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

Correspondence to: Lei Cheng, lei.cheng@whu.edu.cn
Diego Miralles, Diego.Miralles@UGent.be

Highlights:

1. A Budyko-constrained machine learning approach is developed for estimating long-term mean runoff and baseflow.
2. The hybrid approach performs well in terms of both the runoff coefficient (R^2=0.93) and the baseflow coefficient (R^2=0.84).
3. Globally, 30% of the precipitation is partitioned into runoff, with baseflow contribution estimated to be twice the quickflow (20.6% vs. 9.7%).
4. Primary drivers of runoff partitioning vary in space with topography and soil properties as dominant factors.

Abstract: Understanding the partitioning of runoff into baseflow and quickflow is crucial for informed decision-making in water resources management, guiding the implementation of flood mitigation strategies, and supporting the development of drought resilience measures. Methods that combine the physically-based Budyko framework with machine learning (ML) have shown promise in estimating global runoff. However, such 'hybrid' approaches have not been used for baseflow estimation. Here, we develop a Budyko-constrained ML approach for baseflow estimation by incorporating the Budyko-based baseflow coefficient (BFC) curve as a physical constraint. We estimate the parameters of the original Budyko curve and the newly developed BFC curve based on 13 climatic and physiographic characteristics using boosted regression trees (BRT). BRT models are trained and tested in 1226 catchments worldwide and subsequently applied to the entire global land surface at a 0.25° grid scale. The catchment-trained models exhibit strong performance during the testing phase, with R^2 values of 0.96 and 0.88 for runoff and baseflow, respectively. Results reveal that, on average, 30.3% (spatial standard deviation std=26.5%) of the continental precipitation is partitioned into runoff, of which 20.6% (std=22.1%) is baseflow and 9.7% (std=10.3%) is quickflow. Among the 13 climatic and physiographic characteristics, topography and soil-related characteristics generally emerge as the most important drivers, although significant regional variability is observed. Comparisons with previous datasets suggest that global runoff partitioning is still highly uncertain and warrants further research.

Keywords: runoff partitioning, baseflow, quickflow, Budyko, machine learning

1 Introduction

Accurate partitioning of runoff (Q) into its main components – baseflow (Q_b) and quickflow (Q_q) – is crucial for water management and emergency planning during droughts (Apurv & Cai, 2020) and floods (Roxy et al., 2017). Baseflow, sometimes referred to as 'slow flow', provides most of the water for sustaining river flows during dry periods (Miller et al., 2016). It originates from groundwater and other delayed sources, such as wetlands, lakes, melting of snow and ice (Hall, 1968). Quickflow is directly responsible for flood generation (Yin et al., 2018), and is a result of fast processes such as saturation or infiltration of excess overland flow and fast subsurface flow, i.e., processes where precipitation is not retained in the soil (Beven & Kirkby, 1979). Process-based models play an important role in accurately estimating global runoff, quickflow and baseflow; this includes among others land surface models (LSMs) and global hydrological models (GHMs). Nonetheless, LSMs and GHMs struggle with runoff partitioning resulting in poor performances in terms of baseflow index (BFI= Q_b/Q) – see e.g. Beck et al. (2017).

To complement process-oriented models, data-driven machine learning (ML) techniques have been developed to assess runoff partitioning regionally and globally without the biases induced by process-based models. For instance, Huang et al. (2021) adopted a random forest model (RF) and multiple linear regression approach to estimate the baseflow index (BFI = Q_b/Q) in the United States. Beck et al. (2013, 2015) achieved satisfactory performance for BFI estimation globally (R^2=0.65) and provided global BFI datasets by using neural networks to relate BFI to climatic/physiographic characteristics. ML has the potential to build effective relationships between inputs and outputs, even if underlying physical processes are unknown. That is why ML has been growing in popularity in hydrological sciences beyond runoff (Xie et al., 2021), being used to predict evaporation (Jung et al., 2010) and precipitation (Sadeghi et al., 2019; Beck et al., 2010) as well. Despite the strength of pure ML models, the major limitation is their “black box” nature, and hence their lack of physical constraints and limited interpretability. The combination of physically-based models and ML methods, i.e., 'hybrid' approaches, can retain both of their individual strengths (de Bézenac et al., 2019; Koppa et al., 2022; Kraft et al., 2022; Zhao et al., 2019). Hence, these physically-constrained ML methods can potentially improve the realism of the runoff partitioning estimates globally.

Previous studies have illustrated the advantage of the Budyko (1961) framework as a physical constraint for pure ML to estimate runoff (Bai et al., 2020; Liu & You, 2021), while no such attempt has been made for baseflow estimation yet. Recently, the Budyko framework was expanded by Cheng et al. (2021) to partition baseflow from precipitation with the Budyko-based baseflow coefficient (BFC) curve. This enables the Budyko framework to provide consistent physical constraints for both runoff and baseflow estimation. Both the Budyko and BFC curves depend on the aridity index and use lumped parameters (parameter α in Fu's equation and $Q_{b,p}$ in the BFC curve, see Section 2.1) to incorporate climatic and physiographic properties such as vegetation, soil, topography, and human activities (Mianabadi et al., 2020; Potter et al., 2005; Tang & Wang, 2017; Zhang et al., 2001). Zhang et al. (2001) revealed the impact of vegetation change on long-term evaporation and suggested α equal to 0.5 and 2.0 for herbaceous plants and trees, respectively. Besides vegetation, Liu et al. (2018) indicated that climate seasonality also plays an important role on the Budyko parameter α. More complex relationships have also been proposed for small catchments (Bai et al., 2020). These different relationships between α and catchment properties indicate that a detailed understanding of the Budyko parameter is yet to be achieved (Padrón et al., 2017). ML models have the strength to achieve better regionalization of the parameters within the Budyko and BFC curves.

In this study, we design a framework for the long-term partitioning of global runoff by adopting the Budyko and Budyko-based BFC curves as physical constraints for ML models. The hybrid Budyko–ML approach makes full use of the available data, and enables physical consistency by obeying the Budyko limits of water and energy conservation. The primary objectives are to (1) develop Budyko constrained ML models to estimate individual runoff components globally, (2) assess the accuracy of the different estimated components, (3) analyse their spatial patterns, and (4) identify and quantify the primary drivers of runoff partitioning. The derived dataset includes gridded total runoff (Q), baseflow (Q_b), quickflow (Q_q), runoff coefficient (RFC=Q/P), baseflow coefficient (BFC=Q_b/P), and quickflow coefficient (QFC=RFC–BFC) at 0.25° resolution. The structure of the paper is as follows: Section 2 describes the data development process, Section 3 describes the input datasets, and Sections 4 and 5 present the results and discussions, respectively.

2 Methods

2.1 Budyko curve for runoff estimation

The Budyko framework is a first order approach that partitions long-term mean precipitation into runoff and actual evaporation (Budyko, 1961). According to this framework, both fluxes are limited by the water supply (typically precipitation, P) and the energy demand on evaporation (typically potential evaporation, E_p). This framework assumes that long-term soil water storage changes are negligible. Hence, the water balance can be written as:

$$P = E_a + Q \tag{1}$$

where P is precipitation, Q is runoff and E_a is actual evaporation.

As the original Budyko (1961) equation does not consider climatic and physiographic properties, several studies proposed alternative equations that introduce a single parameter to incorporate these properties (Choudhury, 1999; Yang et al., 2007; Zhang et al., 2001). The formulation proposed by Fu (1981) and Zhang et al. (2004) is adopted in this study:

$$\frac{E_a}{P} = 1 + \frac{E_p}{P} - \left[1 + (\frac{E_p}{P})^{\alpha}\right]^{\frac{1}{\alpha}} \tag{2}$$

By combining Eq.1 and Eq.2, the equation for Q estimation can be written as:

$$\frac{Q}{P} = 1 - \frac{E_a}{P} = -\frac{E_p}{P} + \left[1 + (\frac{E_p}{P})^{\alpha}\right]^{\frac{1}{\alpha}} \tag{3}$$

where α is a parameter reflecting the secondary controls such as climate variability, vegetation, soil and topography, and can range from 1 to ∞ (Zhang et al., 2004). High α values result in low runoff and high actual evaporation for specific precipitation and potential evaporation values (Figure 1a). Figures in this study visualise P/E_p instead of E_p/P to put more focus on humid regions with larger variability of Q/P.

2.2 BFC curve for baseflow estimation

The BFC curve was developed by Cheng et al. (2021) to estimate long-term mean baseflow (Q_b) based on the Budyko framework with suitable modifications (see supplementary material). The BFC equation (Eq. 4) indicates that the baseflow not only depends on P and potential evaporation E_p, but also on the potential baseflow

($Q_{b,p}$). This latter parameter is newly introduced in this study and indicates the amount of baseflow that would occur if sufficient water were available. Hence, it is an upper limit for the baseflow, analogous to the concept of E_p for the case of evaporation. See the supplementary material for the derivation of the equation, which is slightly modified compared to Cheng et al. (2021). The final equation of the BFC curve is as follows:

$$\frac{Q_b}{P} = \frac{Q_{b,p}}{P} + [1 + (\frac{E_p}{P})^{\alpha}]^{\frac{1}{\alpha}} - [1 + (\frac{E_p}{P} + \frac{Q_{b,p}}{P})^{\alpha}]^{\frac{1}{\alpha}} \quad (4)$$

where α is a parameter (identical to the one in Eq. 3). Q_b/P increases with increasing P/E_p and $Q_{b,p}/P$. High $Q_{b,p}$ values result in high baseflow and low quickflow for specific precipitation and potential evaporation values (Figure 1b).

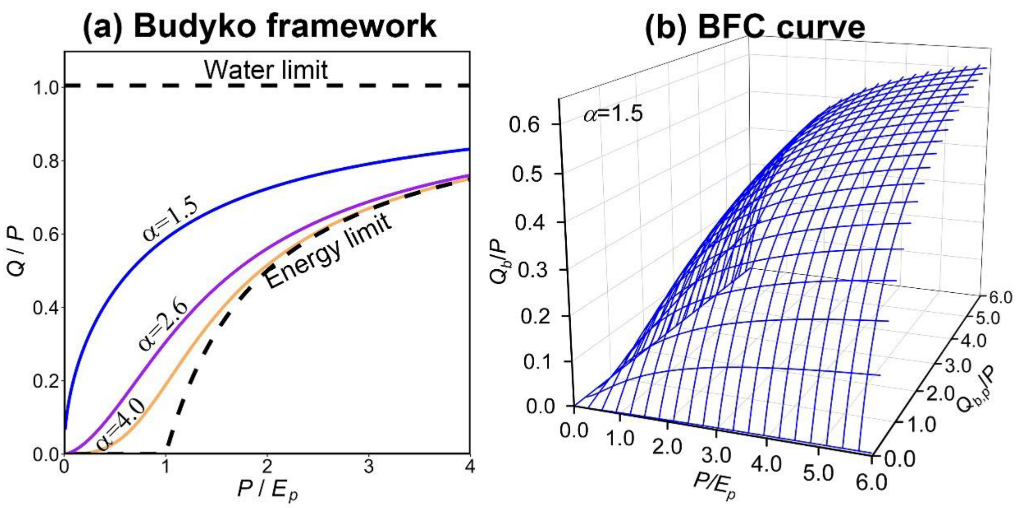


Figure 1. Visualization of the physical constraints for (a) runoff, Budyko curve (Eq. 3), and (b) baseflow, BFC curve (Eq. 4).

2.3 Calibration of parameters

The Budyko and BFC curves include the following two parameters: α and $Q_{b,p}$. The latter is also parameterized, since its value cannot be determined with any available dataset. For individual catchments, the parameter α is calibrated first by using the Budyko curve (Eq. 3) and observed long-term mean Q, P, and E_p. Then, $Q_{b,p}$ is calibrated using the BFC curve (Eq. 4) and observed long-term mean Q_b, P, E_p, and α (as calibrated in the previous step).

2.4 Machine learning to relate parameters to climatic and physiographic properties

The parameters (α and $Q_{b,p}$) are regionalized as functions of climatic and physiographic properties using ML. The calibrated α and $Q_{b,p}$ in each catchment (see Section 2.3) serve as a benchmark for training ML models. The catchment-trained ML models are then used to regionalize α and $Q_{b,p}$ globally at grid scale.

This study uses Boosted Regression Trees (BRT), which combines the strengths of a regression tree algorithm and a boosting algorithm (Elith et al., 2008). BRT differs fundamentally from conventional techniques that aim to produce a single "best" parsimonious model, as it constructs multiple regression models in the algorithm (Elith et al., 2008). The process of training a BRT model includes two parts: regression trees and a boosting algorithm. First, multiple regression trees are built by minimizing the prediction errors. Second, the boosting algorithm combines the regression trees to give improved predictive performance. An effective strategy for fitting a single decision tree is to grow a large tree, and then to prune it by collapsing the weakest links as identified through cross-validation (Franklin, 2008). The first regression tree is grown using recursive binary splits, that is, a binary split is repeatedly applied to its own output until the loss function is maximally reduced. The second tree is fitted to the residuals of the first tree, and the second tree can contain quite different variables and split points. Consequently, multiple trees are fitted additively based on the residuals of the previous tree. For multiple fitted trees, the boosting algorithm averages trees to increase model performance. The dominant drivers for the parameters α and $Q_{b,p}$ are estimated through method local interpretable model-agnostic explanations (LIME) (Ribeiro et al., 2016).

Several hyper-parameters in BRT can be adjusted, including tree complexity (tc), learning rate (lr) and bag fraction (bf). To find the most robust model for our analysis, combinations of the following parameter values are tested using a 10-fold cross-validation strategy: $tc \in \{4, 7, 10, 12\}$, $lr \in \{0.0005, 0.005, 0.01\}$ and $bf \in \{0.4, 0.5, 0.6, 0.8\}$ (Elith et al., 2008). The combination of hyper-parameter values with the highest test performance is tc=12, lr=0.01, and bf=0.50. 10-fold cross-validation strategy is also used for training models. The training is conducted ten times. Each time, 10 groups of catchments are randomly formed, of which nine groups are used for training and one for testing. Ten BRT models are finally constructed at catchment scale and then applied at grid scale globally. 10 maps of α and $Q_{b,p}$ are produced, and hence 10 maps of Q and Q_b. The mean values of the 10 maps are computed as a final result, with the uncertainty calculated as the standard deviation of the 10 BRT models shown in supplementary Figure S3.

2.5 Overview of the modelling process

The methods and input datasets used in this study are summarized in Figure 2. Global runoff partitioning maps are developed with the following steps: First, parameters α and $Q_{b,p}$ in Eq.3 and Eq.4 are calibrated at catchment level (see Section 2.3). This step uses Q, Q_b, P and E_p data at the catchment level. Q_b is extracted from Q using the digital filter method (see Section 3.1). Next, BRT models are developed at the catchment level to relate α and $Q_{b,p}$ to various climatic and physiographic properties. This step applies a 10-fold cross-validation strategy, resulting in 10 BRT models. The BRT models are then used to estimate α and $Q_{b,p}$ globally at grid scale. Finally, Q and Q_b are estimated globally using Eq. 3–4, P and E_p data, and the BRT-derived parameters α and $Q_{b,p}$ as inputs.

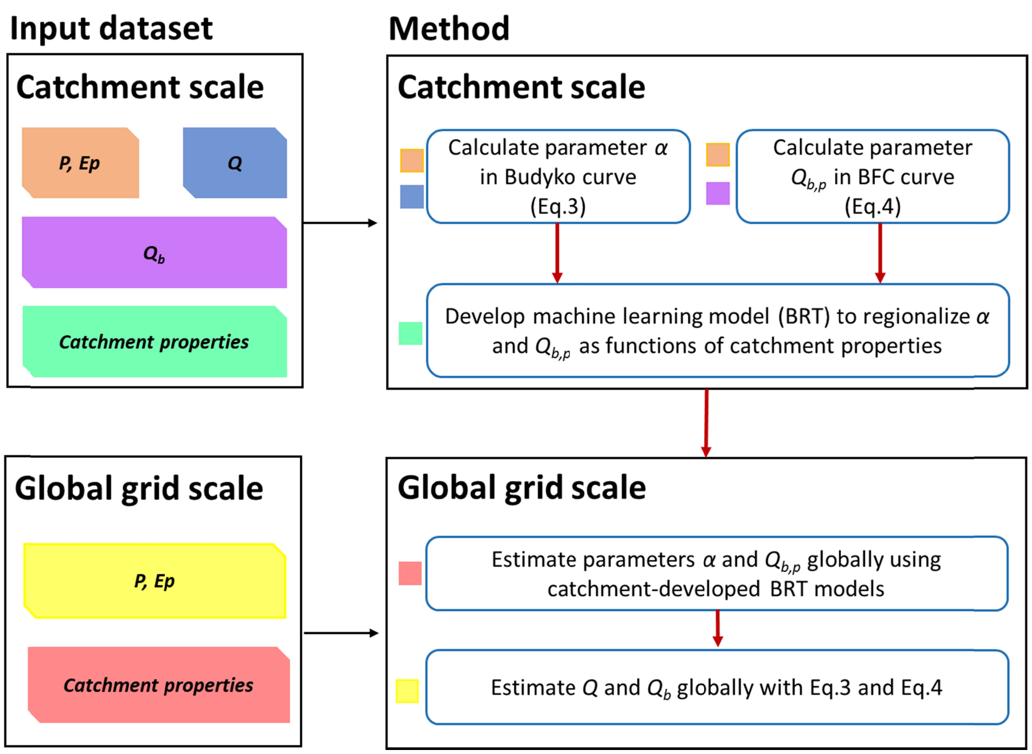


Figure 2. Overview of the modelling process and input datasets. The input datasets required in each step are indicated by the colored squares; they correspond to the input datasets listed on the left.

3 Data

3.1 Observed runoff, baseflow and quickflow

Observed daily discharge data from 3274 gauge-stations are obtained from the Global Runoff Data Centre (GRDC) dataset together with the corresponding catchment boundaries (https://www.bafg.de/GRDC/). A set of 1314 gauge-stations and their corresponding catchments are selected from the initial dataset based on the following requirements. First, the record length needed to be at least 10 years to allow analyses on long-term mean values. Second, the missing data rate should be smaller than 20% to warrant the representativeness of the mean values. Third, the water balance should close, i.e., $|\frac{P-E_a-Q}{P}| < 0.1$, to exclude stations with too large data uncertainties, or regional groundwater export/import as this is not included in the Budyko framework. The spatial distribution of the selected 1314 catchments is shown in Figure 3c.

Daily Q_b and quickflow (Q_q=Q–Q_b) are separated from daily Q using a digital filter technique, more specifically the Lyne–Hollick (LH) method (Lyne & Hollick, 1979). Different digital filter methods have no significant influence on the long-term estimation of Q_b and Q_s (Chen et al., 2023). The LH method has the advantage of being minimally parameterized, and thus is easily applicable to a large sample of catchments. The filter parameter f_1, also called recession constant, affects the degree of attenuation. The number of passes determines the degree of smoothing, with the backward pass nullifying the phase distortion from the forward pass. Here, the LH method is applied in a conventional way with three passes (forward, backward, and forward again) and the filter parameter f_1 is set to 0.925 (Nathan & McMahon, 1990).

Long-term mean Q, Q_b and Q_q are estimated from their daily values and used to estimate the catchment runoff coefficient (RFC=Q/P), baseflow coefficient (BFC=Q_b/P), and quickflow coefficient (QFC=RFC–BFC). Note that the time period for P, E_p, Q, Q_b and Q_q are consistent within each catchment by selecting their crossing period. The available data lengths of the 1314 catchments vary from 10 to 41 years.

3.2 Climatic and Physiographic Characteristics

Table 1 lists 16 climatic and physiographic variables that are used in this study, including the respective references, original spatial resolution and temporal coverage. In this study, analysis is done at 0.25° resolution; hence, observations are resampled to 0.25° using bilinear interpolation method when needed. Among the 16 variables, precipitation (P), potential evaporation (E_p) and evaporation (E_a) are direct inputs in the Budyko and BFC curves (Eq. 3 and 4). The remaining 13 variables are predictors for parameters α and $Q_{b,p}$ during the ML step (see Section 2.4). Three of these characteristics are related to climate, four to vegetation, three to topography, two to soil and one is related to human activities.

Table 1. Gridded climatic and physiographic characteristics used directly in the Budyko and BFC curve (P, E_a and E_p), or to predict runoff and baseflow with ML (remaining variables).

Subcategory	Data	Description	Data Sources	Original Resolution	Temporal coverage
Climate	P	Precipitation	MSWEP v1.1 (Beck et al., 2017)	0.25°	1980–2020
	E_p	Potential Evaporation	TerraClimate (Abatzoglou et al., 2018)	1/24°	
	E_a	Actual evaporation	GLEAM v3.6 (Martens et al., 2017)	0.25°	
	TC	Air temperature	ERA5 (Hersbach et al., 2020)	0.25°	
	SAI	Seasonality and asynchrony index	Calculated from daily P and E_p (Liu et al., 2018)	0.25°	
	SWE	Snow water equivalent	GLOBSNOW L3av2 and NSIDC v0.1 (Armstrong, 2005; Luojus et al., 2013)	0.25°	
Vegetation	NDVI	Normalized difference vegetation index	MODIS (https://modis.gsfc.nasa.gov/)	0.05°	2000–2014
	WUE	Water use efficiency			
	LAI	Leaf area index			
	RD	Maximum rooting depth	Fan et al. (2017)	~1km	Static
Topography	CTI	Topographic index	Marthews et al. (2015)	500m	Static
	ELEV	Mean	Yamazaki et al. (2019)	90m	

		elevation			
	SLO	Slope	Amatulli et al. (2018)	1km	
Soil	STHI	Average soil and sedimentary deposit thickness	Pelletier, J.D, et.al, (2016)	1km	Static
	SPO	Soil porosity	SoilGrids 2.0 (Poggio et al., 2021)	250 m	
Human activities	HFP	Human influence index	Eric W. Sanderson et al. (2002)	1km	1995–2004

4 Results

4.1 Calibrated parameters at catchment scale

Figure 3 visualises the values of the parameters α and $Q_{b,p}$ for all 1314 catchments as calibrated with Eq. 3 and 4. Due to the large variability of catchment conditions, the parameter values vary such that their 5% and 95% quantiles range between α = 1.94–5.55 and $Q_{b,p}$ = 104–2242 mm yr^{-1}, resulting in different Budyko and BFC curves as shown in Figure 3a and b with the purple and orange lines. The median values of the catchment-specific α and $Q_{b,p}$ are 2.83 and 547 mm yr^{-1}, respectively. The large variability of α and $Q_{b,p}$ is spatially shown in Figure 3c and d, respectively. Large α values (i.e., $\alpha > 4$) mainly appear in the Great Plains in North America, the east coast of Australia and South America. A mix of low and median α values (i.e., $\alpha < 3$) appear in Europe and in the east of the United States. For the spatial distribution of $Q_{b,p}$, Australia, southern Africa and the middle of the United States show low $Q_{b,p}$ values. There are no obvious spatial patterns elsewhere. Of the 1314 catchments, 88 stations show extreme parameter values (i.e., α>6.0 and $Q_{b,p}$ >3000; see the tails in the density plots in Figure 3c and d). In addition, 38 stations of these 88 stations fall outside the Budyko limits (i.e., $\alpha \rightarrow \infty$) as shown in Figure 3a. These 88 stations with extreme parameter values are considered as outliers and are therefore excluded, such that 1226 stations are left for the remainder of the analysis. These remaining stations are further tested by analysing results during calibration and validation periods (see Figure S1). For each catchment, the first 20 years are used to calibrate the parameters α and $Q_{b,p}$, and the remaining years are used for validation. Catchments with less than 30 years of data are not validated. The validated runoff and baseflow show a correlation (R^2) of 0.95 and 0.94, respectively. The high validation performance illustrates that the Budyko and BFC curve with the calibrated α and $Q_{b,p}$ can reproduce the spatial variability of Q and Q_b well at the selected catchments.

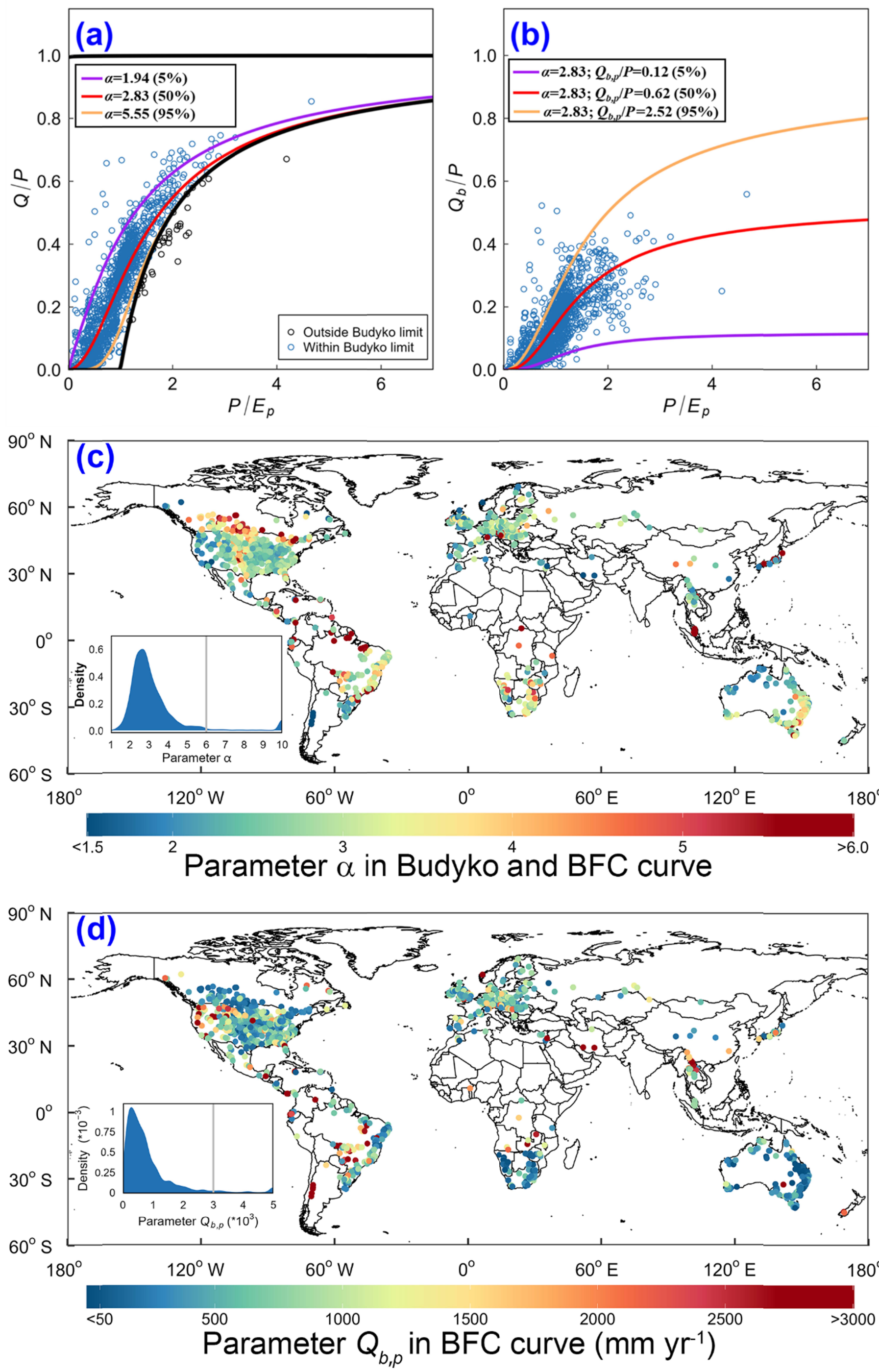


Figure 3. Scatterplots of (a) Q/P and (b) Q_b/P versus P/E_p. The lines in panel (a) and (b) are the Budyko and BFC curves, respectively, using the following quantile values for the parameters α and $Q_{b,p}$: 5% (purple), 50% (red), and 95% (orange). Note, that in (b) α is fixed to its median value of 2.83 to focus on $Q_{b,p}$ changes. Spatial distribution and density plots of (c) α and (d) $Q_{b,p}$. Extreme values beyond the grey lines in the density plots are considered as outliers and are removed.

4.2 Catchment-scale performance of trained models

The first row in Figure 4 shows the performances of BRT-derived α, Q/P and Q using Eq.3 and BRT-derived α (Budyko–ML). During training, all three variables agree well with observations with RMSE for Q/P and Q equal to 0.02 and 17 mm yr^{-1}, respectively (Figure 4b and c). During testing, the performances decrease as expected, especially for α, with R^2 = 0.50. The performances of Q/P and Q remain high though, with R^2 = 0.93 and 0.96 and RMSE = 0.04 and 46 mm yr^{-1}, respectively.

Similar to runoff, the second row in Figure 4 shows the performances of BRT-derived $Q_{b,p}$, Q_b/P and Q_b estimated with Eq. 4 and BRT-derived α and $Q_{b,p}$. The performance for these three variables is high during training, with R^2 = 0.92, 0.96 and 0.97, respectively (Figure 4d, e and f). The performance of $Q_{b,p}$ decreases during the testing phase (R^2=0.41). As a result, also Q_b/P and Q_b perform slightly worse during testing, but their performances are still acceptable, with R^2 for Q_b/P and Q_b equal to 0.84 and 0.88, respectively. The good performances of runoff and baseflow during the testing phase indicates that the trained ML models for α and $Q_{b,p}$ at catchment scale are reliable for estimating runoff and baseflow globally.

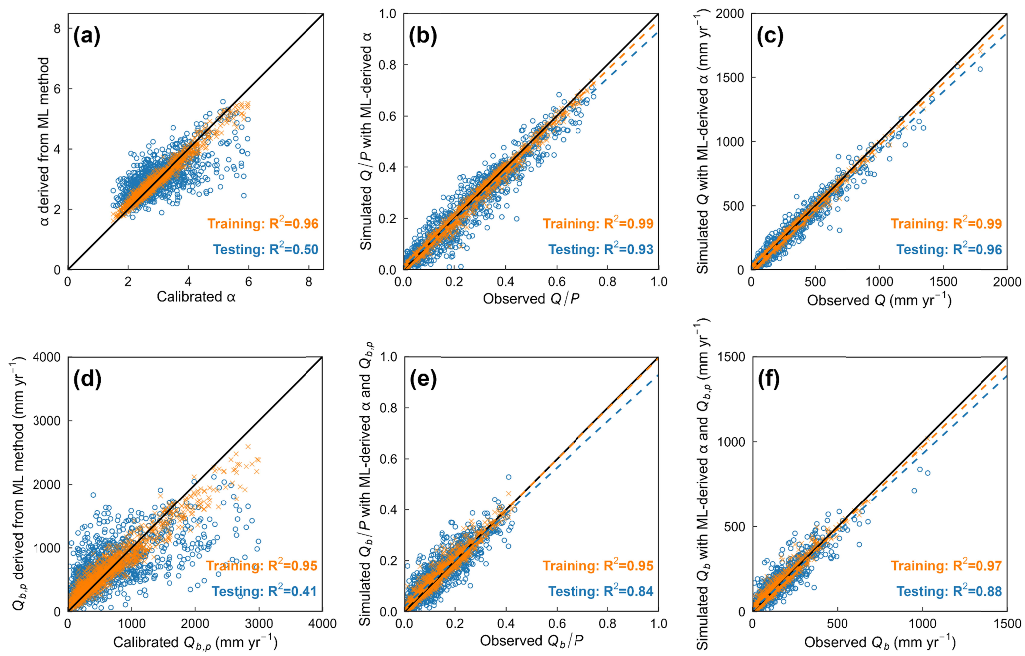


Figure 4. Performance of parameters α and $Q_{b,p}$ (a and d) from BRT models, Q/P (b), Q_b/P (e), Q (c) and Q_b (f) using the Budyko framework and BFC curve at catchment scale. Each point represents a catchment (n=1226 in each panel). The orange and blue lines are linear regression lines and the black lines mark the 1:1 relation.

4.3 Global map of runoff partitioning

The catchment-trained BRT models are applied globally to estimate parameters α and $Q_{b,p}$ on grid-scale. The spatial distribution of α and $Q_{b,p}$ is shown in Figure S2. Figure 5a, c and e illustrate the global, Budyko-based Q, Q_b and Q_q. This is compared to observations in Figure 5b, d, and f, respectively. The estimated global values show very similar spatial patterns to the observations. High flows appear at medium latitudes near the equator (30°N–30°S), while low flows are located at high latitudes (30°N–90°N and 30°S–60°S). Northern Africa and western Asia are exceptions, as they show low values at medium latitudes. Similarly, exceptions at high latitudes appear in regions near the coast of Chile, Europe and Canada where high runoff is found in the estimated maps (Figure 5a, c and e). The uncertainty for each variable (α, $Q_{b,p}$, Q/P, Q_b/P, Q, and Q_b), which is considered equal to the standard deviation in the 10 BRT models, is shown in Figure S3. The spatial uncertainty is equal to 0.13 for α, 114 mm yr^{-1} for $Q_{b,p}$, 0.009 for Q/P, 0.01 for Q_b/P, 5.78 mm yr^{-1} for Q, and 10.9 mm yr^{-1} for Q_b. The uncertainty is larger for Q_b than Q since Q only relies on one parameter (α), while Q_b relies on two parameters (α and $Q_{b,p}$). Overall, the global, long-term mean annual Q is on average 274 (std=418) mm yr^{-1}, Q_b 151 (std=181) mm yr^{-1} and Q_q 123 (std=270) mm yr^{-1}. The results illustrate that the global river supply relies more on baseflow than quickflow.

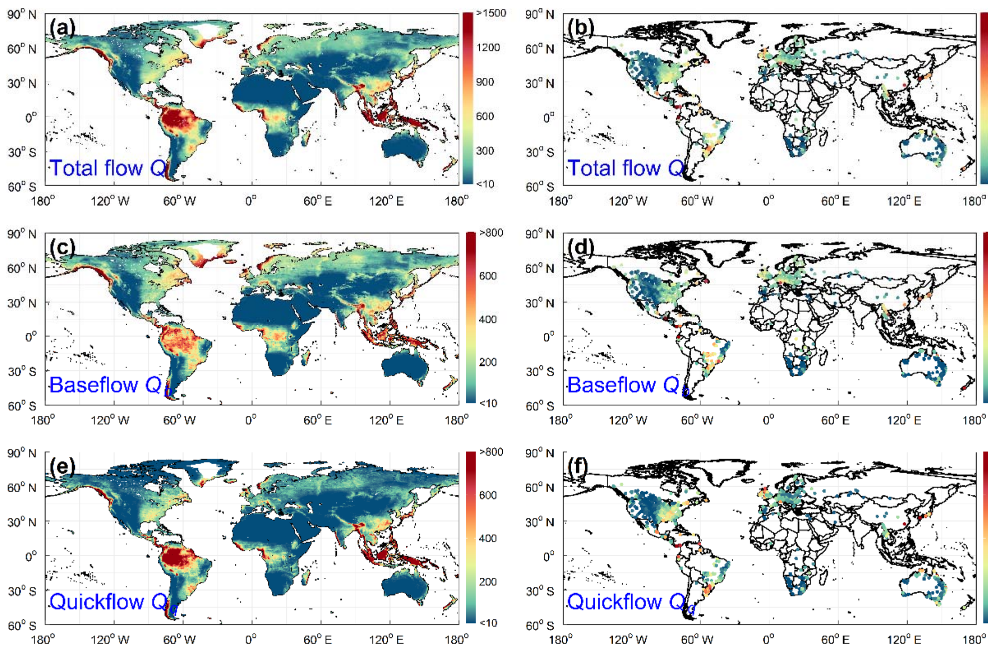


Figure 5. Global maps of estimated (a) Q, (c) Q_b and (e) Q_q. Station-based observed (b) Q, (d) Q_b and (f) Q_q.

Figure 6 shows the global map of gridded RFC, BFC, QFC, baseflow index (BFI=Q_b/Q) and quickflow index (QFI=Q_q/Q). RFC, BFC and QFC have similar spatial patterns with low values in western America, northern Africa, southern Africa, western Asia, central Asia, and Australia. However, regions with high RFC values, show both high and low BFC and QFC values depending on the region. For example, Q is partitioned more into Q_q than Q_b in the Amazon and southeast Asia, while in Canada and Russia, more Q_b is generated than Q_q. This is also illustrated with Figure 6f–h, which show RFC, BFC and QFC for different quantiles as a function of the latitude. High RFC values are generally located at 60°N–85°N, 5°N–5°S and 45°S–60°S. In these latitudinal intervals, the partitioning of Q into Q_b and Q_q differs in space. As shown in Figure 6i and j, between 60°N–85°N, the majority of Q are mostly partitioned into Q_b (the median values in this latitudinal interval is on average 80.7%), and less Q_q (19.3%). Between 5°N–5°S and 45°S–60°S, the BFC and QFC are quite similar to each other with Q_b of 42.1% and 54.9%, respectively, and Q_q of 57.9% and 45.1%, respectively. Across all latitudes, the mean difference between the 5th and 95th quantiles is larger for RFC (0.50) and BFC (0.35) than for QFC (0.26). The spread of QFC is more pronounced near the equator (5°N–5°S) and in the Southern Hemisphere high latitudes (45°S–60°S). Overall, average 30.3% (std=26.5%) of P is partitioned into Q, of which 20.6% (spatial standard deviation std=22.1%) is Q_b and 9.7% (std=10.3%) Q_q.

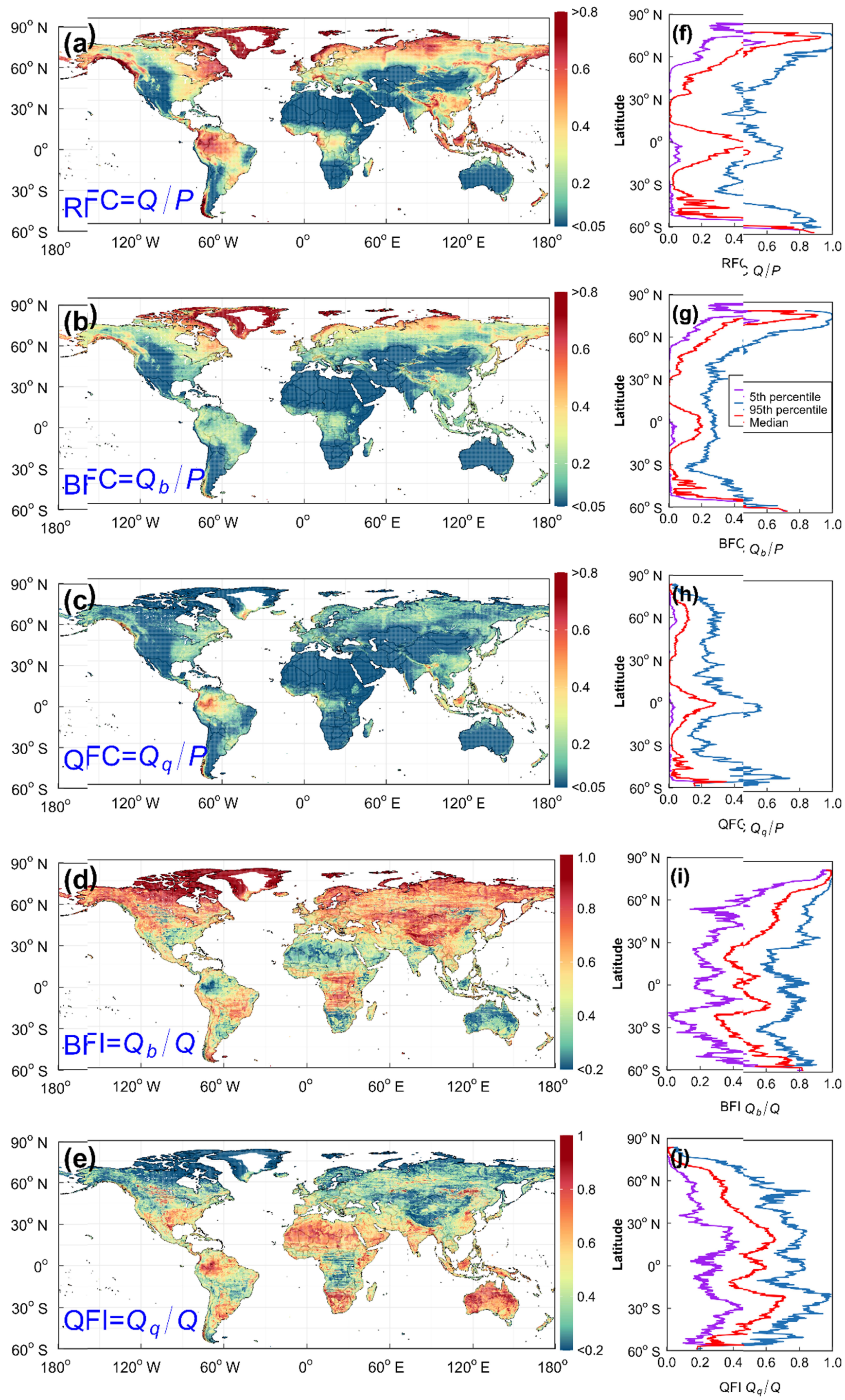


Figure 6. Global map of flow coefficients: (a) runoff coefficient (RFC) estimated with the Budyko curve (Eq. 3), (b) baseflow coefficient (BFC) estimated with the BFC curve (Eq. 4), (c) quickflow coefficient (QFC=RFC–BFC), (d) baseflow index (BFI=Q_b/Q) and (e) quickflow index (QFI=Q_q/Q). Subplots (f), (g), (h), (i) and (j) show the median latitudinal variation of the respective variables (in red), and the 5th (purple) and 95th (blue) percentiles.

4.4 Dominant drivers of runoff partitioning

The most important drivers for α and $Q_{b,p}$ vary across regions (Figure 7). Topography and soil properties are most important for most catchments (75.5% catchments for α; 67.0% catchments for $Q_{b,p}$, see blue points in Figure 7). Slope (SLO) is identified most frequently as the dominant driver for both α (48.7% catchments) and $Q_{b,p}$ (34.1% catchments). The second driver is elevation (ELEV) for α dominant in 25.6% catchments, and soil thickness (STHI) for $Q_{b,p}$ in 29.1% catchments. Climate related factors are recognized as the most important driver at a smaller number of catchments (12.4% for α and 27.3% for $Q_{b,p}$) as also vegetation related factors (12.1% for α and 5.7% for $Q_{b,p}$).

The main drivers for the parameters are region specific. This provides another perspective on why there is no universally accepted relationship yet (Padrón et al., 2017), besides the complex interaction between the drivers (Ning et al., 2019) and uncertainties in P, E_p, and Q (Koppa et al., 2021). Previous studies have investigated the main drivers to Budyko parameters. These identified dominant drivers depend on the region of interest and are different from each other. Considering large basins globally, studies illustrated the main dominant property is vegetation (Li et al., 2013; Zhang et al., 2001) or climate seasonality (Liu et al., 2018). In small catchments regionally, the influence of climatic and physiographic properties on α becomes more variable as other factors need to be considered including soil properties (Shen et al., 2017), topography (Shao et al., 2012), human activities (Xing et al., 2018) and a combination of various controls (Yang et al., 2007). According to the regions identified in Figure 7, topography and soil related factors should be first considered for regionalizing α in most catchments. Climate related factors are important in eastern South America and the coastline of Australia. Vegetation is the most dominant driver in western America and the United Kingdom.

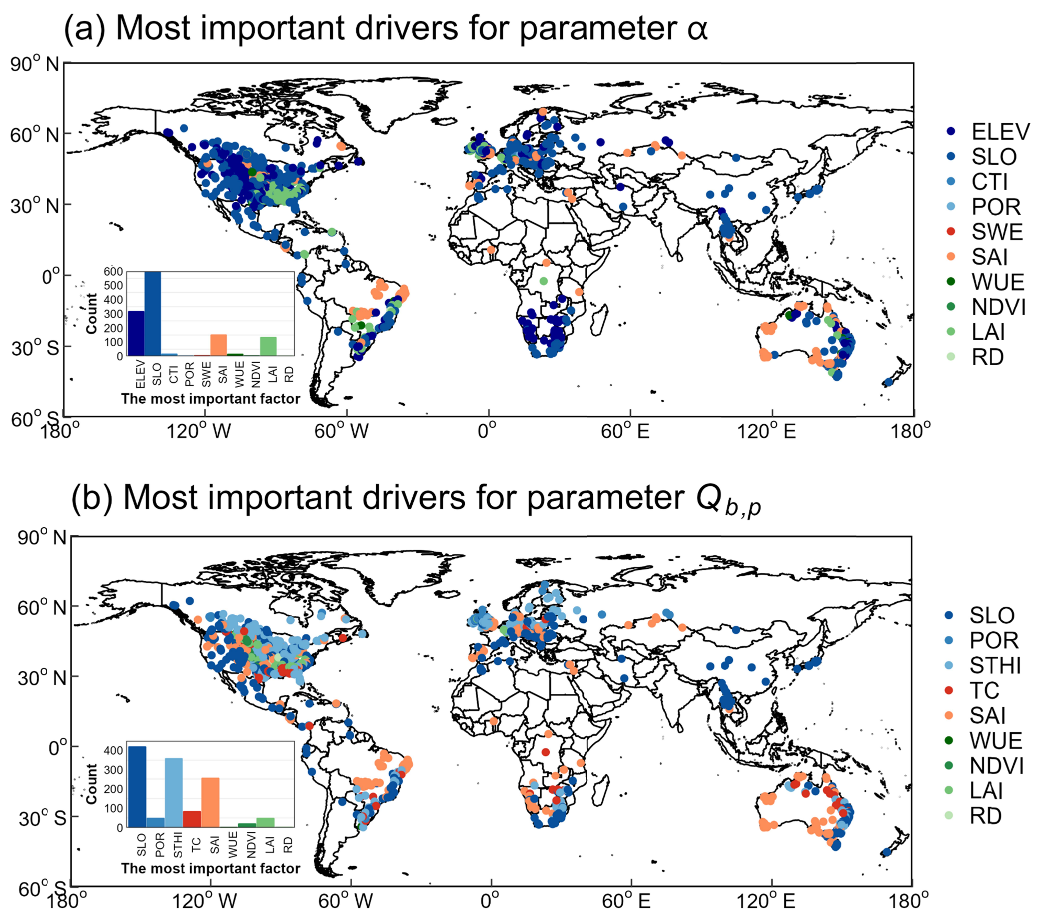


Figure 7. Spatial distribution of the most important driver to: (a) parameter α in Budyko and BFC curves, and (b) parameter $Q_{b,p}$ in BFC curve. The explanation of the abbreviations is provided in Table 1.

5 Discussion

5.1 Comparison to existing global datasets

The global, gridded baseflow dataset developed in this study is compared to baseflow products from previous studies. This includes the Global Streamflow Characteristics Dataset (GSCD) (Beck et al., 2015) and ERA5-Land dataset (Muñoz Sabater, 2019). The baseflow is directly available in the ERA5-Land dataset, and indirectly with GSDC as only the baseflow index (BFI) and runoff (Q_{mean}) are provided. To remain consistent with this study, “BFI1” in the GSCD dataset is selected where the baseflow is also calculated from the digital filter method. The long-term mean baseflow of this study (151 (spatial standard deviation std=181) mm yr^{-1}) is smaller than GSCD-based baseflow ($Q_{b,GSCD}$ = 241 (std=321) mm yr^{-1}, Figure 8a and c), and larger than ERA5-Land-based baseflow ($Q_{b,ERA5\text{-}Land}$ = 79 (std=145) mm yr^{-1}, Figure 8b and d). The baseflow estimated in this study has a larger spatial correlation with $Q_{b,GSCD}$ (R^2=0.84) than $Q_{b,ERA5\text{-}Land}$ (R^2=0.51), while a smaller RMSE is found relative to $Q_{b,ERA5\text{-}Land}$ (RMSE=143 mm yr^{-1}) compared to $Q_{b,GSCD}$ (RMSE=198 mm yr^{-1}). This means the spatial variability of our Budyko-ML-based baseflow is more similar to GSCD, while the magnitudes are closer to ERA5-Land values. The baseflow coefficient (BFC=Q_b/P) is also compared to estimates according to GSCD and ERA5-Land. Figure S4 shows that the correlation of BFC is higher relative to GSCD (R^2=0.73) than ERA5-Land (R^2=0.07).

These three datasets are a result of different types of methods with each their strengths and weaknesses: physically constrained ML in this study, pure ML for GSCD and the land surface model H-TESSEL for ERA5_Land. The GSCD dataset is not physically constrained, but available at a higher spatial resolution (0.05°) and based on significantly more catchments (3394). The H-TESSEL used for ERA5_Land is not specifically developed for runoff estimation, but for land-atmosphere interactions. Also, H-TESSEL uses "expert opinion" based parameterization instead of being calibrated. This may be the cause for the poor correlation results and confirms opinions that land surface models such as the H-TESSEL poorly estimate baseflow and groundwater-surface water interactions (Clark et al., 2015, Beck et al., 2017).

As the baseflow index (BFI=Q_b/Q) is more sensitive to both baseflow and runoff uncertainties (Gnann et al., 2019), the BFI is only compared to GSCD. As shown in Figure S5, the performance of BFI at catchment scale is acceptable with R^2 of 0.54. However, the performance decreases at global grid scale (R^2=0.18). This low correlation of BFI may be attributed to several aspects. First, getting accurate BFI estimates from separately estimated Q and Q_b is difficult (Beck et al., 2017). The spatial variability of the individual variables Q and Q_b is much larger than the difference between Q and Q_b within each catchment (BFI). High similarities in the spatial distribution of Q and Q_b does not guarantee similar BFI values. The BFI is especially sensitive in dry regions since very low Q values (in the denominator) would result in high BFI values. Second, the study period of the GSCD dataset and this study is different. The GSCD dataset adopted average P and E_p datasets with daily Q datasets, while we kept their period consistent by selecting the crossing period. Third, the different digital filters used, Lyne–Hollick in this study and method from Van Dijk (2010) for GSCD, may cause a slight difference in the observed baseflow used to train models.

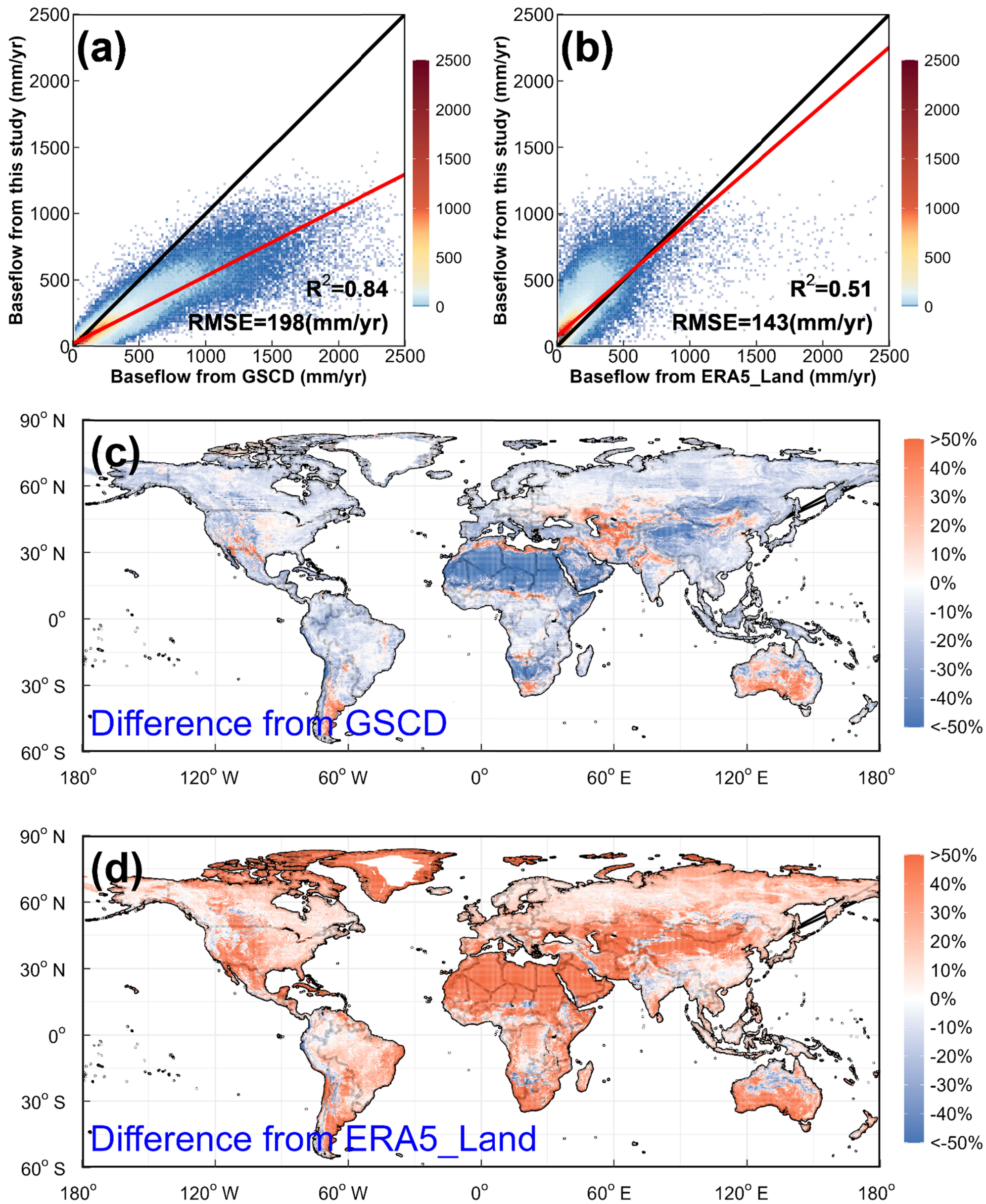


Figure 8. Estimated global baseflow compared to GSCD (a and c) and ERA5-Land (b and d). Subplots (a) and (d) show the pixel-by-pixel scatter plot with the red lines representing the fitted curve, and the black lines the 1:1 line. Subplots (c) and (d) show the spatial distribution of their difference, calculated as their difference divide by their average (i.e., ($Q_{b,this\ study}$ –

$Q_{b,GSCD}$)/(0.5 * ($Q_{b,this\ study}$ + $Q_{b,GSCD}$)) for plot (c), and replace $Q_{b,GSCD}$ with $Q_{b,ERA5\text{-}Land}$ for plot (d).

5.2 Partitioning in northern latitudes

The runoff partitioning dataset developed in this study has difficulties in representing high northern latitudes correctly (HNL). First, precipitation datasets tend to be underestimated at latitudes higher than 60 °N (snow-dominant regions) at a long-term scale (Beck et al., 2017). The MSWEP dataset used in this study attempted to correct this underestimation bias by the catch-ratio equation (Goodison et al., 1998). But the precipitation in HNL still has uncertainties for the gauge under-catch. By using the Budyko constrained ML framework, precipitation is a dominant forcing for the runoff and baseflow estimation, such that the uncertainty of precipitation could bring large uncertainties. Second, there are no catchments in HNL when applying the selection criteria as described in Section 3.1, which means relations based on trained ML models may not be accurate there. As a data-oriented method, ML relies on training data sources with all representative data expected to be included (Ma et al., 2020). Third, the partitioning of baseflow and quickflow from snow-melt runoff is different from precipitation-generated runoff. Based on the definition of baseflow from Hall (1968), snow-melt runoff is grouped under baseflow. However, this contradicts the digital filter technique that considers high frequency parts as quickflow. Snow-melt runoff can have quickflow features of high frequency and peaks during summer, even generating floods (Benn et al., 2012). As shown in Figure 6d and 6i, the baseflow index in high northern latitudes have high values, which means runoff comes more from baseflow rather than quickflow. It makes sense if we consider snow-melt runoff as baseflow following the definition of Hall (1968). However, we recommend further investigation on runoff partitioning in NHL regions.

6 Conclusion

This study extends the Budyko constrained machine learning (Budyko–ML) approach to develop global datasets for the runoff (Q), baseflow (Q_b), quickflow (Q_q) and the respective coefficients relative to the precipitation (Q/P, Q_b/P and Q_q/P) at 0.25°resolution. This hybrid approach, combining the Budyko-based framework (Budyko and BFC curve) with ML (BRT), retains the advantage of both the physical part and ML to achieve good performances within physical boundaries. This advantage is illustrated by the good performance of both Q and Q_b as they performed well compared to field observations during the testing phase with R^2=0.96 and R^2=0.88 for Q and Q_b, respectively. BRT estimates parameters globally, despite their unknown relationship with climatic and physiographic properties. Among the 13 climatic and physiographic properties considered, the main drivers are region specific, with topography and soil related factors being predominant in most catchments. Findings indicate that global runoff amounts to 274 (spatial standard deviation std=418) mm yr^{-1}, which means 30.3% (std=26.5%) of the precipitation, of which 20.6 (std=22.1%) is baseflow and 9.7% (std=10.3%) is quickflow. Our baseflow estimates are lower than GSCD estimates (241 (std=321) mm yr^{-1}) but larger than ERA5-Land estimates (79 (std=145) mm yr^{-1}). These large differences illustrate the large uncertainty that remains in runoff partitioning at global scales, and the required efforts to improve it further.

Data Availability Statement

The developed global 0.25° datasets including runoff (Q), baseflow (Q_b), runoff coefficient (Q/P), baseflow coefficient (Q_b/P) and baseflow index (Q_b/Q) are available at Global runoff partitioning based on Budyko-constrained machine learning | Zenodo.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data 2018 5:1*, *5*(1), 1–12. https://doi.org/10.1038/sdata.2017.191

Amatulli, G., Domisch, S., Tuanmu, M. N., Parmentier, B., Ranipeta, A., Malczyk, J., & Jetz, W. (2018). Data Descriptor: A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific Data*, *5*. https://doi.org/10.1038/sdata.2018.40

Apurv, T., & Cai, X. (2020). Impact of Droughts on Water Supply in U.S. Watersheds: The Role of Renewable Surface and Groundwater Resources. *Earth's Future*, *8*(10). https://doi.org/10.1029/2020EF001648

Bai, P., Liu, X., Zhang, D., & Liu, C. (2020). Estimation of the Budyko model parameter for small basins in China. *Hydrological Processes*, *34*(1), 125–138. https://doi.org/10.1002/hyp.13577

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.

Beck, H. E., Van Dijk, A. I. J. M., De Roo, A., Dutra, E., Fink, G., Orth, R., & Schellekens, J. (2017). Global evaluation of runoff from 10 state-of-the-art hydrological models. *Hydrology and Earth System Sciences*, *21*(6), 2881–2903. https://doi.org/10.5194/hess-21-2881-2017

Beck, H. E., Van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & De Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, *21*(1), 589–615. https://doi.org/10.5194/HESS-21-589-2017

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., Van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., & Wood, E. F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, *21*(12), 6201–6217. https://doi.org/10.5194/hess-21-6201-2017

Benn, D. I., Bolch, T., Hands, K., Gulley, J., Luckman, A., Nicholson, L. I., Quincey, D., Thompson, S., Toumi, R., & Wiseman, S. (2012). Response of debris-covered glaciers in the Mount Everest region to recent warming, and implications for outburst flood hazards. In *Earth-Science Reviews* (Vol. 114, Issues 1–2, pp. 156–174). https://doi.org/10.1016/j.earscirev.2012.03.008

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, *24*(1), 43–69. https://doi.org/10.1080/02626667909491834

de Bézenac, E., Pajot, A., & Gallinari, P. (2019). Deep learning for physical processes: incorporating prior scientific knowledge. *Journal of Statistical Mechanics: Theory and Experiment*, *2019*(12), 124009. https://doi.org/10.1088/1742-5468/ab3195

Budyko, M. I. (1961). The Heat Balance of the Earth's Surface. *Soviet Geography*, *2*(4), 3–13. https://doi.org/10.1080/00385417.1961.10770761

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). https://doi.org/10.1029/2020WR029529

Cheng, S., Cheng, L., Qin, S., Zhang, L., Liu, P., Liu, L., et al. (2022). Improved Understanding of How Catchment Properties Control Hydrological Partitioning Through Machine Learning. *Water Resources Research*, *58*(4). https://doi.org/10.1029/2021WR031412

Choudhury, BhaskarJ. (1999). Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model. *Journal of Hydrology*, *216*(1), 99–110. https://doi.org/https://doi.org/10.1016/S0022-1694(98)00293-5

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., et al. (2015). Improving the representation of hydrologic processes in Earth System Models.

Water Resources Research, *51*(8), 5929–5956. https://doi.org/10.1002/2015WR017096

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

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Sanderson Eric W., Malanding Jaiteh, Marc A. Levy, Kent H. Redford, Antoinette V. Wannebo, & Gillian Woolmer. (2002). The Human Footprint and the Last of the Wild. *BioScience*, *52*(10), 891–904.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(40), 10572–10577. https://doi.org/10.1073/pnas.1712381114

Franklin, J. (2008). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer 2005 27:2*, *27*(2), 83–85. https://doi.org/10.1007/BF02985802

Fu, B. F. (1981). On the calculation of the evaporation from land surface. *Chinese Journal of Atmospheric Sciences*, *5*(1), 23–31.

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

Good, S. P., Moore, G. W., & Miralles, D. G. (2017). A mesic maximum in biological water use demarcates biome sensitivity to aridity shifts. *Nature Ecology and Evolution*, *1*(12), 1883–1888. https://doi.org/10.1038/s41559-017-0371-8

Goodison, B., Louie, P., & Yang, D. (1998). WMO Solid Precipitation Measurement Intercomparison. *World Meteorological Organization-Publications-WMO TD*, *67*.

Hall, F. R. (1968). Base-Flow Recessions—A Review. *Water Resources Research*, *4*(5), 973–983. https://doi.org/10.1029/WR004I005P00973

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Huang, S., Dong, Q., Zhang, X., & Deng, W. (2021). Catchment natural driving factors and prediction of baseflow index for Continental United States based on Random Forest technique. *Stochastic Environmental Research and Risk Assessment*, *35*(12), 2567–2581. https://doi.org/10.1007/s00477-021-02057-2

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., De Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Zhang, K. (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, *467*(7318), 951–954. https://doi.org/10.1038/nature09396

Kraft, B., Jung, M., Körner, M., Koirala, S., & Reichstein, M. (2022). Towards hybrid modeling of the global hydrological cycle. *Hydrology and Earth System Sciences*, *26*(6), 1579–1614. https://doi.org/10.5194/hess-26-1579-2022

Koppa, A., Alam, S., Miralles, D. G., & Gebremichael, M. (2021). Budyko-Based Long-Term Water and Energy Balance Closure in Global Watersheds From Earth Observations. *Water Resources Research*, *57*(5). https://doi.org/10.1029/2020WR028658

Koppa, A., Rains, D., Hulsman, P., Poyatos, R., & Miralles, D. G. (2022). A deep learning-based hybrid model of global terrestrial evaporation. *Nature Communications*, *13*(1), 1912. https://doi.org/10.1038/s41467-022-29543-7

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

Liu, J., Zhang, Q., Singh, V. P., Song, C., Zhang, Y., Sun, P., & Gu, X. (2018). Hydrological effects of climate variability and vegetation dynamics on annual fluvial water balance in global large river basins. *Hydrology and Earth System Sciences*, *22*(7), 4047–4060. https://doi.org/10.5194/HESS-22-4047-2018

Lyne, V., & Hollick, M. (1979). Stochastic time-variable rainfall-runoff modelling. In *Institute of Engineers Australia National Conference*. Perth. Retrieved from https://www.academia.edu

Ma, K., Feng, D., Lawson, K., Tsai, W. P., Liang, C., Huang, X., et al. (2021). Transferring Hydrologic Data Across Continents – Leveraging Data-Rich Regions to Improve Hydrologic Prediction in Data-Sparse Regions. *Water Resources Research*, *57*(5). https://doi.org/10.1029/2020WR028600

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., et al. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. https://doi.org/10.5194/GMD-10-1903-2017

Marthews, T. R., Dadson, S. J., Lehner, B., Abele, S., & Gedney, N. (2015). High-resolution global topographic index values for use in large-scale hydrological modelling. *Hydrology and Earth System Sciences*, *19*(1), 91–104. https://doi.org/10.5194/hess-19-91-2015

Mianabadi, A., Davary, K., Pourreza-Bilondi, M., & Coenders-Gerrits, A. M. J. (2020, September 1). Budyko framework; towards non-steady state conditions. *Journal of Hydrology*. Elsevier B.V. https://doi.org/10.1016/j.jhydrol.2020.125089

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*(5), 3547–3562. https://doi.org/10.1002/2015WR017963

Miralles, D. G., Holmes, T. R. H., de Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, *15*(2), 453–469. https://doi.org/10.5194/hess-15-453-2011

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J. N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research*, *26*(7), 1465–1473. https://doi.org/10.1029/WR026I007P01465

Ning, T., Zhou, S., Chang, F., Shen, H., Li, Z., & Liu, W. (2019). Interaction of vegetation, climate and topography on evapotranspiration modelling at different time scales within the Budyko framework. *Agricultural and Forest Meteorology*, *275*, 59–68. https://doi.org/10.1016/j.agrformet.2019.05.001

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

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *SOIL*, *7*(1), 217–240. https://doi.org/10.5194/soil-7-217-2021

Porporato, A., Daly, E., & Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. *American Naturalist*, *164*(5), 625–632. https://doi.org/10.1086/424970/ASSET/IMAGES/LARGE/FG3.JPEG

Potter, N. J., Zhang, L., D Milly, P. C., McMahon, T. A., & Jakeman, A. J. (2005). Effects of rainfall seasonality and soil moisture capacity on mean annual water balance for Australian catchments, 6007. https://doi.org/10.1029/2004WR003697

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you, explaining the predictions of any classifier. KDD 2016. https://doi.org/10.1145/2939672.2939778

Roxy, M. K., Ghosh, S., Pathak, A., Athulya, R., Mujumdar, M., Murtugudde, R., et al. (2017). A threefold rise in widespread extreme rain events over central India. *Nature Communications 2017 8:1*, *8*(1), 1–11. https://doi.org/10.1038/s41467-017-00744-9

Sadeghi, M., Asanjan, A. A., Faridzad, M., Nguyen, P. H. U., Hsu, K., Sorooshian, S., & Braithwaite, D. A. N. (2019). PERSIANN-CNN: Precipitation estimation from remotely sensed information using artificial neural networks–convolutional neural networks. *Journal of Hydrometeorology*, *20*(12), 2273–2289. https://doi.org/10.1175/JHM-D-19-0110.1

Shao, Q., Traylen, A., & Zhang, L. (2012). Nonparametric method for estimating the effects of climatic and catchment characteristics on mean annual evapotranspiration. *Water Resources Research*, *48*(3), 3517. https://doi.org/10.1029/2010WR009610

Shen, C. (2018). A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists. *Water Resources Research*, *54*(11), 8558–8593. https://doi.org/10.1029/2018WR022643

Shen, Q., Cong, Z., & Lei, H. (2017). Evaluating the impact of climate and underlying surface change on runoff within the Budyko framework: A study across 224 catchments in China. *Journal of Hydrology*, *554*, 251–262. https://doi.org/10.1016/j.jhydrol.2017.09.023

Tan, X., Liu, B., & Tan, X. (2020). Global Changes in Baseflow Under the Impacts of Changing Climate and Vegetation. *Water Resources Research*, *56*(9). https://doi.org/10.1029/2020WR027349

Tang, Y., & Wang, D. (2017). Evaluating the role of watershed properties in long-term water balance through a Budyko equation based on two-stage partitioning of precipitation. *Water Resources Research*, *53*(5), 4142–4157. https://doi.org/10.1002/2016WR019920

Xie, K., Liu, P., Zhang, J., Han, D., Wang, G., & Shen, C. (2021). Physics-guided deep learning for rainfall-runoff modeling by considering extreme events and monotonic relationships. *Journal of Hydrology*, *603*, 127043. https://doi.org/10.1016/J.JHYDROL.2021.127043

Xing, W., Wang, W., Shao, Q., & Yong, B. (2018). Identification of dominant interactions between climatic seasonality, catchment characteristics and agricultural activities on Budyko-type equation parameter estimation. *Journal of Hydrology*, *556*, 585–599. https://doi.org/10.1016/j.jhydrol.2017.11.048

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset. *Water Resources Research*, *55*(6), 5053–5073. https://doi.org/10.1029/2019WR024873

Yang, D., Sun, F., Liu, Z., Cong, Z., Ni, G., & Lei, Z. (2007). Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of China using the Budyko hypothesis. *Water Resources Research*, *43*(4). https://doi.org/10.1029/2006WR005224

Yin, J., Gentine, P., Zhou, S., Sullivan, S. C., Wang, R., Zhang, Y., & Guo, S. (2018). Large increase in global storm runoff extremes driven by climate and anthropogenic changes. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06765-2

Zhang, L., Dawes, W. R., & Walker, G. R. (2001). Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water Resources Research*, *37*(3), 701–708. https://doi.org/10.1029/2000WR900325

Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., & Briggs, P. R. (2004). A rational function approach for estimating mean annual evapotranspiration. *Water Resources Research*, *40*(2). https://doi.org/10.1029/2003WR002710

Zhao, W. L., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S., Wen, Y., Lin, C., Li, X., & Qiu, G. Y. (2019). Physics-Constrained Machine Learning of Evapotranspiration. *Geophysical Research Letters*, *46*(24), 14496–14507. https://doi.org/10.1029/2019GL085291

Figure1.


(a) Budyko framework
Water limit
α=1.5
α=2.6
α=4.0
Energy limit
Q / P
P / E_p
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
(b) BFC curve
α=1.5
Q_b/P
P/E_p
Q_b,p/P
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1.0
2.0
3.0
4.0
5.0
6.0

Figure2.


Input dataset
Method
Catchment scale
P, Ep
Q
Q_b
Catchment properties
Catchment scale
Calculate parameter α in Budyko curve (Eq.3)
Calculate parameter $Q_{b,p}$ in BFC curve (Eq.4)
Develop machine learning model (BRT) to regionalize α and $Q_{b,p}$ as functions of catchment properties
Global grid scale
P, Ep
Catchment properties
Global grid scale
Estimate parameters α and $Q_{b,p}$ globally using catchment-developed BRT models
Estimate Q and Q_b globally with Eq.3 and Eq.4

Figure3.

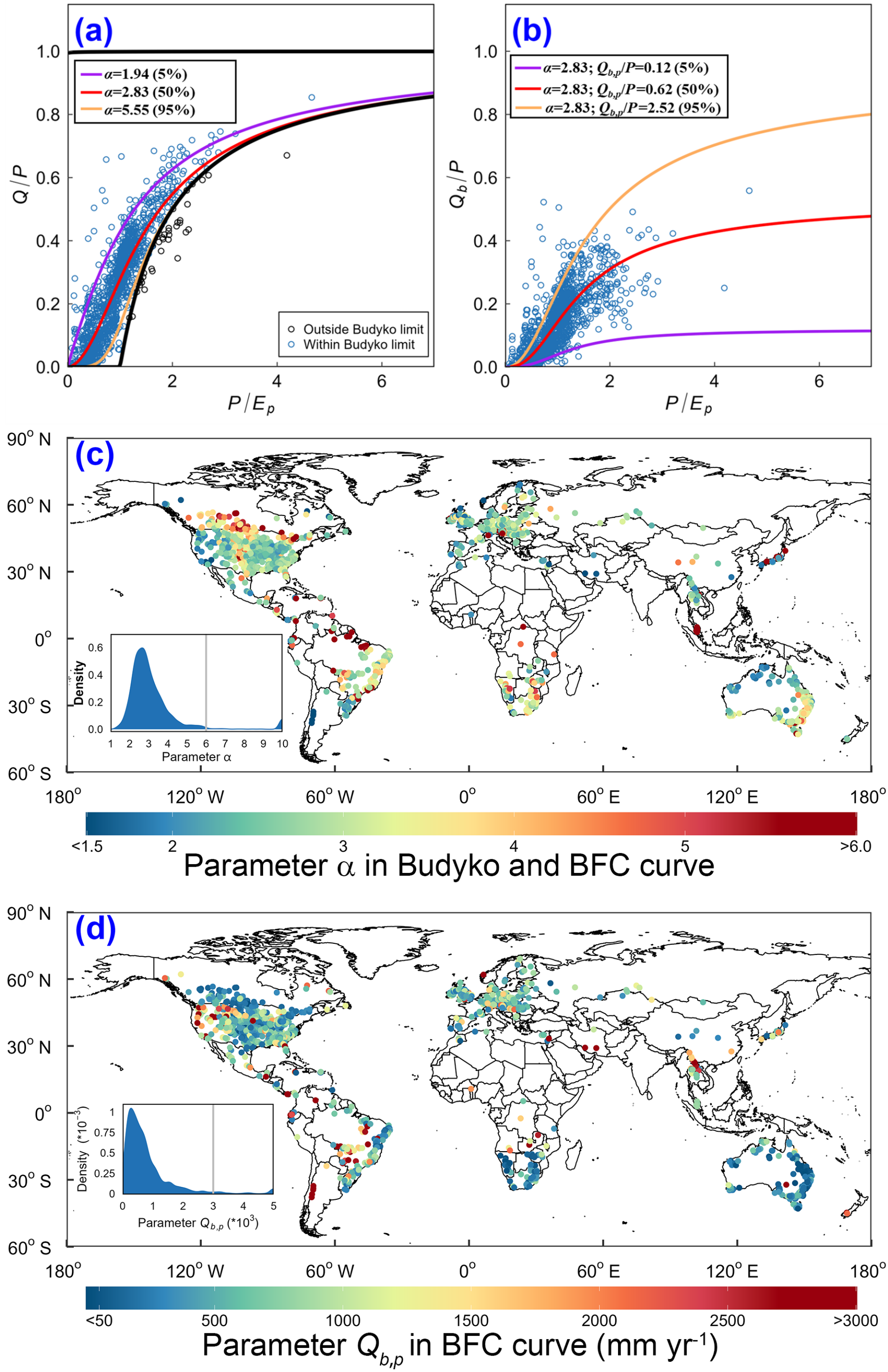
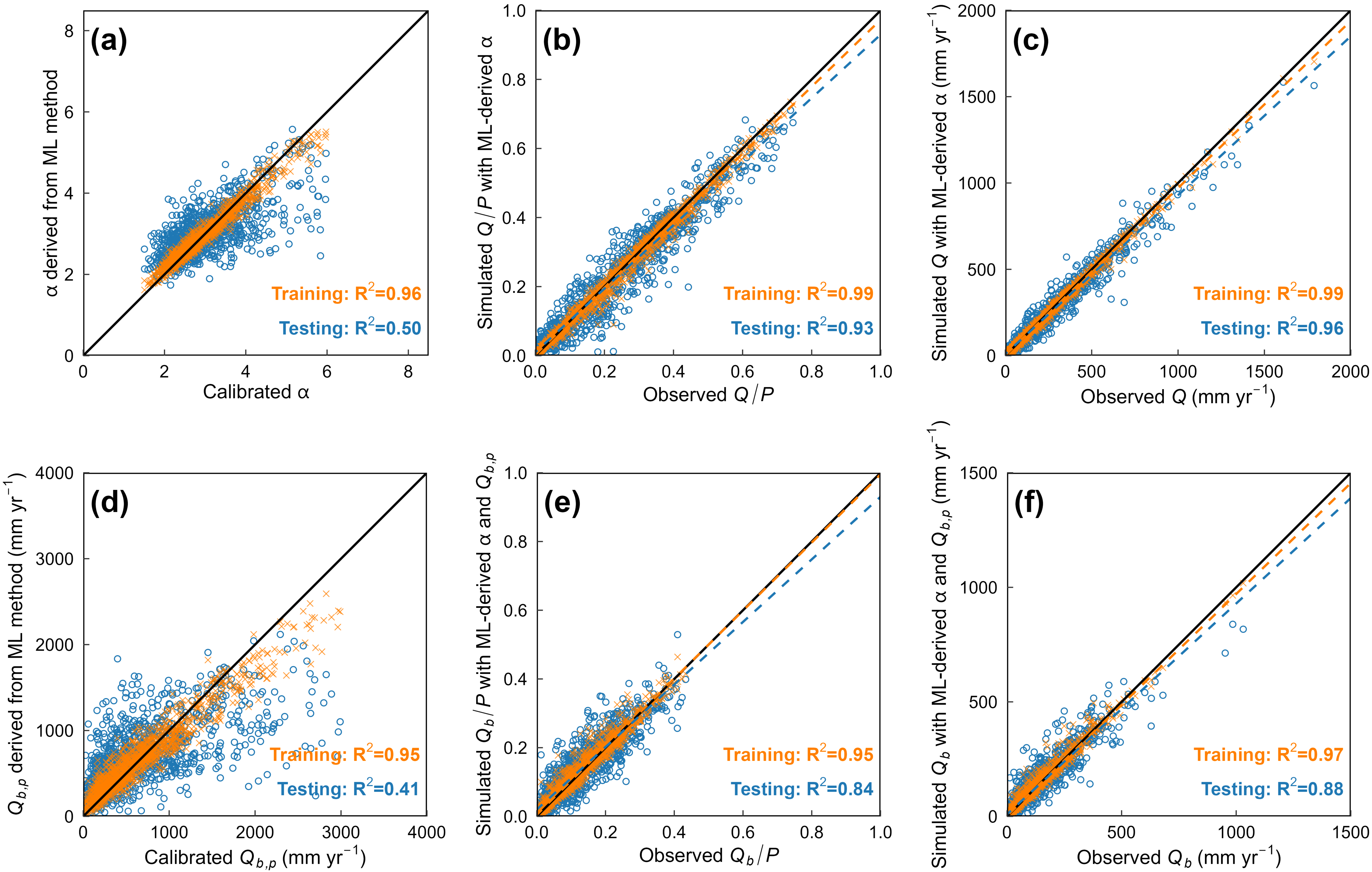
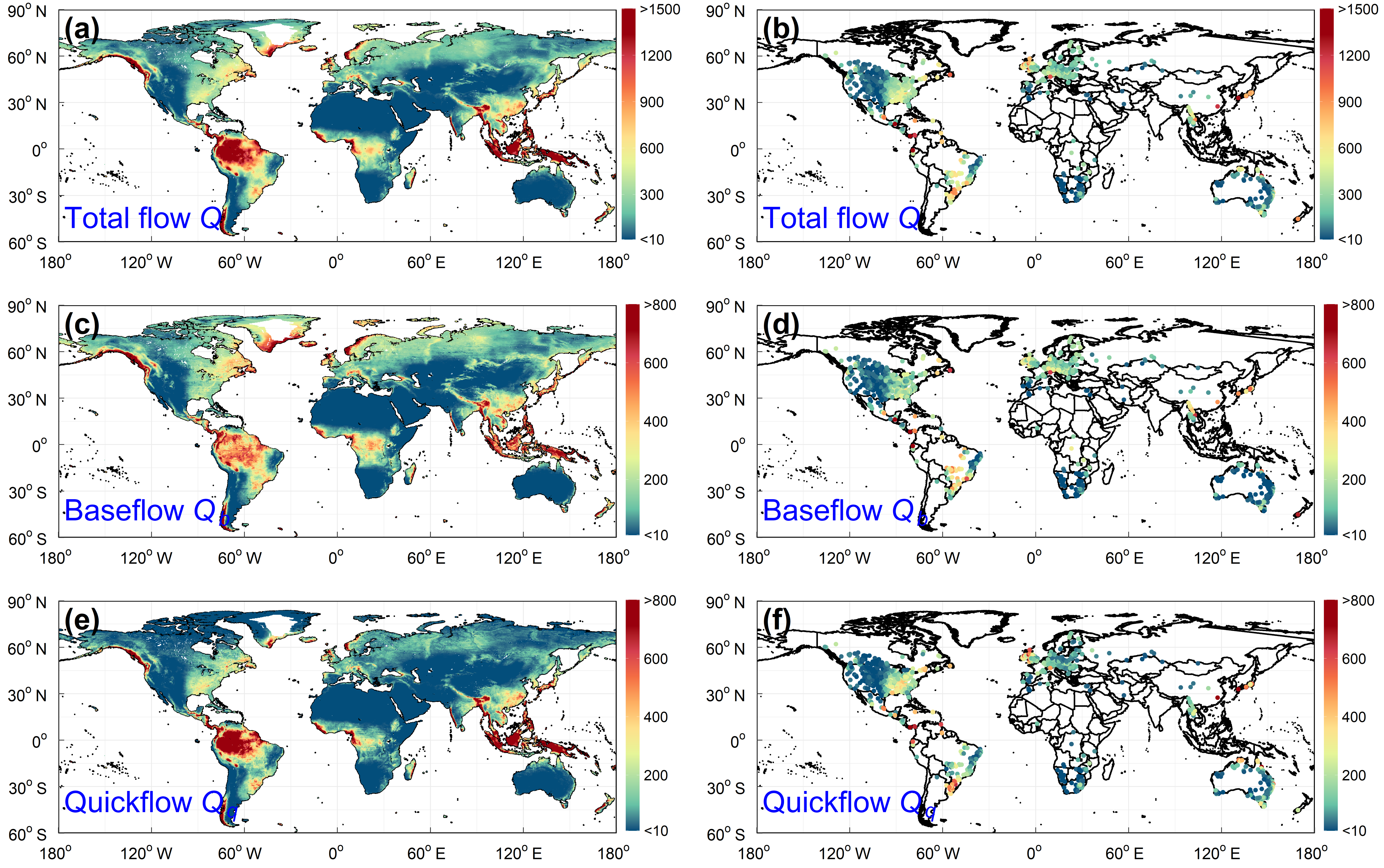

(a)
α=1.94 (5%)
α=2.83 (50%)
α=5.55 (95%)
Outside Budyko limit
Within Budyko limit
Q/P
P/E_p
(b)
α=2.83; Q_b,p/P=0.12 (5%)
α=2.83; Q_b,p/P=0.62 (50%)
α=2.83; Q_b,p/P=2.52 (95%)
Q_b/P
P/E_p
(c)
Density
Parameter α
Parameter α in Budyko and BFC curve
<1.5
>6.0
(d)
Density (*10^-3)
Parameter Q_b,p (*10^3)
Parameter Q_b,p in BFC curve (mm yr^-1)
<50
>3000
90° N
60° N
30° N
0°
30° S
60° S
180°
120° W
60° W
60° E
120° E

Figure4.



(a)
Training: R2=0.96
Testing: R2=0.50
α derived from ML method
Calibrated α
(b)
Training: R2=0.99
Testing: R2=0.93
Simulated Q/P with ML-derived α
Observed Q/P
(c)
Training: R2=0.99
Testing: R2=0.96
Simulated Q with ML-derived α (mm yr−1)
Observed Q (mm yr−1)
(d)
Training: R2=0.95
Testing: R2=0.41
Qb,p derived from ML method (mm yr−1)
Calibrated Qb,p (mm yr−1)
(e)
Training: R2=0.95
Testing: R2=0.84
Simulated Qb/P with ML-derived α and Qb,p
Observed Qb/P
(f)
Training: R2=0.97
Testing: R2=0.88
Simulated Qb with ML-derived α and Qb,p (mm yr−1)
Observed Qb (mm yr−1)

Figure5.



(a)
Total flow Q
(b)
Total flow Q
(c)
Baseflow Q_b
(d)
Baseflow Q_b
(e)
Quickflow Q_q
(f)
Quickflow Q_q
>1500
1200
900
600
300
<10
>800
600
400
200
<10
90° N
60° N
30° N
0°
30° S
60° S
180°
120° W
60° W
0°
60° E
120° E
180°

Figure6.

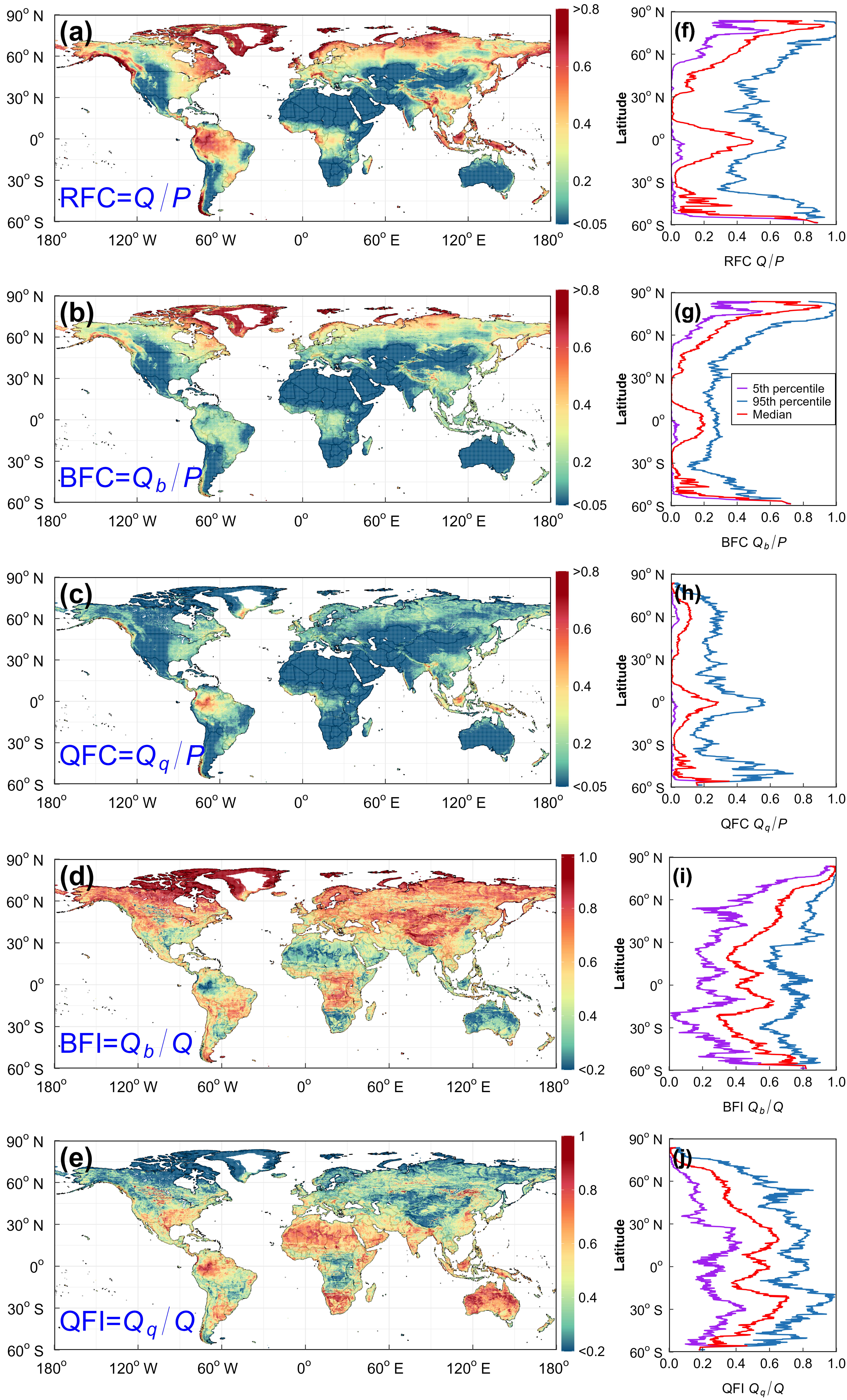

(a)
RFC=Q / P
(b)
BFC=Q_b / P
(c)
QFC=Q_q / P
(d)
BFI=Q_b / Q
(e)
QFI=Q_q / Q
90° N
60° N
30° N
0°
30° S
60° S
180°
120° W
60° W
0°
60° E
120° E
180°
>0.8
0.6
0.4
0.2
<0.05
1.0
0.8
0.6
0.4
<0.2
1
(f)
(g)
(h)
(i)
(j)
Latitude
5th percentile
95th percentile
Median
0.0
0.2
0.4
0.6
0.8
1.0
RFC Q / P
BFC Q_b / P
QFC Q_q / P
BFI Q_b / Q
QFI Q_q / Q

Figure7.

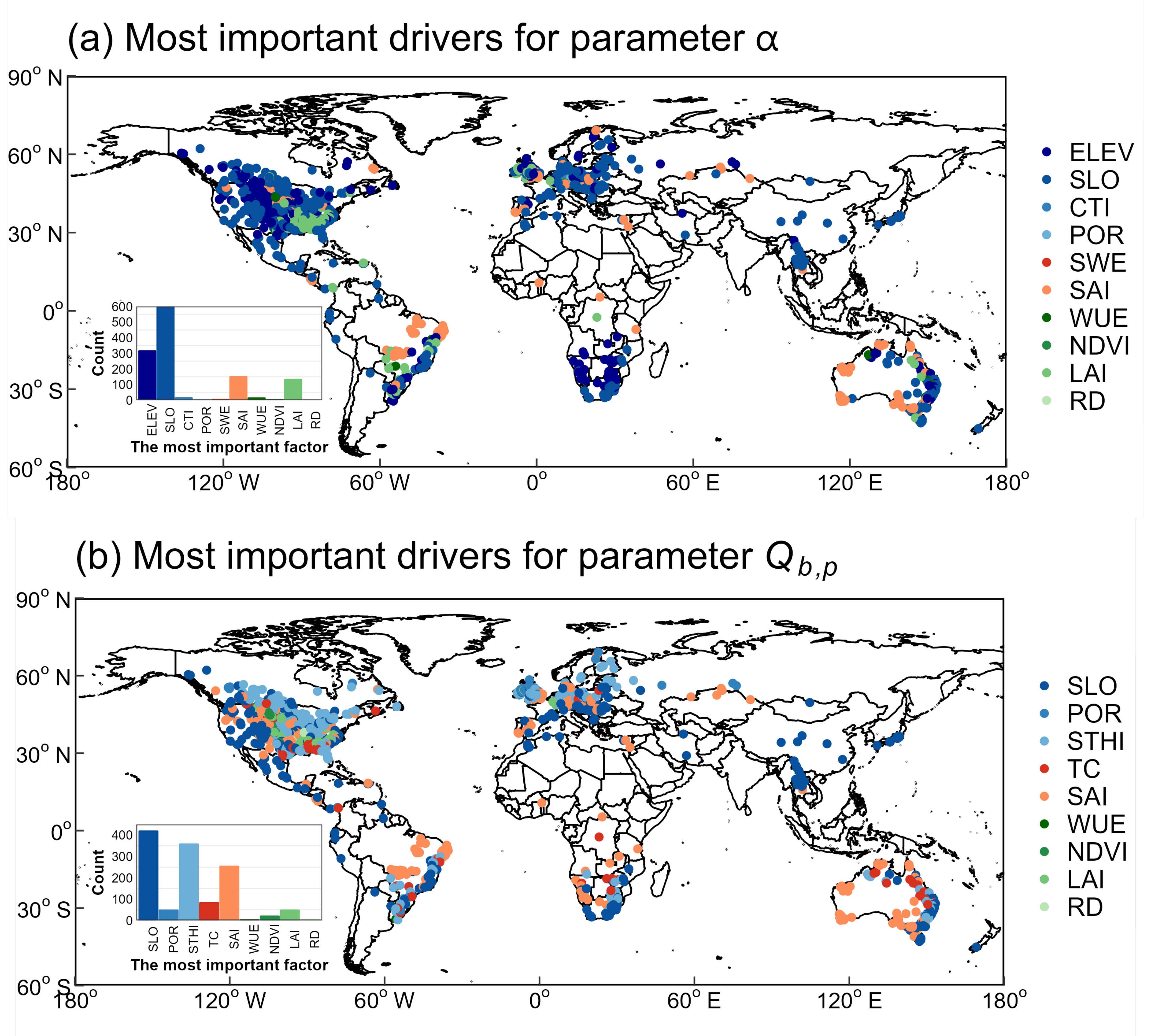
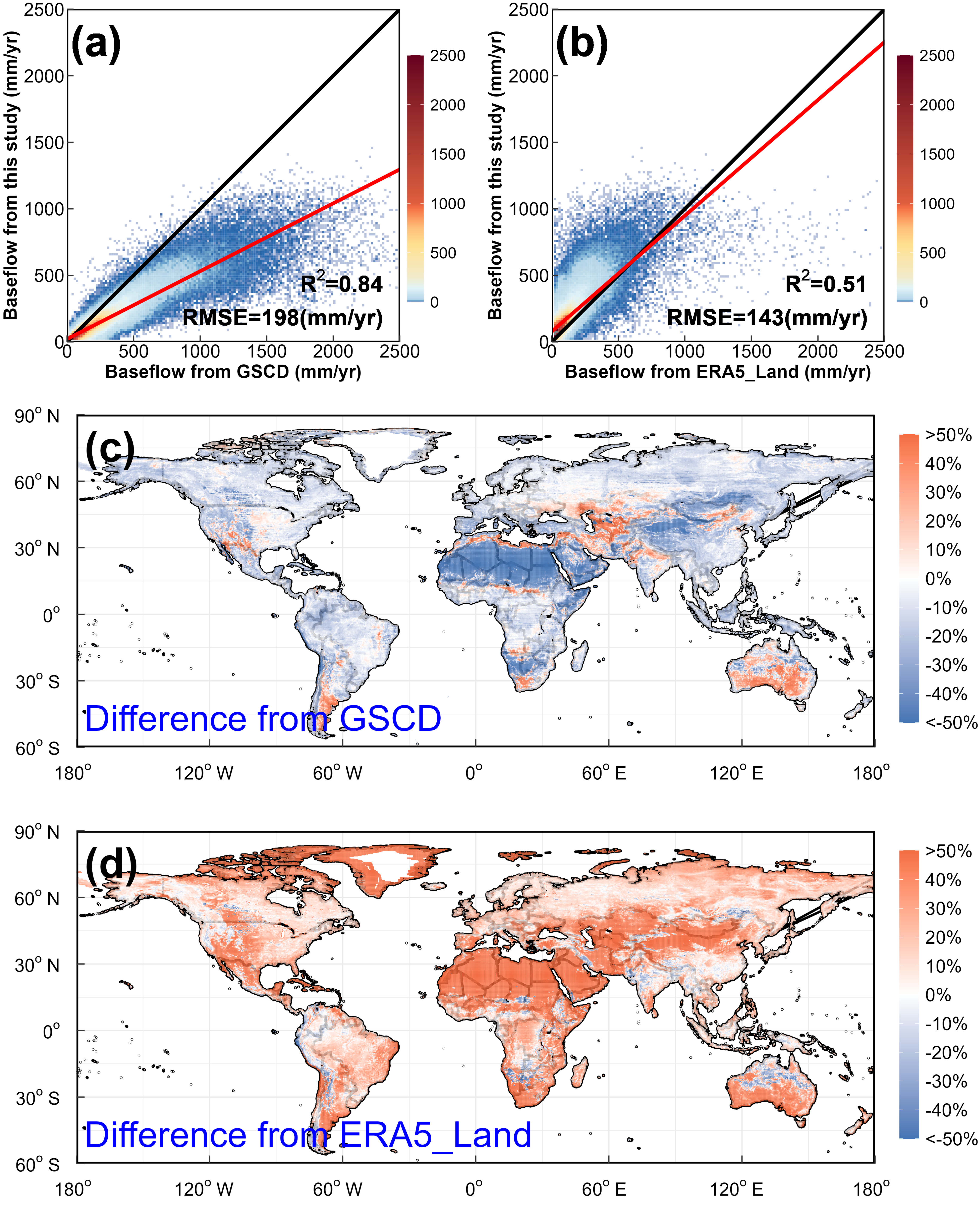

(a) Most important drivers for parameter α
90° N
60° N
30° N
0°
30° S
60° S
180°
120° W
60° W
0°
60° E
120° E
180°
ELEV
SLO
CTI
POR
SWE
SAI
WUE
NDVI
LAI
RD
Count
0
100
200
300
400
500
600
ELEV SLO CTI POR SWE SAI WUE NDVI LAI RD
The most important factor
(b) Most important drivers for parameter $Q_{b,p}$
90° N
60° N
30° N
0°
30° S
60° S
180°
120° W
60° W
0°
60° E
120° E
180°
SLO
POR
STHI
TC
SAI
WUE
NDVI
LAI
RD
Count
0
100
200
300
400
SLO POR STHI TC SAI WUE NDVI LAI RD
The most important factor

Figure8.


(a)
Baseflow from this study (mm/yr)
Baseflow from GSCD (mm/yr)
R^2=0.84
RMSE=198(mm/yr)
(b)
Baseflow from this study (mm/yr)
Baseflow from ERA5_Land (mm/yr)
R^2=0.51
RMSE=143(mm/yr)
(c)
Difference from GSCD
(d)
Difference from ERA5_Land
>50%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
<-50%