[Water Resources Research]

Supporting Information for

**[Global runoff partitioning based on Budyko-constrained machine learning]**[Shujie Cheng[1,2], Petra Hulsman[2], Akash Koppa[2,] Hylke E. Beck[3], Lei Cheng[1], and Diego G. Miralles[2]]

[[1]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

[2]Hydro-Climate Extremes Lab (H-CEL), Ghent University, Ghent, Belgium

[3]King Abdullah University of Science and Technology, Thuwal, Saudi Arabia]

# 1 Method

## Baseflow curve based on limit concept

In general, the water balance can be written as:

$$\frac{dS}{dt} = P - E_a - Q \tag{S1}$$

where $S$ is water stored in underground, $P$ is precipitation, $E_a$ is actual evaporation, $Q$ is discharge which can be partitioned into $Q_b$ is baseflow and $Q_q$ is quick flow ($Q = Q_b + Q_q$).

The basic limit concept of the Budyko framework for estimating $E_a$ is: $E_a/P \to 1\ as\ E_p/P \to \infty$ for very dry conditions, and $E_a \to E_p\ as\ E_p/P \to 0$ for very wet conditions, where $E_p$ is potential evaporation. The demand limit of $E_a$ is $E_p$ and the supply limit is $P$. Fu (1981) proposed $E_a$ can be calculated with:

$$\frac{E_a}{P} = 1 + \frac{E_p}{P} - \left[1 + (\frac{E_p}{P})^{a_1}\right]^{1/a_1} \tag{S2}$$

Assuming $\frac{dS}{dt} \approx 0$ on long term time scales and with the catchment retention defined as

$$CR = E_a + Q_b \tag{S3}$$

Equation S1 can be expressed as:

$$P = CR + Q_q \tag{S4}$$

The demand limit for $CR$ is $CR_0 = E_p + Q_{b,p}$ . The $E_p$ and $Q_{b,p}$ are the potential values for $E$ and $Q_b$, respectively. According to Zhang et al. (2008), the limits concept of Budyko can also be applied to $CR$ such that: $CR/P \to 1\ as\ CR_0/P \to \infty$ for very dry conditions, and$CR \to CR_0\ as\ CR_0/P \to 0$ for very wet conditions. Then $CR$ can be estimated as:

$$\frac{CR}{P} = 1 + \frac{CR0}{P} - \left[1 + (\frac{CR0}{P})^{a_2}\right]^{1/a_2} \tag{S5}$$

Combining Eq. S2, Eq. S3 and Eq. S5:

$$\frac{Q_b}{P} = \frac{Q_{b,p}}{P} + [1 + (\frac{E_p}{P})^{a_1}]^{1/a_1} - [1 + (\frac{E_p + Q_{b,p}}{P})^{a_2}]^{1/a_2} \tag{S6}$$

Under very limited storage capacity conditions (for instance an impervious catchment), no/limited water is stored in the subsurface such that the baseflow also approaches zero (i.e., $Q_b/P \to 0$ if $Q_{b,p}/P \to 0$). Under that condition, Eq. S11 changes to $0 \approx [1 + (\frac{E_p}{P})^{a_1}]^{1/a_1} - \left[1 + \left(\frac{E_p}{P}\right)^{a_2}\right]^{1/a_2}$. This equation can only be satisfied if $a_1 = a_2$. Thus Eq. S11 can be written as:

$$\frac{Q_b}{P} = \frac{Q_{b,p}}{P} + [1 + (\frac{E_p}{P})^{\alpha}]^{1/\alpha} - [1 + (\frac{E_p + Q_{b,p}}{P})^{\alpha}]^{1/\alpha} \tag{S7}$$

## 2 Figures

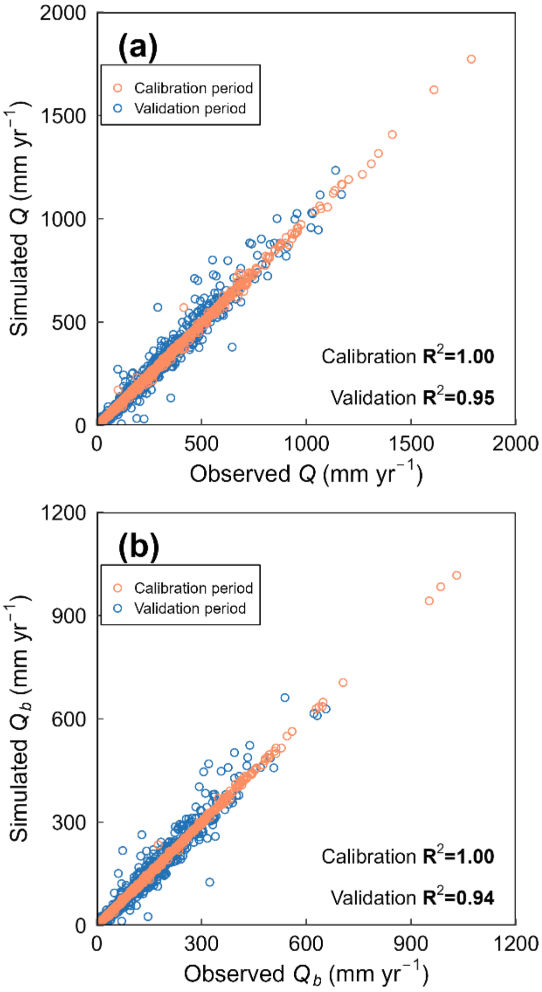


Figure S1. Performance of (a) $Q$ and (b) $Q_b$ at catchment scale during the calibration (orange) and validation (blue) periods.

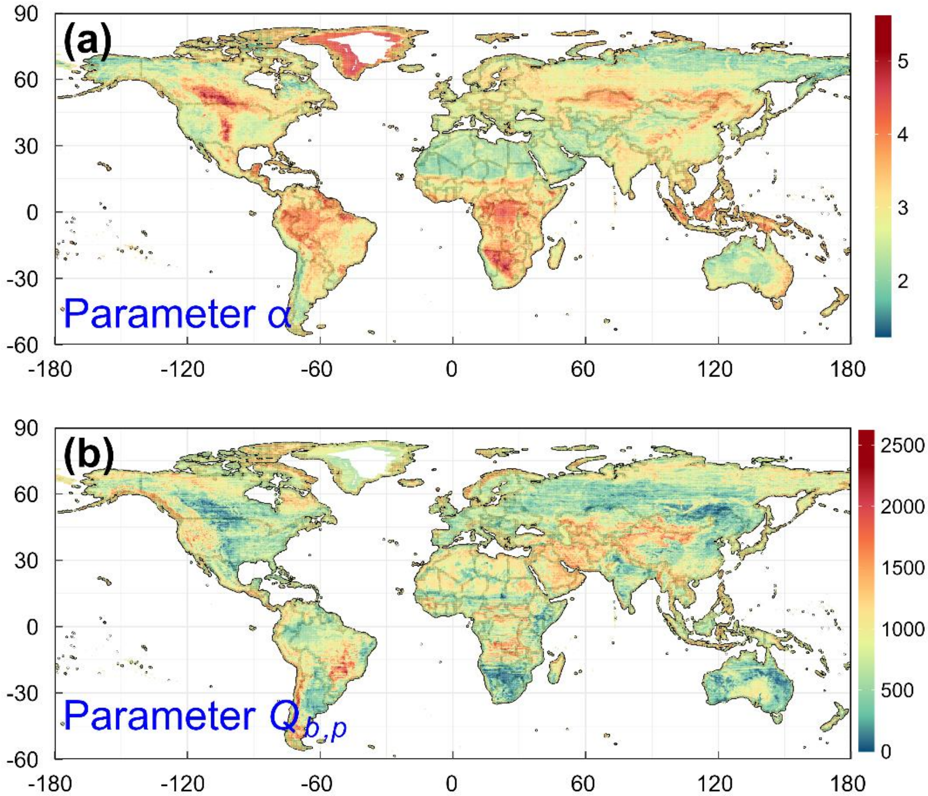


Figure S2. Global maps of (a) parameter $\alpha$ in the Budyko curve (Eq. 3) and BFC curve (Eq. 4), and (b) parameter $Q_{b,p}$ in BFC curve estimated as the mean of 10 BRT models.

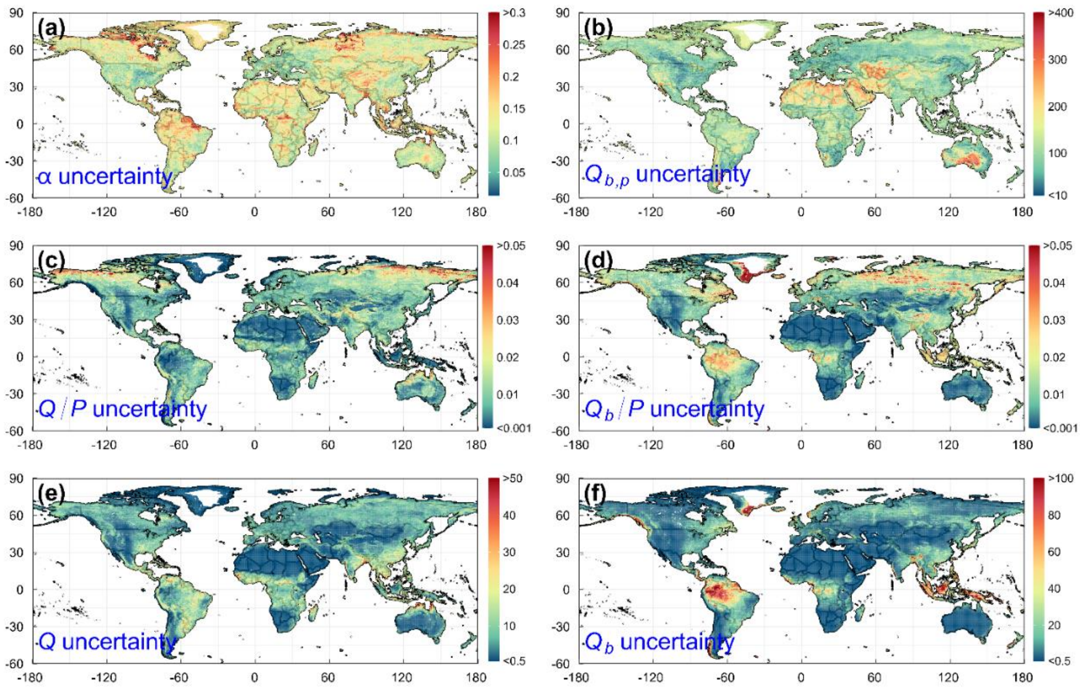


Figure S3. Global map of the uncertainty of (a) parameter α, (b) parameter $Q_{b,p}$, (c) runoff coefficient (RC=$Q/P$), (d) baseflow coefficient (BFC=$Q_b/P$), (e) runoff ($Q$), and (f) baseflow ($Q_b$). These uncertainty values are equal to the standard deviation of the 10 trained BRT models using the 10-fold cross-validation strategy.

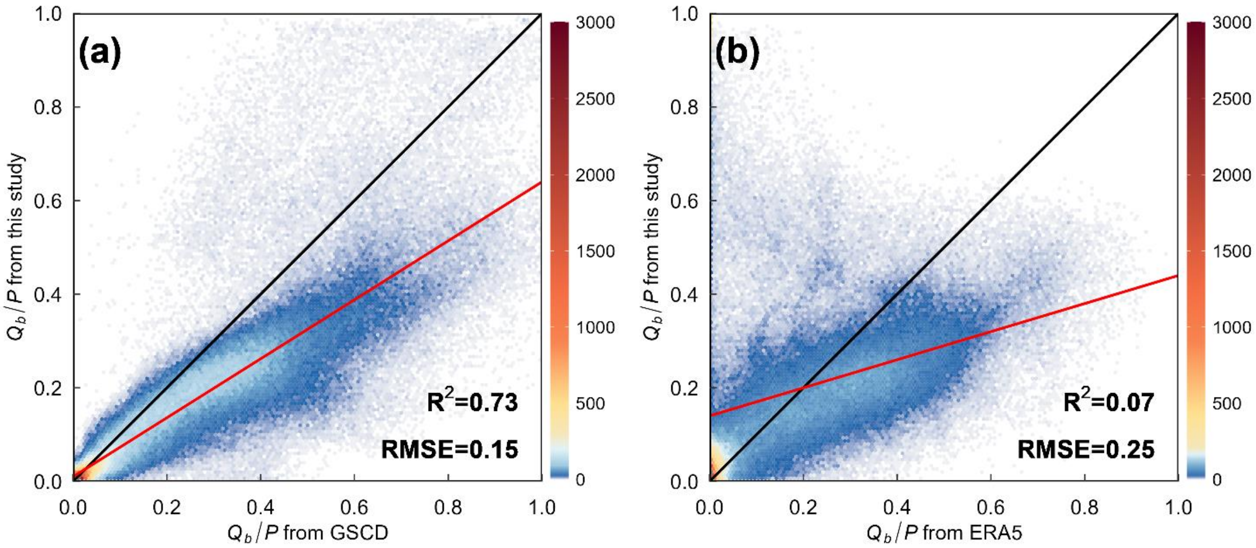


Figure S4. Comparison of the baseflow coefficient ($Q_b/P$) from this study with estimates according to (a) GSCD and (b) ERA5-Land.

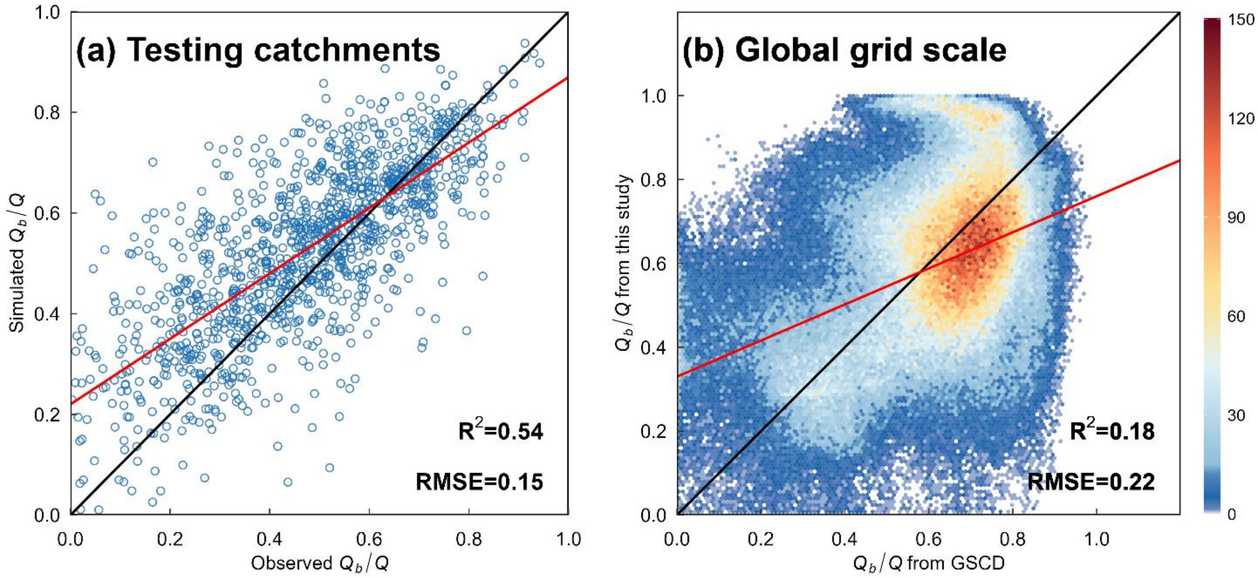


Figure S5. Comparison of the baseflow index ($Q_b/Q$) as estimated in this study with (a) field observations and (b) GSCD estimates.