering Q_B and Q_D are displayed in the supplementary material (Figures S5 e S6) and the results are summarized in Figure 7. Following de Lavenne and Andréassian (2018), for verification purposes, we computed the same correlations for a reduced sample, representing catchments in which we assume a "true" closed water balance to occur. This "control sample" contains 114 catchments that show little influence of IGF, thus having similar values of effective and topographic areas ($0.95 < A_{\mathrm{eff}}/\ A_{\mathrm{topo}} < 1.05$) (Figure S7).

The relationships are characterized by a high degree of uneven variability and nonlinearity. The A_{eff} correction significantly altered the correlation (magnitude, significance and signal) between deviations and some attributes, indicating that, indeed, the consideration of IGF affects the identification of influencing factors on streamflow components. As also observed in Beck et al. (2013,2015) and Padrón et al. (2017), the correlations were all rather weak (< 0.4), emphasizing the role of aridity as the main control on the long-term water balance (Budyko, 1974). The main differences between correlations (no-correction versus A_{eff} corrected) were found for (1) land cover, (2) topography, and (3) groundwater attributes, and (4) climate seasonality.

For the land-cover attributes, the consideration of A_{eff} resulted in correlations shifting from negligible and/or insignificant to moderate and significant, suggesting that not considering IGF in our formulations may have masked the relationships between streamflow and land cover. The positive correlation found for forest cover suggests a decrease in Q for the same ϕ values and increasing values of the attribute when IGF is considered, which is consistent with the general conclusions of previous studies (Silveira & Alonso, 2009; Zhou et al., 2015). In fact, as pointed out by Padrón et al. (2017) and Gan et al. (2021), catchments with large forest rates tend to favor ET over Q, since it increases the water retention potential and may decelerate water movement as a consequence of longer flow paths and poriferous soils. Similar patterns can be observed for Q_B and Q_D, although with lower magnitude. The opposite rationale can be applied for the negative correlation found for crops cover: catchments with a predominance of agricultural cover tends to amplify Q and reduce ET (Dias et al., 2015), which might be explained by decreasing in the catchment's water retention ability (Zhou et al., 2015). The control sample (Figure S7), shows even higher absolute correlation values for both land cover attributes.

We also found significant correlations between catchments' mean slope and Q-deviations (Figure 6) for both scenarios, which supports previous findings indicating slope as a major control mechanism of streamflow (Beck et al., 2013, 2015; Padrón et al., 2017; Trancoso et al., 2016). For non-corrected area values, the negative correlation suggests that increasing mean slope favors streamflow production for a fixed value of ϕ (Zhou et al., 2015). In fact, hilly catchments tend to gain water through IGF (Schwamback et al. (2022); Liu et al. (2020)), which in turn favors the generation of streamflow components, and might explain why such behavior was observed in the uncorrected case. After the A_{eff} correction, the correlation shifted the signal, suggesting the opposite: catchments with steeper slopes disfavor the generation of Q, the same happening with Q_B. This somewhat counterintuitive result (one might have expected that IGF inclusion would lead to null correlation, and not an inverse signal) requires a more in-depth analysis. Brazilian watersheds with stepper slopes have, in general, deeper water table levels (Figure S7), which might lead to having less interactions between groundwater and surface water, disfavoring Q_B. Since Q is dominated by Q_B in Brazilian basins (Figure S1), this might explain the observed effect of slope on these components when A_{eff} was considered (Figure 7). This is also seen in the control sample results (Figure S7), in which a positive and significant correlation for Q_B and Q ($\rho \sim 0.3$) was found. That is, indeed Brazilians' steeper watersheds are prone to disfavor Q_B due to the distance to the water table. The low correlation values for the relationship between slope and Q_D for both corrected and uncorrected scenarios, makes it hard to draw a meaningful conclusion out of the analyzed data.

Looking at the other two topographic attributes when no IGF is accounted, we found a negative correlation for catchment's mean elevation ($\rho = -0.28$) and a positive correlation for catchment area ($\rho = 0.33$), suggesting that elevated and/or small-sized watersheds favors Q (Figure 6). Nevertheless, after applying the proposed correction, we noted that the correlation of both attributes shifted from moderate and significant to negligible and insignificant, suggesting that the effects of these attributes on streamflow components are mainly related to inter-catchment water exchanges. Similar results were found for Q_B. Catchments are prone to gain water through IGF (Schwamback et al., 2022) with increasing elevation, while large catchments are prone to lose water, given their higher water retention capacities and lower hydrological responses (Zhou et al., 2015). Indeed, Brazilian's catchments with larger areas are inclined to present an A_{eff} lower than A_{topo} (median $\frac{A_{\text{eff}}}{A_{\text{topo}}} = 0.82$ for the 100 largest catchments). The opposite happens with elevated catchments (median $\frac{A_{\text{eff}}}{A_{\text{topo}}} = 1.26$ for the 100 highest basins). Moreover, elevation and area should exhibit a similar IGF behavior as in mean slope, as these variables are relatively well correlated (Figure S8). Similar to the slope analysis, the correlation values found between Q_D and these two topographic attributes are rather weak. Finally, the control sample further suggests the incorporation of IGF to explain the previously suggested relationships between topographic indices and streamflow components (Figure S7).

The positive correlation between streamflow components and WTD in the no-correction condition indicates that catchments with a deepest water table tend to show larger values of Q and its components for the same ϕ (Figure 6 and 7). Nevertheless, just as observed for the catchments' mean slope, not considering effects of IGF may have masked this relationship. Catchments with higher absolute values of WTD tend do exhibit a gaining-water condition (median $A_{\text{eff}}/A_{\text{topo}} = 1.30$ for the 100 watersheds with deepest water table), which may influence the results. The consideration of the A_{eff} shifted the signal of the correlations between Q and Q_B, maintaining its magnitude, suggesting the opposite: catchments with deeper water table tend to show lower values of Q and Q_B for the same due to the lower interaction between groundwater and surface water. Once more, the control sample corroborates with this conclusion

(Figure S7). Regarding Q_D , the correction led to insignificant correlations, suggesting that WTD is not a main control mechanism of quickflow in Brazilian catchments.

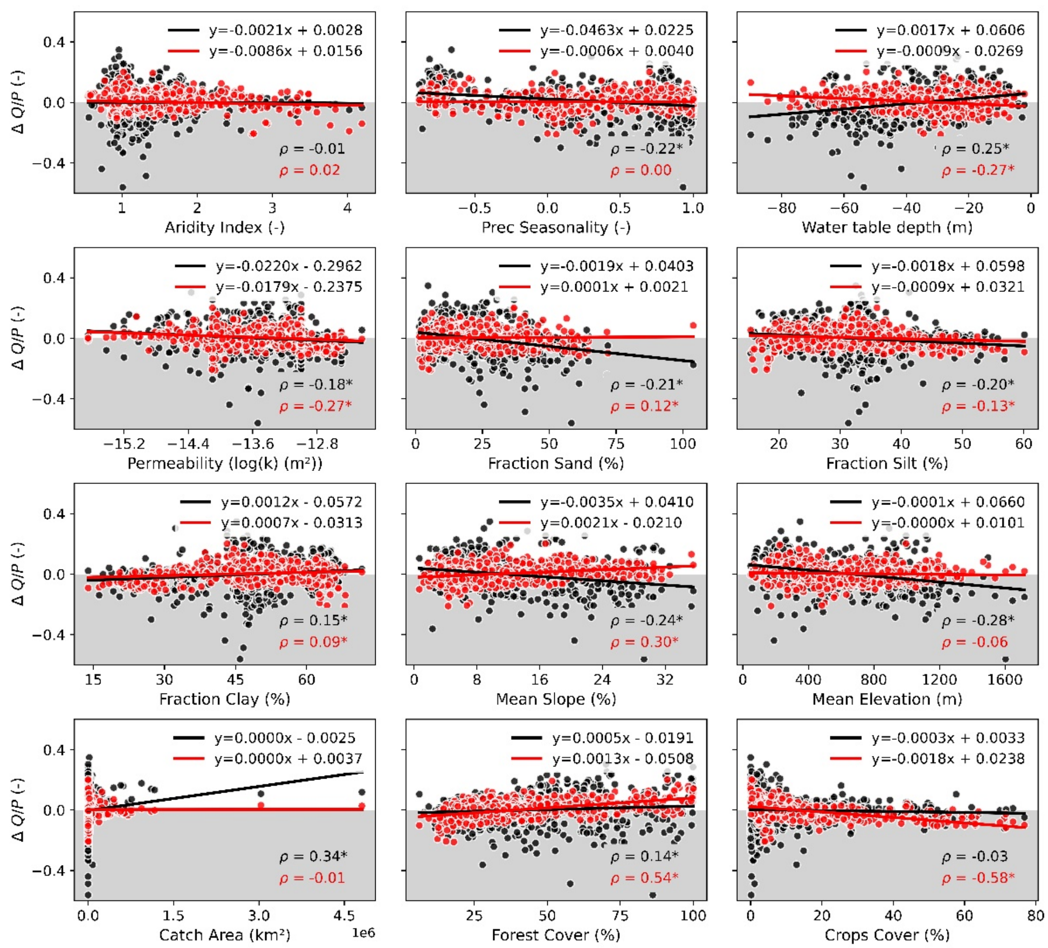

Figure 6 - Scatter plots, linear regression and regression coefficients of the no correction (black) and Aeff-corrected (red) streamflow ratio' deviations ($\frac{\mathbf{Q}}{\mathbf{P}}$) and the 12 climatic and physiographic attributes. Spearman's correlation of the no correction and Aeff corrected conditions are shown in black and red, respectively. Statistically significant correlations are indicated with *.

We did not find a significant correlation between ϕ and deviations for both considered condition (Figure 7), since the functional forms modeled Q, Q_B and Q_D as a function of it, and, therefore, have already accounted for aridity. For the precipitation seasonality (no-correction condition), we found a negative correlation for Q and Q_B, and a positive correlation for Q_D (Figure 7), suggesting that Q and Q_B are favored in catchments with summer-dominant rainfall, while Q_D is disfavored. The A_{eff} correction reduced the correlations for the first two, shifting it from $\rho = 0.22$ and 0.33 to $\rho = 0$ and 0.15, respectively, and increased from $\rho = 0.08$ to 0.30 for Q_D, indicating that total streamflow is not directly affected by precipitation seasonality due to its opposite impact on streamflow components: precipitation seasonality seems to favor baseflow and disfavor quickflow in Brazilian catchments.

Our findings, also observed in the control sample, differ from the reported in previous studies that investigated the role of seasonality on the water-balance (Gan et al., 2021; de Lavenne & Andréassian, 2018; Padrón et al., 2017; Yokoo et al., 2008), with exception of Potter et al. (2005), which also observed the same behavior in Australian catchments. In fact, understanding the role of seasonality on water-balance partitioning is a complex task. Yokoo et al. (2008) showed that its effect on long-term water-balance closely depends on other physiographic attributes, such as soil and topographic characteristics. Gan et al (2021) suggested precipitation seasonality' effect as dependent of vegetation conditions, pointing the coupling of vegetation and seasonality as crucial to understand the long-term water balance. Actually, the seasonality index neglects fine-scale variations, such as daily and intra-monthly rainfall characteristics, which may hamper capturing the role of rainfall temporal distribution on long-term water-balance partitioning. Moreover, our results may be affected by the uneven distribution of precipitation seasonality exhibited by our evaluated catchments, since most of them (~85%) exhibits a summer precipitation cycle, despite their diverse hydroclimatic conditions (Almagro et al., 2021a). In view of the aforementioned, we argue that assessing the role of seasonality on long-term water-balance requires a more in-depth analysis, which is not the main purpose of this study, considering water and energy temporal distributions at finer scales and their co-dependence with landscape characteristics.

Considering the soil attributes, we found overall low correlation values ($\rho \leq 0.2$). For uncorrected conditions, our results suggest that clay/silt fractions disfavor/favor streamflow generation, while after the correction the strength of the correlations were reduced, which may indicate that part of their contribution to streamflow generation are related to IGF. In fact, according to Fan (2019) and Gnann et al. (2021), soil properties are a key control of the subsurface hydraulic properties, which affects streamflow generation. However, the low correlation values prevent us from pursuing a discussion on possible mechanisms that may affect streamflow components and the water-balance partitioning. Interestingly, when we look at the control sample (Figure S7) such relationships receive higher correlation values (especially for Q_D), showing that such soil properties might actually have a stronger influence in streamflow generation. For sand fraction the correction switched the relationships from favoring total streamflow and its components to disfavoring Q_B, Q while still favoring Q_D, although with lower correlation values. That is, it indicates that before the IGF consideration, the streamflow favoring-condition of high-sand content catchments may be attributed to its high susceptibility to inter-catchment water exchanges. Actually, Brazilian catchments with high sand content shows, in general, a gaining water condition (median $A_{\text{eff}}/A_{\text{topo}} = 1.15$ for the 100 watersheds with higher sand fraction).

Interestingly, the A_{eff} correction strengthened the correlation between deviations and permeability (Figure 7). For both conditions, we found a negative correlation, which indicates that with greater the permeability, streamflow components will be favored. This finding reinforces permeability as a major influencing factor of the streamflow components' characteristics. The fact that the proposed correction increased the correlation magnitude indicates that permeability appears as an important control when IGF processes are included in the water balance. The very low correlations found for Q_D highlights, as expected, the absence of a link between quick flow production and geological settings.

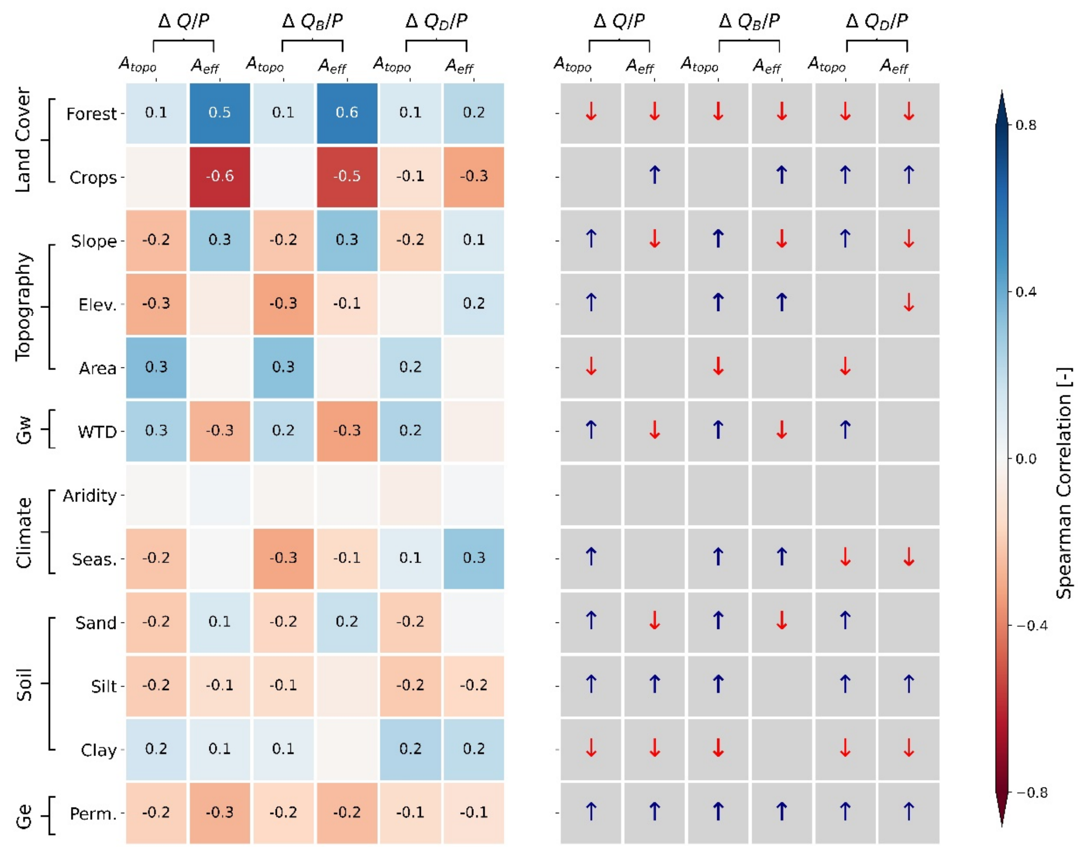

Figure 7 – Relationship between the 12 selected attributes and the deviations (Δ) of the streamflow components. Attributes from top to bottom: forest cover, crops cover, mean slope, mean elevation, catchment area, water table depth, aridity index, precipitation seasonality, sand fraction, silt fraction, clay fraction, and permeability. Left panel: spearman's correlation heatmap. Values with annotation exhibited a significant correlation (p-value < 0.05). Right panel: direction of the relationship. Up and blue (down and red) arrows means that the attribute favors (disfavors) the respective streamflow comoponent.

1. Limitations and broader implications of the study for hydrological processes

   (a) Limitations

Despite using a large-sample hydrology dataset, which avoid wrong conclusions that may derive from few anomalous catchments with atypical hydrologic behavior (Addor et al., 2020; Gnann et al., 2021; Gupta et al., 2014), it is important to recall some limitations that may arise from this study. First, measurement errors in both hydroclimatic and physiographic catchments attributes can significantly affect our results (Westerberg & McMillan, 2015). Although the CABra's dataset provides diverse reliable attributes for more than 700 Brazilian catchments, some of them are derived from gridded data (Almagro et al., 2021a), which may not represent well hydrological process of smaller scales (Gudmundsson and Seneviratne, 2015). Moreover, other sources of uncertainty in the observed/estimated data, such as errors associated with the different mathematical methods to estimate PET and ET, non-representative rating curves to derive Q, and inherent uncertainties of the digital filters to partition Q into Q_B and Q_D (Meira Neto et al., 2020), may also affect our findings.

Second, to relax the closed water-balance assumption, we used the A_{eff} methodology proposed by Liu et al. (2020). However, this simple method may not be effective for some catchments with unusual hydrological behavior, such those found in the northeastern part of Brazil, characterized by (i) a non-perennial regime with little groundwater contribution that challenge the A_{eff} computation and/or (2) high hydrological disturbances, which may exacerbate the deviations of the $A_{\text{eff}}/A_{\text{topo}}$ ratio and violate the Budyko hypothesis (Berghuijs et al., 2021; Schwamback et al., 2022). In fact, there are other ways to consider inter-catchment groundwater transfers: apply correcting or scaling factors to others climatic input data, such as precipitation; or even explicitly consider IGF as a parameter in hydrologic models to compute groundwater exchanges (Mouelhi et al., 2006; Perrin et al., 2003). As pointed by Le Moine et al. (2007), these different options may result in significantly different IGF values, which in turn, could impact our findings.

1. Broader implications

In this paper, we have focused our attention in the definition of the main control mechanisms of Q, Q_B, and Q_D under an open water-balance perspective. The findings reported here, such as the importance of considering IGF, can be incorporated in the development of hydrological conceptual models to predict the catchments' hydrological response to future changes in the evaluated attributes (Fenicia et al., 2014; Gnann et al., 2021). In fact, models need to adequately represent hydrological processes to predict changes accurately (Clark et al., 2017). If we correctly understand the main drivers of change and how they are expected to change, we can use this information for the development of models that ally model complexity, realism and performance (Atkinson et al., 2002). In addition to predict catchments' long-term hydrological response under change, which is still a controversial approach (Berghuijs et al., 2020), our results may represent a step forward for the characterization of water-balance components in ungauged basins, associating ϕ, IGF, and other landscapes and climate parameters that proved to be important for the characterization of the catchments' hydrological behavior. Finally, we highlight other questions that can be arise from our study: (i) How does the role of influencing factors, such as groundwater and land use, vary in different types of climates? (ii) How can human interventions be considered in the characterization and prediction of water balance components and in the accounting of hydraulic connectivity? (iii) What are the uncertainties associated with the different methods of considering inter-catchment water transfers?

1. Conclusion

In this study we asked what is the impact of a closed water-balance assumption on the understanding of the factors that control the long-term streamflow components. To do that, we investigated the performance of recently proposed functional forms to describe Q, Q_B, and Q_D as a function of ϕ considering both an open and close water-balance assumptions and further examined 12 climatic and physiographic attributes that may also contribute to explain the spatial variability of these water balance components, applying our framework to a set of more than 700 Brazilian catchments.

Our findings demonstrate that: (i) streamflow components can be modeled as a function of ϕ. The functional forms were able to explain most of the mean-annual spatial variability of Q, Q_B, and Q_D (R^2 = 0.79, 0.78 and 0.67, respectively) and to capture, in general, the relationships between the L'vovich two-step water balance components; (ii) An open water-balance consideration, by means of the proposed A_{eff} correction, improved the performance of the functional forms to describe Q and Q_B (R^2 = 0.89 and 0.90, respectively), suggesting that, in many catchments, inter-catchment groundwater transfers can not be considered negligible. Furthermore, the proposed correction contributes to disentangle our understanding related to the role of factors controlling the long-term streamflow; (iii) Finally, other factors than ϕ and IGF may also control the streamflow components. In particular, land cover, climate seasonality, groundwater and topography attributes appear to significantly impact the water balance partitioning. Our findings may also contribute to the comprehension and development of hydrological models, seeking to improve our knowledge related to long-term streamflow behavior, its controlling mechanisms, and how they are expected to change.

Acknowledgements

The authors acknowledge the Graduate Program in Hydraulic Engineering and Sanitation – PPGSHS (USP-EESC) and Graduate Program in Environmental Technologies – PPGTA (UFMS-FAENG) for the scientific support. This study was supported by grants from the Ministry of Science, Technology, and Innovation (MCTI) and National Council for Scientific and Technological Development (CNPq) (grants numbers 441289/2017-7 and 309752/2020-5), São Paulo Research Support Foundation (FAPESP, grants numbers 2015/03806-1, 2019/24292-7, and 2020/08140-0), and Coordination for the Improvement of Higher Education Personnel (CAPES) - Finance Code 001 and CAPES PrInt.

Data Availability Statement

The CABra dataset is freely available online at https://doi.org/10.5281/zenodo.4070146 (Almagro et al., 2021b). The long-term mean annual variables and the climatic and physiographic attributes for each catchment evaluated in this study are accessible through https://doi.org/10.5281/zenodo.5895056 (Ballarin et al., 2022).

References

Addor, N., Do, H. X., Alvarez-Garreton, C., Coxon, G., Fowler, K., & Mendoza, P. A. (2020). Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. *Hydrological Sciences Journal*, 65(5), 712–725. https://doi.org/10.1080/02626667.2019.1683182

Almagro, A., Oliveira, P. T., Meira Neto, A. A., Roy, T., & Troch, P. (2021a). CABra: a novel large-sample dataset for Brazilian catchments. *Hydrology and Earth System Sciences*, 1–40. https://doi.org/10.5194/hess-2020-521.

Almagro, A., Oliveira, P. T. S., Meira Neto, A. A., Roy, T., and Troch, P. (2021b). CABra: a novel large-sample dataset for Brazilian catchments (Version 4) [Data set], Zenodo, https://doi.org/10.5281/zenodo.4655204.

Atkinson, S. E., Woods, R. A., & Sivapalan, M. (2002). Climate and landscape controls on water balance model complexity over changing timescales. *Water Resources Research*, 38(12), 50-1-50–17. https://doi.org/10.1029/2002wr001487

Ballarin, A. S, Marchezepe, B. K, Godoi, R., Campos, A. M., Almagro, A., Meira Neto, A. A., Oliveira, P. T. O. (2022). The Impact of a Closed Water-balance Assumption on the Understanding of the Factors that Control the Long-term Streamflow Components (Version 1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.5895056.

Barnhart, T. B., Molotch, N. P., Livneh, B., Harpold, A. A., Knowles, J. F., & Schneider, D. (2016). Snowmelt rate dictates streamflow. *Geophysical Research Letters*, 43(15), 8006–8016. https://doi.org/10.1002/2016GL069690

Beck, H. E., Van Dijk, A. I. J. M., Miralles, D. G., De Jeu, R. A. M., Bruijnzeel, L. A., McVicar, T. R., & Schellekens, J. (2013). Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research*, *49*(12), 7843–7863. https://doi.org/10.1002/2013WR013918

Beck, H. E., de Roo, A., & van Dijk, A. I. J. M. (2015). Global maps of streamflow characteristics based on observations from several thousand catchments. *Journal of Hydrometeorology*, *16*(4), 1478–1501. https://doi.org/10.1175/JHM-D-14-0155.1

Berghuijs, W. R., Gnann, S. J., & Woods, R. A. (2020). Unanswered questions on the Budyko framework. *Hydrological Processes*, *34*(26), 5699–5703. https://doi.org/10.1002/hyp.13958

Beven, K. J. (2000). Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences*. https://doi.org/10.5194/hess-4-203-2000

Bouaziz, L., Weerts, A., Schellekens, J., Sprokkereef, E., Stam, J., Savenije, H., & Hrachowitz, M. (2018). Redressing the balance: Quantifying net intercatchment groundwater flows. *Hydrology and Earth System Sciences*, *22*(12), 6415–6434. https://doi.org/10.5194/hess-22-6415-2018

Cheng, S., Cheng, L., Liu, P., Qin, S., Zhang, L., Xu, C. Y., et al. (2021). An Analytical Baseflow Coefficient Curve for Depicting the Spatial Variability of Mean Annual Catchment Baseflow. *Water Resources Research*, 57(8), 1–19. https://doi.org/10.1029/2020WR029529

Clark, M. P., Bierkens, M. F. P., Samaniego, L., Woods, R. A., Uijlenhoet, R., Bennett, K. E., et al. (2017). The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism. *Hydrology and Earth System Sciences*, 21(7), 3427–3440. https://doi.org/10.5194/hess-21-3427-2017

de Lavenne, A., & Andréassian, V. (2018). Impact of climate seasonality on catchment yield: A parameterization for commonly-used water balance formulas. *Journal of Hydrology*, 558, 266–274. https://doi.org/10.1016/j.jhydrol.2018.01.009

Dias, L. C. P., Macedo, M. N., Costa, M. H., Coe, M. T., & Neill, C. (2015). Effects of land cover change on evapotranspiration and streamflow of small catchments in the Upper Xingu River Basin, Central Brazil. *Journal of Hydrology: Regional Studies*, *4*(PB), 108–122. https://doi.org/10.1016/j.ejrh.2015.05.010

Eckhardt, K. (2005). How to construct recursive digital filters for baseflow separation. *Hydrological Processes*, *19*(2), 507–515. https://doi.org/10.1002/hyp.5675

Fan, Y. (2019). Are catchments leaky? *WIREs Water*, *6*(6), 1–25. https://doi.org/10.1002/wat2.1386

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

Gharari, S., Hrachowitz, M., Fenicia, F., & Savenije, H. H. G. (2011). Hydrological landscape classification: Investigating the performance of HAND based landscape classifications in a central European meso-scale catchment. *Hydrology and Earth System Sciences*, 15(11), 3275–3291. https://doi.org/10.5194/hess-15-3275-2011

Gnann, S. J., Woods, R. A., & Howden, N. J. K. (2019). Is There a Baseflow Budyko Curve? *Water Resources Research*, *55*(4), 2838–2855. https://doi.org/10.1029/2018WR024464

Gnann, S. J., McMillan, H. K., Woods, R. A., & Howden, N. J. K. (2021). Including Regional Knowledge Improves Baseflow Signature Predictions in Large Sample Hydrology. *Water Resources Research*, *57*(2), 1–22. https://doi.org/10.1029/2020WR028354.

Gudmundsson, L., & Seneviratne, S. I. (2015). Towards observation-based gridded runoff estimates for Europe. *Hydrology and Earth System Sciences*, 19(6), 2859–2879. https://doi.org/10.5194/hess-19-2859-2015

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., & Andréassian, V. (2014). Large-sample hydrology: A need to balance depth with breadth. *Hydrology and Earth System Sciences*, 18(2), 463–477. https://doi.org/10.5194/hess-18-463-2014

Institute ofHydrology (1980). Low flow studies report No. 1: Research report. Wallingford, UK: Institute of Hydrology.

Le Moine, N., Andréassian, V., Perrin, C., & Michel, C. (2007). How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments. *Water Resources Research*, *43*(6), 1–11. https://doi.org/10.1029/2006WR005608

Li, D., Pan, M., Cong, Z., Zhang, L., & Wood, E. (2013). Vegetation control on water and energy balance within the Budyko framework. *Water Resources Research*, *49*(2), 969–976. https://doi.org/10.1002/wrcr.20107

Liu, Y., Wagener, T., Beck, H. E., & Hartmann, A. (2020). What is the hydrologically effective area of a catchment? *Environmental Research Letters*, *15*(10). https://doi.org/10.1088/1748-9326/aba7e5

Liu, J., & You, Y. (2021). The Roles of Catchment Characteristics in Precipitation Partitioning Within the Budyko Framework. *Journal of Geophysical Research: Atmospheres*, 126(16), 1–20. https://doi.org/10.1029/2021JD035168

Lucas, M. C., Kublik, N., Rodrigues, D. B. B., Meira Neto, A. A., Almagro, A., Melo, D. de C. D., et al. (2021). Significant baseflow reduction in the sao francisco river basin. *Water (Switzerland),* 13(1), 1–17. https://doi.org/10.3390/w13010002

McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., et al. (2007). Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research*, *43*(7), 1–6. https://doi.org/10.1029/2006WR005467

Meira Neto, A. A., Roy, T., de Oliveira, P. T. S., & Troch, P. A. (2020). An Aridity Index-Based Formulation of Streamflow Components. *Water Resources Research*, *56*(9), 1–14. https://doi.org/10.1029/2020WR027123

Miller, M. P., Buto, S. G., Susong, D. D., & Rumsey, C. A. (2016). The importance of base flow in sustaining surface water flow in the Upper Colorado River Basin. *Water Resources Research*, (52), 3547–3562. https://doi.org/10.1002/2015WR017963.

Mouelhi, S., Michel, C., Perrin, C., & Andréassian, V. (2006). Stepwise development of a two-parameter monthly water balance model. *Journal of Hydrology*, 318(1–4), 200–214. https://doi.org/10.1016/j.jhydrol.2005.06.014

Muñoz, E., Arumí, J. L., Wagener, T., Oyarzún, R., & Parra, V. (2016). Unraveling complex hydrogeological processes in Andean basins in south-central Chile: An integrated assessment to understand hydrological dissimilarity. *Hydrological Processes*, *30*(26), 4934–4943. https://doi.org/10.1002/hyp.11032

Nascimento, V. F., & Ribeiro Neto, A. (2017). Characterization of reservoirs for water supply in Northeast Brazil using high resolution remote sensing. *Brazilian Journal of Water Resources*, *22*(0). https://doi.org/10.1590/2318-0331.0217170060

Oliveira, P. T. S., Lucas, M. C., de Faria Godoi, R., & Wendland, E. (2021). *Groundwater recharge and sustainability in Brazil. Global Groundwater*. INC. https://doi.org/10.1016/b978-0-12-818172-0.00029-3

Padrón, R. S., Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017). Large-Scale Controls of the Surface Water Balance Over Land: Insights From a Systematic Review and Meta-Analysis. *Water Resources Research*, *53*(11), 9659–9678. https://doi.org/10.1002/2017WR021215

Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1–4), 275–289. https://doi.org/10.1016/S0022-1694(03)00225-7

Potter, N. J., Zhang, L., Milly, P. C. D., McMahon, T. A., & Jakeman, A. J. (2005). Effects of rainfall seasonality and soil moisture capacity on mean annual water balance for Australian catchments.

Water Resources Research, 41(6), 1–11. https://doi.org/10.1029/2004WR003697

Priestley, C. H. B., & Taylor, R. J. (1972). On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Monthly Weather Review*, 100(2), 81–92. https://doi.org/10.1175/1520-0493(1972)100%3c0081:otaosh%3e2.3.co;2

Price, K. (2011). Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review. *Progress in Physical Geography*, *35*(4), 465–492. https://doi.org/10.1177/0309133311402714

Safeeq, M., Bart, R. R., Pelak, N. F., Singh, C. K., Dralle, D. N., Hartsough, P., & Wagenbrenner, J. W. (2021). How realistic are water-balance closure assumptions? A demonstration from the southern sierra critical zone observatory and kings river experimental watersheds. *Hydrological Processes*, 35(5). https://doi.org/10.1002/hyp.14199

Salinas, J. L., Laaha, G., Rogger, M., Parajka, J., Viglione, A., Sivapalan, M., & Blöschl, G. (2013). Comparative assessment of predictions in ungauged basins-Part 2: Flood and low flow studies. *Hydrology and Earth System Sciences*, *17*(7), 2637–2652. https://doi.org/10.5194/hess-17-2637-2013

Santhi, C., Allen, P. M., Muttiah, R. S., Arnold, J. G., & Tuppad, P. (2008). Regional estimation of base flow for the conterminous United States by hydrologic landscape regions. *Journal of Hydrology*, *351*(1–2), 139–153. https://doi.org/10.1016/j.jhydrol.2007.12.018

Schaller, M. F., & Fan, Y. (2009). River basins as groundwater exporters and importers: Implications for water cycle and climate modeling. *Journal of Geophysical Research - Atmospheres*, 114(4). https://doi.org/10.1029/2008JD010636

Seber, G. A. F., & Wild, C. J. (2003). Nonlinear regression. Hoboken, NJ: Wiley-Interscience.

Silveira, L., & Alonso, J. (2009). Runoff modifications due to the conversion of natural grasslands to forests in a large basin in Uruguay. *Hydrological Processes*, (23), 320–329. https://doi.org/10.1002/hyp.7156

Singh, S. K., Pahlow, M., Booker, D. J., Shankar, U., & Chamorro, A. (2019). Towards baseflow index characterisation at national scale in New Zealand. *Journal of Hydrology*, 568(July 2018), 646–657. https://doi.org/10.1016/j.jhydrol.2018.11.025

Sivapalan, M., Yaeger, M. A., Harman, C. J., Xu, X., & Troch, P. A. (2011). Functional model of water balance variability at the catchment scale: 1. Evidence of hydrologic similarity and space-time symmetry. *Water Resources Research*, *47*(2), 1–18. https://doi.org/10.1029/2010WR009568

Trancoso, R., Larsen, J. R., McVicar, T. R., Phinn, S. R., & McAlpine, C. A. (2017). CO2-vegetation feedbacks and other climate changes implicated in reducing base flow. *Geophysical Research Letters*, *44*(5), 2310–2318. https://doi.org/10.1002/2017GL072759

Wada, Y., Flörke, M., Hanasaki, N., Eisner, S., Fischer, G., Tramberend, S., et al. (2016). Modeling global water use for the 21st century: The Water Futures and Solutions (WFaS) initiative and its approaches. *Geoscientific Model Development*, *9*(1), 175–222. https://doi.org/10.5194/gmd-9-175-2016

Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment Classification and Hydrologic Similarity. *Geography Compass*, *1*(4), 901–931. https://doi.org/10.1111/j.1749-8198.2007.00039.x

Wang, C., Wang, S., Fu, B., & Zhang, L. (2016). Advances in hydrological modelling with the Budyko framework: A review. *Progress in Physical Geography*, *40*(3), 409–430. https://doi.org/10.1177/0309133315620997

Westerberg, I. K., & McMillan, H. K. (2015). Uncertainty in hydrological signatures. *Hydrology and Earth System Sciences*, 19(9), 3951–3968. https://doi.org/10.5194/hess-19-3951-2015

Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C., et al. (2012). Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers. *Water Resources Research*, *48*(6), 1–13. https://doi.org/10.1029/2011WR011586

Winter, T. C. (2001). The concept of hydrologic landscapes. *Journal of the American Water Resources Association*, *37*(2), 335–349. https://doi.org/10.1111/j.1752-1688.2001.tb00973.x

Woods, R. A. (2009). Analytical model of seasonal climate impacts on snow hydrology: Continuous snowpacks. *Advances in Water Resources*, 32(10), 1465–1481. https://doi.org/10.1016/j.advwatres.2009.06.011

Yao, L., Sankarasubramanian, A., & Wang, D. (2021). Climatic and Landscape Controls on Long-Term Baseflow. *Water Resources Research*, 57(6), 1–20. https://doi.org/10.1029/2020WR029284

Yokoo, Y., Sivapalan, M., & Oki, T. (2008). Investigating the roles of climate seasonality and landscape characteristics on mean annual and monthly water balances. *Journal of Hydrology*, 357(3–4), 255–269. https://doi.org/10.1016/j.jhydrol.2008.05.010

Zhang, Y. K., & Schilling, K. E. (2006). Increasing streamflow and baseflow in Mississippi River since the 1940 s: Effect of land use change. *Journal of Hydrology*, *324*(1–4), 412–422. https://doi.org/10.1016/j.jhydrol.2005.09.033

Zhou, G., Wei, X., Chen, X., Zhou, P., Liu, X., Xiao, Y., et al. (2015). Global pattern for the effect of climate and land cover on water yield. *Nature Communications*, 6, 1–9. https://doi.org/10.1038/ncomms6918