# Is there a baseflow Budyko curve?

**Sebastian J. Gnann[1], Ross A. Woods[1], Nicholas J. K. Howden[1]**

[1]Department of Civil Engineering, University of Bristol, Bristol, UK

**Key Points:**

- The fraction of precipitation that becomes baseflow cannot be estimated using the aridity index alone
- In humid catchments the baseflow fraction is limited by a catchment's wetting potential (storage capacity)
- In arid catchments the baseflow fraction is limited by high vaporisation amounts

**Plain Language Summary** Baseflow originates from stored water (e.g. groundwater) and sustains river flow in dry periods, which makes it an important water resource. Baseflow is known to vary with climate and landscape properties such as geology or vegetation, but there is no universal theory to explain this variability. To explore baseflow variability, we use data from several hundred catchments in the US and the UK. We investigate whether a catchment's baseflow fraction, i.e., the fraction of rainfall that becomes baseflow, can be attributed primarily to the aridity index, a commonly used climate index. The aridity index is defined as the ratio between potential evapotranspiration (available energy) and precipitation (available water). We find that in humid catchments (low aridity index), baseflow cannot be attributed primarily to the aridity index. Rather, a catchment's capacity to store water determines how much precipitation becomes baseflow. In arid catchments (high aridity index), the aridity index can be seen as the primary determinant of baseflow fraction. It strongly influences how much of the precipitation can be evaporated back to the atmosphere and thus cannot become baseflow. These results might help to assess how water availability (in the form of baseflow) changes with changing climate and land use.

---

Corresponding author: Sebastian J. Gnann, `sebastian.gnann@bristol.ac.uk`

**Abstract**

There is no general theory to explain differences in baseflow between catchments, despite evidence that it is mainly controlled by climate and landscape. One hypothesis is that baseflow fraction (the ratio between baseflow and precipitation) can be primarily attributed to the aridity index (the ratio between potential evapotranspiration and precipitation), i.e. that there is a "baseflow Budyko curve". Comparing catchment data from the US and the UK shows, however, that aridity is not always a good predictor of baseflow fraction. We use the revised Ponce-Shetty annual water balance model to show that there is no single "baseflow Budyko curve", but rather a continuum of curves emerging from a more universal model that incorporates both climate and landscape factors. In humid catchments, baseflow fraction is highly variable due to variations in a catchment's wetting potential, a parameter that describes catchment storage capacity. In arid catchments, vaporisation limits its baseflow generation which leads to lower variability in baseflow fraction. Generally, when the magnitude of precipitation is important, the aridity index only partly explains baseflow response. Adapting the model to explain variability of the baseflow index (the ratio between baseflow and total streamflow) shows that the aridity index is generally a poor predictor of baseflow index. While the wetting potentials and other parameters are obtained by fitting the Ponce-Shetty model to annual catchment data, their links to physical properties remain to be explored. This currently limits the model's applicability to gauged catchments with sufficiently long records.

## 1 Introduction

Baseflow is defined as flow derived from groundwater and other delayed sources and thus sustains streamflow also during dry periods [*Hall*, 1968; *Smakhtin*, 2001]. Understanding how baseflow varies with changing climate and landscape properties is crucial for various issues related to water quantity and quality [e.g. *Smakhtin*, 2001; *Price*, 2011; *Beck et al.*, 2013; *Buttle*, 2018]. Population growth is linked to an increase in freshwater demand for agriculture, industry and human consumption and water shortages pose a threat even in humid regions [*Price*, 2011]. Baseflow is essential for ecosystem functioning and provides habitat for stream biota [*Poff et al.*, 1997; *Price*, 2011]. Furthermore, baseflow is important with respect to water quality issues (chemistry, temperature) such as effluent-load from wastewater treatment plants [*Smakhtin*, 2001; *Ficklin et al.*, 2016]. If we want to understand how humans impact baseflow, we need to understand what determines baseflow under (near-)natural conditions.

Many studies found that baseflow is correlated with climate and landscape properties such as soils, geology, topography and vegetation, but a universal relationship or general theory is yet to be found [*Price*, 2011]. Geology was found to be the key variable in various regional studies [e.g. *Neff et al.*, 2005; *Longobardi and Villani*, 2008; *Bloomfield et al.*, 2009]. Similarly, soil classes (which are correlated with geology) were used to explain baseflow variability in the UK [*Boorman et al.*, 1995] and Europe [*Schneider et al.*, 2007]. *Schneider et al.* [2007] found that soils were less influential towards southern Europe, which might be attributed to differences in topography and climate. *Van Dijk* [2010] explored catchments in Australia and concluded that climate was the most important control on baseflow, while *Lacey and Grayson* [1998] found that for southeastern Australia vegetation-geology groups were the main influence. In summary, the studies that found landscape properties to be most influential were usually of regional nature and thus investigated catchments with relatively similar climates. Continental studies and the first global study by *Beck et al.* [2013] led to somewhat inconclusive results. While some key landscape and climate characteristics could be identified, the underlying processes remain to be explained. The influence of lakes [*Neff et al.*, 2005] and snow [*Beck et al.*, 2013], i.e. baseflow generating mechanisms different than groundwater discharge, further complicates the analysis.

Baseflow is usually quantified by the baseflow index (BFI), the long-term ratio between baseflow and total streamflow. Alternatively, we can use the baseflow fraction $K_B$ [*Sivapalan et al.*, 2011], defined as the ratio between mean annual baseflow $\overline{Q_b}$ and precipitation $\overline{P}$ (cf. to the runoff ratio, the ratio between total streamflow $\overline{Q}$ and precipitation $\overline{P}$). $K_B$ has the advantage that it relates baseflow to precipitation, a climate input that is (mostly) independent of catchment form. The similarity to the runoff ratio allows us to investigate $K_B$ in the context of the Budyko hypothesis. A disadvantage of $K_B$ is that we need both streamflow and rainfall data.

The Budyko hypothesis [*Budyko*, 1974] is a widely applied empirical top-down approach in catchment hydrology [*Wang et al.*, 2016]. It hypothesises that the ratio between mean annual actual evapotranspiration $\overline{E_a}$ and precipitation $\overline{P}$ is primarily a function of the ratio between mean annual potential evapotranspiration $\overline{E_p}$ and precipitation $\overline{P}$ , i.e. the aridity index $\varphi = \overline{E_p}/\overline{P}$. As $\overline{E_a}$ is usually not available, $\overline{Q}$ might be used instead [*Andréassian and Perrin*, 2012]. Figure 1a shows a Budyko-type plot for catchments in the US and the UK (data sources will be explained in Section 2.2). The catchments fall relatively close to a single curve, the so called the Budyko curve, for which various model equations exist [see e.g. review by *Wang et al.*, 2015]. Is there a similar behaviour for baseflow, i.e. a baseflow Budyko curve? That is, is the aridity index the primary control on baseflow fraction? *Wang and Wu* [2013] modelled the relationship between baseflow fraction and aridity by means of a Budyko-type curve that approaches unity for increasing humidity. Similarly, *Sivapalan et al.* [2011] reported "that the fraction of precipitation partitioned to slow flow is highest in wet catchments (as high as 0.7) and decreases with increasing aridity". Both studies analysed MOPEX data [*Duan et al.*, 2006], that is data from the contiguous US. Redoing this analysis with data from the US and the UK reveals a different behaviour. We can see from Figure 1b that the fraction of precipitation that becomes baseflow does not always increase with decreasing aridity index but decreases for many humid catchments.

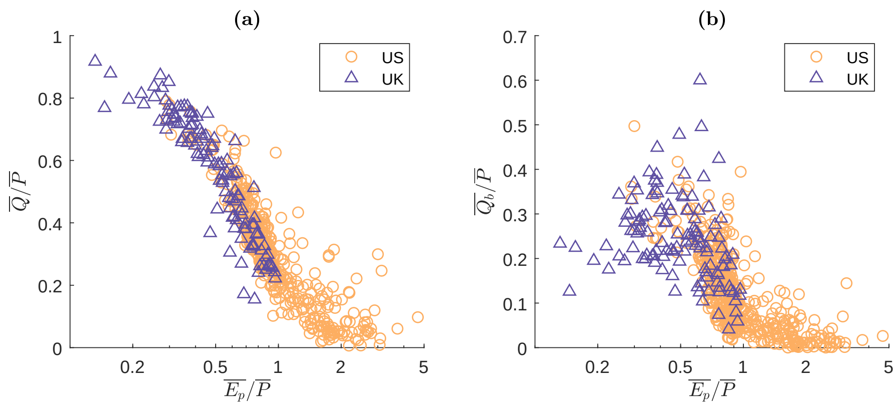



**Figure 1.** Budyko-type curves relating **(a)** mean annual runoff ratio $\overline{Q}/\overline{P}$ to mean aridity index $\overline{E_p}/\overline{P}$ and **(b)** mean annual baseflow fraction $\overline{Q_b}/\overline{P}$ to mean aridity index $\overline{E_p}/\overline{P}$. US catchments are denoted by orange circles, UK catchments are denoted by purple triangles. Catchments with significant snow fractions were removed.

The data presented in Figure 1 suggest that the influence of climate aridity on baseflow fraction is not straightforward or universal. This reinforces the variability in the literature on the relative importance of climate and landscape characteristics. Is there a way to quantify and/or parametrise these relative importances? Can we disentangle the influences of different causal factors such as forcing and catchment form? How can we model baseflow variability in a process-based way? As a framework for addressing these questions, we will use the revised Ponce-Shetty model [*Ponce and Shetty*, 1995a,b; *Sivapalan et al.*, 2011] to model catchment water balance at the annual scale. The Ponce-Shetty model has been described as a functional model [*Sivapalan et al.*, 2011] as it focuses on how water is partitioned, stored and released, i.e. a catchment's functions [*Black*, 1997; *Wagener et al.*, 2007]. This approach is promising as it goes beyond mere empiricism by representing processes such as the partitioning of water at the annual scale. The processes and the respective parameters are arguably highly abstracted and connecting emergent parameters to catchment characteristics remains challenging [*Sivapalan et al.*, 2011]. This approach, however, allows us to investigate large samples of catchments and thus enables us to explore catchment (dis-)similarity and patterns which eventually might be synthesised to new catchment-scale theory [*Sivapalan*, 2005; *McDonnell et al.*, 2007; *Wagener et al.*, 2007; *Harman and Troch*, 2014]. In the face of environmental change [*Milly et al.*, 2008], process-based models that allow for extrapolation are more needed than ever [*Wagener et al.*, 2010].

We will use the revised Ponce-Shetty annual water balance model to obtain and investigate a theoretical model of baseflow fraction (and baseflow index) as a function of mean annual climate variables [*Sivapalan et al.*, 2011]. We will fit the Ponce-Shetty model to catchments in the US and the UK to obtain catchment-scale parameter values defining how water is partitioned at the annual scale (Ponce-Shetty parameters; described in Section 2). We will then assess whether this approach has the potential to explain the variability in baseflow fraction (and baseflow index) shown in Figure 1b and the apparently differing behaviour exhibited by the catchments in the UK.

# 2 Theory and Data

## 2.1 Theory

### *2.1.1 Annual Water Balance Model*

The revised Ponce-Shetty model [*Sivapalan et al.*, 2011] is a functional approach to water balance modelling following *Horton* [1933], *L'vovich* [1979] and *Ponce and Shetty* [1995a,b]. A catchment's annual water balance is conceptualised as a two-stage partitioning process. First, precipitation $P$ is partitioned into fast flow $Q_f$ (direct runoff and fast subsurface flow) and wetting $W$ (water that is being stored). The stored water is then further partitioned into vaporisation $V$ (water returned to the atmosphere) and baseflow (slow flow) $Q_b$. Fast flow and baseflow combined yield total streamflow $Q$. Inter-annual water storage change and other water gains or losses such as inter-catchment groundwater flows are assumed to be negligible. Figure 2 shows a schematic of the model.

The balance equations for the two partitioning stages are:

$$P = Q_f + W \tag{1}$$
$$W = Q_b + V \tag{2}$$

The balance equations for the whole catchment are:

$$P = V + Q \tag{3}$$
$$Q = Q_f + Q_b \tag{4}$$

These balance equations are used to determine $V$ (from Equation (4)) and $W$ (from Equation (1)). Data sources for $Q$ and $P$ and the estimation of $Q_f$ and $Q_b$ are described in the following subsections.

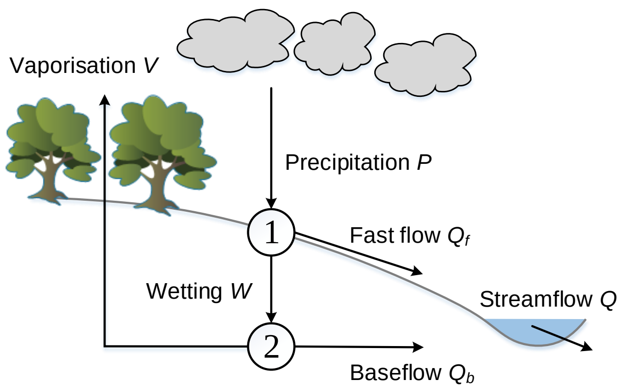



**Figure 2.** Schematic of the Ponce-Shetty model indicating the two partitioning stages (1) and (2).

#### *2.1.2 Baseflow Estimation*

To obtain an estimate of fast flow and baseflow we perform a hydrograph separation using digital filtering techniques. Following *Troch et al.* [2009] who reported that the choice of the filter has no significant influence on annual water balance metrics (they analysed the Horton index), many subsequent studies used only one hydrograph separation technique [e.g. *Sivapalan et al.*, 2011; *Harman et al.*, 2011]. Since in the original *Troch et al.* [2009] paper only 33 catchments were analysed, we perform a comparative analysis of baseflow separation methods for all the catchments investigated here. We use the one-parameter Lyne-Hollick digital filter [*Lyne and Hollick*, 1979] which is applied forwards, backwards and forwards again using a filter parameter of 0.925. As an alternative, we test the UK Institute of Hydrology (UKIH) smoothed minima method [*Institute of Hydrology*, 1980]. Both filters have the advantage of being only minimally parameterised (one parameter) and thus being easily applied to a large sample of catchments. Knowing $P$, $Q$ (both measured), $Q_f$, $Q_b$ (both estimated), we can then calculate $V$ and $W$.

#### *2.1.3 Ponce-Shetty Equations*

Based on empirical observations *Ponce and Shetty* [1995a] presented a mathematical model of the two-stage partitioning which was re-introduced by *Sivapalan et al.* [2011]. The form of the equations follows the curve number runoff equation [*NRCS*, 2004], which is an empirical equation that satisfies conservation of mass. The idea of two competing processes (here: fast flow vs. wetting and baseflow vs. vaporisation) was later generalised by means of the so called proportionality hypothesis and the Maximum Entropy Production (MEP) principle was identified as a possible thermodynamic basis for this mathematical form [*Wang and Tang*, 2014; *Wang et al.*, 2015; *Zhao et al.*, 2016].

The first partitioning stage is modelled as follows:

$$Q_f = \begin{cases} 0, & \text{if } P \le \lambda_P W_\mathrm{p} \\ \frac{(P-\lambda_P W_p)^2}{P+(1-2\lambda_P)W_p}, & \text{if } P > \lambda_P W_\mathrm{p} \end{cases} \tag{5}$$

$$W = \begin{cases} P, & \text{if } P \le \lambda_P W_\mathrm{p} \\ P - \frac{(P-\lambda_P W_p)^2}{P+(1-2\lambda_P)W_p}, & \text{if } P > \lambda_P W_\mathrm{p} \end{cases} \tag{6}$$

$$P \to \infty,\ Q_f \to P - W_p,\ W \to W_p \tag{7}$$

where $\lambda_P$ is the fast flow initial abstraction coefficient and $W_p$ is the wetting potential. Their product $\lambda_P W_p$ is the fast flow generation threshold. This form is convenient as $\lambda_P$ ranges between zero and unity [*Ponce and Shetty*, 1995a]. The second partitioning stage is modelled as follows:

$$Q_b = \begin{cases} 0, & \text{if } W \leq \lambda_W V_\mathrm{p} \\ \frac{(W-\lambda_W V_p)^2}{W+(1-2\lambda_W)V_p}, & \text{if } W > \lambda_W V_\mathrm{p} \end{cases} \tag{8}$$

$$V = \begin{cases} W, & \text{if } W \leq \lambda_W V_\mathrm{p} \\ W - \frac{(W-\lambda_W V_p)^2}{W+(1-2\lambda_W)V_p}, & \text{if } W > \lambda_W V_\mathrm{p} \end{cases} \tag{9}$$

$$W \to \infty,\ Q_b \to W - V_p,\ V \to V_p \tag{10}$$

where $\lambda_W$ is the baseflow initial abstraction coefficient and $V_p$ is the vaporisation potential. Their product $\lambda_W V_p$ is the baseflow generation threshold.

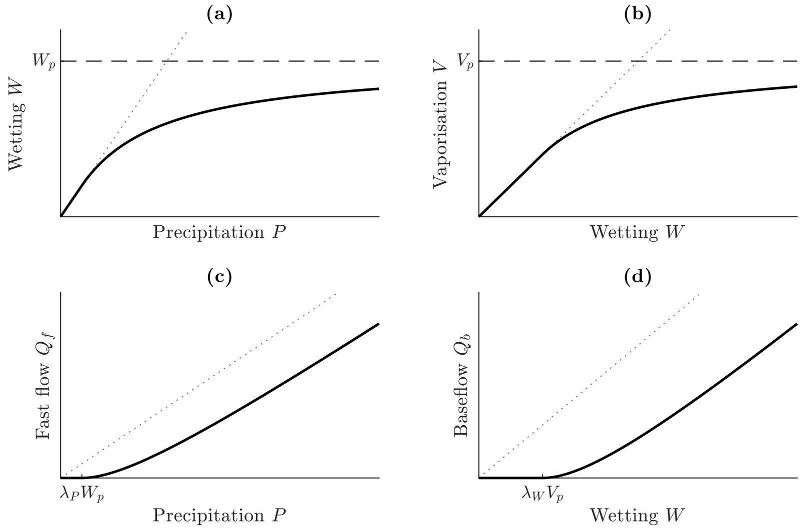



**Figure 3.** Example L'vovich-type curves: **(a)** precipitation-wetting curve (Equation (6)), **(b)** wetting-vaporisation curve (Equation (9)), **(c)** precipitation-fast flow curve (Equation (5)), **(d)** wetting-baseflow curve (Equation (8)). The dotted lines indicate the lines through the origin, which (in theory) cannot be exceeded. The dashed lines indicate the potentials. The ticks indicate the thresholds.

Figure 3 shows curves derived from the Ponce-Shetty model equations. Both the $P$-$W$-plot (Figure 3a) and the $W$-$V$-plot (Figure 3c) start at the origin and approach a limit (their potentials). The wetting potential $W_p$ can be seen as some sort of storage capacity of a catchment. The vaporisation potential $V_p$ can be seen as some sort of energy limit (somewhat analagous to potential evapotranspiration). The $P$-$Q_f$-plot (Figure 3b) and the $W$-$Q_b$-plot (Figure 3d) start to rise after a certain threshold and then rise without a (theoretical) limit. The precipitation threshold is a minimum amount of rainfall required to generate fast slow. The baseflow threshold is a minimum amount of wetting required to generate baseflow. This reflects the idea that if there is only little rain (or wetting), the water will not reach the stream and evaporate (e.g. interception). The physical meaning of these parameters is somewhat ambiguous as they are emergent parameters representing processes over a large area (catchment) and over a long time (years). Links to physical (observable) catchment characteristics remain to be explored, but will be discussed qualitatively in Section 4.

#### 2.1.4 Rescaled Form of the Ponce-Shetty Equations

In order to compare between catchments the (mean annual) Ponce-Shetty variables can be normalised using the Ponce-Shetty parameters [*Sivapalan et al.*, 2011]. We define two rescaled driving variables: rescaled (mean annual) precipitation $\tilde{P}$ and a rescaled vaporisation potential $\tilde{V}_p$.

$$\tilde{P} = \frac{\overline{P} - \lambda_P W_p}{(1 - \lambda_P) W_p} \tag{11}$$

$$\tilde{V}_p = \frac{V_p - \lambda_W V_p}{(1 - \lambda_P) W_p} \tag{12}$$

#### 2.1.5 Catchment Indices

We define two catchment indices: the baseflow fraction $K_B$ (note that this definition is slightly different from the usual definition as it includes the parameter $\lambda_W V_p$) and the baseflow index BFI.

$$K_B = \frac{\overline{Q_b}}{\overline{P} - \lambda_W V_p} \tag{13}$$

$$\mathrm{BFI} = \frac{\overline{Q_b}}{\overline{Q}} \tag{14}$$

We can approximate these indices using the rescaled driving variables (Equations (11) and (12)) [for the full derivation of $K_B$ see *Sivapalan et al.*, 2011, and for the derivation of BFI see Appendix A: ]:

$$K_B = \frac{\tilde{P}}{(1 + \tilde{P})(\tilde{P} + \tilde{V}_p + \tilde{V}_p \tilde{P})} \tag{15}$$

$$\mathrm{BFI} = \frac{1}{(1 + \tilde{P})(1 + \tilde{V}_p)} \tag{16}$$

These expressions can be used to model the observed catchment indices (Equations (13) and (14)). These equations are functions of two variables ($\tilde{V}_p$ and $\tilde{P}$) and not just a single variable such as aridity (which might be defined here as rescaled aridity index $\tilde{\varphi} = \frac{\tilde{V}_p}{\tilde{P}}$). Note that in the derivation of these equations we assume a parameter $K = \frac{\lambda_P W_p - \lambda_W V_p}{(1-\lambda_P) W_p}$ (not presented here for brevity) to be zero. This assumption led to insignificant differences which is consistent with *Sivapalan et al.* [2011].

### 2.2 Data

We use data from the contiguous US and Great Britain. CAMELS [*Newman et al.*, 2015; *Addor et al.*, 2017a] includes daily precipitation, potential evapotranspiration and streamflow data as well as a wide range of catchment attributes for 671 catchments in the contiguous US. The UK Benchmark Network (UKBN2) [*Harrigan et al.*, 2017] describes catchments in the UK that are near-natural. It consists of 146 catchments whereof 8 catchments in Northern Ireland are not considered. The data is obtained from different sources. Daily streamflow data, catchment characteristics and catchment boundaries are obtained from the NRFA [*National River Flow Archive*, 2018], precipitation data from CEH-GEAR [*Tanguy et al.*, 2016], and potential evapotranspiration data from CHESS-PE [*Robinson et al.*, 2016]. We trim the daily data to contain only full water years (starting 1 October). We then aggregate daily data to obtain annual data, which are used to calibrate the Ponce-Shetty model for each catchment. For all other calculations we use mean annual data, i.e. data averaged over all full water years. To obtain a suitable dataset we remove some of the catchments according to the following criteria:

- Catchments with areas smaller than 10 $km^2$ as measurement errors and catchment delineation errors tend to be significant for very small catchments.
- Catchments with records shorter than 15 years as calibrating the Ponce-Shetty model requires many annual values. This threshold is chosen to remove some rather short and thus potentially unreliable records, while trying to keep enough catchments for the ongoing analysis.
- Catchments where snow and lakes are influential, as these processes are not considered in the Ponce-Shetty model. We remove catchments with fractions of precipitation falling as snow > 0.2 and catchments with significant surface water bodies. The latter is done by removing UKBN2 catchments with FARL < 0.8 (a parameter quantifying the influence of lakes and reservoirs) and CAMELS catchments with frac_water > 0.05.
- Catchments with runoff ratios larger than unity in any year of record ($Q/P > 1$), resulting in negative vaporisation values ($V < 0$), as this indicates significant water balance issues and thus violates the assumptions of the Ponce-Shetty model.

The final dataset consists of 571 out of 817 catchments.

## 3 Results

### 3.1 Baseflow Estimation

Table 1 shows several metrics comparing results obtained using the Lyne-Hollick filter [*Lyne and Hollick*, 1979] and the UKIH method [*Institute of Hydrology*, 1980]. The two methods show good agreement. While the choice of filter might have a significant impact on individual catchments, it does not alter the overall results. We continue using the baseflow estimates obtained by using the Lyne-Hollick filter.

**Table 1.** Comparison of mean annual baseflow $\overline{Q_b}$, Ponce-Shetty parameters, $K_B$ and BFI using different baseflow separation techniques (Lyne-Hollick filter and UKIH method). The relative error (RE) is defined as RE = $\left|1 - \frac{x_a}{x_b}\right|$. The absolute error (AE) is defined as AE = $|x_a - x_b|$.

| | $\overline{Q_b}$ [mm] | $W_p$ [mm] | $\lambda_P$ [-] | $V_p$ [mm] | $\lambda_W$ [-] | $K_B$ [-] | BFI [-] |
|---|---|---|---|---|---|---|---|
| Pearson correlation | 1.00 | 0.84 | 0.98 | 0.95 | 0.97 | 0.99 | 0.93 |
| Spearman correlation | 1.00 | 0.99 | 0.96 | 0.99 | 0.95 | 0.99 | 0.96 |
| Median RE | 0.07 | 0.05 | 0.17 | 0.05 | 0.31 | 0.07 | 0.07 |
| Median AE | 11 | 159 | 0.01 | 147 | 0.00 | 0.01 | 0.03 |

### 3.2 Parameter Estimation and Uncertainty

The Ponce-Shetty parameters are fitted to each individual catchment by means of a non-linear least squares fitting algorithm, whereby $\lambda_P$ and $\lambda_W$ are restricted to be between zero and unity (their theoretical limits), and $W_p$ and $V_p$ are restricted to be between 0 mm and an upper limit. We choose an (arbitrary) upper limit of 50000 mm which is deemed high enough to not affect the parameter estimation. An even higher limit does not affect the estimated parameter values except for very few catchments with $W_p$ and/or $V_p$ values which are (almost) at the limit. The problem that some of the obtained parameter values are at the upper limit is discussed in the next paragraph. We can use two values for the wetting $W$ to fit the second partitioning stage. Either the observed $W$ obtained from Equation (1) or the modelled $W$ following from the fitted model for the first partitioning stage

(Equation (6)). Following [*Sivapalan et al.*, 2011] we use the modelled $W$ to obtain an internally consistent water balance.

To fit a meaningful parameter set, the catchments should exhibit their functional behaviour [*Sivapalan et al.*, 2011]. If the vaporisation values (wetting values) are far away from the vaporisation potential (wetting potential), we will have a roughly linear relationship and hence fitting the functional form is not possible (see Figure 4a). This can be seen especially for $V_p$ in arid catchments (e.g. in the middle of the US). In these catchments, the obtained potentials are at the specified upper limit (50000 mm). Similarly, being at the potential all the time does not allow us to fit a functional relationship either; this can be seen especially for $V_p$ in humid catchments (e.g. along the west coast of the UK). In these catchments the obtained initial abstraction coefficient is unity (see Figure 4b). We remove these catchments from the analysis because the Ponce-Shetty model is unable to describe them adequately.

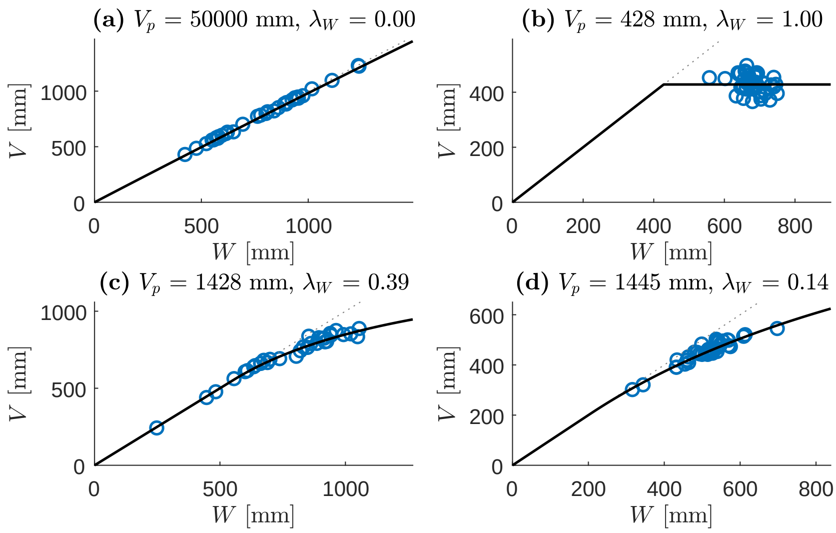



**Figure 4.** Examples of catchments (station numbers in brackets) with fitted $W$-$V$-curves. **(a)** Coleto Creek, Texas (08176900): extremely high $V_p$, $V_p$ not identifiable. **(b)** Aire, Yorkshire (27035): $V$ always approximately equal to $V_p$, $\lambda_W$ not identifiable. **(c)** Bear Creek, Texas (08158810): $V_p$ and $\lambda_W$ identifiable. **(d)** Pincey Brook, Essex (38026): $V_p$ and $\lambda_W$ identifiable.

Table 2 shows overall statistics for the parameter estimation after having removed the catchments described in the last paragraph. The parameter uncertainty (in the form of 95% confidence intervals) is particularly high for extremely large values for either of the potentials ($\gg$ 10000 mm). These large values are consistently uncertain, which coincides with *Sivapalan et al.* [2011] who found that for some catchments the (apparently very high) potentials could not be properly identified. The confidence intervals for $\lambda_P$ and $\lambda_W$ need careful interpretation, as these two parameters have heavily skewed distributions (most catchments have parameter values close to zero). We do not remove catchments with high uncertainty from the analysis as a threshold would necessarily be subjective, which leaves us with 545 catchments for the ongoing analysis.

### 3.3 Maps of Ponce-Shetty Parameters and Baseflow Metrics

Figure 5 shows maps of the fitted parameters for CAMELS catchments. The patterns agree well with *Sivapalan et al.* [2011] who used MOPEX catchments. High wetting

**Table 2.** Parameter statistics and uncertainty for all catchments used in the analysis. *CI 95%* denotes the 95% confidence interval. *Rel. CI 95%* denotes the relative confidence limits, i.e. the confidence limits normalised by the parameter values. *Spearman* denotes the Spearman correlation of the relative confidence limits with the parameter values.

| | Min | Median | Max | Median CI 95% | Median Rel. CI 95% | Spearman |
|---|---|---|---|---|---|---|
| $W_p$ [mm] | 756 | 3044 | 42857 | 1591 | 0.50 | 0.32 |
| $\lambda_P$ [-] | 0 | 0.05 | 0.64 | 0.12 | >1 | -0.91 |
| $V_p$ [mm] | 316 | 2911 | 44652 | 2264 | 0.74 | 0.49 |
| $\lambda_W$ [-] | 0 | 0.02 | 0.91 | 0.13 | >1 | -0.91 |

potentials $W_p$ can be seen in the middle of the US (Great Plains), in the east (southern parts of the Appalachians), south east (around Florida) and in parts of the central north (Michigan). High vaporisation potentials $V_p$ can be seen in the middle of the US (Great Plains) and in all southern regions. The fast flow thresholds $W_p\lambda_P$ are high in the south, the south east and in the middle of the US except for the north. The baseflow thresholds $V_p\lambda_W$ are similarly high in most of these areas, but also in some catchments along the west coast. The spatial similarity of the thresholds is reflected by a significant rank correlation of 0.61 between $W_p\lambda_P$ and $V_p\lambda_W$.

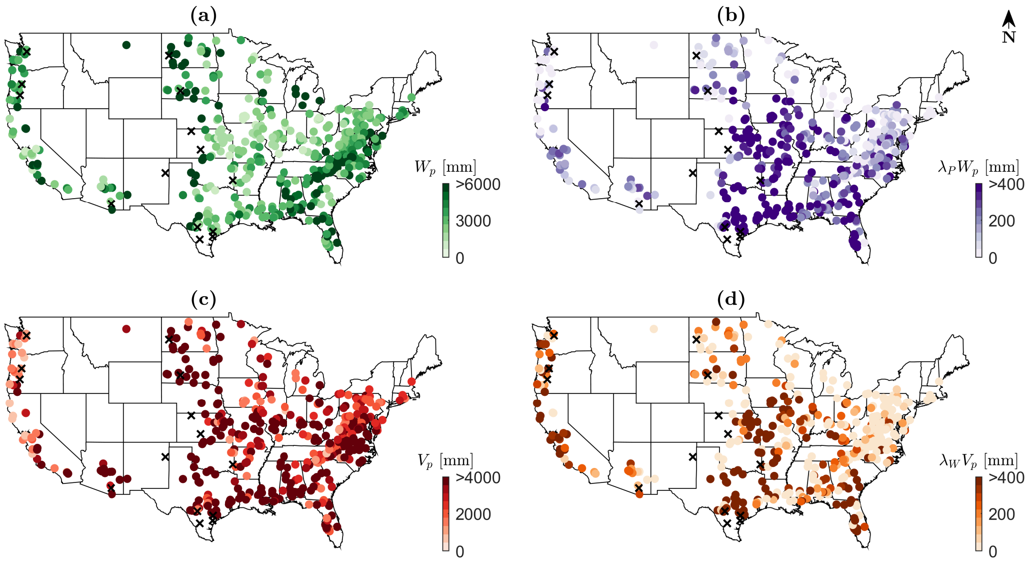



**Figure 5.** The fitted parameters for CAMELS catchments: wetting potential **(a)**, fast flow threshold **(b)**, vaporisation potential **(c)**, and baseflow threshold **(d)**. Crosses denote catchments where some of the parameters could not be identified properly.

Figure 6 shows maps of the fitted parameters for UKBN2 catchments. On average, the values are lower than for the CAMELS catchments, especially for $V_p$, which is consistent with generally lower vaporisation intensities (cf. to $E_p$). High wetting potentials $W_p$ can be found in the south west, the south, the middle (the Midlands) and along the south eastern coast. The vaporisation potentials $V_p$ are high in the south, especially in the south east. High $W_p\lambda_P$ can be found in the south east and for a few catchments in the north.

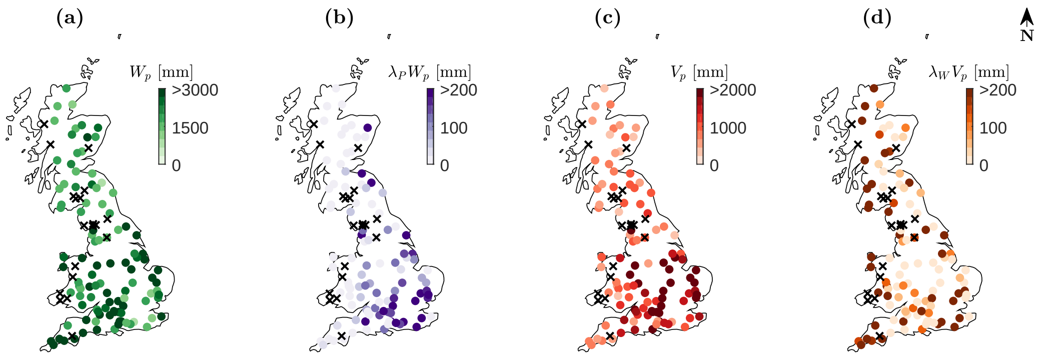



**Figure 6.** The fitted parameters for UKBN2 catchments: wetting potential **(a)**, fast flow threshold **(b)**, vaporisation potential **(c)**, and baseflow threshold **(d)**. Crosses denote catchments where some of the parameters could not be identified properly.

High $V_p \lambda_W$ can be found in catchments scattered throughout the UK, most notably all along the west coast and in the south east.

Figure 7 shows maps of $K_B$ and BFI for CAMELS and UKBN2 catchments. Generally, $K_B$ is lower than BFI as it compares $\overline{Q_b}$ to $\overline{P}$ rather than $\overline{Q}$, which is always lower than $\overline{P}$. This is reflected in the ranges of values shown in Figure 7. While in some regions both $K_B$ and BFI are rather high (e.g. in the eastern US and in the south west of the UK), in other regions BFI can be high while $K_B$ is rather low (e.g. in the southern US and in the middle of the US and in the south east of the UK) , which broadly agrees with *Santhi et al.* [2008] who found that catchments with high BFI can still have low baseflow volumes. This coincides with the maps showing the Ponce-Shetty parameters (Figures 5 and 6). Catchments with high $K_B$ generally have a high $W_p$, low $W_p \lambda_P$ and low $V_p$. Catchments with high BFI also occur in areas with high $V_p$.

### 3.4 Baseflow Variability with Climate Variables

Figure 8 shows how the baseflow fraction varies with the rescaled climate variables. To show the dependence of $K_B$ on both $\tilde{P}$ and $\tilde{V}_p$ we make use of a contour plot (see Figure 8a). We plot $\tilde{P}$ and $\tilde{V}_p$ on the $x$- and $y$- axes, respectively, and use contours to represent the model for $K_B$ (Equation (15)) and coloured dots to represent the observed $K_B$ values (Equation (13)). Figure 8b, shows an equivalent plot using the ratio between $\tilde{P}$ and $\tilde{V}_p$ (rescaled aridity index $\tilde{\varphi}$) with some example model curves with either fixed $\tilde{P}$ or $\tilde{V}_p$, respectively – this is comparable to common Budyko-type plots. To get a better understanding it is useful to recall how a contour plot of the rescaled aridity index would look like, which is shown in Figure 8c. The line through the origin represents a rescaled aridity index of unity, above that line (top left) are humid catchments, below that line (bottom right) are arid catchments. Note that we are using rescaled variables and hence we are not looking at the common aridity index. $\tilde{P}$ is a relative rainfall amount and $\tilde{V}_p$ is a relative vaporisation potential, both rescaled by their thresholds and the wetting potential of the catchment. The general notion that low $\tilde{\varphi}$ indicates humid (energy-limited) catchments and that high $\tilde{\varphi}$ indicates arid (water-limited) catchments is still valid.

The contours in Figure 8a start parallel to the line through the origin and thus parallel to the rescaled aridity index. They start to bend for higher values of $\tilde{P}$ (humid side of the plot) and become perpendicular to the rescaled aridity index. This demonstrates that a catchment having a certain rescaled aridity index can have very different values of $K_B$. Roughly, if both $\tilde{P}$ and $\tilde{V}_p$ are low, we get a rather high $K_B$ and if both are high, we get

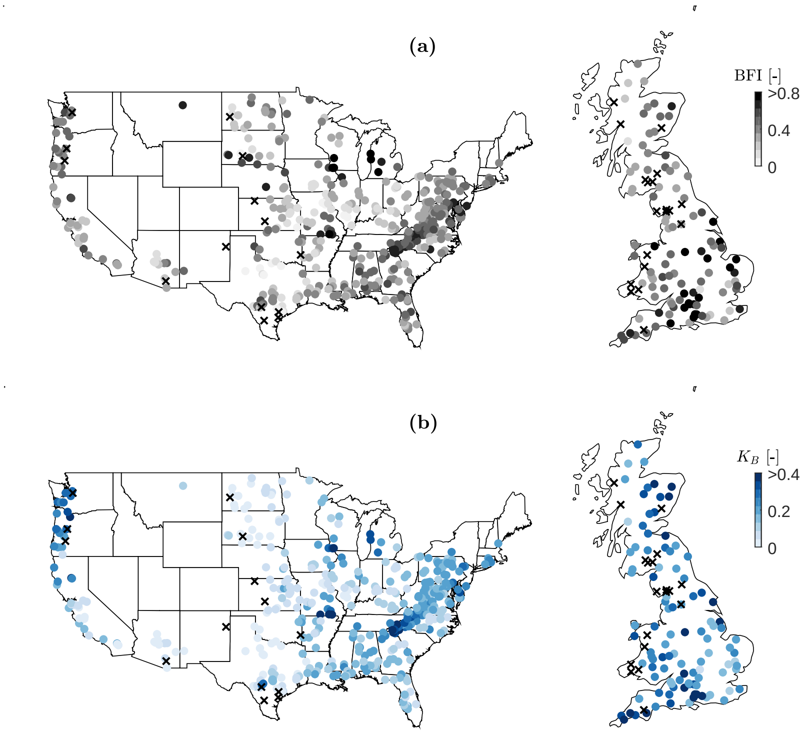



**Figure 7.** $K_B$ **(a)** and BFI **(b)** for CAMELS and UKBN2 catchments. Note that the colour scales are different to reflect the range of the values. Crosses denote catchments where some of the parameters could not be identified properly. Note that the maps of the US and the UK are not to the same scale.

a rather low $K_B$. The contours are not just bending on the humid side (top left), they are also indicating higher gradients and thus a high variability in $K_B$. In contrast, there is relatively little variation on the arid side (bottom right), i.e. most of the catchments have a similar $K_B$. The observed values (represented by coloured dots) agree well with the model contours (median absolute error = 0.02, median relative error = 0.14). This can be expected, since the model has sufficient degrees of freedom to fit the data well (the Ponce-Shetty model is fitted to each individual catchment). The Budyko-type plot shown in Figure 8b reflects these observations with a tight ensemble of curves for arid catchments and a spread out ensemble of curves for humid catchments. The observed values agree with this general behaviour, they are tight for arid catchments and scattered for humid catchments.

Figure 9 shows how BFI varies with $\tilde{P}$ and $\tilde{V}_p$. The contours shown in Figure 9a are symmetric around the line through the origin. The BFI is highest for low values of both $\tilde{P}$ and $\tilde{V}_p$ and gets lower for both higher $\tilde{P}$ and $\tilde{V}_p$. The observed values agree well with the model contours (median absolute error = 0.05, median relative error = 0.14). Again, this can be expected, since the model has sufficient degrees of freedom to fit the data well. Figure 9b shows that there is no clear relationship between BFI and the rescaled aridity index. This is in agreement with the observed values, which are scattered over most areas of the plot.

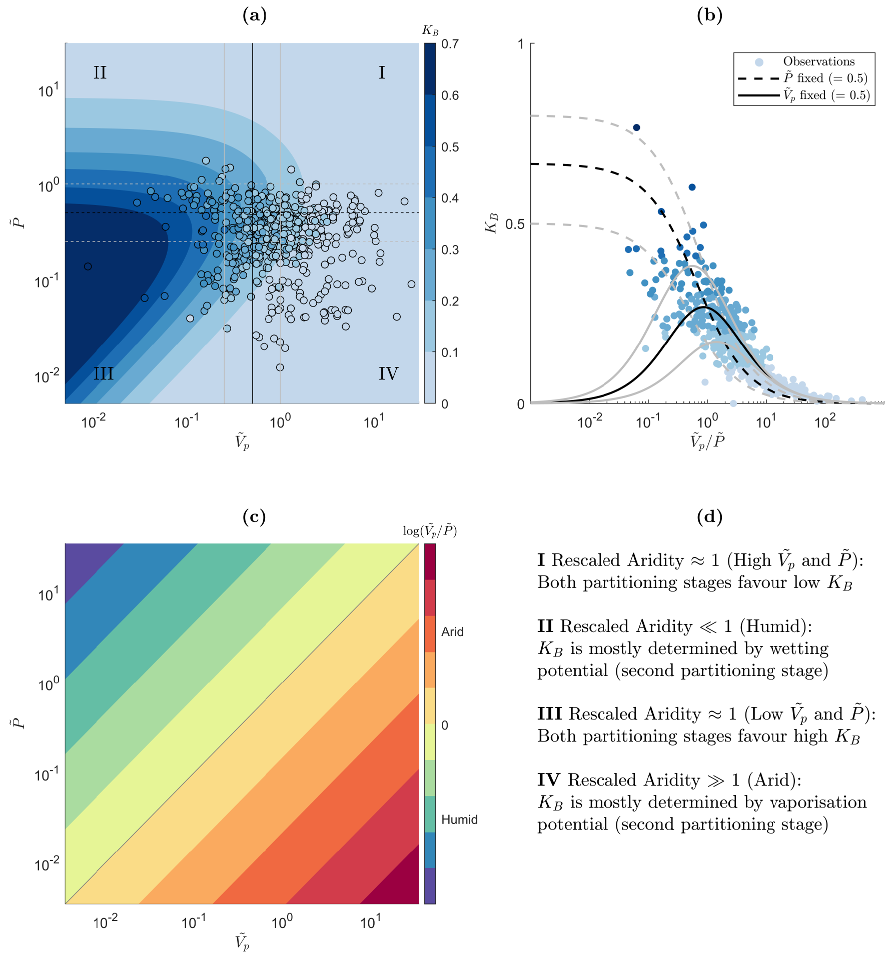



**Figure 8.** **(a)** Contour plot of $K_B$ as a function of the rescaled vaporisation potential $\tilde{V}_p$ and rescaled precipitation $\tilde{P}$ (Equation (15)). The dots indicate the observed values (Equation (13)). **(b)** $K_B$ as function of the ratio between $\tilde{V}_p$ and $\tilde{P}$ (i.e. rescaled aridity index $\tilde{\varphi}$). The black and grey lines (solid and dashed) are example model curves with either fixed $\tilde{V}_p$ or $\tilde{P}$. The dots indicate the observed values. **(c)** Logarithm of the rescaled aridity index $\tilde{\varphi}$ as a function of $\tilde{V}_p$ and $\tilde{P}$. The grey line denotes a rescaled aridity index of unity (log equals zero). **(d)** Different regions of the $K_B$ contour plot are annotated, a more detailed explanation is given in Section 4.

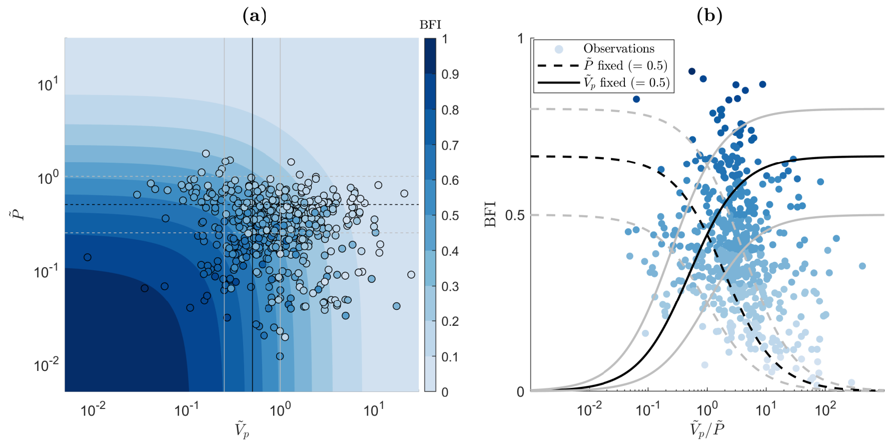



**Figure 9.** **(a)** Contour plot of BFI as a function of the rescaled vaporisation potential $\tilde{V}_p$ and rescaled precipitation $\tilde{P}$ (Equation (16)). The dots indicate the observed values (Equation (14)). **(b)** BFI as function of the ratio between $\tilde{V}_p$ and $\tilde{P}$ (i.e. rescaled aridity index $\tilde{\varphi}$). The black and grey lines (solid and dashed) are example model curves with either fixed $\tilde{V}_p$ or $\tilde{P}$. The dots indicate the observed values.

## 4 Discussion

The ranges of the parameter values (see Table 2) are in general agreement with *Sivapalan et al.* [2011] who also used a non-linear least squares method, and *Harman et al.* [2011] who used a Bayesian framework. The high parameter uncertainty for some catchments and problems in parameter identifiability might have two reasons. As described before, it could simply be a consequence of not having sufficient data to meaningfully fit the Ponce-Shetty model. It could, however, also indicate that the Ponce-Shetty model is not adequate for certain catchments. Even a good fit does not necessarily mean that the model is correctly representing the processes, which are arguably very simplified. We assume inter-annual water storage change as well as other water gains and losses to be negligible. This might not be a valid assumption for every catchment investigated here, and hence adds uncertainty to the parameter estimation. To assess the influence of inter-annual water storage change we alternatively calculated 3-year averages and calibrated the Ponce-Shetty model to these. This leads to overall similar parameter values (Pearson correlations: $W_p$: 0.86, $\lambda_P W_p$: 0.81, $V_p$: 0.79, $\lambda_W V_p$: 0.67). There are, however, problems associated with averaging. Extreme years, which are especially important to fit the Ponce-Shetty model, are averaged out and thus information is lost. Furthermore, by averaging and fitting a non-linear function, we introduce some bias ["the average of the function will not be the function of the average inputs", see *Rouholahnejad Freund and Kirchner*, 2017]. This makes it difficult to tell whether inter-annual water storage change is the cause for the deviations in the parameter values. For now we argue that the model fits our data sufficiently well for the purpose of this work. Being capable of explaining the observed variations in baseflow further corroborates the model's suitability. For specific places, however, the uncertainty might be very large and conclusions or predictions should therefore be made with care. It would be interesting to see whether more detailed modelling approaches would lead to the emergent behaviour inherent in the Ponce-Shetty theory and/or similar parameter values.

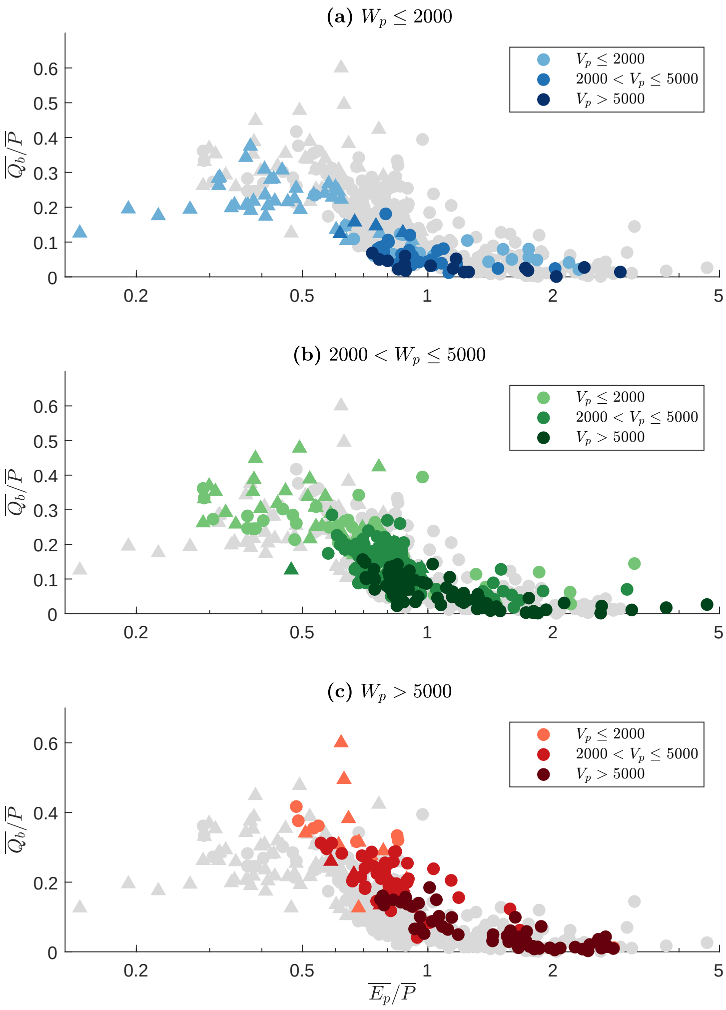
From Figure 8 we can see how $K_B$ varies with $\tilde{P}$ and $\tilde{V}_p$. Generally, $K_B$ cannot be described by a single Budyko-type curve, but by a continuum of curves that depend on the catchment's (Ponce-Shetty) parameters. $K_B$ is consistently low for high rescaled aridity values, which can be attributed to relatively high amounts of vaporisation ($K_B$ is dominated by the second partitioning stage, i.e. $V_p$). The behaviour of $K_B$ is more complicated for humid catchments. Starting at the origin of Figure 8a and moving along the $y$-axis towards more humid catchments, $K_B$ first increases, then reaches a peak and decreases again. This decrease can be attributed to an exhausted wetting potential leading to "saturation excess fast flow" ($K_B$ is dominated by the first partitioning stage, i.e. $W_p$). This was already recognised by *Milly* [1994] who stated that finite water storage capacity and finite permeability are possible causes for runoff. In such humid catchments, an increase in precipitation thus mainly leads to an increase in fast flow, which agrees with *Harman et al.* [2011] who found that fast flow elasticities are clearly larger than baseflow elasticities in humid catchments. Similarly, *Trancoso et al.* [2017] found that "higher precipitation in tropical regions may be generating more overland flow, which tends to reduce the slow component [...]".  . Baseflow fraction can hence be low for both arid and humid catchments, but for different reasons. This may help to explain the diversity of results from empirical studies on controls on baseflow.

Figure 9 shows how the BFI varies with $\tilde{P}$ and $\tilde{V_p}$. The magnitude of $\tilde{P}$ and $\tilde{V_p}$ rather than the ratio between them determines the BFI. If both $\tilde{P}$ and $\tilde{V_p}$ are low, BFI is high. That means that at the first partitioning stage precipitation becomes mainly wetting, and at the second partitioning stage this wetting becomes mainly baseflow. If either $\tilde{P}$ and $\tilde{V_p}$ are high, we obtain a lower BFI. In the first case, most of the precipitation becomes fast flow and thus the BFI is low. In the second case, most of the precipitation becomes wetting, but most of that wetting evaporates, so that $Q_b$ and thus the BFI will be rather low. In comparison to $K_B$, BFI is highly variable also for high rescaled aridity. Low amounts of baseflow (compared to precipitation) can lead to a high BFI if the amount of fast flow is even lower. This explains most of the differences between $K_B$ and BFI (see Figures 5 and 6 and the description in Section 3.3).

The results show that $K_B$ (and BFI) is influenced by the magnitude of $\tilde{P}$ and $\tilde{V_p}$ and not just their ratio. This explains the scatter especially for humid catchments (see Figure 8b). While an aridity index is certainly useful, it can be restrictive in cases where the magnitude of precipitation is important. This agrees for example with *Berghuijs et al.* [2017] who found that runoff is most sensitive to changes in precipitation and this sensitivity is not captured by only looking at the aridity index. Similarly, the ratio between precipitation and the wetting potential ($\approx \tilde{P}$) explains most of the variability in baseflow fraction which the aridity index could not explain (see Figure 8a, especially region II, and Figure 10).

Especially in humid catchments, the ratio of precipitation to a catchment's wetting potential can be a major control on baseflow. Given the same climate, a catchment with a higher wetting potential will have a higher baseflow fraction and BFI. This is a possible explanation for the partly inconclusive results found in studies before. Regional studies with similar climate could relate the amount of baseflow to a catchment's form, mostly soils [*Boorman et al.*, 1995] and geology [*Neff et al.*, 2005; *Longobardi and Villani*, 2008; *Bloomfield et al.*, 2009]. These attributes are parametrised by the Ponce-Shetty parameters (especially $W_p$), yet in a rather abstract way which so far eludes a quantitative linking to landscape characteristics. Continental [*Schneider et al.*, 2007; *Van Dijk*, 2010; *Trancoso et al.*, 2017] and global studies [*Beck et al.*, 2013, 2015] found catchment form to be less influential and often couldn't come to conclusive results, as it is neither climate nor form alone that lead to a certain catchment response, but their interaction.

Figure 10 shows the $\overline{Q_b}/\overline{P}$ vs. $\overline{E_p}/\overline{P}$ plot (from Figure 1) with catchments stratified and coloured according to their wetting and vaporisation potentials, respectively. Three different ranges of $W_p$ are shown and they form three somewhat distinct point clouds. The remaining variation can be attributed to differences in the thresholds, the rather broadly defined categories and differences in the magnitude of $\overline{E_p}$ and $\overline{P}$. The cloud with the lowest $W_p$ exhibits the lowest baseflow fraction and vice versa. High values of $K_B$ are usually associated with low values of $V_p$ (indicated by the lightness of the colours). We can also see that CAMELS and UKBN2 catchments do not generally behave differently, but since certain catchment types occur predominantly in the US or the UK, the CAMELS




483 **Figure 10.** Scatter plots of mean annual baseflow fraction $\overline{Q_b}/\overline{P}$ vs. mean aridity index $\overline{E_p}/\overline{P}$. CAMELS
484 catchments are denoted by circles, UKBN2 catchments are denoted by triangles. Catchments are highlighted
485 according to their wetting potential $W_p$: **(a)** low wetting potentials, **(b)** medium wetting potentials, and **(c)**
486 high wetting potentials. Darker shading indicates higher vaporisation potential $V_p$. All units are in mm.

and UKBN2 point clouds appear to be different. Very humid catchments with rather low $W_p$ are mostly located in the UK and they are most clearly deviating from the point cloud representing CAMELS catchments (see also Figure 1).

We did not include catchments with significant snow fraction or lakes. While these catchments might be seen as having an "extended" wetting potential (storage), they represent conceptually different processes, for which additional explanatory variables might be needed. These processes might be added as an additional partitioning stage to the model to make it more universal. Especially the snowy catchments show an increase in $K_B$ for increasing humidity almost up to unity (not shown here), which could explain e.g. why *Wang and Wu* [2013] used a baseflow Budyko model that approaches unity. Snowy catchments might be considered to have virtually unlimited storage potential as the snowpack can grow continuously, and thus baseflow fractions in these catchments can get very high.

The Ponce-Shetty parameters are emergent, rather abstract properties and relating them to catchment characteristics might not be straightforward. The Ponce-Shetty parameters are lumping a variety of processes and characteristics, notably soils, geology, vegetation, topography and climate seasonality. This means that for now, the presented model can only explain and predict annual baseflow variability in gauged catchments where the model was calibrated. It might be used to investigate the effects of a changing climate (e.g. changing precipitation) on baseflow in different types of (gauged) catchments [cf. *Buttle*, 2018]. A transfer to ungauged catchments requires a regionalisation procedure. Qualitatively, links between parameters and catchment characteristics can be seen. $V_p$ is correlated with energy availability (comparable to potential evapotranspiration), yet it rather emerges from the interaction of the available energy with vegetation and other catchment characteristics. Large wetting potentials can be seen in moorland and wetland areas (e.g. south west UK, Florida) and in the presence of major aquifers (e.g. Chalk in southern England, Great Plains aquifer). A quantitative linking of the Ponce-Shetty parameters to landscape properties or other regionalisation approaches are, however, beyond the scope of this work.

## 5 Conclusions

The present work shows that there is no single baseflow Budyko curve, that is, in general baseflow fraction cannot be modelled as a function of an aridity index alone. Even if samples of catchments seem to form a single curve, this might be misleading as many of them might actually sit on different curves (see Figure 9b). The influence of catchment water storage on long-term water balance has long been recognised [e.g. *Milly*, 1994]. The approach employed here incorporates that in a simple way by modelling baseflow fraction as a function of two variables. A rescaled precipitation, that is the ratio between precipitation and a catchment's wetting potential, and a rescaled vaporisation potential. These two variables reflect the two-stage partitioning underlying the Ponce-Shetty model, namely the partitioning between fast flow and wetting, and the subsequent partitioning between slow flow and vaporisation. Depending on the climatic regime, one of these partitioning stages dominates. In arid catchments, baseflow fraction is mainly limited by high amounts of vaporisation. In humid catchments, baseflow fraction is mainly limited by the storage capacity of a catchment.

The differences between CAMELS (US) and UKBN2 (UK) catchments shown in Figure 1b and Figure 10 have two main causes. Firstly, using aridity as a ratio is restrictive. Catchments with a similar aridity index usually have lower precipitation and vaporisation intensities in the UK than in the US. Secondly, the wetting potentials in the UK differ from the ones in the US. Most of the very humid catchments in the UK have rather low wetting potentials, i.e. they are (almost) fully saturated and a large fraction of precipitation runs off quickly to the stream. This difference is, however, not a clear distinction as it can be seen from Figure 10. Catchments in the US and the UK do not behave fundamentally differently, they rather happen to have predominantly different characteristics.

Baseflow (a catchment function) can be seen as the result of climate interacting with landscape [forcing acting on form, cf. *Wagener et al.*, 2007]. To explain baseflow variability in a process-based way, we should try to disentangle forcing and form, knowing that this might only be partially possible as catchment form (and function) may reflect a co-evolution with climate forcing. The Ponce-Shetty approach partly disentangles forcing and form, yet in a rather abstract way. Furthermore, the parameters still lump together a variety of processes that are not only reflecting catchment form (e.g. topography, geology, vegetation, etc.), but also climate (e.g. seasonality, storminess). Intra-annual climate variability can have a significant impact on such lumped parameters [*Roderick and Farquhar*, 2011; *Berghuijs and Woods*, 2016].

Using large samples of catchments allows us to detect and explain (dis-)similarities and patterns and to synthesise already available data [*Falkenmark and Chapman*, 1989; *Sivapalan*, 2005; *Harman and Troch*, 2014]. While large sample hydrology arguably neglects many details, synthesising data to find new theory has proven to be a fruitful approach that – besides improved understanding – might help to constrain models [*Shafii et al.*, 2017], to transfer knowledge to ungauged catchments [*Hrachowitz et al.*, 2013] and to deal with predictions under change [*Wagener et al.*, 2010; *Ehret et al.*, 2014]. It is essential to include a variety of catchments, both in terms of climate and landscape characteristics, which is exemplified by the "unexpected behaviour" of UK catchments in this work. Even more data are needed to corroborate the theory, to understand more of the details (e.g. Ponce-Shetty parameters) or to detect limitations of the presented approach, which eventually advances our understanding.

Simple approaches such as the Ponce-Shetty model are useful as they are easily applied to large samples. They also allow us to better understand the model's dynamics and stop us from being lost in the calibration stage. We acknowledge that there is a danger in being too simple or simple due to lack of understanding (cf. *Schwartz et al.*, 2017), which might partly be true for the hydrograph separation approach and the Ponce-Shetty model here. We are confident, however, that the chosen methods are appropriate for the present work as they are capable of explaining the observed phenomena and thus help to improve our understanding of how baseflow varies with climate and landscape.

## A: Appendix

To obtain an equation for the BFI we make use of another catchment index presented in *Sivapalan et al.* [2011], the runoff ratio $K_R$:

$$K_R = \frac{\overline{Q_f} + \overline{Q_b}}{\overline{P} - \lambda_W V_p} \tag{A.1}$$

$K_R$ can be approximated theoretically by:

$$K_R = \frac{\tilde{P}(1+\tilde{V}_p)}{\tilde{P} + \tilde{V}_p + \tilde{V}_p\tilde{P}} \tag{A.2}$$

We can write the BFI using $K_B$ and $K_R$:

$$\text{BFI} = \frac{\overline{Q_b}}{\overline{Q_f} + \overline{Q_b}} = \frac{K_B}{K_R} \tag{A.3}$$

$$\text{BFI} = \frac{\tilde{P}(1+\tilde{P})^{-1}}{\tilde{P} + \tilde{V}_p + \tilde{V}_p\tilde{P}} \left( \frac{\tilde{P}(1+\tilde{V}_p)}{\tilde{P} + \tilde{V}_p + \tilde{V}_p\tilde{P}} \right)^{-1} \tag{A.4}$$

$$\text{BFI} = \frac{1}{(1+\tilde{P})(1+\tilde{V}_p)} \tag{A.5}$$

**Acknowledgments**

This work is funded as part of the Water Informatics Science and Engineering Centre for Doctoral Training (WISE CDT) under a grant from the Engineering and Physical Sciences Research Council (EPSRC), grant number EP/L016214/1. Map colours are based on `www.ColorBrewer.org`, by Cynthia A. Brewer, Penn State. The CAMELS dataset [*Newman et al.*, 2014; *Addor et al.*, 2017b] is available from `https://ral.ucar.edu/solutions/products/camels`. Information about the UK Benchmark Network can be obtained from `https://nrfa.ceh.ac.uk/benchmark-network`. Streamflow data and catchments attributes are available from `https://nrfa.ceh.ac.uk`. CEH-GEAR precipitation data are available from `https://doi.org/10.5285/33604ea0-c238-4488-813d-0ad9ab7c51ca`. CHESS-PE potential evapotranspiration data are available from `https://doi.org/10.5285/8baf805d-39ce-4dac-b224-c926ada353b7`. Thanks to Gemma Coxon and Jim Freer for discussions and for assisting with the UKBN2 data. We thankfully acknowledge the input from the Associate Editor and three anonymous reviewers whose comments have helped to clarify and improve this manuscript.

**References**

Addor, N., A. J. Newman, N. Mizukami, and M. P. Clark (2017a), The CAMELS data set: catchment attributes and meteorology for large-sample studies, *Hydrology and Earth System Sciences*, *21*(10), 5293–5313, doi:10.5194/hess-21-5293-2017.

Addor, N., A. Newman, N. Mizukami, and M. Clark (2017b), Catchment attributes for large-sample studies, Boulder, CO: UCAR/NCAR, doi:10.5065/D6G73C3Q.

Andréassian, V., and C. Perrin (2012), On the ambiguous interpretation of the Turc-Budyko nondimensional graph, *Water Resources Research*, *48*(10), 1–5, doi:10.1029/2012WR012532.

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

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

Berghuijs, W. R., and R. A. Woods (2016), Correspondence: Space-time asymmetry undermines water yield assessment, *Nature Communications*, *7*, 1–2, doi:10.1038/ncomms11603.

Berghuijs, W. R., J. R. Larsen, T. H. van Emmerik, and R. A. Woods (2017), A Global Assessment of Runoff Sensitivity to Changes in Precipitation, Potential Evaporation, and Other Factors, *Water Resources Research*, *53*(10), 8475–8486, doi:10.1002/2017WR021593.

Black, P. E. (1997), Watershed functions, *Journal of the American Water Resources Association*, *33*(1), 1–11, doi:10.1111/j.1752-1688.1997.tb04077.x.

Bloomfield, J. P., D. J. Allen, and K. J. Griffiths (2009), Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK, *Journal of Hydrology*, *373*(1-2), 164–176, doi:10.1016/j.jhydrol.2009.04.025.

Boorman, D. B., J. M. Hollis, and A. Lilly (1995), Hydrology of soil types: a hydrologically-based classification of the soils of United Kingdom, *Institute of Hydrology*, *IH Report.*(126), 137, doi:10.1029/98GL02804.

Budyko, M. I. (1974), *Climate and Life: English Ed. edited by David H. Miller*, Academic Press.

Buttle, J. M. (2018), Mediating stream baseflow response to climate change: The role of basin storage, *Hydrological Processes*, *32*(3), 363–378, doi:10.1002/hyp.11418.

Duan, Q., J. Schaake, V. Andréassian, S. W. Franks, G. Goteti, H. V. Gupta, Y. Gusev, F. Habets, A. Hall, L. E. Hay, T. Hogue, M. Huang, G. Leavesley, X. Liang, O. Nasonova, J. Noilhan, L. Oudin, S. Sorooshian, T. Wagener, and E. F. Wood (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *Journal of Hydrology*, *320*(1-2), 3–17, doi:10.1016/j.jhydrol.2005.07.031.

Ehret, U., H. V. Gupta, M. Sivapalan, S. V. Weijs, S. J. Schymanski, G. Blöschl, A. N. Gelfan, C. Harman, A. Kleidon, T. A. Bogaard, D. Wang, T. Wagener, U. Scherer, E. Zehe, M. F. Bierkens, G. Di Baldassarre, J. Parajka, L. P. H. van Beek, A. Van Griensven, M. C. Westhoff, and H. C. Winsemius (2014), Advancing catchment hydrology to deal with predictions under change, *Hydrology and Earth System Sciences*, *18*(2), 649–671, doi:10.5194/hess-18-649-2014.

Falkenmark, M., and T. Chapman (1989), *Comparative hydrology: an ecological approach to land and water resources*, The Unesco Press.

Ficklin, D. L., S. M. Robeson, and J. H. Knouft (2016), Impacts of recent climate change on trends in baseflow and stormflow in United States watersheds, *Geophysical Research Letters*, *43*(10), 5079–5088, doi:10.1002/2016GL069121.

Hall, F. R. (1968), Base-flow recessions – a review, *Water Resources Research*, *4*(5), 973–983, doi:10.1029/WR004i005p00973.

Harman, C., and P. A. Troch (2014), What makes Darwinian hydrology "Darwinian"? Asking a different kind of question about landscapes, *Hydrology and Earth System Sciences*, *18*(2), 417–433, doi:10.5194/hess-18-417-2014.

Harman, C., P. A. Troch, and M. Sivapalan (2011), Functional model of water balance variability at the catchment scale: 2. Elasticity of fast and slow runoff components to precipitation change in the continental United States, *Water Resources Research*, *47*(2), 1–12, doi:10.1029/2010WR009656.

Harrigan, S., J. Hannaford, K. Muchan, and T. J. Marsh (2017), Designation and trend analysis of the updated UK Benchmark Network of river flow stations: the UKBN2 dataset, *Hydrology Research*, *49*(2), 552–567, doi:10.2166/nh.2017.058.

Horton, R. E. (1933), The role of infiltration in the hydrological cycle, *Eos, Transactions American Geophysical Union*, pp. 446–460.

Hrachowitz, M., H. H. G. Savenije, G. Blöschl, J. J. McDonnell, M. Sivapalan, J. Pomeroy, B. Arheimer, T. Blume, M. P. Clark, U. Ehret, F. Fenicia, J. E. Freer, A. Gelfan, H. V. Gupta, D. Hughes, R. Hut, A. Montanari, S. Pande, D. Tetzlaff, P. A. Troch, S. Uhlenbrook, T. Wagener, H. C. Winsemius, R. A. Woods, E. Zehe, and C. Cudennec (2013), A decade of Predictions in Ungauged Basins (PUB) – a review, *Hydrological Sciences Journal*, *58*(6), 1198–1255, doi:10.1080/02626667.2013.803183.

Institute of Hydrology (1980), *Low Flow Studies Report No. 1: Research Report*, Institute of Hydrology.

Lacey, G. C., and R. B. Grayson (1998), Relating baseflow to catchment properties in south-eastern Australia, *Journal of Hydrology*, *204*(1-4), 231–250, doi:10.1016/S0022-1694(97)00124-8.

Longobardi, A., and P. Villani (2008), Baseflow index regionalization analysis in a mediterranean area and data scarcity context: Role of the catchment permeability index, *Journal of Hydrology*, *355*(1-4), 63–75, doi:10.1016/j.jhydrol.2008.03.011.

L'vovich, M. I. (1979), *World water resources and their future*, American Geophysical Union, doi:10.1029/SP013.

Lyne, V. D., and M. Hollick (1979), Stochastic time-variable rainfall runoff modelling, in *Hydrology and Water Resources Symposium*, pp. 82–92.

McDonnell, J. J., M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, R. P. Hooper, J. W. Kirchner, M. L. Roderick, J. S. Selker, and M. Weiler (2007), Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, *Water Resources Research*, *43*(7), 1–6, doi:10.1029/2006WR005467.

Milly, P. C. D. (1994), Climate, soil water storage, and the average annual water balance, *Water Resources Research*, *30*(7), 2143–2156, doi:10.1029/94WR00586.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Stationarity is dead: Whither water management?, *Science*, *319*(5863), 573–574, doi:10.1126/science.1151915.

National River Flow Archive (2018), `https://nrfa.ceh.ac.uk`, NERC CEH, Wallingford.

Neff, B., S. Day, A. Piggott, and L. Fuller (2005), Baseflow in the Great Lakes Basin. Scientific Investigations Report 2005–5217, *US. Geological Survey*, *23*.

Newman, A., K. Sampson, M. Clark, A. Bock, R. Viger, and D. Blodgett (2014), A large-sample watershed-scale hydrometeorological dataset for the contiguous USA, Boulder, CO, UCAR/NCAR, doi:10.5065/D6MW2F4D.

Newman, A. J., M. P. Clark, K. Sampson, A. W. Wood, L. E. Hay, A. Bock, R. J. Viger, D. Blodgett, L. D. Brekke, J. R. Arnold, T. Hopson, and Q. Duan (2015), 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, *Hydrology and Earth System Sciences*, *19*(1), 209–223, doi: 10.5194/hess-19-209-2015.

NRCS (2004), National Engineering Handbook: Part 630 - Hydrology, *USDA Soil Conservation Service: Washington, DC, USA*.

Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks, and J. C. Stromberg (1997), The natural flow regime, *BioScience*, *47*(11), 769–784, doi:10.2307/1313099.

Ponce, V., and A. Shetty (1995a), A conceptual model of catchment water balance: 1. Formulation and calibration, *Journal of Hydrology*, *173*(1-4), 27–40, doi:10.1016/0022-1694(95)02739-C.

Ponce, V., and A. Shetty (1995b), A conceptual model of catchment water balance: 2. Application to runoff and baseflow modeling, *Journal of Hydrology*, *173*(1-4), 41–50, doi:10.1016/0022-1694(95)02745-B.

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, doi:10.1177/0309133311402714.

Robinson, E., E. Blyth, D. Clark, E. Comyn-Platt, J. Finch, and A. Rudd (2016), Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (1961-2015) [CHESS-PE], doi:10.5285/8baf805d-39ce-4dac-b224-c926ada353b7.

Roderick, M. L., and G. D. Farquhar (2011), A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties, *Water Resources Research*, *47*(12), 1–11, doi:10.1029/2010WR009826.

Rouholahnejad Freund, E., and J. W. Kirchner (2017), A Budyko framework for estimating how spatial heterogeneity and lateral moisture redistribution affect average evapotranspiration rates as seen from the atmosphere, *Hydrology and Earth System Sciences*, *21*(1), 217–233, doi:10.5194/hess-21-217-2017.

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

Schneider, M. K., F. Brunner, J. M. Hollis, and C. Stamm (2007), Towards a hydrological classification of European soils: preliminary test of its predictive power for the baseflow index using river discharge data, *Hydrology and Earth System Sciences*, *11*, 1501–1513, doi:10.5194/hess-11-1501-2007.

Schwartz, F. W., G. Liu, P. Aggarwal, and C. M. Schwartz (2017), Naïve simplicity: The overlooked piece of the complexity-simplicity paradigm, *Groundwater*, *55*(5), 703–711, doi:10.1111/gwat.12570.

Shafii, M., N. Basu, J. R. Craig, S. L. Schiff, and P. Van Cappellen (2017), A diagnostic approach to constraining flow partitioning in hydrologic models using a multiobjective optimization framework, *Water Resources Research*, *53*(4), 3279–3301, doi:10.1002/

2016WR019736.

Sivapalan, M. (2005), Pattern, process and function: Elements of a unified theory of hydrology at the catchment scale, in *Encyclopedia of Hydrological Sciences*, John Wiley & Sons, Ltd, Chichester, UK, doi:10.1002/0470848944.hsa012.

Sivapalan, M., M. A. Yaeger, C. Harman, X. Xu, and P. A. Troch (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, doi: 10.1029/2010WR009568.

Smakhtin, V. U. (2001), Low flow hydrology: A review, *Journal of Hydrology*, *240*(3-4), 147–186, doi:10.1016/S0022-1694(00)00340-1.

Tanguy, M., H. Dixon, I. Prosdocimi, D. G. Morris, and V. D. J. Keller (2016), Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2015) [CEH-GEAR], doi:10.5285/33604ea0-c238-4488-813d-0ad9ab7c51ca.

Trancoso, R., S. Phinn, T. R. McVicar, J. R. Larsen, and C. A. McAlpine (2017), Regional variation in streamflow drivers across a continental climatic gradient, *Ecohydrology*, *10*(3), e1816, doi:10.1002/eco.1816.

Troch, P. A., G. F. Martinez, V. R. N. Pauwels, M. Durcik, M. Sivapalan, C. Harman, P. D. Brooks, H. Gupta, and T. Huxman (2009), Climate and vegetation water use efficiency at catchment scales, *Hydrological Processes*, *23*(16), 2409–2414, doi: 10.1002/hyp.7358.

Van Dijk, A. I. (2010), Climate and terrain factors explaining streamflow response and recession in Australian catchments, *Hydrology and Earth System Sciences*, *14*(1), 159–169, doi:10.5194/hess-14-159-2010.

Wagener, T., M. Sivapalan, P. A. Troch, and R. A. Woods (2007), Catchment classification and hydrologic similarity, *Geography Compass*, *1*(4), 901–931, doi:10.1111/j.1749-8198.2007.00039.x.

Wagener, T., M. Sivapalan, P. A. Troch, B. L. McGlynn, C. Harman, H. V. Gupta, P. Kumar, P. S. C. Rao, N. Basu, and J. S. Wilson (2010), The future of hydrology: An evolving science for a changing world, *Water Resources Research*, *46*(5), 1–10, doi: 10.1029/2009WR008906.

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

Wang, D., and Y. Tang (2014), A one-parameter Budyko model for water balance captures emergent behavior in Darwinian hydrologic models, *Geophysical Research Letters*, *41*(13), 4569–4577, doi:10.1002/2014GL060509.

Wang, D., and L. Wu (2013), Similarity of climate control on base flow and perennial stream density in the Budyko framework, *Hydrology and Earth System Sciences*, *17*(1), 315–324, doi:10.5194/hess-17-315-2013.

Wang, D., J. Zhao, Y. Tang, and M. Sivapalan (2015), A thermodynamic interpretation of Budyko and L'vovich formulations of annual water balance: Proportionality hypothesis and maximum entropy production, *Water Resources Research*, *51*(4), 3007–3016, doi: 10.1002/2014WR016857.

Zhao, J., D. Wang, H. Yang, and M. Sivapalan (2016), Unifying catchment water balance models for different time scales through the maximum entropy production principle, *Water Resources Research*, *52*(9), 7503–7512, doi:10.1002/2016WR018977.