AGU ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2025WR040578

Regionalization of Optimal Baseflow Separation Using Catchment-Scale Characteristics

Yongen Lin[1], Yiwen Mei[1], Dagang Wang[1], Jinxin Zhu[1], Huan Wu[2], Shuo Wang[3], Gaohong Yin[4], Liang Gao[5], and Emmanouil N. Anagnoustou[6]

[1]Carbon-Water Observation and Research Station in Karst Regions of Northern Guangdong, School of Geography and Planning, Sun Yat-sen University, Guangzhou, China, [2]School of Atmospheric Science, Sun Yat-sen University, Zhuhai, China, [3]Department of Land Surveying and Geo-informatics, Hong Kong Polytechnic University, Hong Kong Special Administrative Region, China, [4]School of New Energy and Environment, Jilin University, Changchun, China, [5]State Key Laboratory of Internet of Things for Smart City and Department of Ocean Science and Technology, University of Macau, Macau, China, [6]Department of Civil and Environmental Engineering, University of Connecticut, Storrs, CT, USA

Key Points:

- A random forest-based framework is developed to predict baseflow parameters using catchment characteristics
- Catchment area dominates, ahead of climatic indices, soil hydraulic attributes, and human-influenced factors
- The model significantly improves baseflow separation accuracy compared to the area-based power function

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Mei,
meiyw3@mail.sysu.edu.cn

Citation:

Lin, Y., Mei, Y., Wang, D., Zhu, J., Wu, H., Wang, S., et al. (2026). Regionalization of optimal baseflow separation using catchment-scale characteristics. *Water Resources Research*, *62*, e2025WR040578. https://doi.org/10.1029/2025WR040578

Received 26 MAR 2025
Accepted 1 MAR 2026

Abstract Empirical baseflow filters are widely used for baseflow separation. These filters rely on ad-hoc parameters that introduce significant uncertainties in the calculation. A recent study by Mei et al. (2024, https://doi.org/10.1029/2023wr036386) optimized these parameters using environmental tracer data for 1,100 catchments across the Contiguous United States (CONUS). However, optimized parameters are unavailable for most CONUS catchments lacking tracer data. To address this gap, we developed regionalization models for the filter parameters using the random forest (RF) algorithm and 82 catchment-scale characteristics, including geomorphology, climate, soil properties, and human activities. We demonstrated this approach for the block length parameter N of the smooth minima baseflow filter, one of the optimized filters in Mei et al.'s study, across 855 catchments. Our results show that the prediction of N achieves an R^2 of 0.80. Predictor importance analysis identified catchment area as the most influential factor, followed by climate, hydrology, soil, and water usage characteristics. Using the RF-predicted N in baseflow separation improves daily baseflow accuracy, with the median Kling-Gupta Efficiency increasing from 0.62 to 0.80 compared to the literature-suggested area-based power function. This study enhances the accuracy of baseflow separation, providing a robust foundation for understanding streamflow partitioning and supporting improved hydrological modeling.

Plain Language Summary Baseflow, the slow and steady component of river flow, comes mainly from groundwater and soil moisture. It sustains river levels during periods with no rainfall, supporting ecosystems and human activities such as agriculture and industry. Unlike streamflow, which can be measured directly, baseflow is typically estimated using tools called "baseflow filters." These filters use rules (parameters) to calculate baseflow from streamflow data, and assigning accurate values to these parameters is critical for reliable results. In this study, we developed a data-driven framework using random forest models to predict a key filter parameter based on factors including catchment area, climate, soil properties, and human activities. Testing our framework on 855 catchments across the Contiguous United States, we found that it accurately predicts the parameter, with catchment area being the most important factor. Using our predicted parameter improves the accuracy of daily baseflow estimates compared to the empirical function based on catchment area alone. This helps us better understand how much of a river's flow comes from baseflow, supporting better decisions in managing water resources.

1. Introduction

Baseflow is an integral component of streamflow, primarily originating from deep soil storage and groundwater sources (Beck et al., 2013; Hare et al., 2021; J. Xie et al., 2024). It exhibits a delayed response, typically occurring over several days or longer after rainfall or snowmelt events (Hare et al., 2021; Jasechko et al., 2016). This delayed contribution of baseflow is critical for maintaining river flows during rainless period, thereby playing an indispensable role in supporting aquatic ecosystems, water supplies, and ecological balance (Lauenroth & Bradford, 2012; Q. Liu et al., 2021; Xu et al., 2023). Quantifying the proportion of baseflow in streamflow is vital for understanding dynamic composition of streamflow and variation of water transit times (Blöschl et al., 2019; Hare et al., 2021; Loague et al., 2010; McDonnell & Beven, 2014). The reduction in baseflow can lead to adverse environmental conditions, including elevated water temperatures, decreased dissolved oxygen levels, and nutrient

enrichment, potentially resulting in harmful algal blooms and detrimental effects on endemic aquatic species (Brookfield et al., 2021; Hare et al., 2021; Price, 2011). Conversely, increasing baseflow may elevate flood risks and strain hydraulic infrastructure (Bhaskar et al., 2016; Evans et al., 2020). Thus, understanding baseflow dynamics is critical for water resource management and aquatic environments protection (Hellwig & Stahl, 2018; Kourakos et al., 2019; Xu et al., 2023; Zarnaghsh & Husic, 2021), particularly in the face of evolving climatic pressures (Brown et al., 2019; Evans et al., 2020; L. Li et al., 2024) and growing resource demands (Alzraiee et al., 2024; Flörke et al., 2018; Knox et al., 2018).

Since direct measurement of baseflow remains challenging, empirical baseflow filters are commonly used to separate baseflow from streamflow (Humphrey et al., 2022; Mei, Wang, et al., 2024). Despite their straightforward application, these methods face scrutiny for potentially unclear results, primarily stemming from the ad-hoc filter parameters, whose determination is fraught with uncertainty (Eckhardt, 2012; Stoelzle et al., 2020; Yang et al., 2021). To address this issue, some studies utilize tracer data to guide the selection of filter parameters (Cartwright, 2022; Longobardi et al., 2016). Specifically, these tracer-guided methods identify the optimal filter parameters that capture the temporal dynamics of baseflow, which are consistent with the continuous measurements of conservative chemical concentration in streamflow. Due to the additional constraint imposed by the chemical concentration dynamics on the selection of filter parameters, this approach is considered more physically relevant and objective. However, tracer-guided methods are not applicable to most sites worldwide, as reliable tracer data with sufficient spatiotemporal coverage are rare (Halder et al., 2015; Hou et al., 2024; Putman & Bowen, 2019). To address this challenge and fully leverage tracer information from available sites, regionalization of empirical baseflow filter parameters based on catchment-scale characteristics may present a viable solution. Yet, the potential of such a framework for quantifying baseflow contributions has not been explicitly explored in the literature, highlighting the need for further investigation.

Empirical baseflow filters are simple conceptions of the catchment system on the integrated runoff generation and routing processes (X. Chen & Sivapalan, 2020; Eckhardt, 2005; Stewart, 2015). As such, the filter parameters can be regarded as hydrological signatures that reflect baseflow dynamics (Beck et al., 2013; Singh et al., 2019; Stewart, 2015; Stoelzle et al., 2020). From this perspective, the filter parameters are related to catchment-scale characteristics, such as climate, hydrology, soil properties, vegetation, land cover, and human activities (Addor et al., 2018; L. Li et al., 2014; Stoelzle et al., 2013; J. Zhang et al., 2020). This enable their prediction from catchment attributes much like other baseflow signatures, such as baseflow index (BFI), recession rate, and runoff coefficient (Beck et al., 2013; Singh et al., 2019; J. Zhang et al., 2020). For example, the block length parameter of the smooth minima method reflects catchment response and recession rate (Aksoy et al., 2008; Stoelzle et al., 2020), which are influenced by aquifer geometry, climate, landscape heterogeneity, hydraulic conductivity, and river channel topology (Beck et al., 2013; Jachens et al., 2020; Stoelzle et al., 2013). Similarly, the maximum BFI parameter of the Eckhardt (2005) recursive digital filter represents the long-term baseflow contribution to streamflow (Eckhardt, 2005, 2012) and has been associated with catchment size and human-influenced factors (Bloomfield et al., 2021; Singh et al., 2019; J. Zhang et al., 2020). While prior studies have not directly predicted filter parameters from catchment characteristics, the relationships they establish provide a strong rationale for developing predictive models for parameter regionalization.

A technical difficulty in regionalizing optimal baseflow filter parameters is the scarcity of large-sample, unified catchment databases. This gap may have been addressed by Mei, Wang, et al. (2024), who optimized the parameters of four empirical baseflow filters using specific electrical conductance (SEC) and turbidity data for 1,100 catchments across the Contiguous United States (CONUS). Their accompanying database (Mei, Zhu, & Wang, 2024), which includes both optimal parameters and derived daily baseflow time series, presents a timely opportunity to test the prediction of filter parameters from catchment characteristics. Another challenge is the absence of explicit functional relationships between these parameters and catchment characteristics. Existing attempts are limited; to our knowledge, only a power-law relationship between catchment area and the block length parameter of SMM has been proposed (Aksoy et al., 2008; Stoelzle et al., 2020; J. Zhang et al., 2017), an approach constrained by its fixed form and reliance on a single predictor. Machine learning models, by contrast, offer a powerful alternative by capturing complex, nonlinear interactions among multiple catchment descriptors (Nearing et al., 2021). Specifically, the random forest (RF) algorithm has proven effective in hydrology for modeling such interacting effects among climate, landscape, and hydrological variables, owing to its robustness to multicollinearity and minimal distributional assumptions (Ali et al., 2025; Y. Shen et al., 2022). Therefore, this study develops RF models to predict optimal baseflow filter parameters. Our central objective is to determine if

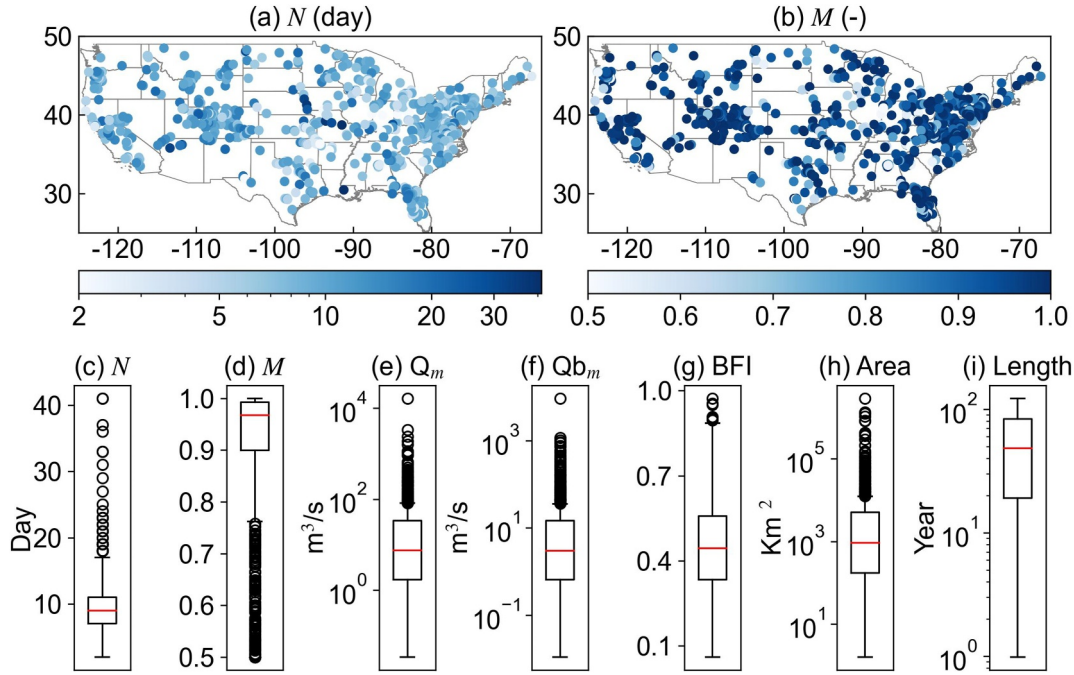


Figure 1. Spatial distribution of the 855 selected gages with the optimal SMM parameters N (a) and M (b) superimposed. Statistical distributions of the optimal N (c), M (d), average streamflow (e), average baseflow (f), BFI (g), catchment area (h), and flow record length (i) are also shown.

these parameters can be accurately extended to catchments lacking environmental tracer data, thereby enabling more objective and reliable baseflow quantification. The developed regionalization framework could be adapted to other regional monitoring networks where tracer data are limited.

2. Study Catchments and Data

2.1. The CONUS-Wise Baseflow Database

The Mei, Zhu, and Wang (2024) data set contains 16 optimized daily baseflow time series for 1,100 gages across CONUS. These series were generated by 16 hybrid baseflow filters, which were assembled from two tracers, two concentration estimation methods, and four empirical baseflow filters. Based on the performance evaluation by Mei, Wang, et al. (2024), SEC was identified a superior tracer to turbidity, and extreme value interpolation was found more suitable than a power-function relationship for tracer concentration estimation. Furthermore, the smooth minima method (SMM) demonstrated more consistent behavior than other digital filters. Consequently, this study utilizes the subset of 987 gages from the data set where the SMM parameters were optimized using SEC and extreme value interpolation for developing regionalization models. The SMM has two parameters: the block length parameter (N) and the local minima criteria parameter (M), detailed in Section 3.1. The corresponding daily baseflow time series for these gages, also sourced from Mei, Wang, et al. (2024), are used for validation (Section 3.4).

The selected gages were filtered based on optimization performance and data availability. Specifically, gages with suboptimal Kling-Gupta Efficient (KGE) between estimated and observed SEC (KGE < 0.5) were excluded (Mei, Wang, et al., 2024). Gages with incomplete catchment characteristics were also removed from the analysis (Section 2.2). After applying these criteria, a total of 855 gages were retained. The spatial distributions of these gages overlaid by their respective parameter values for N and M are shown in Figures 1a and 1b. Most gages are concentrated in the eastern CONUS compared to the central and the western parts. Within western CONUS, gages are primarily clustered near the coast and within the upper Colorado River basin. The parameters N and M show no dominant spatial pattern, although higher N s tend to located in the mountainous regions with drier climate and greater snow persistence (Figure 1a). For most catchments, N is less than 11 days, while M exceeds 0.9 (Figures 1c and 1d). The statistical distributions of average streamflow, average baseflow, BFI, catchment area,

and baseflow record length collectively demonstrate considerable diversity among the selected catchments (Figures 1e–1i).

2.2. Catchment Attributes

To predict the optimal SMM parameters, a wide range of catchment attributes were used. These attributes represent hydrology, climate, soil texture and hydraulics, geomorphology and geology, vegetation and land cover, and water usage of the catchments (Table 1), totaling 82 predictors. These attributes were derived from different spatiotemporal data products using similar methodologies. Most of the hydrology, climate, vegetation and land cover, and water usage variables are spatiotemporal (highlighted in red under the "Data sources and calculation processes" column of Table 1). They were aggregated over the catchment boundary to form corresponding time series, which were then used to calculate different statistics, including the mean, standard deviation (SD), *i*-th percentiles, and seasonality index (SI) (Walsh and Lawler, 1981). For the temporal-only streamflow and potential evapotranspiration (PE) variables (highlighted in blue), the different statistics were calculated directly from the time series. Spatial-only variables (unhighlighted) were processed by calculating the arithmetic mean of grids within each catchment boundary.

3. Methods

3.1. Recaps on the Smooth Minima Baseflow Separation Method

The SMM family is a widely used category of baseflow separation methods. This study employs the Aksoy et al. (2008) SMM variant, specifically designed to handle intermittent streamflow data. SMM operates under the assumption that baseflow constitutes 100% of streamflow during low-flow periods (Aksoy et al., 2008; Piggott et al., 2005). This is based on the observation that during extended rainless periods, surface runoff and rapid subsurface flow cease, leaving slow subsurface flows (e.g., groundwater and deep interflow) as the sole contributors. The implementation of SMM involves four main steps: Split the streamflow time series into consecutive, non-overlapping blocks of N days. Identify the local minima of streamflow, Q_x, within each N-day block. Filter out minima exceeding M times the discharge of their two immediate neighbors,

$$Q_{x,n-1} \geq \mathrm{MQ}_{x,n} \leq Q_{x,n+1}, \tag{1}$$

where n is the index of the block. Connect the remaining minima with straight lines to form the baseflow time series. Baseflow is set equal to streamflow whenever the calculated value exceeds the observed streamflow.

The integer block length parameter N represents an average streamflow delay and is typically defaulted to 5 days (Hare et al., 2021; J. Xie et al., 2024). N is sensitive to the partitioning of baseflow; generally, using smaller values of N result in a higher proportion of baseflow in streamflow (Aksoy et al., 2008; Stoelzle et al., 2020). N can also be calculated as a function of the drainage area, A in km^2 (Aksoy et al., 2008; J. Zhang et al., 2017):

$$N = \left[1.6 \times A^{0.2}\right]. \tag{2}$$

where $[*]$ represent round up to the closest integer. Parameter M determines if a streamflow minimum qualifies a local minimum. Larger values (usually not exceeding 1) indicate more strict criteria for Q_x to retain. Unlike N, M is less sensitive to baseflow separation results and is typically set to a default value of 0.9 (Aksoy et al., 2008; Stoelzle et al., 2020). Additionally, our attempts to model the optimal M showed poor performance (see Text S1 in Supporting Information S1 for the modeling procedures and Figure S1 in Supporting Information S1 for the performance). Therefore, the target parameter for our prediction model is N (Section 3.2).

3.2. Random Forest-Based Prediction Framework

The N prediction framework comprises four random forest (RF) models designed for different purposes (Figure 2 parts a and b). RF is an ensemble algorithm that handles nonlinear relationships via randomized feature and data subsets per tree, reducing overfitting and enhancing generalization (Breiman, 2001). The first RF regression model (RF_1, part a of the figure) evaluates predictor importance for all catchment attributes (Table 1). To minimize uncertainty in predictor importance ranking, RF_1 was tuned for three hyperparameters, namely trees

Table 1
Catchment Attributes Used to Develop Prediction Models for the Optimal SMM Parameters

Category	Attributes	Symbol (unit)	Data sources and calculation processes
Hydrology	Normalized mean annual streamflow	Q_{yr} (mm/yr)	USGS, NWIS
	SD of normalized daily streamflow	Q_{SD} (mm/day)	
	10-th percentile of normalized daily streamflow	Q_{10} (mm/day)	
	30-th percentile of normalized daily streamflow	Q_{30} (mm/day)	
	Median of normalized daily streamflow	Q_{50} (mm/day)	
	SI of normalized streamflow	Q_{SI} (−)	$Q_{SI} = \frac{1}{Q_{yr}} \sum_{m \in M_a} \left\lvert \overline{Q}_m - \frac{Q_{yr}}{\lvert M_a \rvert} \right\rvert$
	Mean daily SWE	$\overline{SWE}$ (mm)	SNODAS (NOHRSC, 2004)
	SD of daily SWE	SWE_{SD} (mm)	
	Proportion of days with SWE larger than 0.05 mm	f_{SWE} (−)	$f_{SWE} = \frac{1}{N_Y} \sum_{N_Y}^{y} \left[\frac{1}{\lvert D \rvert} \sum_{d \in D} I_d(SWE_d > 0.05) \right]$
	SI of SWE	SWE_{SI} (−)	$SWE_{SI} = \frac{1}{\lvert M_w \rvert \overline{SWE}} \sum_{m \in M_w} \lvert \overline{SWE}_m - \overline{SWE} \rvert$
	Mean annual PE	PE_{yr} (mm/yr)	Calculated based on Droogers and Allen (2002) using catchment-averaged temperature and precipitation from Daymet v4 (Thornton et al., 2021) as inputs
	SD of daily PE	PE_{SD} (mm/day)	
	Maximum daily PE	PE_x (mm/day)	
	Median of daily PE	PE_{50} (mm/day)	
	Minimum daily PE	PE_n (mm/day)	
	SI of PE	PE_{SI} (−)	$PE_{SI} = \frac{1}{PE_{yr}} \sum_{m \in M_a} \left\lvert \overline{PE}_m - \frac{PE_{yr}}{\lvert M_a \rvert} \right\rvert$
	Groundwater table depth	D_{gw} (m b.g.)	MODFLOW-6 (Zell and Sanford, 2020)
Climate	Mean daily minimum temperature	$\overline{TN}$ (°C)	Daymet v4 (Thornton et al., 2021)
	Maximum daily minimum temperature	TN_x (°C)	
	Median of daily minimum temperature	TN_{50} (°C)	
	Minimum daily minimum temperature	TN_n (°C)	
	SI of minimum temperature	TN_{SI} (−)	$TN_{SI} = \frac{1}{\lvert M_a \rvert \overline{TN}} \sum_{m \in M_a} \lvert \overline{TN}_m - \overline{TN} \rvert$
	Mean daily maximum temperature	$\overline{TX}$ (°C)	Daymet v4 (Thornton et al., 2021)
	Maximum daily maximum temperature	TX_x (°C)	
	Median of daily maximum temperature	TX_{50} (°C)	
	Minimum daily maximum temperature	TX_n (°C)	
	SI of maximum temperature	TX_{SI} (−)	$TX_{SI} = \frac{1}{\lvert M_a \rvert \overline{TX}} \sum_{m \in M_a} \lvert \overline{TX}_m - \overline{TX} \rvert$
	Mean annual precipitation	P_{yr} (mm/yr)	Daymet v4 (Thornton et al., 2021)
	95-th percentile of daily precipitation larger than 0.01 mm	P_{95} (mm/day)	
	Median of daily precipitation larger than 0.01 mm	P_{50} (mm/day)	
	5-th percentile of daily precipitation larger than 0.01 mm	P_5 (mm/day)	
	Proportion of days with precipitation larger than 0.01 mm	f_P (%)	$f_P = \frac{1}{N_Y} \sum_{N_Y}^{y} \left[\frac{1}{\lvert D \rvert} \sum_{d \in D} J_d(P_d > 0.01) \right]$
	SI of precipitation	P_{SI} (−)	$P_{SI} = \frac{1}{P_{yr}} \sum_{m \in M_a} \left\lvert \overline{P}_m - \frac{P_{yr}}{\lvert M_a \rvert} \right\rvert$
Soil texture and hydraulics	Volumetric fraction of clay	f_{cly} (−)	GSDE (Shangguan et al., 2014)
	Volumetric fraction of silt	f_{slt} (−)	
	Volumetric fraction of sand	f_{snd} (−)	
	Volumetric fraction of soil organic matters	f_{OM} (−)	
	Volumetric fraction of gravels	f_{grv} (−)	

Table 1
Continued

Category	Attributes	Symbol (unit)	Data sources and calculation processes
	Soil susceptibility factor to erosion	K (−)	Calculated following Benavidez et al. (2018) using f_{cly}, f_{snd}, f_{slt}, and f_{OM} as the inputs
	Residual moisture content	W_{rsd} (%)	Global soil hydraulic and thermal properties data set (Dai et al., 2019)
	Wilting point	W_{wlt} (%)	Calculated based on the pedotransfer functions of Saxton and Rawls (2006) using f_{cly}, f_{snd}, f_{slt}, and f_{OM} as the inputs
	Field capacity	W_{FC} (%)	
	Saturated water capacity	W_{sat} (%)	
	Coefficients of moisture-tension	λ (−)	
	Saturated hydraulic conductivity	K_{sat} (mm/hr)	
	Available water capacity	ΔW (%)	$\Delta W = W_{FC} - W_{wlt}$
Geomorphology and geology	Catchment area	A (km^2)	USGS, NWIS
	Latitude	Lat (°N)	
	Longitude	Lon (°W)	
	Relief ratio	R_R (−)	Global distributed basin morphometric data set (X. Shen et al., 2017)
	Elongation ratio	R_E (−)	
	Mean elevation	$\overline{Z}$ (m a.s.l.)	HydroSHEDS v1 (Lehner et al., 2008)
	SD of elevation (m a.s.l.)	Z_{SD} (m a.s.l.)	
	Mean drop (m)	$\overline{Z_o}$ (m)	$\overline{Z_o} = \frac{1}{\lvert N_P \rvert} \sum_{b \in N_P} (Z_b - Z_o)$
	Porosity of subsoil	ϕ (%)	GLHYMPS v2.0 (Huscroft et al., 2018)
	Permeability of subsoil	k_R (m^2)	
	Occurrence of R horizon	Pr_{RO} (−)	DTB (Shangguan et al., 2017)
Vegetation and land cover	Mean 16-daily LAI	$\overline{LAI}$ (−)	GLOBMAP LAI v3 (Y. Liu et al., 2012)
	SD of 16-daily LAI	LAI_{SD} (−)	
	SI of LAI	LAI_{SI} (−)	$LAI_{SI} = \frac{1}{\lvert M_a \rvert \overline{LAI}} \sum_{m \in M_a} \lvert \overline{LAI}_m - \overline{LAI} \rvert$
	Biomass of root	M_{rt} (Mg/ha)	Global map of root biomass across the world's forests (Huang et al., 2021)
	Biomass of shoot	M_{sh} (Mg/ha)	
	Mean annual areal fraction of water bodies	f_{Wat} (−)	MODIS, MCD12Q2 v6.1 (Friedl et al., 2022)
	Mean annual areal fraction of vegetation cover	f_{Veg} (−)	
	Mean annual areal fraction of urban and barren land	f_{UB} (−)	
	Mean annual areal fraction of impervious surface	$\overline{f_{Im}}$ (−)	NLCD (Dewitz, 2021)
	SD of annual areal fraction of impervious surface	$f_{Im,SD}$ (−)	
	Mean areal fraction of irrigation	$\overline{f_{Ir}}$ (−)	LANID (Y. Xie et al., 2021)
	SD of areal fraction of irrigation	$f_{Ir,SD}$ (−)	
	Population density	$\overline{Pop}$ (cap/km^2)	U.G.L.I. (Swanwick et al., 2022)
Water usage	Mean annual surface water withdrawal	$WW_{S,yr}$ (gal/cap/yr)	Public supply water use reanalysis for the 2000–2020 period by HUC12, month, and year for the conterminous United States (Alzraiee et al., 2024)
	SD of monthly surface water withdrawal	$WW_{S,SD}$ (gal/cap/day)	
	SI of surface water withdrawal	$WW_{S,SI}$ (−)	
	Mean annual groundwater withdrawal	$WW_{G,yr}$ (gal/cap/yr)	
	SD of monthly groundwater withdrawal	$WW_{G,SD}$ (gal/cap/day)	
	SI of groundwater withdrawal	$WW_{G,SI}$ (−)	
	Mean annual IWU	IWU_{yr} (mm/yr)	RZSM-based IWU data (C. Zhang & Long, 2021)

Table 1
Continued

Category	Attributes	Symbol (unit)	Data sources and calculation processes
	SD of monthly IWU	IWU_{SD} (mm/mon)	$IWU_{SI} = \frac{1}{IWU_{yr}} \sum_{m \in M_a} \left\| \overline{IWU}_m - \frac{IWU_{yr}}{\lvert M_a \rvert} \right\|$
	SI of IWU	IWU_{SI} (−)	
	Mean daily storage	$\overline{S}$ ($\times 10^6$ m^3)	ResOpsUS (Steyaert et al., 2022)
	SD of daily storage	S_{SD} ($\times 10^6$ m^3)	
	SI of storage	S_{SI} (−)	

Note. Data products in the "Data sources and calculation processes" column is categorized as follow: unhighlighted (spatial), highlighted blue (temporal), and highlighted red (spatiotemporal). Catchment attributes highlighted in gray in the "Symbol (unit)" column are selected for the development of prediction models (see Section 3.3 for details). Acronyms: *USGS*: United States Geological Surveys; *NWIS*: National Water Information System; *SNODAS*: Snow Data Assimilation System; *SWE*: Snow water equivalent; *GSDE*: Global Soil Data set for Earth System Modeling; *GLHYMPS*: GLobal HYdrogeology MaPS of permeability and porosity; *DTB*: Global depth to bedrock data set; *LAI*: Leaf area index; *MODIS*: Moderate Resolution Imaging Spectroradiometer; *NLCD*: National Land Cover Data; *LANID*: Landsat-based Irrigation Data set; *U.G.L.I.*: Updatable Gridded Lightweight Impervious population data set; *HUC*: Hydrologic Unit Code; *RZSM*: Root-zone soil moisture; *IWU*: Irrigation water used. Quantities: m: month index; M_a: a set of all months; $|\cdot|$: length of the set; $\overline{X}_m$: mean flux or state of month m; M_w: a set of winter months (1, 2, 3, 11, and 12); d: day index; D: a set of all days in a year; N_Y: number of years in the records; SWE_d: SWE of day d; $I_d(SWE_d > 0.05)$: an indicator showing that if SWE_d is larger than 0.05 mm; P_d: precipitation of day d; $J_d(P_d > 0.01)$: an indicator showing that if P_d is larger than 0.01 mm; b: grid index the catchment grids; N_P: a set of all catchment grids; Z_b: elevation of grid b; Z_o: elevation of the catchment outlet grid. k_R values are log-transferred and multiplied by 100 for model training.

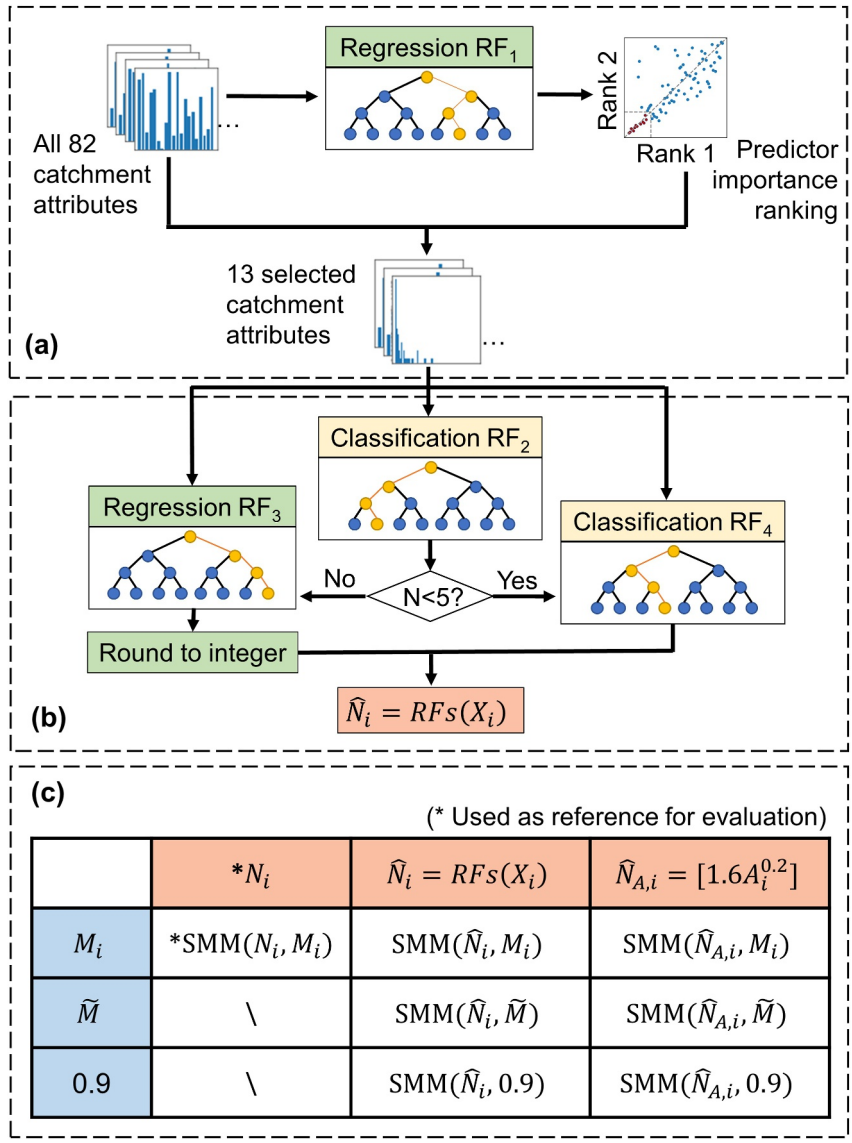
count, number of features for data splitting, and minimum leaf sample, using 10-fold cross validation and grid search (see Text S2 and Table S1 in Supporting Information S1 for the detailed procedures) on 80% of the gages. Ten replicated RF models were then trained with the optimal parameters. Predictive values of the catchment attributes were measured by Gini feature importance and minimal depth (Breiman, 2001; Ishwaran, 2007). Catchment attributes with high predictive values were retained for subsequent models.

Part b of Figure 2 illustrates a cascaded prediction framework for N. A classifier (RF_2) first determines if $N < 5$ days. Predictions of $N \geq 5$ are routed to a regressor (RF_3), the outputs of which are rounded to the nearest integer. Predications of $N < 5$ proceed to classifier RF_4, predicting N as 2, 3, or 4 days ($N \geq 2$ days following Mei, Wang, et al. (2024)). The data set comprises 116 gages with $N < 5$ (label 1) and 739 with $N \geq 5$ (label 0). RF_2 employed stratified sampling on 80% of each group (93 $N < 5$, 592 $N \geq 5$) to address population imbalance during model training. RF_3 was trained on the same 685 gages for continuous N regression, while RF_4 was trained solely on 93 $N < 5$ gages. Although RF_4's training sample is relatively small, the risk of overfitting is minimal due to the model's low complexity, tasked with distinguishing only three discrete outcomes ($N = 2, 3, 4$ days). This cascade mitigates RF3's overestimation in predicting N values smaller than 5 days (as shown later in Figure 4). Besides, RF_3 was preserved without retraining on the $N \geq 5$ training subset to maintain its performance for large N.

3.3. Measurements of Predictor Importance

Predictor importance is measured at two points in the framework (Figure 2), the 10 replicated RF_1 for predictor selection (part a) and collectively for the RF_2 to RF_4 prediction models (part b). For RF_1, predictor importance is evaluated using the Gini feature importance index and minimal depth technique (Breiman, 2001; Ishwaran, 2007). The Gini index measures node purity; feature importance depends on split frequency and impurity reduction. Minimal depth is the shortest distance from the root of the tree to the root of a feature's maximum subtree. Features with smaller minimal depths (appear earlier in the tree structure) are more influential (Ishwaran, 2007; Ishwaran et al., 2010). A key advantage of the minimal depth method is its ability to test for statistically significant predictors. The average minimal depth among all features serves as the threshold to identify these important variables (Ishwaran et al., 2010).

Figure 3 presents the feature importance ranking from the two approaches, which show strong consistency, especially for the top-ranked attributes. Based on the ranking, 13 catchment attributes were retained for prediction model development. These attributes consistently rank within the top 15 for both feature importance rankings and pass the minimal depth significance test (Ishwaran et al., 2010). Their names are displayed in the lower-right corner of the scatter plot (Figure 3) and highlighted in gray in Table 1 under the "Symbol (unit)" column. Among the selected attributes, catchment area (A) is the most important with an average minimal depth of 0.67



	$*N_i$	$\widehat{N}_i = RFs(X_i)$	$\widehat{N}_{A,i} = [1.6A_i^{0.2}]$
M_i	$*\mathrm{SMM}(N_i, M_i)$	$\mathrm{SMM}(\widehat{N}_i, M_i)$	$\mathrm{SMM}(\widehat{N}_{A,i}, M_i)$
$\widetilde{M}$	\	$\mathrm{SMM}(\widehat{N}_i, \widetilde{M})$	$\mathrm{SMM}(\widehat{N}_{A,i}, \widetilde{M})$
0.9	\	$\mathrm{SMM}(\widehat{N}_i, 0.9)$	$\mathrm{SMM}(\widehat{N}_{A,i}, 0.9)$

Figure 2. The RF-based prediction framework for the optimal *N* parameter of SMM.

and a Gini index of 0.35, followed by proportion of precipitation days (f_P). While the magnitudes of the Spearman correlation coefficient differ from the RF-based indices, the 13 selected attributes are characterized by comparatively high values.

To interpret the relative importance of the catchment attributes in the cascaded prediction framework, one of the permutation feature importance techniques measuring the decrease in coefficient of determination (R^2) was employed. This technique was chosen over the Gini index or minimal depth for compatibility with cascaded RF models. The process involves randomly permuting the values of a catchment attribute while holding others unchanged and evaluating R^2 reduction. Influential attributes show significant performance drops when permuted (Breiman, 2001). To complement the predictor importance ranking, the functional relationship between the selected catchment attributes and the optimal *N* parameter was evaluated using the partial dependence plots (PDP). This technique demonstrates predictive responses of a predictor by marginalizing the effects of other predictors (Friedman, 2001).

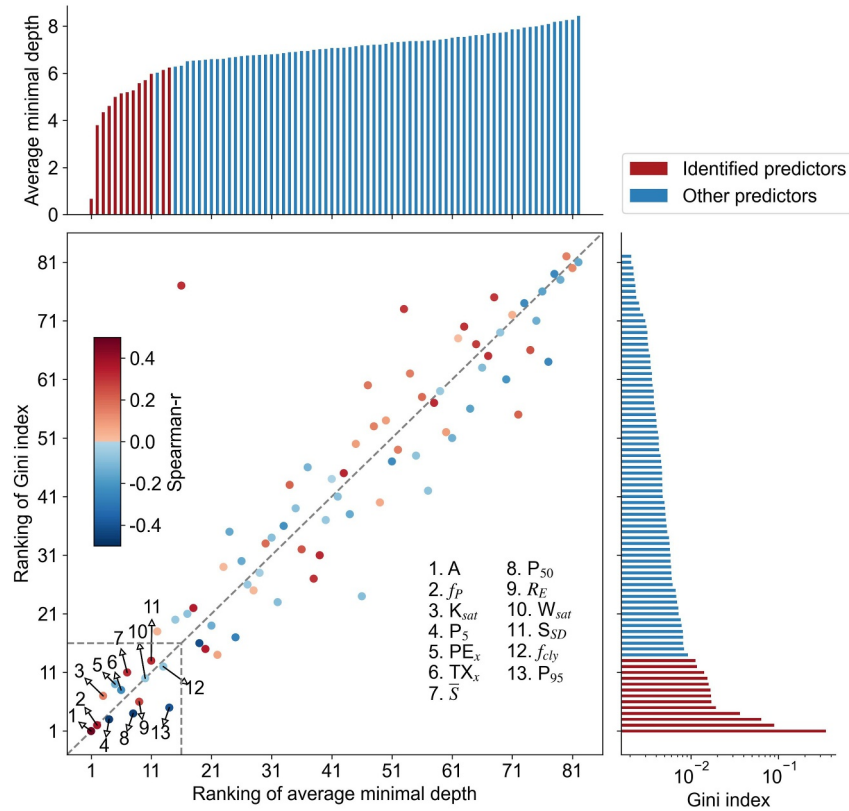


Figure 3. Ranking of predictor importance indices and selection of predictors. The scatter plot shows the ranking of the average minimal depth and the Gini index, while the histograms display the actual values. The color scale of the dots shows the Spearman correlation coefficient between catchment attributes and the reference N.

3.4. Performance Evaluation

There are two strategies to evaluate the performance of the RF-based N prediction (Figure 2 part c). Strategy 1 directly compared predicted N values against optimal values (Mei, Zhu, & Wang, 2024) using bias (B) and the coefficient of determination (R^2):

$$B = \mu_{N,p} - \mu_{N,r}, \tag{3}$$

$$R^2 = \frac{\sum_{i=1}^{C}\left(N_i - \mu_{N,r}\right)^2 - \sum_{i=1}^{C}\left(N_i - \hat{N}_i\right)^2}{\sum_{i=1}^{C}\left(N_i - \mu_{N,r}\right)^2}, \tag{4}$$

where μ_N is the mean of N over the gages, and the subscripts p and r denote the predicted and reference N, respectively. C represents the total number of gages, and N_i and $\hat{N}_i$ are the reference and predicted N for gage i. Positive (negative) values of B indicate over- (under-) estimation, with magnitudes closer to zero indicating better performance. R^2 measures the explained variance, ranging from 0 to 1, where 1 indicates perfect agreement. For classifiers (RF_2 and RF_4), performance was assessed using the Receiver Operating Characteristic (ROC) curve and the area under the ROC curve (AUC). The ROC curve plots the true positive rate against the false positive rate at various thresholds, and the AUC ranges from 0 to 1, with value closer to 1 indicating better classifier performance. To evaluate the additive value of our RF-based prediction, the area-based power function estimation of (denoted as $\hat{N}_{A,i}$) was also included.

Strategy 2 used the different N predictions to parameterize SMM for baseflow separation. Six parameterization schemes were designed, combining two predictions of N (RF-based and power function) with three parameterizations of M: the optimal value (M_i) from Mei, Zhu, and Wang (2024), the median of M_i ($\tilde{M} = 0.97$), and the

10.1029/2025WR040578

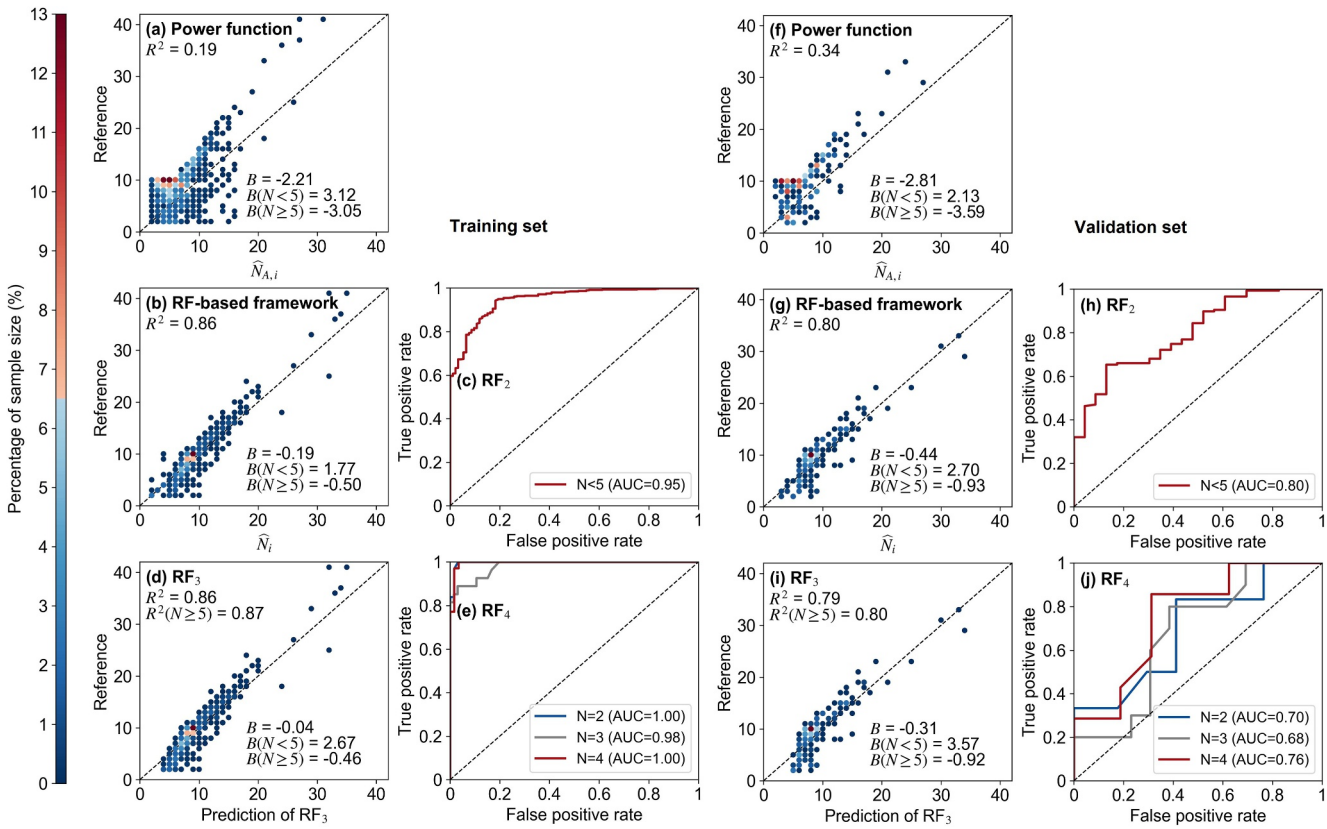


Figure 4. Performance of N prediction using the area-based power function (a, f) and the RF-based prediction framework (b, g). The performance of the RF_2 (c, h), RF_3 (d, i), and RF_4 (e, j), which constitute the RF-based prediction framework, is also shown. Panels (a)–(e) represent the training set, and panels (f)–(j) represent the validation set.

fixed value of 0.9. Baseflow separation using SMM(N_i, M_i) served as the reference. Discrepancies between predicted and reference baseflow time series were measured by KGE (Gupta et al., 2009):

$$\mathrm{KGE} = 1 - \sqrt{(\rho - 1)^2 + \left(\frac{\sigma_p}{\sigma_r} - 1\right)^2 + \left(\frac{\mu_p}{\mu_r} - 1\right)^2}, \tag{5}$$

where ρ is the correlation coefficient, and σ and μ are the standard deviation and mean of the baseflow time series, respectively. KGE values range from $-\infty$ to 1, with higher values indicating better performance. To further validate baseflow accuracy, we derived SEC time series from the predicted baseflow (Text S3 in Supporting Information S1) and compared them to the observed SEC using KGE.

4. Results

4.1. Evaluation of the Optimal N Prediction

Figure 4 compares the performance of the catchment area-based power function (Equation 2) and our RF-based prediction framework in estimating the parameter N for both the training and validation sets. It should be noted that the power function requires no optimization and is applied directly to the training and validation sets separately. The power function relationship demonstrates suboptimal performance with R^2 values of only 0.19 and 0.34 and biases of −2.21 and −2.81 days for the training and validation sets, respectively (panel a, f). The predicted $\hat{N}_{A,i}$ exhibits a clear trend of transitioning from over- to under-estimation as N increases. Specifically, for cases where N is less than 5 days, the biases are 3.12 and 2.13 days for the training and validation sets, respectively, while for cases where N is no less than 5 days, the biases are −3.05 and −3.59 days, respectively. In contrast, our RF-based prediction framework achieves significantly better performance with R^2 values of 0.86 and

0.80 for the training and validation sets, respectively (panel b, g), explaining at least 80% of the variability in N. The bias of $\hat{N}_i$ is −0.19 and −0.44 for the training and validation sets, just 9% and 16% of the biases observed in the $\hat{N}_{A,i}$ cases, respectively. However, overestimation for low N ($N < 5$) values and underestimation for high N ($N \geq 5$) values remain notable with biases of 1.77 and −0.50 days for the training set and 2.70 and −0.93 days for the validation set, respectively. The spatial distributions of both $\hat{N}_i$ and $\hat{N}_{A,i}$ broadly reflect the pattern of N_i (i.e., higher values in mountainous regions). However, the prediction biases for both methods show no systematic spatial structure (Figure S2 in Supporting Information S1).

Figures 4c–4e and 4h–4j illustrate the performances of the three individual RF models (RF_2, RF_3, and RF_4) within the prediction framework for the training and validation sets, respectively. Our analysis focuses on the validation set as it is the primary target of the prediction (panels h–j). The binary prediction model, RF_2, achieves an AUC of 0.80, indicating satisfactory prediction of whether N is less than 5 days or not (panel h). RF_3, trained on the entire training set, is specifically applied to gages identified as $N \geq 5$ by RF_2. This is because RF_3 overestimates $N < 5$ cases by 3.57 days (panel i). For cases where $N \geq 5$, R^2 of 0.8, indicating reasonable prediction accuracy. To address the overestimation issue for $N < 5$, an additional RF classification model, RF_4, was introduced (panel j). While RF_4 achieved moderate AUCs ranging from 0.68 to 0.76, it effectively reduces the overestimation for $N < 5$ from 3.57 to 2.70 days (compare panels g and i). Furthermore, the overall prediction performance improves slightly with the R^2 increasing from 0.79 to 0.80. These results justify the use of a combined prediction framework rather than relying on a single RF regression model for the entire space of N.

4.2. Predictor Importance Analysis

Figure 5a shows permutation feature importance for the 13 selected predictors, sorted by descending predictive importance based on the mean decrease in R^2. Catchment area is the most influential predictor with a mean decrease in R^2 of 0.12 ± 0.01. Distributions of decreases in R^2 for the remaining attributes exhibit some degrees of overlap, suggesting non-strict ranking. Following A, the climatic attribute TX_x (0.04 ± 0.02) and the hydrological attribute PE_x (0.03 ± 0.01) rank next. These attributes are associated with the dryness condition of the catchments. Overall, the important predictors include two geomorphological attributes (A and R_E), five climatic attributes (TX_x, f_P, P_5, P_{50}, and P_{95}), one hydrological attribute (PE_x), three soil texture and hydraulic attributes (f_{cly}, W_{sat}, and K_{sat}), and two water usage attributes (S_{SD} and $\overline{S}$).

The partial dependence plots for the 13 selected catchment attributes are depicted in Figures 5b–5n. The catchment areas attribute exhibits a nearly monotonic power-law-like increasing trend with N, which is similar to the area-based power-law function (Equation 2). However, for A smaller than 1,000 km^2, the mean of N remains at 8 days with minimal fluctuation. Between 1,000 and 100,000 km^2, the partial dependence plot shows a noticeable upward trend, while for A larger than 100,000 km^2, the response to increasing catchment area becomes more rapid, though the trend gradually levels off (Figure 5b). Most of the other catchment attributes also reveal clear trends but with relatively high uncertainty. For instance, the second most important attribute, the maximum daily maximum temperature (TX_x), exhibits five distinct regimes: the values of N remain unchanged for TX_x lower than 32°C, within the range of 34–42°C, and above 44°C. These plateaus are connected by two gradually decreasing trend lines (Figure 5c). Similarly, the maximum daily PE, PE_x, exhibits three regimes with an overall decreasing trend (Figure 5d). In summary, the attributes A, K_{sat}, R_E, and f_P reveal gradually increasing trends, while TX_x, PE_x, P_{95}, W_{sat}, and f_{cly} show decreasing trends. The trends for the remaining catchment attributes are not statistically significant. Detailed explanations for some of the key observed trends are provided in Sections 5.1 and 5.2.

4.3. Evaluation Through the Baseflow Separation Results

The performance of baseflow separation using different N and M parameters is evaluated using KGE (Figure 6). Experiments utilizing RF-predicted $\hat{N}_i$ outperform those using power-function-predicted $\hat{N}_{A,i}$ across M_i, $\tilde{M}$, and $M = 0.9$, with medians of KGE higher by 0.20, 0.15, and 0.18 for the training set and by 0.18, 0.12, and 0.18 for the validation set, respectively. These results indicate a noticeable decline in baseflow separation performance when substituting $\hat{N}_i$ with $\hat{N}_{A,i}$, aligns with the N prediction performance shown in Figure 4. In contrast, the choice of M has a relatively minor influence regardless the choice of N. For the $\hat{N}_i$ cases, employing the median of M ($\tilde{M}$) or default 0.9 instead of the optimal M (M_i) reduces median KGE of 0.04 and 0.07 for the training set and

10.1029/2025WR040578

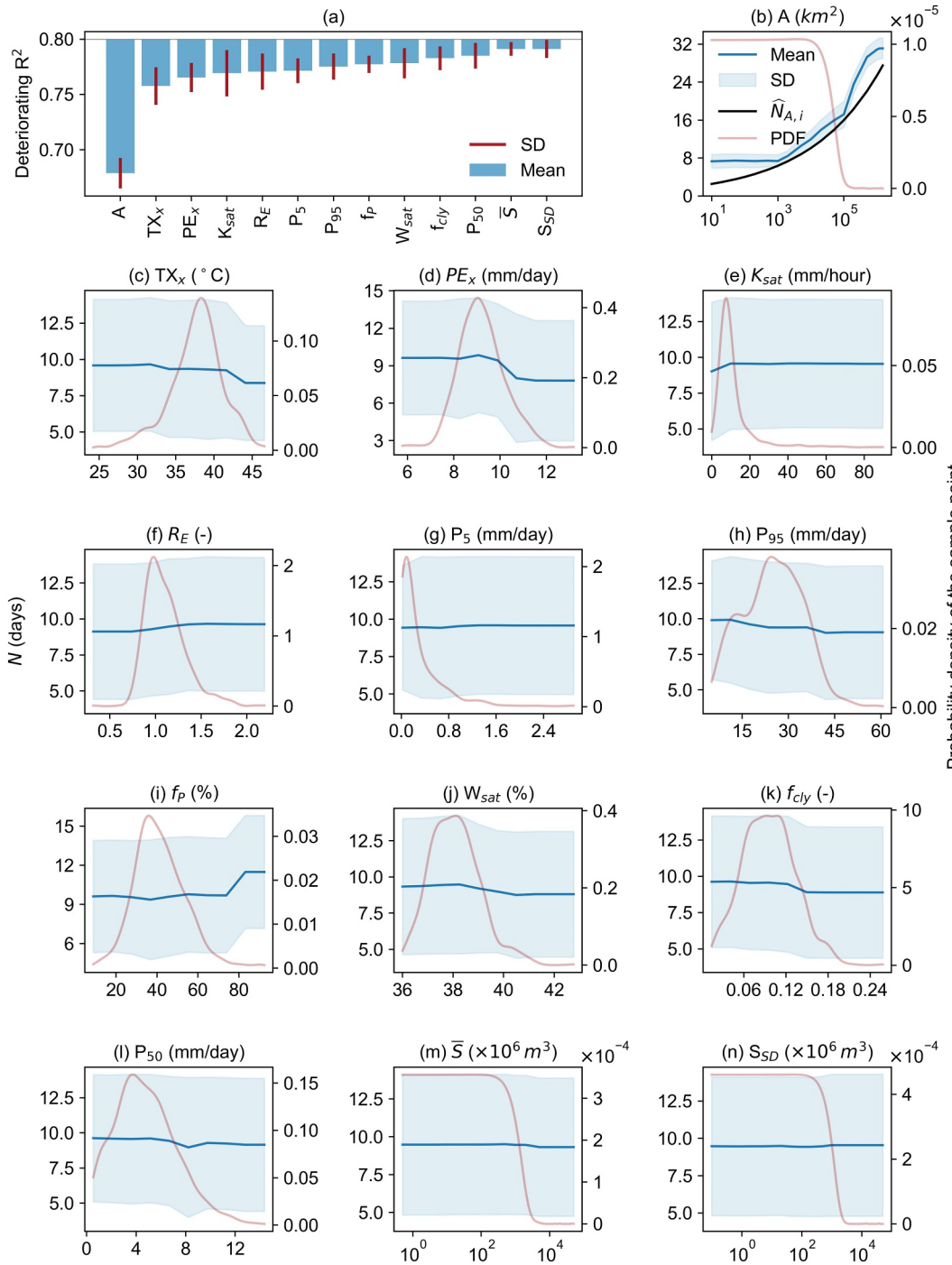


Figure 5. Predictor importance of the 13 selected catchment attributes, measured by the decrease in R^2 (a), and their partial dependence relationships with the optimal N (b–n).

0.04 and 0.08 for the validation set, respectively. For the $\hat{N}_{A,i}$ cases, using $\tilde{M}$ slightly increases median KGE by 0.01 and 0.02 for the training and validation sets, respectively. This confirms the limited influence of M, validating 0.9 as a practical default. Across all six parameterizations, the spatial distribution of KGE reveal no clear pattern (Figure S3 in Supporting Information S1).

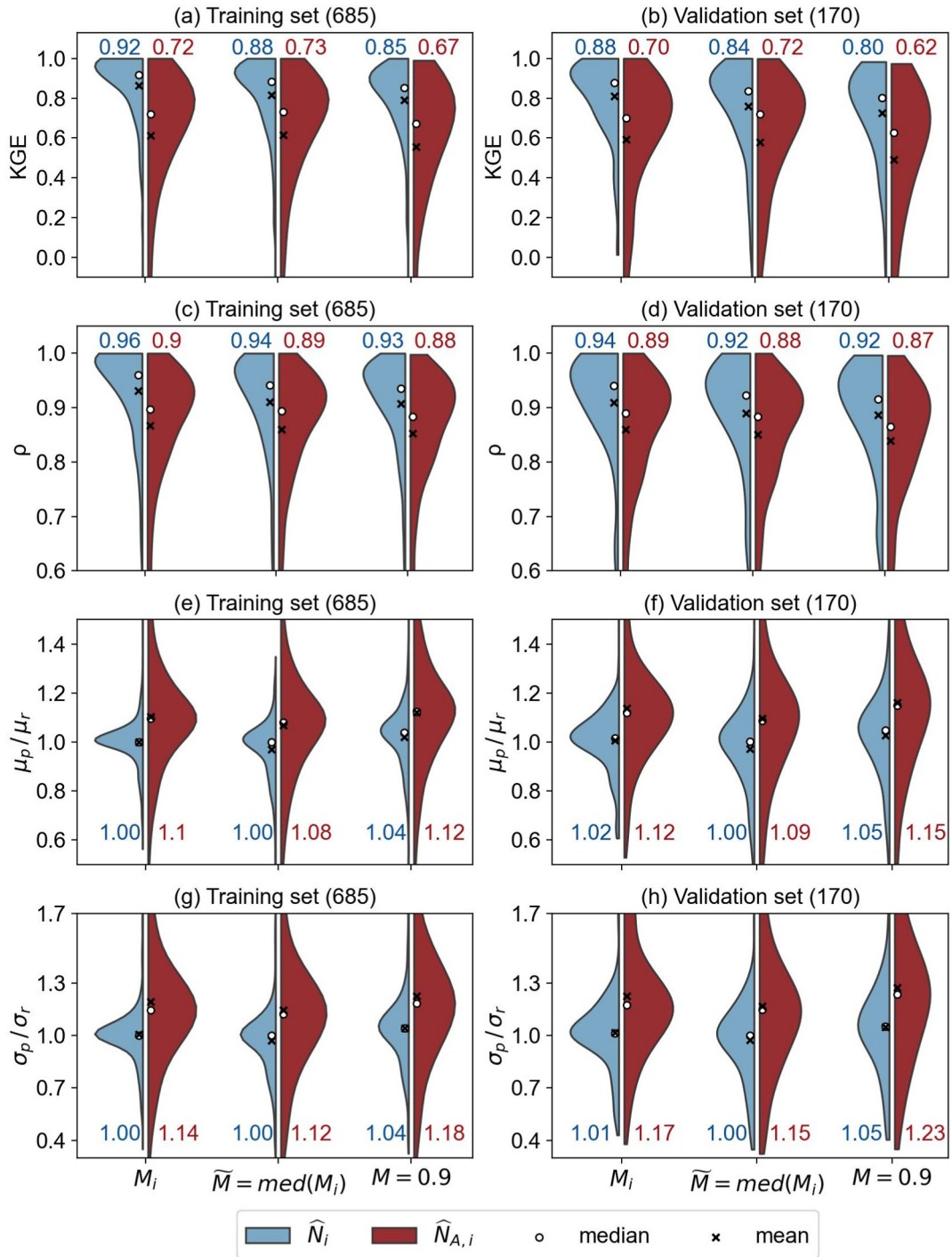


Figure 6. Baseflow separation performance of SMM with different parameterizations measured by KGE (a, b) and its three components: correlation coefficient (c, d), bias ratio (e, f), and variability ratio (g, h). The titles indicate the number of catchments analyzed. The numbers above/below the violins represent the median metric values for each parameterization.

Figures 6c–6h decompose KGE into its components: correlation (ρ), variability ratio ($\frac{\sigma_p}{\sigma_r}$), and bias ratio ($\frac{\mu_p}{\mu_r}$). Both RF-based and power function predictions achieve high ρ. However, they differ markedly in $\frac{\sigma_p}{\sigma_r}$ and $\frac{\mu_p}{\mu_r}$. Power function predictions have median $\frac{\mu_p}{\mu_r}$ values of 1.08–1.15 and median $\frac{\sigma_p}{\sigma_r}$ values of 1.12–1.23, indicating systematic overestimation of both the mean and standard deviation of baseflow time series across most catchments. In

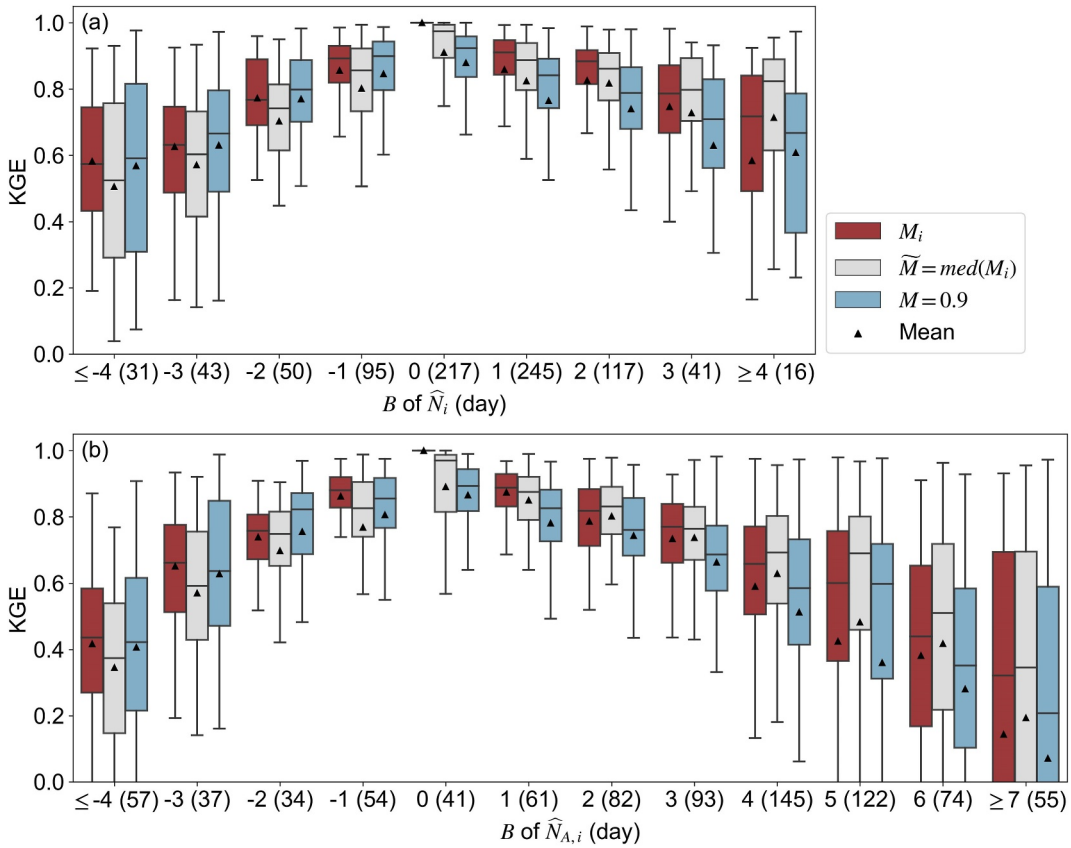


Figure 7. Variation in baseflow separation performance indicated by KGE over different levels of $\hat{N}_i$ bias (a) and $\hat{N}_{A,i}$ bias (b) with different parameterizations of M. The numbers in parentheses on the x-axis tick labels indicate the number of gages within each bin.

contrast, the RF-based predictions yield median $\frac{\sigma_p}{\sigma_r}$ and $\frac{\mu_p}{\mu_r}$ values closer to 1, suggesting substantially better reproduction of both the mean baseflow state and temporal variability. Parameter M is less influential as indicated by the similar performance among different parameterizations for the components, while using the default value results in slightly lower median ρ, $\frac{\sigma_p}{\sigma_r}$, and $\frac{\mu_p}{\mu_r}$.

The baseflow separation performance was also evaluated with respect to the observed SEC across different parameterizations (Figure S4 in Supporting Information S1), revealing a similar pattern as Figures 6a and 6b that SECs estimated based on $\hat{N}_i$ consistently outperform the $\hat{N}_{A,i}$-based ones.

Figure 7 illustrates the performance of the SMM baseflow separation measured by KGE varying for the bias of N. Overall, the use of $\hat{N}_i$ yield better performance than $\hat{N}_{A,i}$ as evidenced by the distribution of sites across different bins of B, aligning with Figure 6. Specifically, 517 sites (65% of the total) fall within the $|B| \le 1$ category for $\hat{N}_i$, compared to only 156 sites (18%) for $\hat{N}_{A,i}$. Additionally, 251 sites (29%) using $\hat{N}_{A,i}$ have $B > 4$ days. Regardless of whether $\hat{N}_i$ or $\hat{N}_{A,i}$ is used, the KGE decreases as the magnitude of B increases. However, the decrease in KGE is less pronounced for positive biases compared to negative biases, and the range of KGE widens with higher magnitudes of B. A detailed examination of the performance within each bias bin reveals variations in the influence of different parameterizations of M. At $0 \le B \le 2$ days, $\hat{N}_i$ paired with M_i yields highest medians of KGE, followed by the $\tilde{M}$ and 0.9 cases. For $B \ge 3$ days, $\tilde{M}$ emerges as the best-performing option, even outperforming M_i. Similar trends are observed for the $\hat{N}_{A,i}$ case with minor differences; M_i performs best for biases in the range $0 \le B \le 1$, while $\tilde{M}$ is optimal for biases in the range $2 \le B \le 4$. However, when B is negative, $\tilde{M}$ yields the lowest median KGEs among the three parameterizations for both $\hat{N}_i$ and $\hat{N}_{A,i}$. In contrast, the default value of 0.9 performs similarly to M_i for biases no smaller than -2 days and even outperforms M_i for more negative biases.

5. Discussion

5.1. In Search of Sufficient Formulas for the Parameter *N*

The parameter N in SMM represents an estimate of streamflow delay and catchment response (Stoelzle et al., 2020). While often default to 5 days, this approach overlooks catchment diversity (Piggott et al., 2005; Price, 2011; Stoelzle et al., 2020). The scale-based power function using drainage area (Equation 2) improve on defaults but remain insufficient (Aksoy et al., 2008; Stoelzle et al., 2020; J. Zhang et al., 2017). While A is identified as the most important prediction, our RF-based prediction framework reveals improved accuracy in estimating N by incorporating additional catchment attributes. This improvement can be attributed to two key factors. First, the delay in streamflow response is not solely controlled by the distance rainfall excess must travel, as reflected by the scale effect captured by A and R_E (Figures 5b and 5f). It is also influenced by the rainfall partitioning into sinks (e.g., evapotranspiration, surface flow, and baseflow), which operate at varying celebrities (McDonnell & Beven, 2014). These processes are reflected in the climate, hydrology, soil, and water usages attributes of the catchment (see Section 5.2 for detailed discussions). Second, the RF-based prediction models offer greater flexibility in incorporating diverse catchment-scale characteristics compared to the power function. This is supported by the nonlinear response of N to A in partial dependence plot, exhibiting an overall increasing trend with three distinct regimes (Figure 5b). For $A < 1000$ km^2, the scale effect is relatively weak and the incorporation of other attributes fills the variation of N; for larger A, the scale effect prevails as spatial integration averages out heterogeneity (H. Li & Sivapalan, 2011; Merz et al., 2009). Together, these factors justify our multivariate, nonparametric approach.

5.2. Sinks of Rainfall Control the Value of *N*

Beyond scale-dependent attributes, additional catchment characteristics were found to be relevant for the prediction of N. These attributes likely influence the average streamflow delay by governing the partitioning of precipitation into distinct pathways with varying turnover time. Based on the long-term catchment water balance, precipitation is primarily partitioned into evapotranspiration and streamflow. Both TX_x and PE_x serve as proxies for the intensity of evapotranspiration processes. Specifically, PE_x reflects the catchment's desiccation potential; higher values indicate greater actual evapotranspiration losses during runoff generation and routing. This is particularly relevant in our study sites, which are predominately located in humid regions where energy, rather than water availability, is the primary limit on evapotranspiration (S. Chen & Ruan, 2023; X. Chen & Sivapalan, 2020). Higher evapotranspiration losses lead to more rapid recession of the streamflow hydrograph, which in turn reduces the streamflow delay and results in smaller N_s (Figure 5d). TX_x functions analogously to PE_x as temperature is positively correlated with potential evapotranspiration (Figure 5c). The predictive importance of these maximum values (TX_x, PE_x) over their means likely stems from their direct quantification of peak summer energy forcing, which governs catchment recession dynamics. While annual means integrate across all seasons, these maxima capture the intense evaporative stress that controls the critical post-spring recession period (Collins et al., 2022; Stoelzle et al., 2013), thereby providing a more informative predictor of the integrated streamflow delay (N) than averages diluted by less influential seasons.

The partial dependence relationship of soil texture and hydraulics attributes—specifically K_{sat}, f_{cly}, and W_{sat}—with N arise from their control over runoff partitioning into surface flow and baseflow paths, governed by differences in flow celerity (Nagy et al., 2024; Stoelzle et al., 2020). Higher K_{sat} indicate greater infiltration rate, directing more rainfall excess into the slower subsurface flow path. This leads to a longer average streamflow delay (larger N) when K_{sat} is relatively small (Figure 5e). In contrast, higher f_{cly} and W_{sat} enhance soil water retention (Addor et al., 2018; Bloomfield et al., 2021; Price, 2011), facilitate a faster flow response as shown in the partial dependence plots (Figures 5j and 5k).

The precipitation-derived attributes f_P, P_5, P_{50}, and P_{95} influence how water is partitioned among evapotranspiration, surface flow, and baseflow. f_P represents catchment wetness, and a higher f_P indicates longer precipitation durations and reduced evapotranspiration losses (X. Chen & Sivapalan, 2020; Lauenroth & Bradford, 2012), leading to prolonged hydrograph recession (larger N, Figure 5i). P_5, P_{50}, and P_{95} were derived from non-zero catchment-averaged precipitation, indicating that P_{95} reflects the extreme precipitation, P_5 is closer to the mean precipitation regime, while P_{50} lies in between. Therefore, P_5 may function similarly as f_P, revealing slight positive correlation with N (Figure 5g). Conversely, higher P_{95} (extreme intensity) increases the likelihood

of infiltration excess, causing a larger proportion of rainfall excess to travel as faster surface flow (X. Chen & Sivapalan, 2020; Loague et al., 2010), pointing to a smaller N (Figures 5h and 5l).

5.3. Limited Influence of Geological Characteristics on the Parameter N

Our feature importance analysis indicates limited predictive significance of geological characteristics for the parameter N. This suggests that the tracer-optimized baseflow component is not a pure proxy for deep groundwater discharge but aligns with the "old water" paradigm, wherein event streamflow comprises substantial pre-event water mobilized diverse subsurface stores (Blöschl et al., 2019; Klaus & McDonnell, 2013). These include not only groundwater discharge but also deeper subsurface flow pathways (e.g., lateral flow above restrictive layers) and bank storage return flow (X. Chen et al., 2006; Kirchner, 2003; Yang et al., 2021). The integration of these pathways, which are less uniquely tied to bedrock geology, likely dilutes the geological signal. This is particularly relevant given that such sources can constitute a larger fraction of streamflow than deep groundwater during high-flow periods (Cartwright, 2022; Stoelzle et al., 2020). Since the parameter N exerts a more pronounced influence on baseflow separation during high-flow periods or precipitation events than during baseflow recession (Aksoy et al., 2008; Stoelzle et al., 2020), the optimized value reflects a wet-period signal dominated by interflow and bank storage return flow (Cartwright, 2022; X. Chen et al., 2006), whose dynamics are governed more by soil properties than by bedrock geology (McGuire et al., 2024). Consequently, soil hydraulic attributes emerge as stronger predictors of N, while the influence of geology is comparatively subdued.

5.4. Limitation and Future Research

While our RF framework provides satisfactory predictions for the parameter N, its compound RF structure and nonparametric nature limits interpretability. This study used permutation feature importance and 1-dimensional PDPs to illustrate the marginal effects of attributes on N, but these techniques do not account for attribute interactions. Furthermore, our PDPs reveal relatively wide standard deviations for most attributes except catchment area (Figure 5). This dominance is largely due to the stronger influence of catchment area compared to other attributes. Future work could explore inherent interpretable techniques, such as symbolic regression, for modeling the relationship between N and catchment attributes, which may improve our understanding of the underlying mechanisms.

Mei, Wang, et al. (2024) reveal that optimal M deviated from the default value of 0.9 and varies across catchments. Consequently, this study attempted to model the optimal M using a RF-based prediction framework similar to that used for N. However, the performance of the M model is unsatisfactory with R^2 of 0.2 (Figure S1 in Supporting Information S1). Despite this, using $M = 0.9$ in conjunction with $\hat{N}_i$ predicted by our RF-based framework still yield acceptable baseflow separation performance (Figure 6). This can be attributed to the relative insensitivity of M to baseflow separation (Aksoy et al., 2008; Stoelzle et al., 2020). Moreover, we found that using $\tilde{M}$ improves performance compared to the default value of 0.9, underscoring the benefits of incorporating network-specific information. Therefore, adopting the median optimal M can serve as a viable alternative when direct prediction of M is not feasible. Overall, the RF-based prediction framework of N and the parameterization of M can be adapted to other regional gage networks with partial tracer data coverage. Future studies could focus on evaluating the applicability of the RF-based prediction framework for such networks and exploring other empirical baseflow filters in addition to SMM.

To support broader applications, an RF-based regionalization model for the SMM parameter N was created in Python and shared on Zenodo along with predicted N values for 855 studied gages (Lin et al., 2025). This data set opens a promising research direction: leveraging the RF-based SMM parameter regionalization model to construct a regional database of optimal N parameter for other CONUS gages that lack reliable environmental tracer measurements. With the optimal N parameter, baseflow separation can be performed for these catchments to generate reliable baseflow estimates. Such a data set would facilitate studies on regional baseflow variations and the response of baseflow to environmental changes.

6. Conclusions

This study develops a parameter regionalization approach for the parameter N (block length) of the smooth minima baseflow separation method, utilizing random forest models and a range of catchment-scale

characteristics. The accuracy of the regionalization model is evaluated by comparing the predicted N values with the specific electrical conductance-optimized N values provided by Mei, Wang, et al. (2024). The importance of the catchment-scale characteristics in predicting N is also assessed. Furthermore, the predicted N values are applied in baseflow separation and the results are evaluated.

The main conclusions are that the RF-based prediction framework demonstrates excellent performance in predicting N with an R^2 of 0.80 and a bias of −0.44 days. Predictor importance analysis reveals that N is primarily controlled by the catchment scale, while other parameters reflecting the allocation of rainfall excess among different water-balance sinks with varying hydrological response times are of secondary importance. For baseflow separation, using the multivariate RF prediction of N improves performance compared to the one using the power function with catchment area, increasing the median KGE for magnitudes of 0.12 and 0.18. Although the prediction of the parameter M is suboptimal ($R^2 = 0.13$), using the default and the median of optimal M across the gages still reasonably captures the baseflow dynamics with median KGEs reaching 0.80 and 0.84, respectively. Future research could explore the application of this prediction model to other regional gage networks to improve baseflow estimation.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

Daily streamflow, baseflow, and the optimal SMM parameters are archived in Mei, Zhu, and Wang (2024). The RF-based regionalization models for parameter N, the 82 catchment attributes, and the predicted N values for the 855 studied gages are available in Lin et al. (2025).

Acknowledgments
This work is supported by the National Key Research and Development (R&D) Program of China (2024YFC3013302), the National Natural Science Foundation of China (52579030), and the Guangdong Natural Science Foundation (2025A1515012264 and 2025A1515011666).

References

Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N., & Clark, M. P. (2018). A ranking of hydrological signatures based on their predictability in space. *Water Resources Research*, *54*(11), 8792–8812. https://doi.org/10.1029/2018wr022606

Aksoy, H., Unal, N. E., & Pektas, A. O. (2008). Smoothed minima baseflow separation tool for perennial and intermittent streams. *Hydrological Processes*, *22*(22), 4467–4476. https://doi.org/10.1002/hyp.7077

Ali, A. M., Imhoff, R. O., & Weerts, A. H. (2025). Machine learning for predicting spatially variable lateral hydraulic conductivity: A step toward efficient hydrological model calibration and global applicability. *Water Resources Research*, *61*(10), e2025WR040108. https://doi.org/10.1029/2025WR040108

Alzraiee, A., Niswonger, R., Luukkonen, C., Larsen, J., Martin, D., Herbert, D., et al. (2024). Next generation public supply water withdrawal estimation for the conterminous United States using machine learning and operational frameworks. *Water Resources Research*, *60*(7), e2023WR036632. https://doi.org/10.1029/2023WR036632

Beck, H. E., van Dijk, A. I. J. M., Miralles, D. G., de Jeu, R. A. M., Sampurno 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

Benavidez, R., Jackson, B., Maxwell, D., & Norton, K. (2018). A review of the (Revised) universal soil loss equation ((R)USLE): With a view to increasing its global applicability and improving soil loss estimates. *Hydrology and Earth System Sciences*, *22*(11), 6059–6086. https://doi.org/10.5194/hess-22-6059-2018

Bhaskar, A. S., Beesley, L., Burns, M. J., Fletcher, T. D., Hamel, P., Oldham, C. E., & Roy, A. H. (2016). Will it rise or will it fall? Managing the complex effects of urbanization on base flow. *Freshwater Science*, *35*(1), 293–310. https://doi.org/10.1086/685084

Bloomfield, J. P., Gong, M., Marchant, B. P., Coxon, G., & Addor, N. (2021). How is baseflow index (BFI) impacted by water resource management practices? *Hydrology and Earth System Sciences*, *25*(10), 5355–5379. https://doi.org/10.5194/hess-25-5355-2021

Blöschl, G., Bierkens, M. F., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., et al. (2019). Twenty-three unsolved problems in hydrology (UPH)—A community perspective. *Hydrological Sciences Journal*, *64*(10), 1141–1158. https://doi.org/10.1080/02626667.2019.1620507

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324

Brookfield, A. E., Hansen, A. T., Sullivan, P. L., Czuba, J. A., Kirk, M. F., Li, L., et al. (2021). Predicting algal blooms: Are we overlooking groundwater? *Science of the Total Environment*, *769*, 144442. https://doi.org/10.1016/j.scitotenv.2020.144442

Brown, T. C., Mahat, V., & Ramirez, J. A. (2019). Adaptation to future water shortages in the United States caused by population growth and climate change. *Earth's Future*, *7*(3), 219–234. https://doi.org/10.1029/2018EF001091

Cartwright, I. (2022). Implications of variations in stream specific conductivity for estimating baseflow using chemical mass balance and calibrated hydrograph techniques. *Hydrology and Earth System Sciences*, *26*(1), 183–195. https://doi.org/10.5194/hess-26-183-2022

Chen, S., & Ruan, X. (2023). A hybrid Budyko-type regression framework for estimating baseflow from climate and catchment attributes. *Journal of Hydrology*, *618*, 129118. https://doi.org/10.1016/j.jhydrol.2023.129118

Chen, X., Chen, D. Y., & Chen, X.-H. (2006). Simulation of baseflow accounting for the effect of bank storage and its implication in baseflow separation. *Journal of Hydrology*, *327*(3), 539–549. https://doi.org/10.1016/j.jhydrol.2005.11.057

Chen, X., & Sivapalan, M. (2020). Hydrological basis of the Budyko Curve: Data-guided exploration of the mediating role of soil moisture. *Water Resources Research*, *56*(10), e2020WR028221. https://doi.org/10.1029/2020WR028221

Collins, M. J., Hodgkins, G. A., Archfield, S. A., & Hirsch, R. M. (2022). The occurrence of large floods in the United States in the modern hydroclimate regime: Seasonality, trends, and large-scale climate associations. *Water Resources Research*, *58*(2), e2021WR030480. https://doi.org/10.1029/2021WR030480

Dai, Y., Xin, Q., Wei, N., Zhang, Y., Shangguan, W., Yuan, H., et al. (2019). A global high-resolution data set of soil hydraulic and thermal properties for land surface modeling. *Journal of Advances in Modeling Earth Systems*, *11*(9), 2996–3023. https://doi.org/10.1029/2019MS001784

Dewitz, J. (2021). National land cover database (NLCD) 2019 Products (ver. 2.0, June 2021). *U.S. Geological Survey data release*. https://doi.org/10.5066/P9KZCM54

Droogers, P., & Allen, R. G. (2002). Estimating reference evapotranspiration under inaccurate data conditions. *Irrigation and Drainage Systems*, *16*(1), 33–45. https://doi.org/10.1023/A:1015508322413

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

Eckhardt, K. (2012). Technical note: Analytical sensitivity analysis of a two parameter recursive digital baseflow separation filter. *Hydrology and Earth System Sciences*, *16*(2), 451–455. https://doi.org/10.5194/hess-16-451-2012

Evans, S. G., Yokeley, B., Stephens, C., & Brewer, B. (2020). Potential mechanistic causes of increased baseflow across northern Eurasia catchments underlain by permafrost. *Hydrological Processes*, *34*(11), 2676–2690. https://doi.org/10.1002/hyp.13759

Flörke, M., Schneider, C., & McDonald, R. I. (2018). Water competition between cities and agriculture driven by climate change and urban growth. *Nature Sustainability*, *1*(1), 51–58. https://doi.org/10.1038/s41893-017-0006-8

Friedl, M., Gray, J., & Sulla-Menashe, D. (2022). MODIS/Terra+Aqua Land Cover Dynamics Yearly L3 Global 500m SIN Grid V061, NASA EOSDIS Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MCD12Q2.061

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Halder, J., Terzer, S., Wassenaar, L. I., Araguás-Araguás, L. J., & Aggarwal, P. K. (2015). The global network of isotopes in rivers (GNIR): Integration of water isotopes in watershed observation and riverine research. *Hydrology and Earth System Sciences*, *19*(8), 3419–3431. https://doi.org/10.5194/hess-19-3419-2015

Hare, D. K., Helton, A. M., Johnson, Z. C., Lane, J. W., & Briggs, M. A. (2021). Continental-scale analysis of shallow and deep groundwater contributions to streams. *Nature Communications*, *12*(1), 1450. https://doi.org/10.1038/s41467-021-21651-0

Hellwig, J., & Stahl, K. (2018). An assessment of trends and potential future changes in groundwater-baseflow drought based on catchment response times. *Hydrology and Earth System Sciences*, *22*(12), 6209–6224. https://doi.org/10.5194/hess-22-6209-2018

Hou, X., Xie, D., Feng, L., Shen, F., & Nienhuis, J. H. (2024). Sustained increase in suspended sediments near global river deltas over the past two decades. *Nature Communications*, *15*(1), 3319. https://doi.org/10.1038/s41467-024-47598-6

Huang, Y., Ciais, P., Santoro, M., Makowski, D., Chave, J., Schepaschenko, D., et al. (2021). A global map of root biomass across the world's forests. *Earth System Science Data*, *13*(9), 4263–4274. https://doi.org/10.5194/essd-13-4263-2021

Humphrey, C. E., Solomon, D. K., Genereux, D. P., Gilmore, T. E., Mittelstet, A. R., Zlotnik, V. A., et al. (2022). Using automated seepage meters to quantify the spatial variability and net flux of groundwater to a stream. *Water Resources Research*, *58*(6), e2021WR030711. https://doi.org/10.1029/2021WR030711

Huscroft, J., Gleeson, T., Hartmann, J., & Börker, J. (2018). Compiling and mapping global permeability of the unconsolidated and consolidated Earth: GLobal HYdrogeology MaPS 2.0 (GLHYMPS 2.0). *Geophysical Research Letters*, *45*(4), 1897–1904. https://doi.org/10.1002/2017GL075860

Ishwaran, H. (2007). Variable importance in binary regression trees and forests. *Electronic Journal of Statistics*, *1*, 519–537. https://doi.org/10.1214/07-ejs039

Ishwaran, H., Kogalur, U. B., Gorodeski, E. Z., Minn, A. J., & Lauer, M. S. (2010). High-dimensional variable selection for survival data. *Journal of the American Statistical Association*, *105*(489), 205–217. https://doi.org/10.1198/jasa.2009.tm08622

Jachens, E. R., Rupp, D. E., Roques, C., & Selker, J. S. (2020). Recession analysis revisited: Impacts of climate on parameter estimation. *Hydrology and Earth System Sciences*, *24*(3), 1159–1170. https://doi.org/10.5194/hess-24-1159-2020

Jasechko, S., Kirchner, J. W., Welker, J. M., & McDonnell, J. J. (2016). Substantial proportion of global streamflow less than three months old. *Nature Geoscience*, *9*(2), 126–129. https://doi.org/10.1038/ngeo2636

Kirchner, J. W. (2003). A double paradox in catchment hydrology and geochemistry. *Hydrological Processes*, *17*(4), 871–874. https://doi.org/10.1002/hyp.5108

Klaus, J., & McDonnell, J. J. (2013). Hydrograph separation using stable isotopes: Review and evaluation. *Journal of Hydrology*, *505*, 47–64. https://doi.org/10.1016/j.jhydrol.2013.09.006

Knox, J. W., Haro-Monteagudo, D., Hess, T. M., & Morris, J. (2018). Identifying trade-offs and reconciling competing demands for water: Integrating agriculture into a robust decision-making framework. *Earth's Future*, *6*(10), 1457–1470. https://doi.org/10.1002/2017EF000741

Kourakos, G., Dahlke, H. E., & Harter, T. (2019). Increasing groundwater availability and seasonal base flow through agricultural managed aquifer recharge in an irrigated basin. *Water Resources Research*, *55*(9), 7464–7492. https://doi.org/10.1029/2018WR024019

Lauenroth, W. K., & Bradford, J. B. (2012). Ecohydrology of dry regions of the United States: Water balance consequences of small precipitation events. *Ecohydrology*, *5*(1), 46–53. https://doi.org/10.1002/eco.195

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94. https://doi.org/10.1029/2008EO100001

Li, H., & Sivapalan, M. (2011). Effect of spatial heterogeneity of runoff generation mechanisms on the scaling behavior of event runoff responses in a natural river basin. *Water Resources Research*, *47*(3). https://doi.org/10.1029/2010WR009712

Li, L., Knapp, J. L. A., Lintern, A., Ng, G. H. C., Perdrial, J., Sullivan, P. L., & Zhi, W. (2024). River water quality shaped by land–river connectivity in a changing climate. *Nature Climate Change*, *14*(3), 225–237. https://doi.org/10.1038/s41558-023-01923-x

Li, L., Maier, H. R., Partington, D., Lambert, M. F., & Simmons, C. T. (2014). Performance assessment and improvement of recursive digital baseflow filters for catchments with different physical characteristics and hydrological inputs. *Environmental Modelling & Software*, *54*, 39–52. https://doi.org/10.1016/j.envsoft.2013.12.011

Lin, Y., Mei, Y., & Wang, D. (2025). Data and code archive for "Regionalization of Optimal Baseflow Separation using Catchment-scale Characteristics" [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.16924118

Liu, Q., Yan, S., Li, M., Ma, X., Liang, L., Zhang, J., & Pan, J. (2021). Baseflow separation and its response to meteorological drought in a temperate water-limited basin, North China. *Chinese Geographical Science*, *31*(5), 867–876. https://doi.org/10.1007/s11769-021-1231-7

Liu, Y., Liu, R., & Chen, J. M. (2012). Retrospective retrieval of long-term consistent global leaf area index (1981–2011) from combined AVHRR and MODIS data. *Journal of Geophysical Research*, *117*(G4). https://doi.org/10.1029/2012JG002084

Loague, K., Heppner, C. S., Ebel, B. A., & VanderKwaak, J. E. (2010). The quixotic search for a comprehensive understanding of hydrologic response at the surface: Horton, Dunne, Dunton, and the role of concept-development simulation. *Hydrological Processes*, *24*(17), 2499–2505. https://doi.org/10.1002/hyp.7834

Longobardi, A., Villani, P., Guida, D., & Cuomo, A. (2016). Hydro-geo-chemical streamflow analysis as a support for digital hydrograph filtering in a small, rainfall dominated, sandstone watershed. *Journal of Hydrology*, *539*, 177–187. https://doi.org/10.1016/j.jhydrol.2016.05.028

McDonnell, J. J., & Beven, K. (2014). Debates—The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph. *Water Resources Research*, *50*(6), 5342–5350. https://doi.org/10.1002/2013WR015141

McGuire, K. J., Klaus, J., & Jackson, C. R. (2024). James buttle review: Interflow, subsurface stormflow and throughflow: A synthesis of field work and modelling. *Hydrological Processes*, *38*(9), e15263. https://doi.org/10.1002/hyp.15263

Mei, Y., Wang, D., Zhu, J., Tang, G., Cai, C., Shen, X., et al. (2024). Optimal baseflow separation through chemical mass balance: Comparing the usages of two tracers, two concentration estimation methods, and four baseflow filters. *Water Resources Research*, *60*(7), e2023WR036386. https://doi.org/10.1029/2023wr036386

Mei, Y., Zhu, J., & Wang, D. (2024). Data and code archive for "Optimal Baseflow Separation through Chemical Mass Balance: Comparing the Usages of Two Tracers, Two Concentration Estimation Methods, and Four Baseflow Filters" [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.8388365

Merz, R., Parajka, J., & Blöschl, G. (2009). Scale effects in conceptual hydrological modeling. *Water Resources Research*, *45*(9). https://doi.org/10.1029/2009wr007872

Nagy, E. D., Szilagyi, J., & Torma, P. (2024). Calibrating the lyne-hollick filter for baseflow separation based on catchment response time. *Journal of Hydrology*, *638*, 131483. https://doi.org/10.1016/j.jhydrol.2024.131483

Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., et al. (2021). What role does hydrological science play in the age of machine learning? *Water Resources Research*, *57*(3), e2020WR028091. https://doi.org/10.1029/2020wr028091

NOHRSC. (2004). In C. U. Boulder (Ed.), *Snow data assimilation system (SNODAS) Data Products at NSIDC, Version 1*. National Snow and Ice Data Center. https://doi.org/10.7265/N5TB14TC

Piggott, A. R., Moin, S., & Southam, C. (2005). A revised approach to the UKIH method for the calculation of baseflow/Une approche améliorée de la méthode de l'UKIH pour le calcul de l'écoulement de base. *Hydrological Sciences Journal*, *50*(5), 920. https://doi.org/10.1623/hysj.2005.50.5.911

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

Putman, A., & Bowen, G. (2019). Technical note: A global database of the stable isotopic ratios of meteoric and terrestrial waters. *Hydrology and Earth System Sciences*, *23*(10), 4389–4396. https://doi.org/10.5194/hess-23-4389-2019

Saxton, K. E., & Rawls, W. J. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Science Society of America Journal*, *70*(5), 1569–1578. https://doi.org/10.2136/sssaj2005.0117

Shangguan, W., Dai, Y., Duan, Q., Liu, B., & Yuan, H. (2014). A global soil data set for Earth system modeling. *Journal of Advances in Modeling Earth Systems*, *6*(1), 249–263. https://doi.org/10.1002/2013MS000293

Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., & Dai, Y. (2017). Mapping the global depth to bedrock for land surface modeling. *Journal of Advances in Modeling Earth Systems*, *9*(1), 65–88. https://doi.org/10.1002/2016MS000686

Shen, X., Anagnostou, E. N., Mei, Y., & Hong, Y. (2017). A global distributed basin morphometric dataset. *Scientific Data*, *4*(1), 160124. https://doi.org/10.1038/sdata.2016.124

Shen, Y., Ruijsch, J., Lu, M., Sutanudjaja, E. H., & Karssenberg, D. (2022). Random forests-based error-correction of streamflow from a large-scale hydrological model: Using model state variables to estimate error terms. *Computers & Geosciences*, *159*, 105019. https://doi.org/10.1016/j.cageo.2021.105019

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

Stewart, M. K. (2015). Promising new baseflow separation and recession analysis methods applied to streamflow at Glendhu Catchment, New Zealand. *Hydrology and Earth System Sciences*, *19*(6), 2587–2603. https://doi.org/10.5194/hess-19-2587-2015

Steyaert, J. C., Condon, L. E., W.D. Turner, S., & Voisin, N. (2022). ResOpsUS, a dataset of historical reservoir operations in the contiguous United States. *Scientific Data*, *9*(1), 34. https://doi.org/10.1038/s41597-022-01134-7

Stoelzle, M., Schuetz, T., Weiler, M., Stahl, K., & Tallaksen, L. M. (2020). Beyond binary baseflow separation: A delayed-flow index for multiple streamflow contributions. *Hydrology and Earth System Sciences*, *24*(2), 849–867. https://doi.org/10.5194/hess-24-849-2020

Stoelzle, M., Stahl, K., & Weiler, M. (2013). Are streamflow recession characteristics really characteristic? *Hydrology and Earth System Sciences*, *17*(2), 817–828. https://doi.org/10.5194/hess-17-817-2013

Swanwick, R. H., Read, Q. D., Guinn, S. M., Williamson, M. A., Hondula, K. L., & Elmore, A. J. (2022). Dasymetric population mapping based on US census data and 30-m gridded estimates of impervious surface. *Scientific Data*, *9*(1), 523. https://doi.org/10.1038/s41597-022-01603-z

Thornton, P. E., Shrestha, R., Thornton, M., Kao, S. C., Wei, Y., & Wilson, B. E. (2021). Gridded daily weather data for North America with comprehensive uncertainty quantification. *Scientific Data*, *8*(1), 190. https://doi.org/10.1038/s41597-021-00973-0

Walsh, R. P. D., & Lawler, D. M. (1981). Rainfall seasonality: Description, spatial patterns and change through time. *Weather*, *36*(7), 201–208. https://doi.org/10.1002/j.1477-8696.1981.tb05400.x

Xie, J., Liu, X., Jasechko, S., Berghuijs, W. R., Wang, K., Liu, C., et al. (2024). Majority of global river flow sustained by groundwater. *Nature Geoscience*, *17*(8), 770–777. https://doi.org/10.1038/s41561-024-01483-5

Xie, Y., Gibbs, H. K., & Lark, T. J. (2021). Landsat-based irrigation dataset (LANID): 30 m resolution maps of irrigation distribution, frequency, and change for the US, 1997-2017. *Earth System Science Data*, *13*(12), 5689–5710. https://doi.org/10.5194/essd-13-5689-2021

Xu, R., Zeng, Z., Pan, M., Ziegler, A. D., Holden, J., Spracklen, D. V., et al. (2023). A global-scale framework for hydropower development incorporating strict environmental constraints. *Nature Water*, *1*(1), 113–122. https://doi.org/10.1038/s44221-022-00004-1

Yang, W., Xiao, C., Zhang, Z., & Liang, X. (2021). Can the two-parameter recursive digital filter baseflow separation method really be calibrated by the conductivity mass balance method? *Hydrology and Earth System Sciences*, *25*(4), 1747–1760. https://doi.org/10.5194/hess-25-1747-2021

Zarnaghsh, A., & Husic, A. (2021). Degree of anthropogenic land disturbance controls fluvial sediment hysteresis. *Environmental Science & Technology*, *55*(20), 13737–13748. https://doi.org/10.1021/acs.est.1c00740

Zell, W. O., & Sanford, W. E. (2020). Calibrated simulation of the long-term average surficial groundwater system and derived spatial distributions of its characteristics for the contiguous United States. *Water Resources Research*, *56*(8), e2019WR026724. https://doi.org/10.1029/2019WR026724

Zhang, C., & Long, D. (2021). Estimating spatially explicit irrigation water use based on remotely sensed evapotranspiration and modeled root zone soil moisture. *Water Resources Research*, *57*(12), e2021WR031382. https://doi.org/10.1029/2021WR031382

Zhang, J., Zhang, Y., Song, J., & Cheng, L. (2017). Evaluating relative merits of four baseflow separation methods in Eastern Australia. *Journal of Hydrology*, *549*, 252–263. https://doi.org/10.1016/j.jhydrol.2017.04.004

Zhang, J., Zhang, Y., Song, J., Cheng, L., Kumar Paul, P., Gan, R., et al. (2020). Large-scale baseflow index prediction using hydrological modelling, linear and multilevel regression approaches. *Journal of Hydrology*, *585*, 124780. https://doi.org/10.1016/j.jhydrol.2020.124780