



Water Resources Research

Supporting Information for

**The Treatment of Uncertainty in Hydrometric Observations:
A Probabilistic Description of Streamflow Records**

Debora Y. de Oliveira¹, and Jasper A. Vrugt^{1,2}

¹Department of Civil and Environmental Engineering, University of California, Irvine, California, USA

²Department of Earth System Science, University of California, Irvine, California, USA

Contents of this file

Texts S1 to S6

Algorithms S1 to S2

Table S1

Figures S1 to S43

Text S1. Variance Estimate Obtained Using First-Order Differencing

Let's consider a n -record of discharge values, $\tilde{\mathbf{y}} = [\tilde{y}_1 \ \tilde{y}_2 \ \dots \ \tilde{y}_n]^\top$. If the discharge record is subject to random errors, the entries of this vector may be written as follows

$$\tilde{\mathbf{y}} = \mathcal{H}(\mathbf{t}) + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}_n(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon), \quad (1)$$

where $\mathcal{H}(t)$ is the data generating process of the actual streamflow at time $t \geq 0$ and the $n \times 1$ vector of errors, $\boldsymbol{\epsilon} = [\epsilon_1 \ \epsilon_2 \ \dots \ \epsilon_n]^\top$, consists of independent variates with zero-mean and variance, $\sigma_{\epsilon_t}^2$, for all $t \in \mathbb{N}_+$. If the errors are independent and identically distributed random variables with zero-mean and nonconstant variance, then we can resort to the nonparametric estimator of Vrugt et al. (2005) to estimate σ_ϵ^2 from the discharge record. This estimator belongs to the class of difference-based variance estimation methods (Hall et al., 1990) and differences the discharge time series, $\tilde{\mathbf{y}}$, k consecutive times to yield a local estimate, $\hat{\sigma}^2$, of the error variance

$$\hat{\sigma}^2 = \binom{2k}{k}^{-1} (\Delta^k(\tilde{y}_t))^2, \quad (2)$$

where $\binom{a}{b} = \frac{a!}{b!(a-b)!}$ is the binomial coefficient and \tilde{y}_t is the t^{th} entry of the n -vector of discharge values, $\tilde{\mathbf{y}}$. If $k = 1$, the difference operator, $\Delta^k(\tilde{y}_t)$, can be written as

$$\Delta^1(\tilde{y}_t) = \tilde{y}_t - \tilde{y}_{t-1} = \begin{cases} \tilde{y}_2 - \tilde{y}_1 \\ \tilde{y}_3 - \tilde{y}_2 \\ \tilde{y}_4 - \tilde{y}_3 \\ \dots \\ \tilde{y}_n - \tilde{y}_{n-1} \end{cases} \quad (3)$$

Substituting \tilde{y}_t with $h_t + \epsilon_t$ (Eq. 1),

$$\Delta^1(\tilde{y}_t) = \begin{cases} (h_2 + \epsilon_2) - (h_1 + \epsilon_1) \\ (h_3 + \epsilon_3) - (h_2 + \epsilon_2) \\ (h_4 + \epsilon_4) - (h_3 + \epsilon_3) \\ \dots \\ (h_n + \epsilon_n) - (h_{n-1} + \epsilon_{n-1}) \end{cases} \quad (4)$$

:

If the time series is smooth enough, $h_t \approx h_{t-1}$,

$$\Delta^1(\tilde{y}_t) = \begin{cases} (\cancel{h_2} + \epsilon_2) - (\cancel{h_1} + \epsilon_1) = \epsilon_2 - \epsilon_1 \\ (\cancel{h_3} + \epsilon_3) - (\cancel{h_2} + \epsilon_2) = \epsilon_3 - \epsilon_2 \\ (\cancel{h_4} + \epsilon_4) - (\cancel{h_3} + \epsilon_3) = \epsilon_4 - \epsilon_3 \\ \dots \\ (\cancel{h_n} + \epsilon_n) - (\cancel{h_{n-1}} + \epsilon_{n-1}) = \epsilon_n - \epsilon_{n-1} \end{cases} \quad (5)$$

Finally,

$$\hat{\sigma}^2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}^{-1} (\Delta^1(\tilde{y}_t))^2 = \begin{cases} \frac{1}{2}\epsilon_2^2 - \epsilon_2\epsilon_1 + \frac{1}{2}\epsilon_1^2 \\ \frac{1}{2}\epsilon_3^2 - \epsilon_3\epsilon_2 + \frac{1}{2}\epsilon_2^2 \\ \frac{1}{2}\epsilon_4^2 - \epsilon_4\epsilon_3 + \frac{1}{2}\epsilon_3^2 \\ \dots \\ \frac{1}{2}\epsilon_n^2 - \epsilon_n\epsilon_{n-1} + \frac{1}{2}\epsilon_{n-1}^2 \end{cases} \quad (6)$$

given that $\begin{pmatrix} 2 \\ 1 \end{pmatrix}^{-1} = \frac{1}{2}$ and $(\Delta^1(\tilde{y}_t))^2$ equals

$$(\Delta^k(\tilde{y}_t))^2 = \begin{cases} (\epsilon_2 - \epsilon_1)^2 = \epsilon_2^2 - 2\epsilon_2\epsilon_1 + \epsilon_1^2 \\ (\epsilon_3 - \epsilon_2)^2 = \epsilon_3^2 - 2\epsilon_3\epsilon_2 + \epsilon_2^2 \\ (\epsilon_4 - \epsilon_3)^2 = \epsilon_4^2 - 2\epsilon_4\epsilon_3 + \epsilon_3^2 \\ \dots \\ (\epsilon_n - \epsilon_{n-1})^2 = \epsilon_n^2 - 2\epsilon_n\epsilon_{n-1} + \epsilon_{n-1}^2 \end{cases} \quad (7)$$

In case of homoscedastic errors, the variance estimate is the average of all entries of the vector defined in Eq. 6,

$$\hat{\sigma}^2 = \frac{1}{n-1} \left(\frac{1}{2}\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2 + \dots + \epsilon_{n-1}^2 + \frac{1}{2}\epsilon_n^2 - (\epsilon_2\epsilon_1 + \epsilon_3\epsilon_2 + \dots + \epsilon_n\epsilon_{n-1}) \right). \quad (8)$$

The true sample variance, s^2 , is

$$s^2 = \frac{\sum_{t=1}^n (\epsilon_t - \bar{\epsilon})^2}{n-1} = \frac{\sum_{t=1}^n \epsilon_t^2}{n-1}. \quad (9)$$

Thus, the approximation error amounts to

$$s^2 - \hat{\sigma}^2 = \frac{1}{n-1} \left(\frac{1}{2}\epsilon_1^2 + \frac{1}{2}\epsilon_n^2 + (\epsilon_2\epsilon_1 + \epsilon_3\epsilon_2 + \dots + \epsilon_n\epsilon_{n-1}) \right). \quad (10)$$

Text S2. Alternative Implementation of the Nonparametric Estimator

If we group the coefficients of the k^{th} -order difference operator in a $(k + 1)$ -vector, \mathbf{d} , then we yield, $\mathbf{d} = [1 \ -3 \ 3 \ -1]^{\top}$, for $k = 3$. If we divide each entry of \mathbf{d} by the Euclidean norm, $E(\mathbf{d}) = \sqrt{\mathbf{d}^{\top} \mathbf{d}}$, of this so-called difference sequence

$$d'_i = d_i / E(\mathbf{d}) \quad (11)$$

then the normalized difference sequence, $\mathbf{d}' = [d'_1 \ d'_2 \ d'_3 \ \dots \ d'_{k+1}]^{\top}$, satisfies the following two conditions (Hall et al., 1990)

$$\sum_{i=1}^{k+1} d'_i = 0 \quad \text{and} \quad \sum_{i=1}^{k+1} d'^2_i = 1 \quad (12)$$

and the hourly discharge error variances may be written without the difference operator, $\Delta^k(\cdot)$, in Equation (3) (of the main manuscript) as follows (Hall et al., 1990; Zhou et al., 2015)

$$\widehat{\sigma}_h^2 = \frac{1}{n-k} \sum_{j=0}^{n-k-1} \left(\sum_{i=1}^{k+1} d'_i \tilde{y}_{i+jh} \right)^2, \quad (13)$$

where \tilde{y}_{i+jh} is the $(i + j)^{\text{th}}$ element of the hourly discharge record, $\tilde{\mathbf{y}}_h$.

The above expression returns an estimate of the error variance for the entire discharge record. This will negate characterization of error heteroscedasticity. A simple solution is to compute local estimates, $\widehat{\sigma}_h^2$, of the hourly error variance

$$\widehat{\sigma}_h^2 = \left(\sum_{i=1}^{k+1} d'_i \tilde{y}_{i+jh} \right)^2, \quad (14)$$

and subsequently investigate the relationship between $\widehat{\sigma}_h^2$ and streamflow magnitude. Indeed, if we embed the normalized difference sequence \mathbf{d}' into a $(n - k) \times n$ matrix \mathbf{D} as follows

$$\mathbf{D} = \begin{bmatrix} d'_1 & d'_2 & d'_3 & \cdots & d'_{m+1} & 0 & \cdots & \cdots & 0 \\ 0 & d'_1 & d'_2 & d'_3 & \cdots & d'_{m+1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & d'_1 & d'_2 & d'_3 & \cdots & d'_{m+1} & 0 \\ 0 & \cdots & \cdots & 0 & d'_1 & d'_2 & d'_3 & \cdots & d'_{m+1} \end{bmatrix}, \quad (15)$$

then the $(n - k) \times 1$ vector of hourly discharge error variances, $\widehat{\sigma}_h^2$ (mm^2/d^2), can be written as the element-wise product, \odot , of the matrix-vector product, $\mathbf{D}\tilde{\mathbf{y}}_h$, as follows

$$\widehat{\sigma}_h^2 = \mathbf{D}\tilde{\mathbf{y}}_h \odot \mathbf{D}\tilde{\mathbf{y}}_h. \quad (16)$$

The Schur product simply squares each element of the $(n - k) \times 1$ vector, $\mathbf{D}\tilde{\mathbf{y}}_h$ to yield the $(n - k) \times 1$ vector of error variance estimates. The corresponding discharge values, \bar{y}_h (mm/d), may be computed using

$$\bar{\mathbf{y}}_h = \mathbf{W}\tilde{\mathbf{y}}_h \quad (17)$$

where the $(n - k) \times n$ weight matrix \mathbf{W} is of the form of Equation (15) except with each entry of d replaced by weights, $w = 1/(k + 1)$. A scatter plot of \bar{y}_h and $\hat{\sigma}_h$ will now reveal the the nature of the random errors of the discharge record.

Text S3. Generation of Simulated and Corrupted Streamflow Time Series

In the first case study, we focus our attention on five watersheds with contrasting hydrologic regimes according to the catchment classification scheme of Brunner et al. (2020). This includes the (i) Leaf River near Collins, MS (strong winter regime), (ii) Cowhouse Creek at Pidcoke, TX (intermittent regime), (iii) Potecasi Creek near Union, NC (weak winter regime), (iv) South Fork Shoshone River near Valley, WY (melt regime) and (v) Nehalem River near Foss, OR (New Year's regime).

For each of the five selected CAMELS watersheds, we simulate a multi-year record, $\mathbf{y}_h = [y_{1h} \ y_{2h} \ \dots \ y_{nh}]^\top$, of hourly discharge by evaluating the Sacramento soil moisture accounting (SAC-SMA) model with calibrated parameter values using measured hourly rainfall rates from NLDAS (Gauch et al., 2020) and hourly estimates of potential evapotranspiration obtained from the formula of Oudin et al. (2005). We use the SAC-SMA model implementation described in Clark et al. (2008) and simulate hourly discharges using a second-order mass-conservative numerical method with adaptive time stepping. Absolute and relative tolerances were fixed at 10^{-3} . This promotes numerical stability and continuity and guarantees that \mathbf{y}_h is without numerical artifacts (Kavetski & Clark, 2010; Clark & Kavetski, 2010; Schoups et al., 2010).

We corrupt the simulated discharge record of each CAMELS watershed with a time series, $\boldsymbol{\epsilon}^*$, of uncorrelated normal variates

$$\tilde{\mathbf{y}}_h = \mathbf{y}_h + \boldsymbol{\epsilon}^*, \quad \boldsymbol{\epsilon}^* \sim \mathcal{N}_n(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon), \quad (18)$$

using the $n \times n$ covariance matrix, $\boldsymbol{\Sigma}_\epsilon$, with error variances, σ_{th}^2 , on the main diagonal equal to

$$\sigma_{th}^2 = (\alpha y_{th} + \beta m_h)^2, \quad (19)$$

where $m_h = \frac{1}{n} \sum_{t=1}^n y_{th}$ (mm/d) is the arithmetic mean of the SAC-SMA simulated hourly discharge record, $\alpha, \beta \geq 0$ are non-negative dimensionless coefficients that determine the nature and magnitude of the errors and $t = (1, 2, \dots, n)$. The use of the mean simulated discharge in the intercept, βm_h , of the error function of Equation (19) warrants the application of a common β value to catchments with widely different flow magnitudes. For $\alpha = 0$, the n -vector of errors will have a constant variance and for $\alpha > 0$ the magnitude of the errors will increase with simulated discharge.

Text S4. Experimental Data

We illustrate our method by application to catchments from the Catchment Attributes and MEteorology for Large-sample Studies (CAMELS) data set (Newman et al., 2015; Addor et al., 2017a). The CAMELS data set provides meteorological data at a daily time step from three different sources including Daymet, Maurer, and NLDAS and daily streamflow time series from the United States Geological Survey (USGS) for 671 U.S. catchments (Newman et al., 2015). This data set was further extended in Addor et al. (2017a) by the addition of catchment attributes divided into six main classes: topography, climate, streamflow, land cover, soil, and geology. In this paper, the analysis was conducted using daily streamflow time series from the entire period available, which for most catchments correspond to the period from 01 October 1980 to 30 September 2014 (34 years). Discharge values were converted from cubic feet per second (cfs) to millimeters per day (mm/d) by using the catchment area from GAGES II (Falcone, 2011), available in the CAMELS data set. We only included in our analysis catchments without any missing streamflow data within the period available in the CAMELS data set, since the proposed replicate generation procedure in its current form (described in section 2.3) requires a complete time series. This first prerequisite leaves us with 621 out of the 671 catchments.

The method of Vrugt et al. (2005) used in this paper to estimate random errors (described in section 2.1) requires streamflow data at a high temporal resolution, such that the sampling frequency is high compared to the typical time scale of the streamflow variation. Preliminary analysis conducted using synthetic daily discharge time series corrupted with different levels of errors revealed that a daily time step can be too coarse to provide an accurate estimation of the error, especially for low levels of errors (results not shown). Therefore, for the estimation of random errors in discharge records we resort to the hourly streamflow time series made available by Gauch et al. (2020). The data set of Gauch et al. (2020) contains hourly time series for a total of 516 CAMELS catchments, starting as early as 1956 to 2007 and extend until May 2020, date at which the data was retrieved from the USGS Water Information System. The proportion of missing data in the hourly time series varies between 0 and 60%, with a median value of 8%. After discarding these missing data, we still have left more than 10 years of (nonconsecutive) hourly data for each catchment (median of 28 years), which was judged enough to infer the error model of each watershed.

We need to make two considerations about the use of the hourly time series provided by Gauch et al. (2020). First, the data set of Gauch et al. (2020) presents streamflow values in millimeters per hour (mm/h), which was converted from cfs using catchment area as published by the

USGS. To make them comparable with the daily time series, we rescaled each hourly time series using the catchment area from GAGES II. Second, the hourly time series provided by Gauch et al. (2020) represent averages of instantaneous streamflow values estimated from stage measurements taken at sub-daily intervals. These instantaneous measurements are currently made by USGS at 15-minute intervals in most gauges, but measurement resolution can vary from 5 to 60 minutes. In some gauges, stage measurements are recorded at variable time intervals. In addition, measurement resolution has varied over the years, being most recent measurements recorded at a higher frequency. For simplicity, we assumed that the hourly discharge data correspond to instantaneous values for all gauges and during the entire 1980–2014 period analyzed here.

We only considered catchments without any missing daily streamflow value, as mentioned previously, and for which the corresponding hourly time series were also available. These two criteria combined resulted in a total of 504 catchments being included in our analysis. We follow the functional classification of Brunner et al. (2020) and split the 504 catchments into the five proposed regime classes: intermittent regime (122 catchments); strong winter regime (176 catchments); weak winter regime (108 catchments); melt regime (44 catchments); and New Year's regime (54 catchments). One catchment of each regime is selected to illustrate our findings, as follows: intermittent regime – Cowhouse Creek at Pidcoke, TX (USGS 08101000); strong winter regime – Leaf River near Collins, MS (USGS 02472000); weak winter regime – Potecasi Creek near Union, NC (USGS 02053200); melt regime – South Fork Shoshone River near Valley, WY (USGS 06280300); and New Year's regime – Nehalem River near Foss, OR (USGS 14301000).

Text S5. Derivation of the Daily Variance Estimate

The daily discharge is computed as the average of the $u = 24$ hourly discharge values. Therefore, we resort to the variance of the sample mean to estimate the variance of the daily discharge. Under the assumption that the u measured values \tilde{y}_i , where $i = (1, 2, \dots, u)$, are identically distributed and measured with no error, the variance of the mean, \bar{y} , can be expressed as follows

$$\begin{aligned}
\text{Var}[\bar{y}] &= \text{Var}\left[\frac{1}{u} \sum_{i=1}^u \tilde{y}_{ih}\right] \\
&= \frac{1}{u^2} \text{Var}\left[\sum_{i=1}^u \tilde{y}_{ih}\right] \\
&= \frac{1}{u^2} [\sigma^2 + \sigma^2 + \dots + \sigma^2] \\
&= \frac{1}{u^2} [u\sigma^2] \\
&= \frac{1}{u} \sigma^2 \\
&\approx \frac{1}{u} s^2,
\end{aligned} \tag{20}$$

where the approximation sign indicates that the quantity $\frac{1}{u} s^2$ is an estimate of the true variance of the mean, since we use in its computation the sample variance, s^2 , instead of the true (unknown) variance, σ^2 .

Now relaxing the assumption that the hourly discharge values are measured with no error, we can estimate the variance of the mean as follows

$$\begin{aligned}
\text{Var}[\bar{y}] &= \text{Var}\left[\frac{1}{u} \sum_{i=1}^u (y_{ih} + \epsilon_{ih})\right] \\
&= \frac{1}{u^2} \text{Var}\left[\sum_{i=1}^u (y_{ih} + \epsilon_{ih})\right],
\end{aligned} \tag{21}$$

where ϵ_{ih} denotes the error associated with the hourly discharge value y_{ih} . If y_{ih} and ϵ_{ih} are mutually independent, the combined variance is the sum of individual variances,

$$\begin{aligned}
\text{Var}[\bar{y}] &= \frac{1}{u^2} \text{Var} \left[\sum_{i=1}^u \tilde{y}_{ih} \right] + \frac{1}{u^2} \text{Var} \left[\sum_{i=1}^u \epsilon_{ih} \right] \\
&= \frac{1}{u^2} [\sigma^2 + \sigma^2 + \dots + \sigma^2] + \frac{1}{u^2} [\sigma_{1h}^2 + \sigma_{2h}^2 + \dots + \sigma_{24h}^2] \\
&= \frac{1}{u^2} [u\sigma^2] + \frac{1}{u^2} \sum_{i=1}^u \sigma_{ih}^2 \\
&= \frac{1}{u} \sigma^2 + \frac{1}{u^2} \sum_{i=1}^u \sigma_{ih}^2 \\
&\approx \frac{1}{u} s^2 + \frac{1}{u^2} \sum_{i=1}^u s_{ih}^2 \\
&\approx \frac{1}{u} \frac{1}{u-1} \sum_{i=1}^u \left(\tilde{y}_{ih} - \frac{1}{u} \sum_{j=1}^u \tilde{y}_{jh} \right)^2 + \frac{1}{u^2} \sum_{i=1}^u s_{ih}^2,
\end{aligned} \tag{22}$$

where s_{ih}^2 denotes the variance of the $u = 24$ hourly discharge values. This concludes our derivation.

Note the similarity between Eq. 22 and Eq. 7 of Raftery et al. (2005). Indeed, our equation would simplify to Eq. 7 of Raftery et al. (2005) if (i) each of the $u = 24$ hourly discharge values were considered to be representative of the daily discharge (instead of the daily discharge being computed as the average of the $u = 24$ hourly discharge values), and (ii) the variance of the $u = 24$ hourly discharge values was constant. Under the first condition, the term $\frac{1}{u}$ would be removed from Eq. 22, and, if the second condition were true, the second term of Eq. 22 would further reduce to s_h^2 .

Text S6. Autocorrelation

To induce autocorrelation between the n entries of the perturbation time series, we revise the covariance matrix, Σ_ϵ , to a covariance matrix of the perturbations, Σ_p , written as a product of the $n \times n$ correlation matrix of an AR(k) process, $\mathbf{R}_{\tilde{y}}$, and the $n \times n$ diagonal matrix of the daily error variances as follows

$$\Sigma_p = \mathbf{R}_{\tilde{y}} \text{diag}(\boldsymbol{\sigma}_\epsilon^2). \quad (23)$$

The main diagonal elements of Σ_p list the variances of the discharge errors, $\boldsymbol{\epsilon} = [\epsilon_1 \dots \epsilon_n]^\top$, and $\mathbf{R}_{\tilde{y}}$ may be written as (Box et al., 2015)

$$\mathbf{R}_{\tilde{y}} = \begin{bmatrix} 1 & r(1) & r(2) & \cdots & r(n-1) \\ r(1) & 1 & r(1) & \cdots & r(n-2) \\ r(2) & r(1) & 1 & \cdots & r(n-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r(n-1) & r(n-2) & r(n-3) & \cdots & 1 \end{bmatrix}. \quad (24)$$

The entries of $\mathbf{R}_{\tilde{y}}$ can be derived from the sample autocorrelation function of the discharge record, $\tilde{\mathbf{y}} = [\tilde{y}_1 \tilde{y}_2 \dots \tilde{y}_n]^\top$. The sample autocorrelation, $\hat{\rho}_{\tilde{y}}(\tau)$, for two streamflow observations, \tilde{y}_i and \tilde{y}_j , a distance (time), $\tau = |i - j|$, apart may be computed using (Box et al., 2015)

$$\begin{aligned} \hat{\rho}_{\tilde{y}}(\tau) &= \frac{\text{Cov}[\tilde{y}_i, \tilde{y}_j]}{\text{Var}[\tilde{y}_i]} = \frac{\hat{\gamma}_{\tilde{y}}(\tau)}{\hat{\gamma}_{\tilde{y}}(0)} = \frac{\frac{1}{n-\tau} \sum_{i=\tau+1}^n (\tilde{y}_i - m_{\tilde{y}})(\tilde{y}_{i-\tau} - m_{\tilde{y}})}{\frac{1}{n-\tau} \sum_{i=\tau+1}^n (\tilde{y}_i - m_{\tilde{y}})^2} \\ &= \frac{\sum_{i=\tau+1}^n (\tilde{y}_i - m_{\tilde{y}})(\tilde{y}_{i-\tau} - m_{\tilde{y}})}{\sum_{i=\tau+1}^n (\tilde{y}_i - m_{\tilde{y}})^2}, \end{aligned} \quad (25)$$

where $m_{\tilde{y}} = \frac{1}{n} \sum_{t=1}^n \tilde{y}_t$ (mm/T) denotes the mean of the n -record of streamflow observations and $\tau \geq 0$. The values of $\hat{\rho}_{\tilde{y}}(\tau)$ for $\tau \in (1, 1, \dots, n-1)$ equal the entries, $r(\tau)$, of the correlation matrix, $\mathbf{R}_{\tilde{y}}$, in Equation (24). Experiments carried out on the CAMELS data set confirmed (not shown) that the use of a normal quantile (NQ) transform of the discharge data improves hydrologic characterization. Therefore, prior to computation of the correlation matrix, $\mathbf{R}_{\tilde{y}}$, the discharge values are replaced by their respective variates of a standard normal distribution.

Algorithm S1. MATLAB implementation of the nonparametric error estimator.

```

function [c,Y_sig,tab] = error_estimation(Y,k,tol,method,m)
%% %%%%%% Function that estimates the error of a time series of data %%
%% Requirements
%% [1] the underlying data-generating function, h(t), is sufficiently smooth %%
%% [2] the sampling interval is high compared to the time-scale of h(t) %%
%% [3] the errors exhibit a constant or heteroscedastic variance %%
%%
%% SYNOPSIS      [c,Y_sig,tab] = error_estimation(Y);
%%                  [c,Y_sig,tab] = error_estimation(Y,k,tol,method,m);
%%
%% INPUTS
%% Y             [required]: n x 1 vector with values of measured signal %%
%% k             [optional]: difference operator applied k times (default k: 3) %%
%% tol           [optional]: value below which estimated errors are discarded %%
%% method        [optional]: 1 use all data                                (default method: 2) %%
%%                      2 use average estimates of sigma                %%
%% m             [optional]: tab window size to left and right (default m: 100) %%
%%
%% OUTPUTS       c: vector with slope and intercept
%%               Y_sig: n-k x 2 matrix with data versus the error sigma
%%               tab: n-k x 2 matrix with sliding average values of sigma
%% %%%%%%
if nargin < 2, k = 3; end          % how many times does one want to difference?
if nargin < 3, tol = 0; end         % remove errors smaller than tol
if nargin < 4, method = 2; end       % fitting method
if nargin < 5, m = 100; end         % window size of moving average

%% Matrix form (Zhou et al., 2015) - see Supporting Information (SI) file
N = numel(Y);
% get d entries
d_pas = pascal(k+1,1); d = d_pas(k+1,1:k+1);
% normalize d entries
d_norm = d./norm(d);           % Eq. (11) of the SI
% Note: sum(d_norm) = 0; and sum(d_norm.^2) = 1
D = zeros(N-k,N); A = D; wghts = 1/(k+1) * ones(1,k+1);
for i = 1:N-k
    D(i,i:i+k) = d_norm;      % entries of D (Eq. (15) of the SI)
    A(i,i:i+k) = wghts;       % to get mean discharge
end
% Now we yield sigma2 estimate
sigma2_est = (D*Y(1:N)).^2; % Eq. (16) of the SI
% Compute mean values of discharge
Y_m = A*Y;                   % Eq. (17) of the SI

%% Estimate slope and intercept
if tol > 0
    % Remove small errors
    idx = sqrt(sigma2_est) > 1e-10; Y_m = Y_m(idx); sigma2_est = sigma2_est(idx);
end
% Prepare return argument - sort first column for moving average in tab
Y_sig = sortrows([ Y_m sqrt(sigma2_est) ],1);
if method == 1
    % Get slope and intercept estimates using all data
    c = polyfit(Y_m,sqrt(sigma2_est),1); % Eq. (14) (all data)
    tab = [];
elseif method == 2
    % Now use moving average with window to get average estimates of sigma
    tab = [ Y_sig(:,1) sqrt(movmean(Y_sig(:,2).^2,[m m])) ];
    tab(end-m+1:end,:) = []; % discard endpoints
    tab(1:m,:) = []; % discard endpoints
    idx = find(isnan(tab(:,2))); tab(idx,:) = [];
    % Now fit line to tabular (averaged) data
    c = polyfit(tab(:,1),tab(:,2),1); % Eq. (14) (moving window)
end

```

Algorithm S2. MATLAB implementation of the replicate generation procedure.

```

function [Yr] = replicate_generation(Y,N,std_Y)
%% Function that generates replicates of time series of data
%%
%% SYNOPSIS      [Yr] = replicate_generation(Y,N,std_Y);
%%
%% INPUTS
%% Y             [required]: n x 1 vector with values of measured signal
%% N             [required]: number of replicates
%% std_Y         [required]: error standard deviation of measured signal
%%
%% OUTPUTS        Yr: n x N matrix with replicates
%%

% How many data points
n = numel(Y);

% Compute empirical probability
% Plotting position can be 'Weibull' or 'Gringorten'
ep = empdis(Y,'Weibull');

% Transformed data, Yt
Yt = norminv(ep);

% Now determine autocorrelation function
[rho] = autocorr(Yt,n-1);

% Create correlation matrix from ACF values
R = toeplitz(rho);

% Now compute covariance matrix from std_Y and correlation matrix
C = corr2cov(std_Y, R);

% Initialize storage matrix
Yr = nan(n,N);
% Now create realizations
for ii = 1:N
    if ii==1
        [r,T] = mvnrnd(zeros(n,1),C);
        Yr(1:n,ii) = Y + r';
    else
        Yr(1:n,ii) = Y + mvnrnd(zeros(n,1),C,[],T)';
    end
    disp(['num2str(ii)' ' of ' num2str(N)])
end

%% Compute the empirical probability
function p = empdis(y,method)

n = length(y);
bp = zeros(n,1);

for i=1:n
    bp(i,1) = sum(y(:,1)<=y(i,1));
end

switch method
    case 'Gringorten'
        p = (bp-0.44)./(n+0.12);
    case 'Weibull'
        p = bp./(n+1);
end

```

Table S1: Summary of the main characteristics of a sample of five watersheds from the CAMELS data base. We list mean values of rainfall, potential evapotranspiration (PE) and discharge for the measurement period of 1 October 1989 to 30 September 2009 (Addor et al., 2017b).

USGS gauge	Regime	Area km ²	Rainfall mm/d	PE mm/d	Discharge mm/d
Leaf River near Collins, MS	strong winter	1927	4.16	3.17	1.33
Cowhouse Creek at Pidcoke, TX	intermittent	1177	2.40	3.30	0.27
Potecasi Creek near Union, NC	weak winter	584	3.33	2.80	0.86
South Fork Shoshone River near Valley, WY	melt	794	1.95	2.54	1.18
Nehalem River near Foss, OR	New Year's	1744	5.82	2.84	3.62

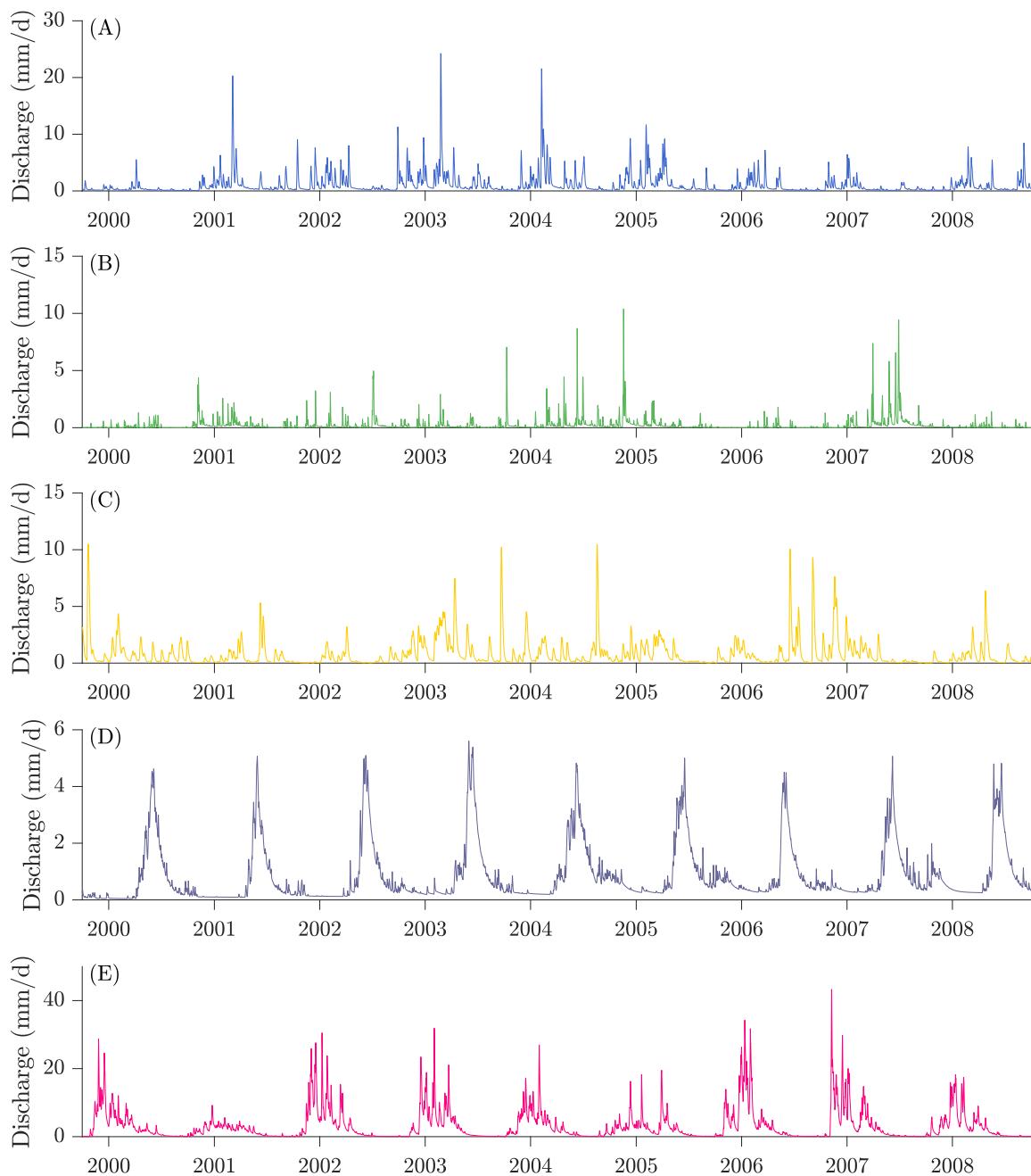


Figure S1: Simulated discharge record of (a) Leaf River near Collins, MS (strong winter regime), (b) Cowhouse Creek at Pidcock, TX (intermittent regime), (c) Potecasi Creek near Union, NC (weak winter regime), (d) South Fork Shoshone River near Valley, WY (melt regime) and (e) Nehalem River near Foss, OR (New Year's regime).

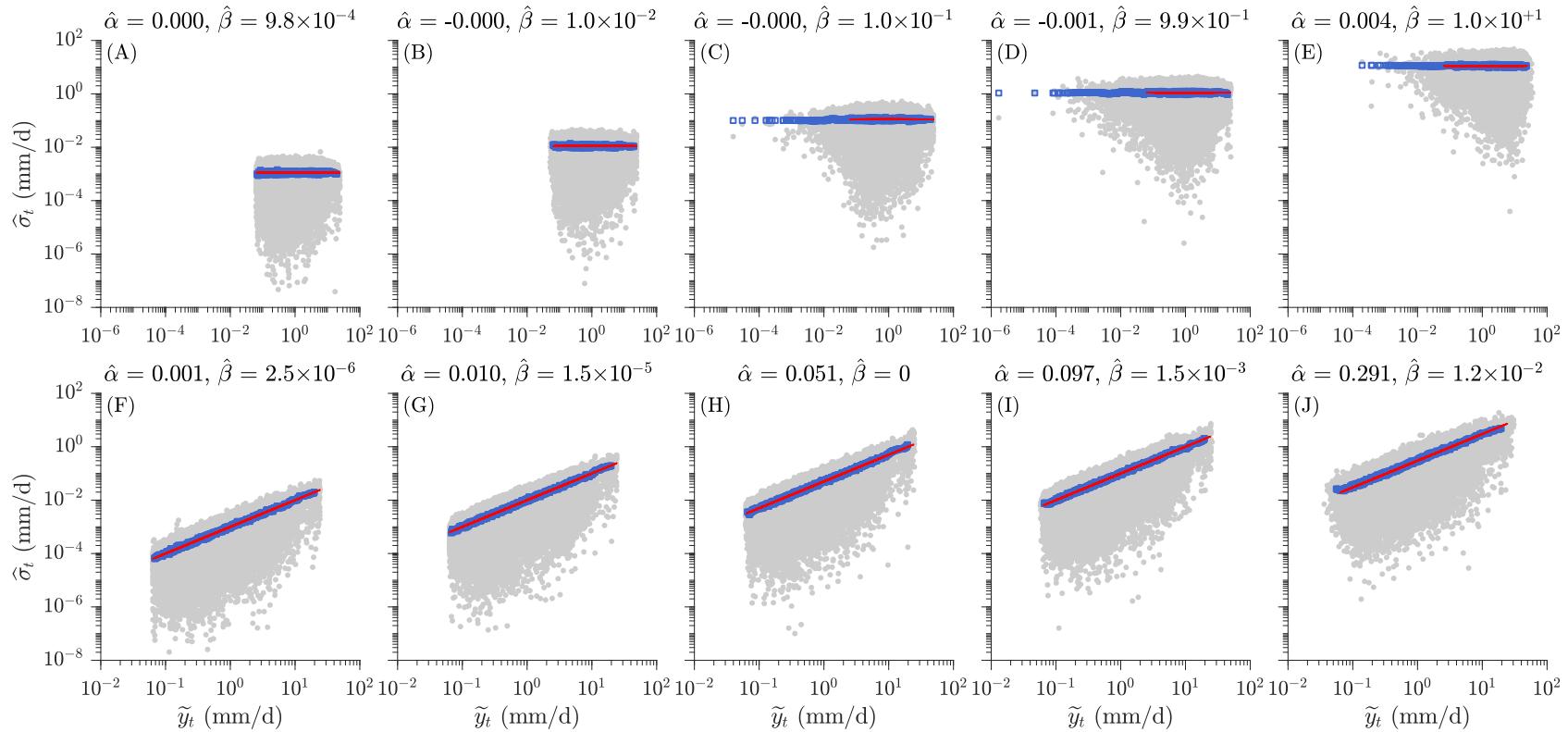


Figure S2: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values obtained using third-order differencing ($k = 3$) for the ten pseudo discharge records of the Leaf River near Collins, MS (USGS 02472000), an example of catchment with a *strong winter regime*. The top and bottom panels correspond to the homoscedastic and heteroscedastic error cases, respectively, and present scatter plots of the $(\bar{y}_h, \hat{\sigma}_h)$ data points for $\alpha = 0$ and (a) $\beta = 0.001$, (b) $\beta = 0.01$, (c) $\beta = 0.1$, (d) $\beta = 1.0$ and (e) $\beta = 10$ and in the heteroscedastic error case with $\beta = 0$ and (f) $\alpha = 0.001$, (g) $\alpha = 0.01$, (h) $\alpha = 0.05$, (i) $\alpha = 0.1$ and (j) $\alpha = 0.3$. Each gray dot signifies a different data pair. The blue squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point. The red line displays the error function of Equation (12) used to corrupt the simulated discharge record. The least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, which is fitted to the blue squares are listed in each graph.

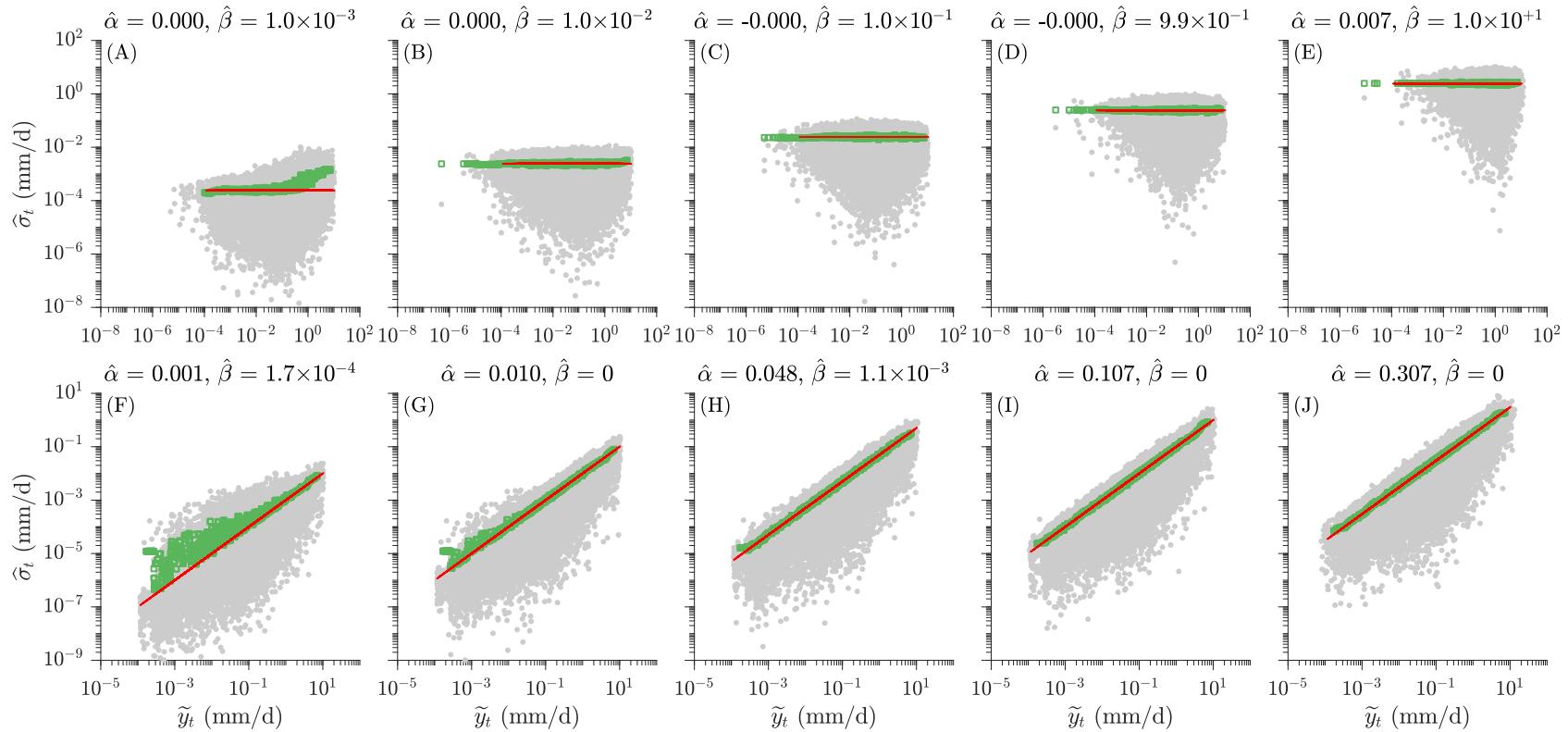


Figure S3: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values obtained using third-order differencing ($k = 3$) for the ten pseudo discharge records of the Cowhouse Creek at Pidcock, TX (USGS 08101000), an example of catchment with the *intermittent regime*. The top and bottom panels correspond to the homoscedastic and heteroscedastic error cases, respectively, and present scatter plots of the $(\bar{y}_h, \hat{\sigma}_h)$ data points for $\alpha = 0$ and (a) $\beta = 0.001$, (b) $\beta = 0.01$, (c) $\beta = 0.1$, (d) $\beta = 1.0$ and (e) $\beta = 10$ and in the heteroscedastic error case with $\beta = 0$ and (f) $\alpha = 0.001$, (g) $\alpha = 0.01$, (h) $\alpha = 0.05$, (i) $\alpha = 0.1$ and (j) $\alpha = 0.3$. Each gray dot signifies a different data pair. The blue squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point. The red line displays the error function of Equation (12) used to corrupt the simulated discharge record. The least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, which is fitted to the blue squares are listed in each graph.

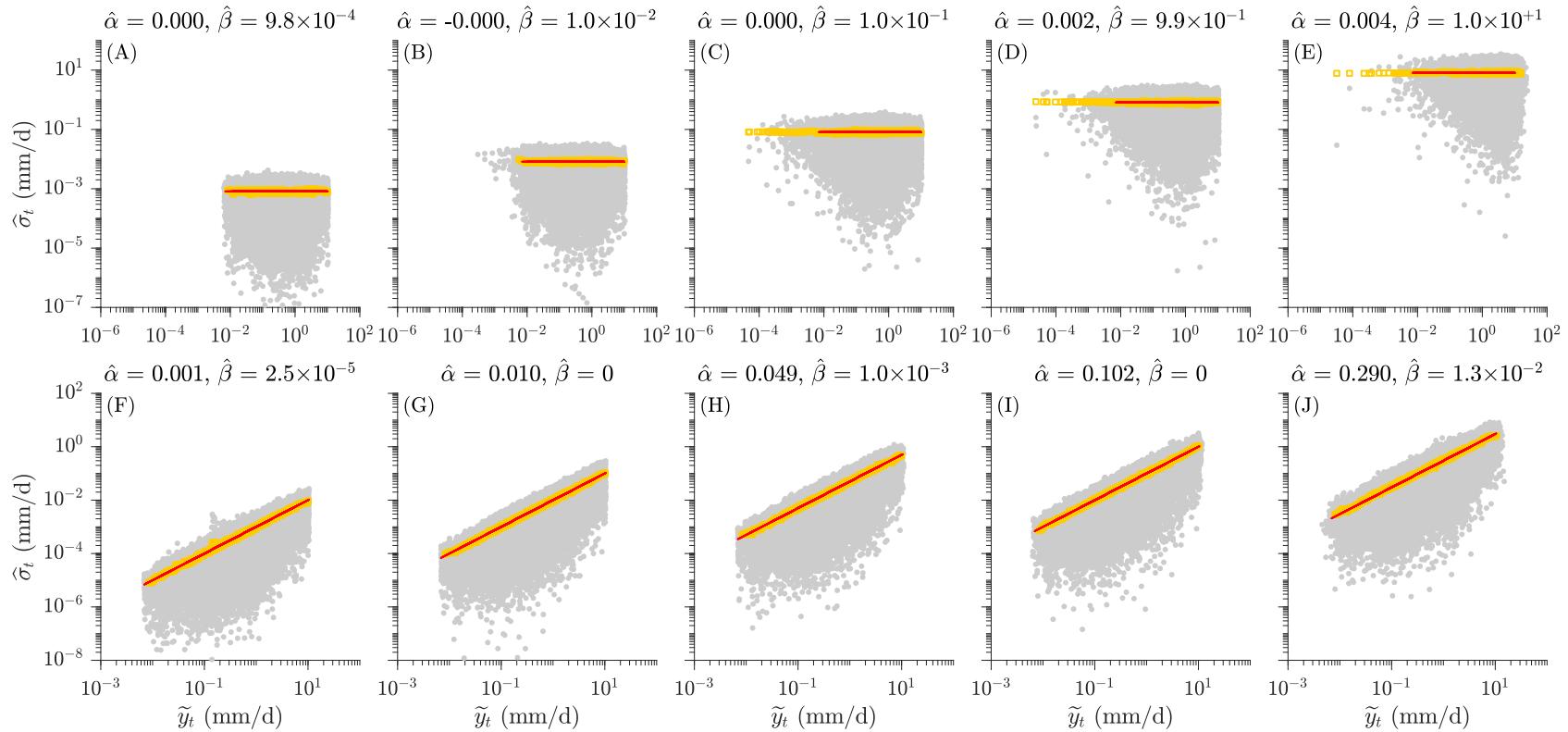


Figure S4: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values obtained using third-order differencing ($k = 3$) for the ten pseudo discharge records of the Potecasi Creek near Union, NC (USGS 02053200), an example of catchment with the *weak winter regime*. The top and bottom panels correspond to the homoscedastic and heteroscedastic error cases, respectively, and present scatter plots of the $(\bar{y}_h, \hat{\sigma}_h)$ data points for $\alpha = 0$ and (a) $\beta = 0.001$, (b) $\beta = 0.01$, (c) $\beta = 0.1$, (d) $\beta = 1.0$ and (e) $\beta = 10$ and in the heteroscedastic error case with $\beta = 0$ and (f) $\alpha = 0.001$, (g) $\alpha = 0.01$, (h) $\alpha = 0.05$, (i) $\alpha = 0.1$ and (j) $\alpha = 0.3$. Each gray dot signifies a different data pair. The blue squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point. The red line displays the error function of Equation (12) used to corrupt the simulated discharge record. The least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha \bar{y}_{th} + \beta m_{y_h}$, which is fitted to the blue squares are listed in each graph.

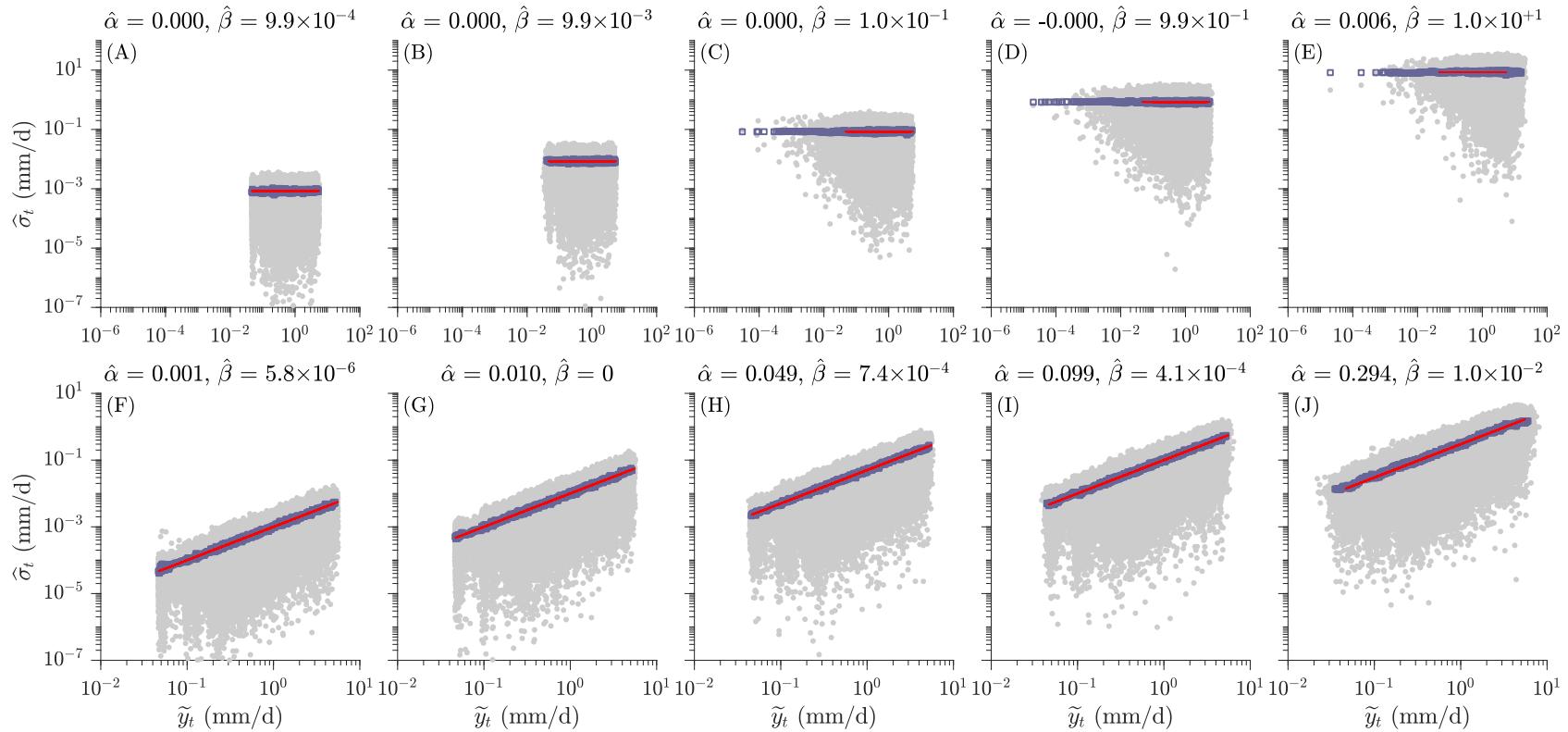


Figure S5: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values obtained using third-order differencing ($k = 3$) for the ten pseudo discharge records of the South Fork Shoshone River near Valley, WY (USGS 06280300), an example of catchment with the *melt regime*. The top and bottom panels correspond to the homoscedastic and heteroscedastic error cases, respectively, and present scatter plots of the $(\bar{y}_h, \hat{\sigma}_h)$ data points for $\alpha = 0$ and (a) $\beta = 0.001$, (b) $\beta = 0.01$, (c) $\beta = 0.1$, (d) $\beta = 1.0$ and (e) $\beta = 10$ and in the heteroscedastic error case with $\beta = 0$ and (f) $\alpha = 0.001$, (g) $\alpha = 0.01$, (h) $\alpha = 0.05$, (i) $\alpha = 0.1$ and (j) $\alpha = 0.3$. Each gray dot signifies a different data pair. The blue squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point. The red line displays the error function of Equation (12) used to corrupt the simulated discharge record. The least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, which is fitted to the blue squares are listed in each graph.

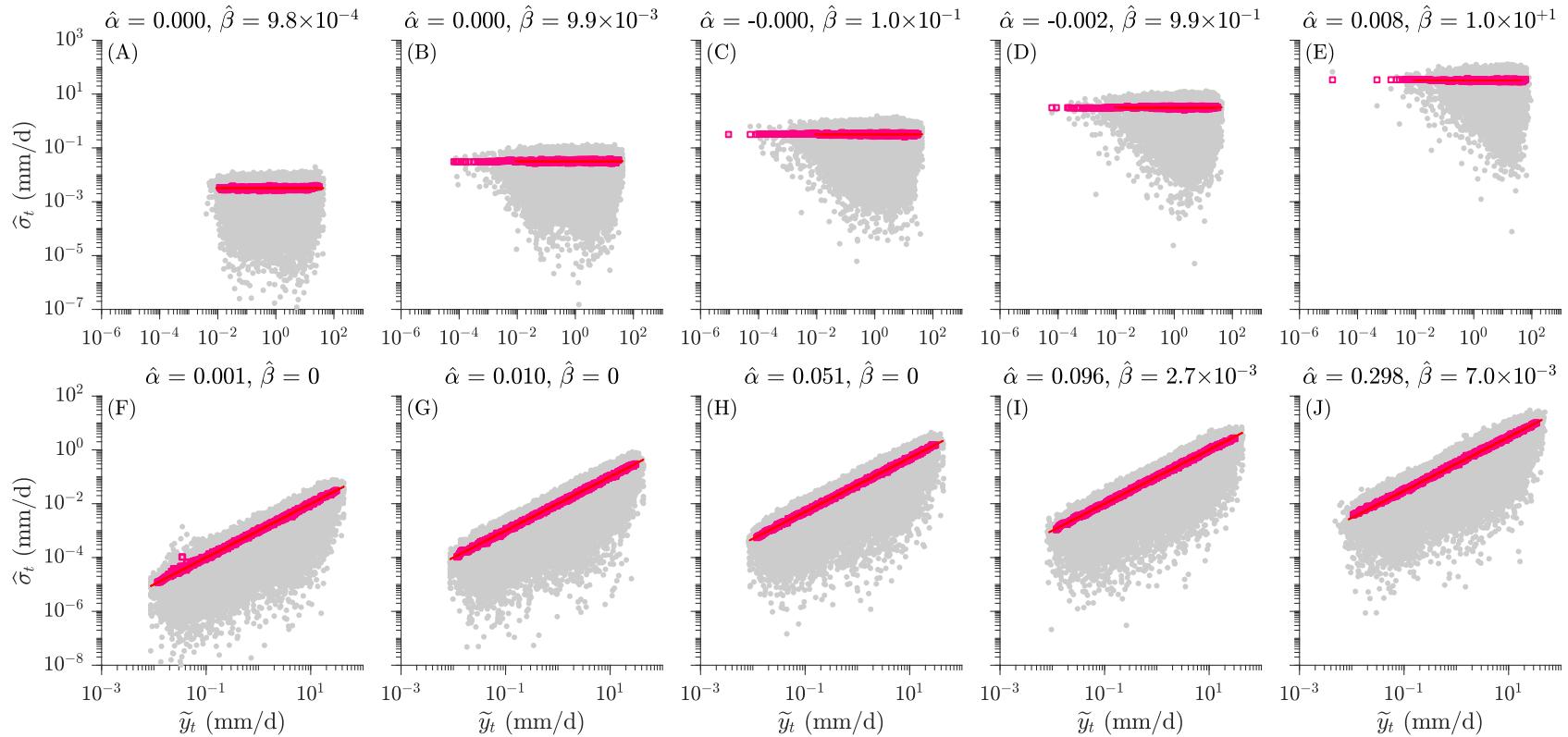


Figure S6: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values obtained using third-order differencing ($k = 3$) for the ten pseudo discharge records of the Nehalem River near Foss, OR (USGS 14301000), an example of catchment with the *New Year's regime*. The top and bottom panels correspond to the homoscedastic and heteroscedastic error cases, respectively, and present scatter plots of the $(\bar{y}_h, \hat{\sigma}_h)$ data points for $\alpha = 0$ and (a) $\beta = 0.001$, (b) $\beta = 0.01$, (c) $\beta = 0.1$, (d) $\beta = 1.0$ and (e) $\beta = 10$ and in the heteroscedastic error case with $\beta = 0$ and (f) $\alpha = 0.001$, (g) $\alpha = 0.01$, (h) $\alpha = 0.05$, (i) $\alpha = 0.1$ and (j) $\alpha = 0.3$. Each gray dot signifies a different data pair. The blue squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point. The red line displays the error function of Equation (12) used to corrupt the simulated discharge record. The least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\hat{\alpha}\bar{y}_{th} + \hat{\beta}m_{y_h}$, which is fitted to the blue squares are listed in each graph.

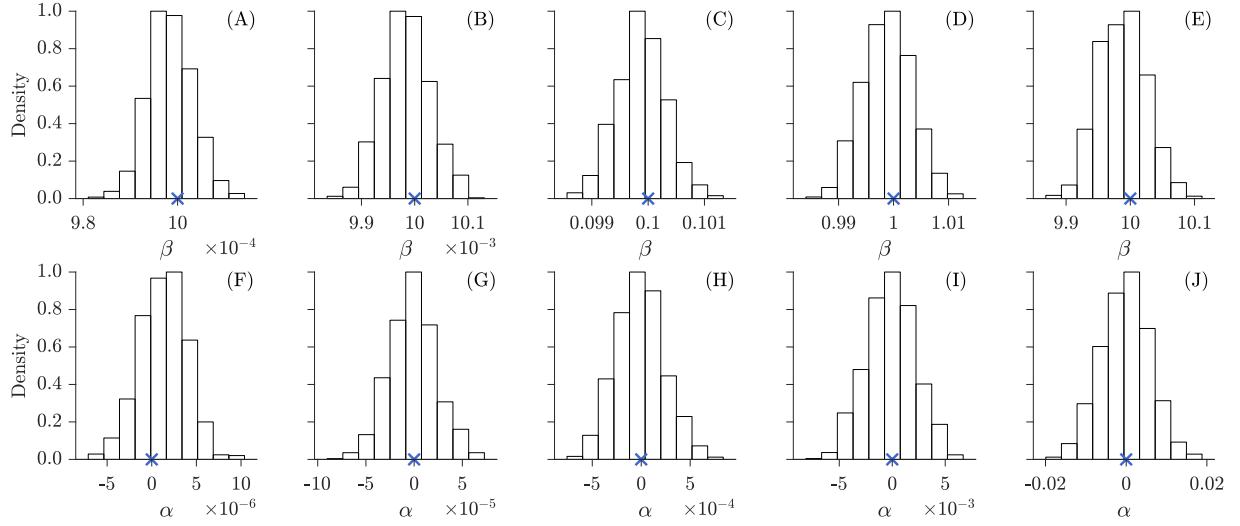


Figure S7: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the Leaf River near Collins, MS (USGS 02472000), an example of catchment with a *strong winter regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (16) and (17) for the homoscedastic error case: $\alpha = 0$ and (a,f) $\beta = 0.001$, (b,g) $\beta = 0.01$, (c,h) $\beta = 0.1$, (d,i) $\beta = 1.0$ and (e,j) $\beta = 10$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

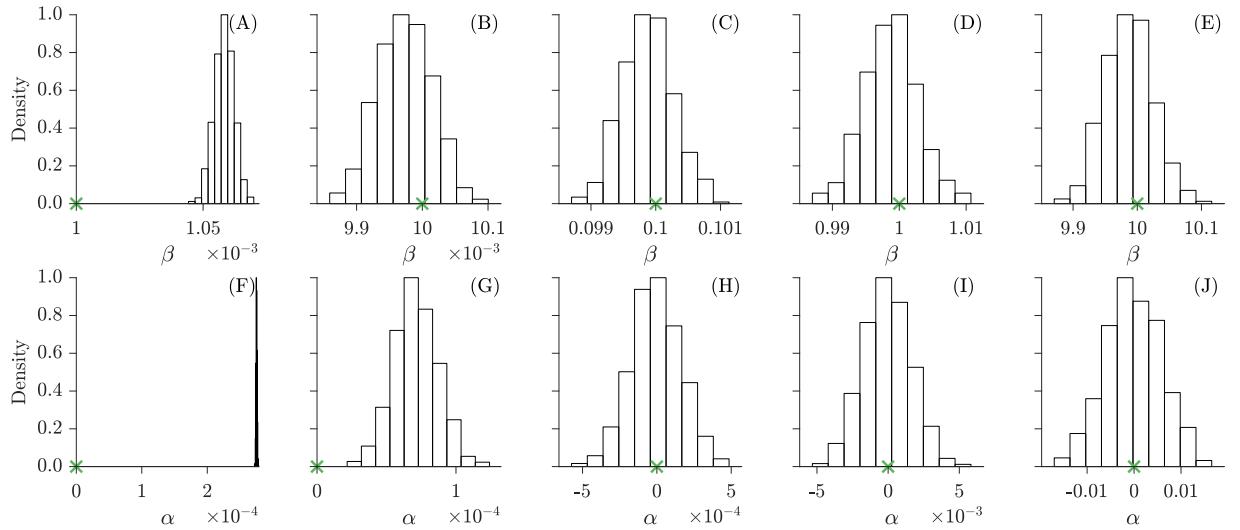


Figure S8: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the Cowhouse Creek at Pidcock, TX (USGS 08101000), an example of catchment with the *intermittent regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (9) and (10) for the homoscedastic error case: $\alpha = 0$ and (a,f) $\beta = 0.001$, (b,g) $\beta = 0.01$, (c,h) $\beta = 0.1$, (d,i) $\beta = 1.0$ and (e,j) $\beta = 10$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

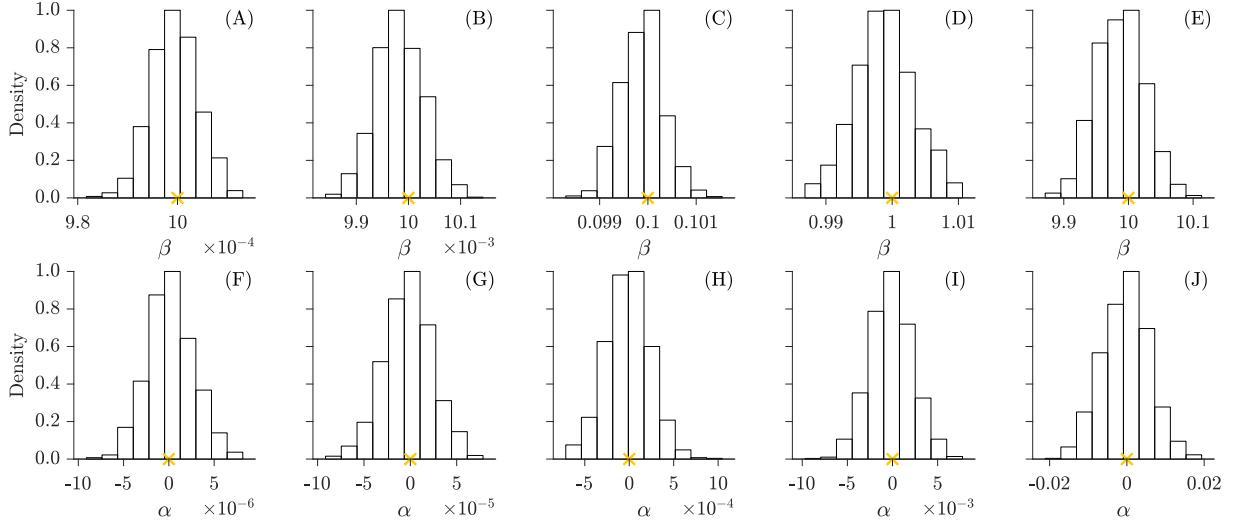


Figure S9: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the Potecasi Creek near Union, NC (USGS 02053200), an example of catchment with the *weak winter regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (9) and (10) for the homoscedastic error case: $\alpha = 0$ and (a,f) $\beta = 0.001$, (b,g) $\beta = 0.01$, (c,h) $\beta = 0.1$, (d,i) $\beta = 1.0$ and (e,j) $\beta = 10$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

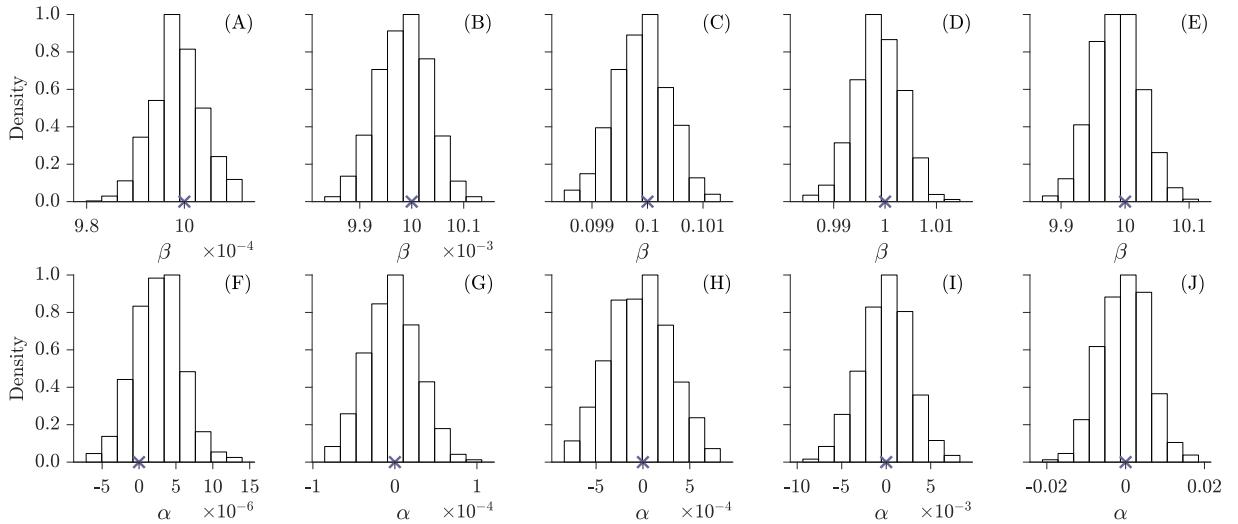


Figure S10: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the South Fork Shoshone River near Valley, WY (USGS 06280300), an example of catchment with the *melt regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (9) and (10) for the homoscedastic error case: $\alpha = 0$ and (a,f) $\beta = 0.001$, (b,g) $\beta = 0.01$, (c,h) $\beta = 0.1$, (d,i) $\beta = 1.0$ and (e,j) $\beta = 10$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

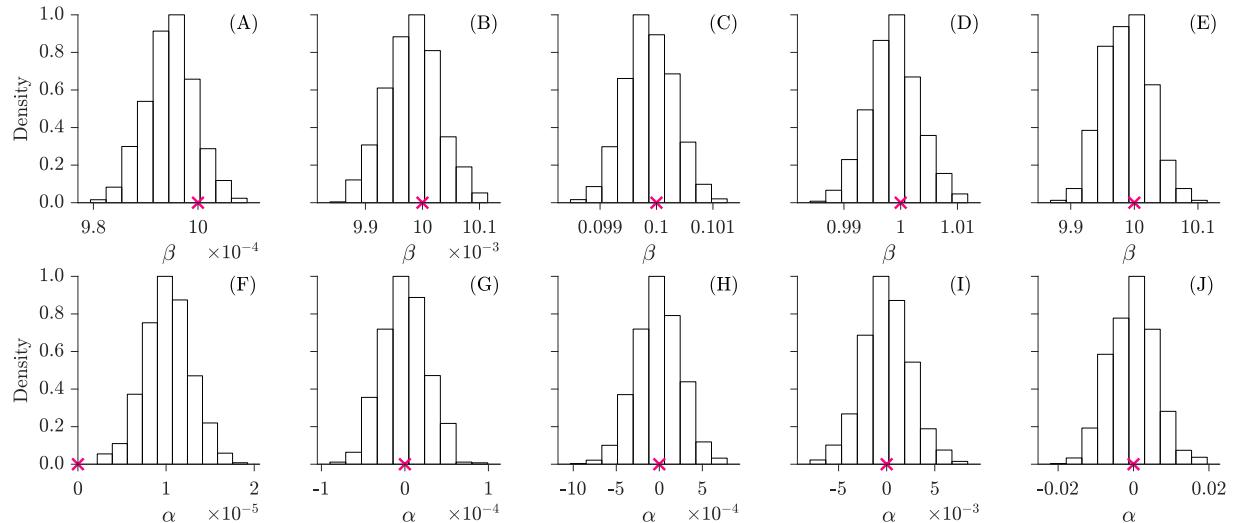


Figure S11: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the Nehalem River near Foss, OR (USGS 14301000), an example of catchment with the *New Year's regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (9) and (10) for the homoscedastic error case: $\alpha = 0$ and (a,f) $\beta = 0.001$, (b,g) $\beta = 0.01$, (c,h) $\beta = 0.1$, (d,i) $\beta = 1.0$ and (e,j) $\beta = 10$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

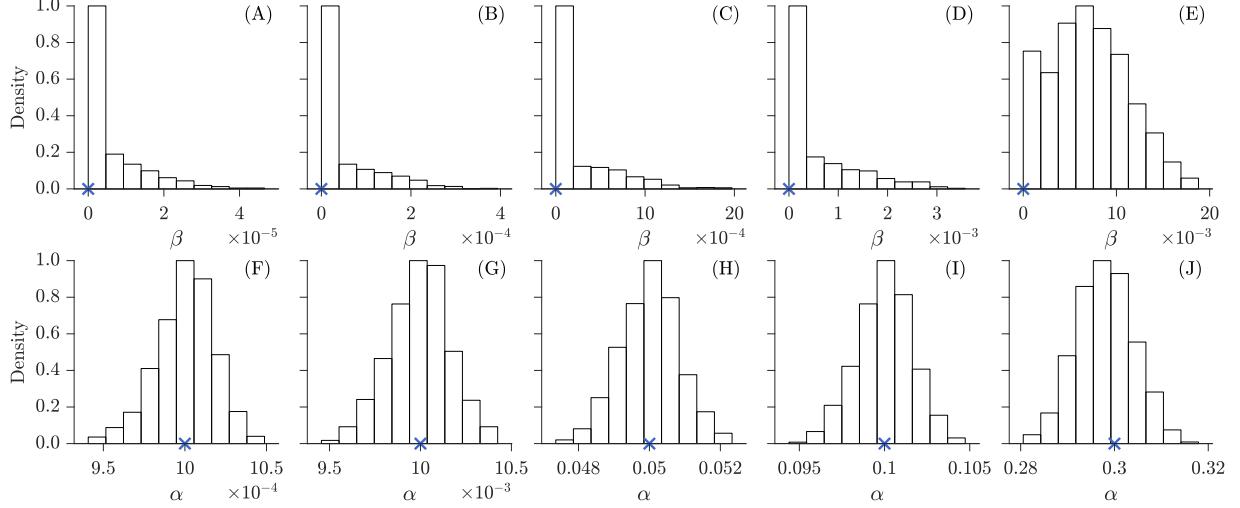


Figure S12: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the Leaf River near Collins, MS (USGS 02472000), an example of catchment with a *strong winter regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (16) and (17) for the heteroscedastic error case with $\beta = 0$ and (a,f) $\alpha = 0.001$, (b,g) $\alpha = 0.01$, (c,h) $\alpha = 0.05$, (d,i) $\alpha = 0.1$ and (e,j) $\alpha = 0.3$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

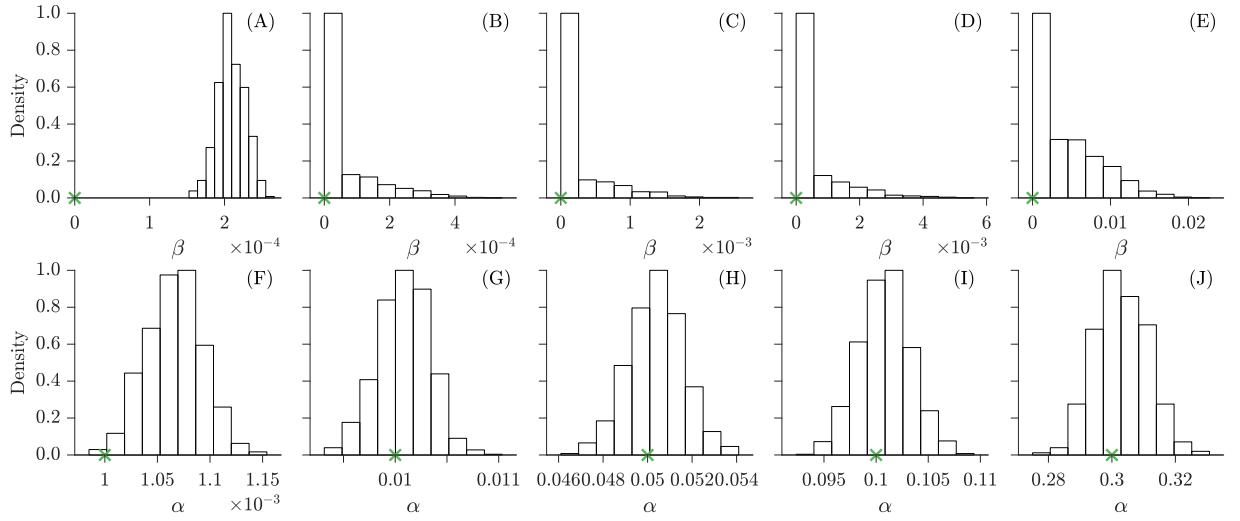


Figure S13: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the Cowhouse Creek at Pidcock, TX (USGS 08101000), an example of catchment with the *intermittent regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (9) and (10) for the heteroscedastic error case with $\beta = 0$ and (a,f) $\alpha = 0.001$, (b,g) $\alpha = 0.01$, (c,h) $\alpha = 0.05$, (d,i) $\alpha = 0.1$ and (e,j) $\alpha = 0.3$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

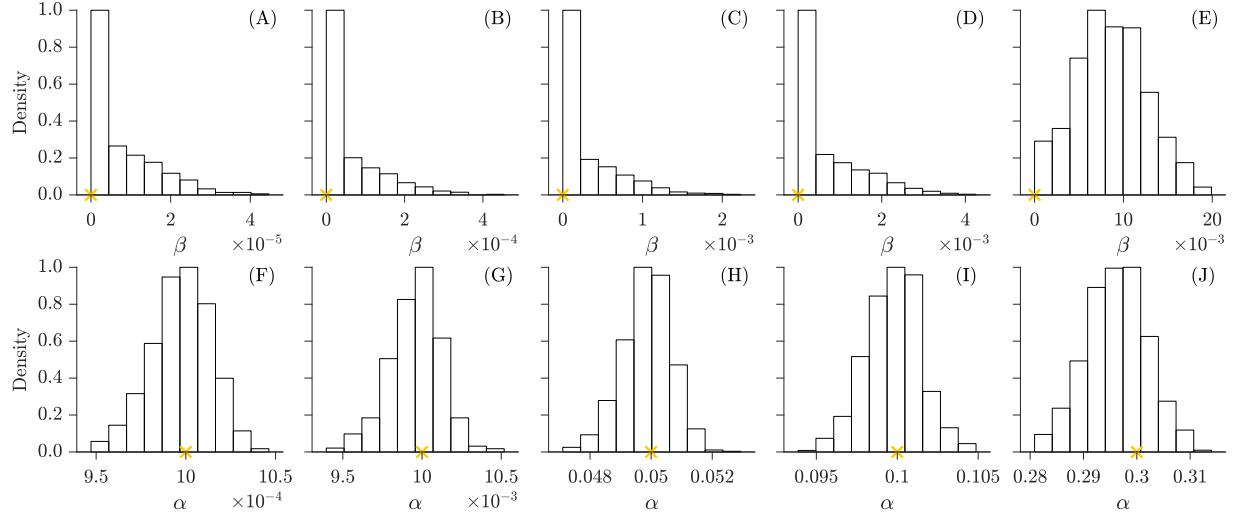


Figure S14: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the Potecasi Creek near Union, NC (USGS 02053200), an example of catchment with the *weak winter regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (9) and (10) for the heteroscedastic error case with $\beta = 0$ and (a,f) $\alpha = 0.001$, (b,g) $\alpha = 0.01$, (c,h) $\alpha = 0.05$, (d,i) $\alpha = 0.1$ and (e,j) $\alpha = 0.3$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

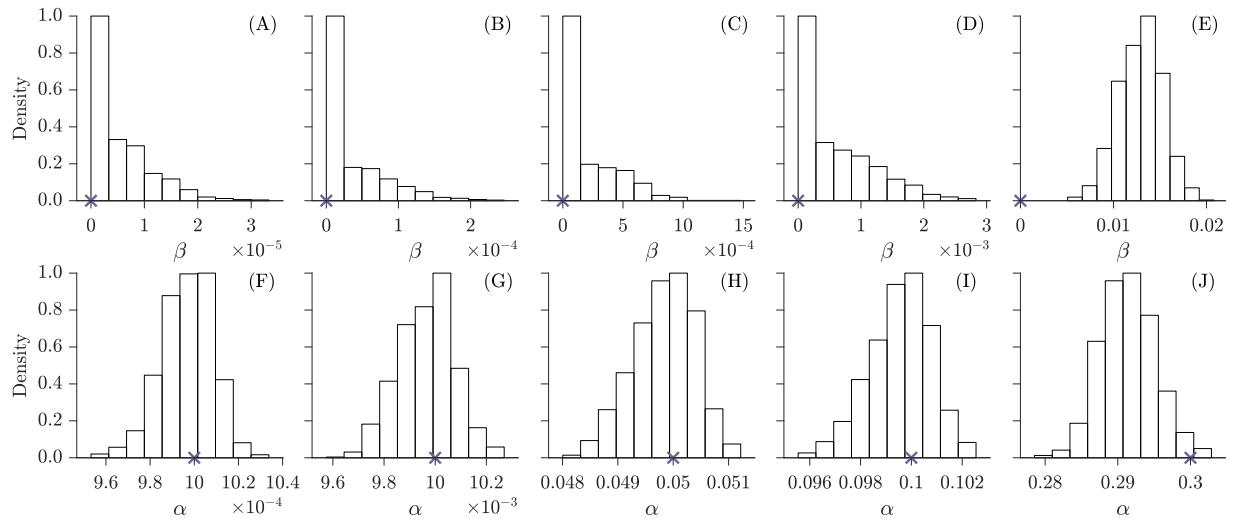


Figure S15: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the South Fork Shoshone River near Valley, WY (USGS 06280300), an example of catchment with the *melt regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (9) and (10) for the heteroscedastic error case with $\beta = 0$ and (a,f) $\alpha = 0.001$, (b,g) $\alpha = 0.01$, (c,h) $\alpha = 0.05$, (d,i) $\alpha = 0.1$ and (e,j) $\alpha = 0.3$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

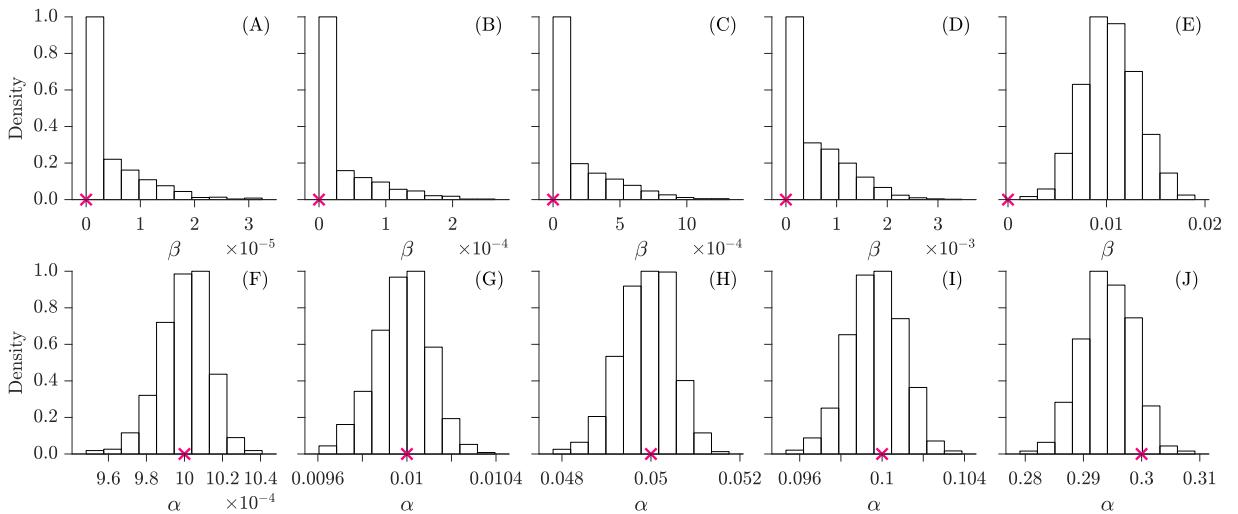


Figure S16: Least squares values of the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the linear regression model, $\alpha\bar{y}_{th} + \beta m_{y_h}$, for the 1,000 realizations of pseudo discharge records of the Nehalem River near Foss, OR (USGS 14301000), an example of catchment with the *New Year's regime*. The top and bottom panels correspond to $\hat{\alpha}$ and $\hat{\beta}$, respectively, derived from Equations (9) and (10) for the heteroscedastic error case with $\beta = 0$ and (a,f) $\alpha = 0.001$, (b,g) $\alpha = 0.01$, (c,h) $\alpha = 0.05$, (d,i) $\alpha = 0.1$ and (e,j) $\alpha = 0.3$. The relative frequencies on the y -axis are normalized to yield a common empirical density between 0 and 1.

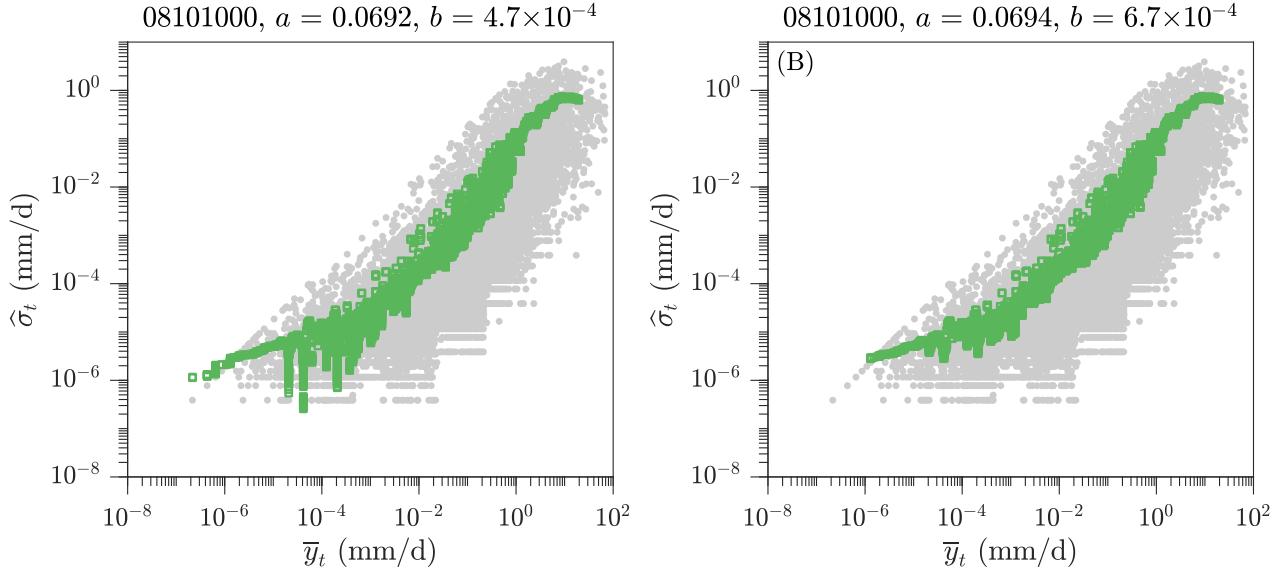


Figure S17: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values for the Cowhouse Creek at Pidcocke, TX (USGS 08101000), an example of catchment with the *intermittent regime*: (a) using all data, and (b) removing from the analysis $\hat{\sigma}_t$ values smaller than 10^{-10} . Each gray dot signifies a different data pair. The green squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point.

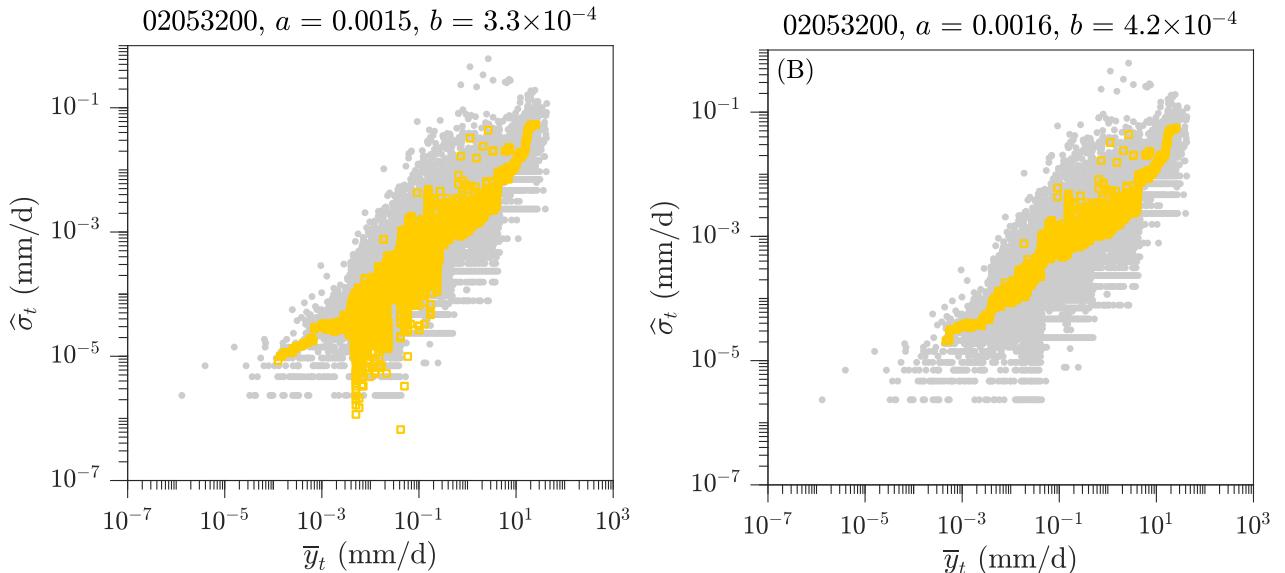


Figure S18: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values for the Potecassi Creek near Union, NC (USGS 02053200), an example of catchment with the *weak winter regime*: (a) using all data, and (b) removing from the analysis $\hat{\sigma}_t$ values smaller than 10^{-10} . Each gray dot signifies a different data pair. The yellow squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point.

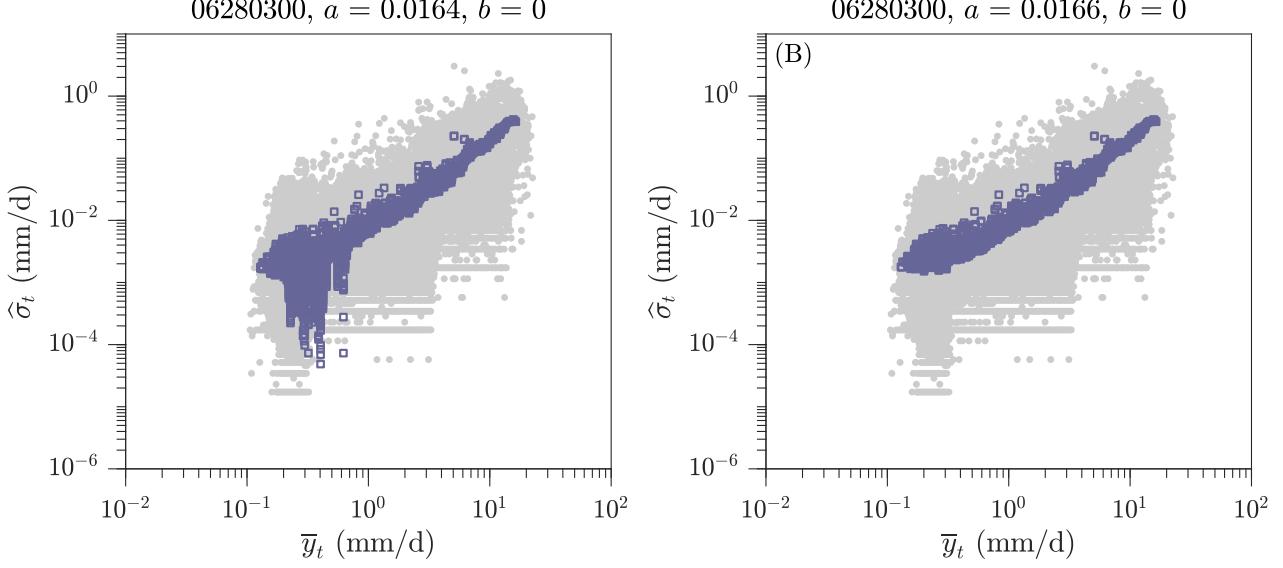


Figure S19: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values for the South Fork Shoshone River near Valley, WY (USGS 06280300), an example of catchment with the *melt regime*: (a) using all data, and (b) removing from the analysis $\hat{\sigma}_t$ values smaller than 10^{-10} . Each gray dot signifies a different data pair. The purple squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point.

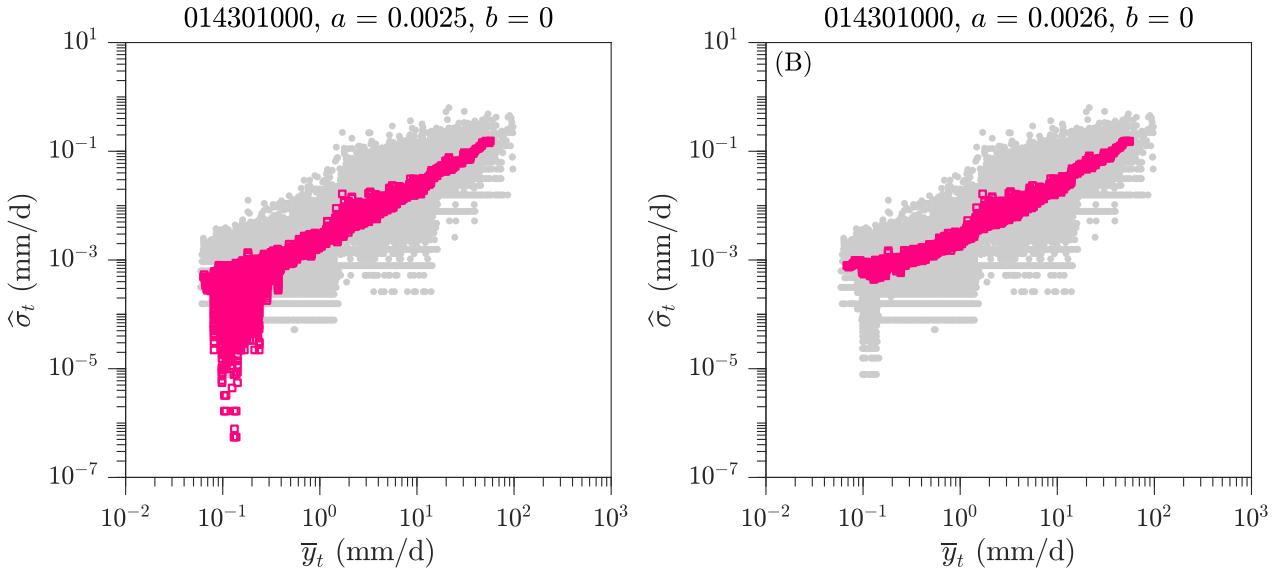


Figure S20: Relationship between the error deviations $\hat{\sigma}_t$ and corresponding hourly discharge values for the Nehalem River near Foss, OR (USGS 14301000), an example of catchment with the *New Year's regime*: (a) using all data, and (b) removing from the analysis $\hat{\sigma}_t$ values smaller than 10^{-10} . Each gray dot signifies a different data pair. The pink squares portray the moving average of the error deviation computed from a window of 100 data pairs on either side of the data point.

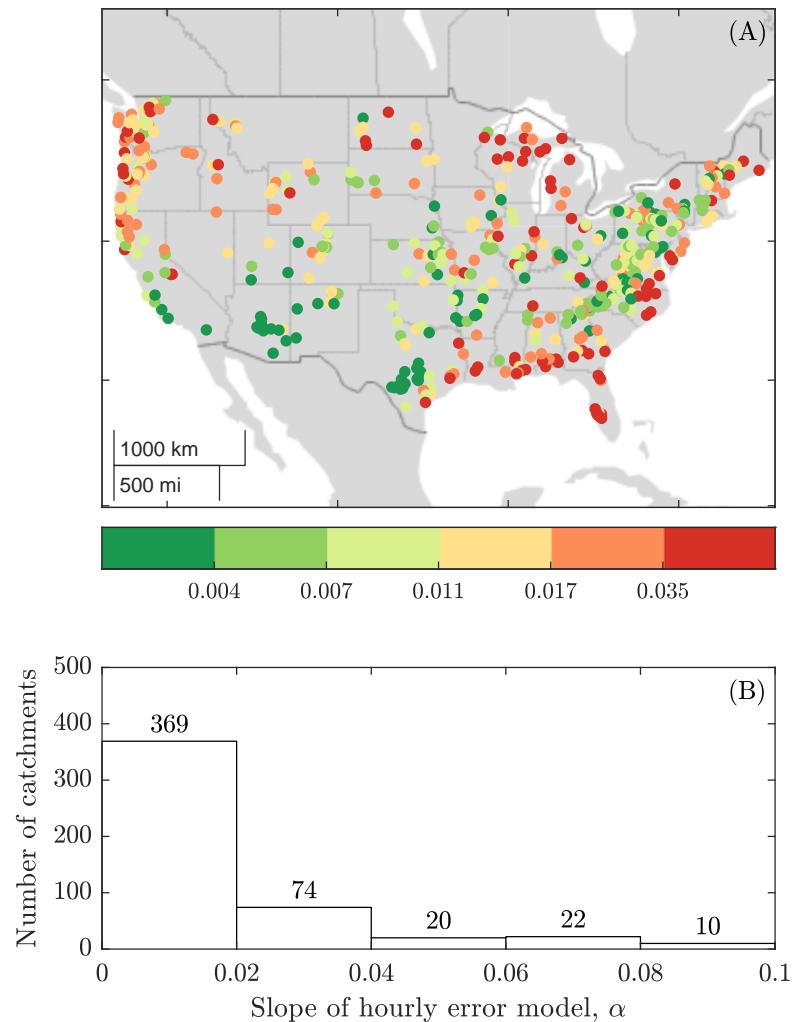


Figure S21: The hourly discharge error model of the CAMELS data set: (a) dimensionless slope, α , for each catchment; (b) frequency distribution of the dimensionless slope, α , of the 504 catchments. In panel (a), each color class encompasses about one sixth of the total number of considered catchments.

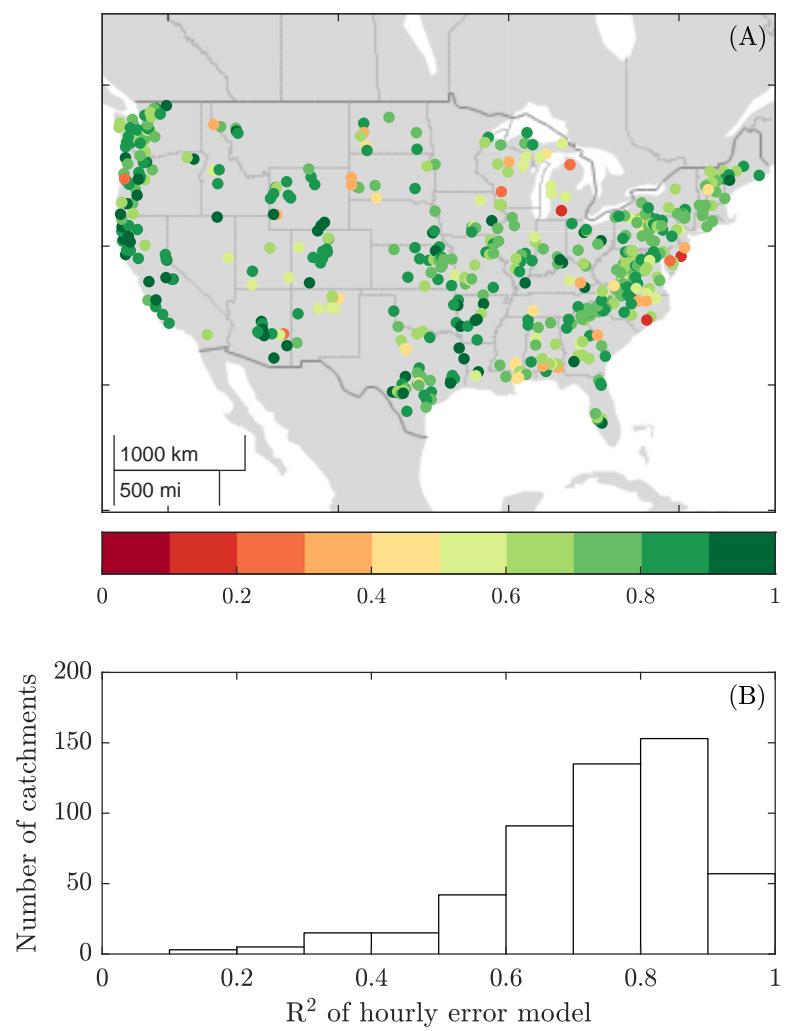


Figure S22: Coefficient of determination (R^2) of the hourly error model for the 504 catchments of the CAMELS data set: (a) spatial distribution, and (b) frequency distribution.

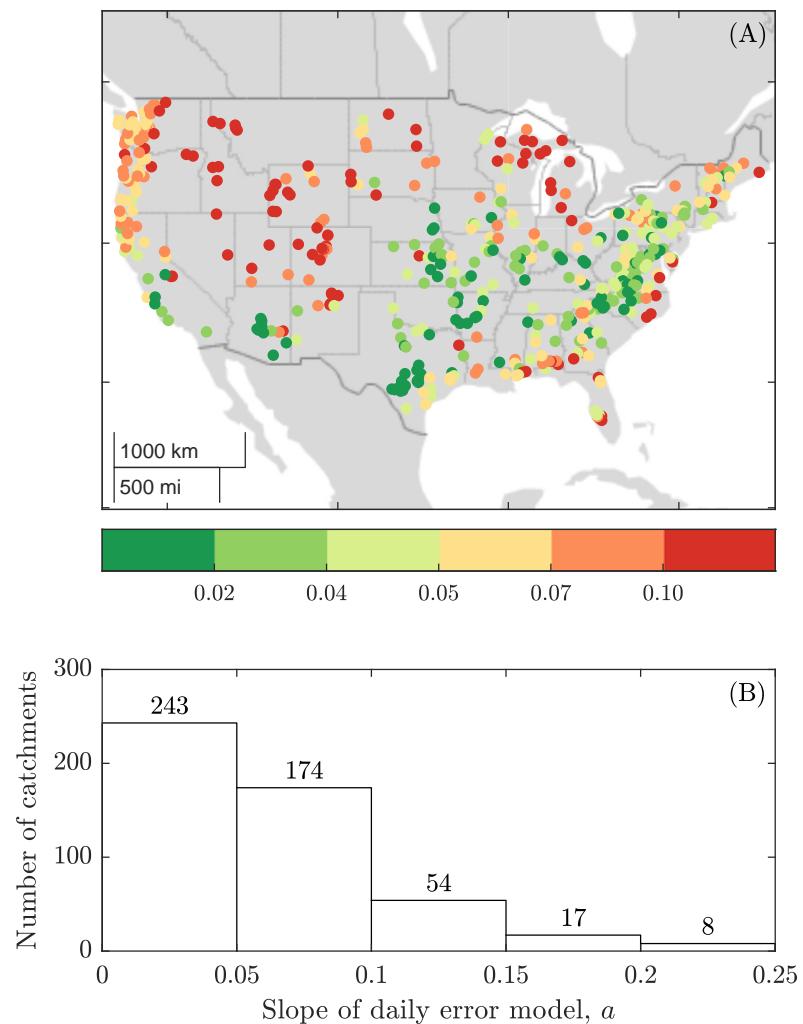


Figure S23: The daily discharge error model of the CAMELS data set: (a) dimensionless slope, a , for each catchment; (b) frequency distribution of the dimensionless slope, a , of the 504 catchments. In panel (a), each color class encompasses about one sixth of the total number of considered catchments.

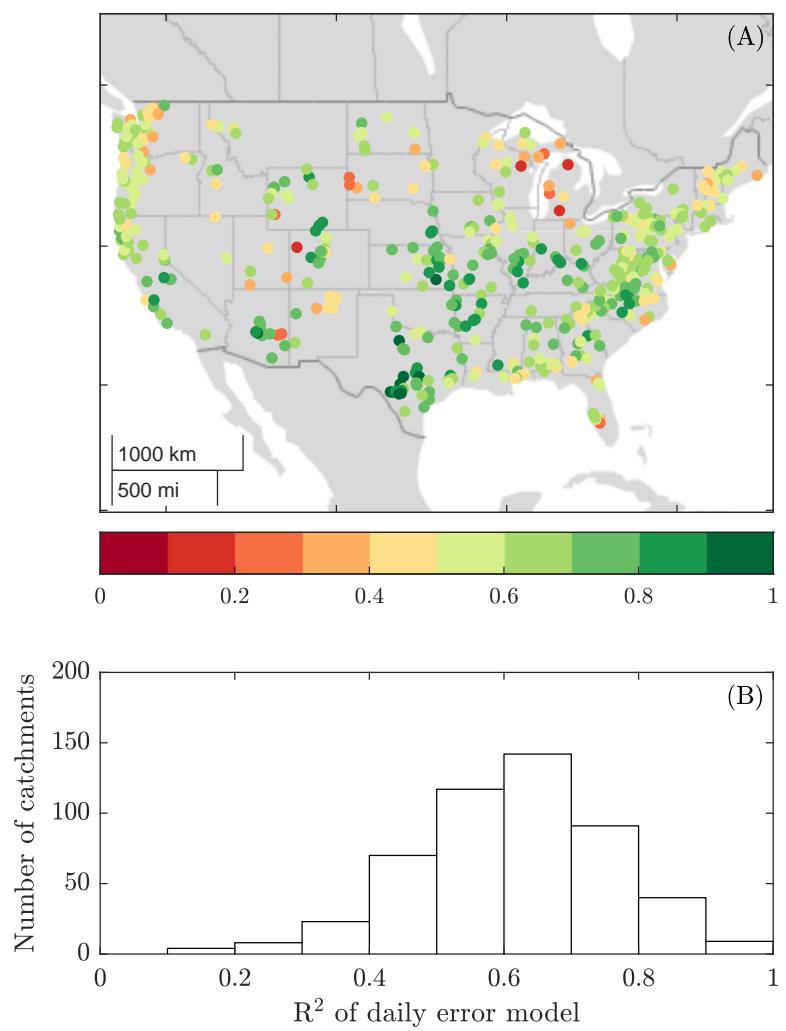


Figure S24: Coefficient of determination (R^2) of the daily error model for the 504 catchments of the CAMELS data set: (a) spatial distribution, and (b) frequency distribution.

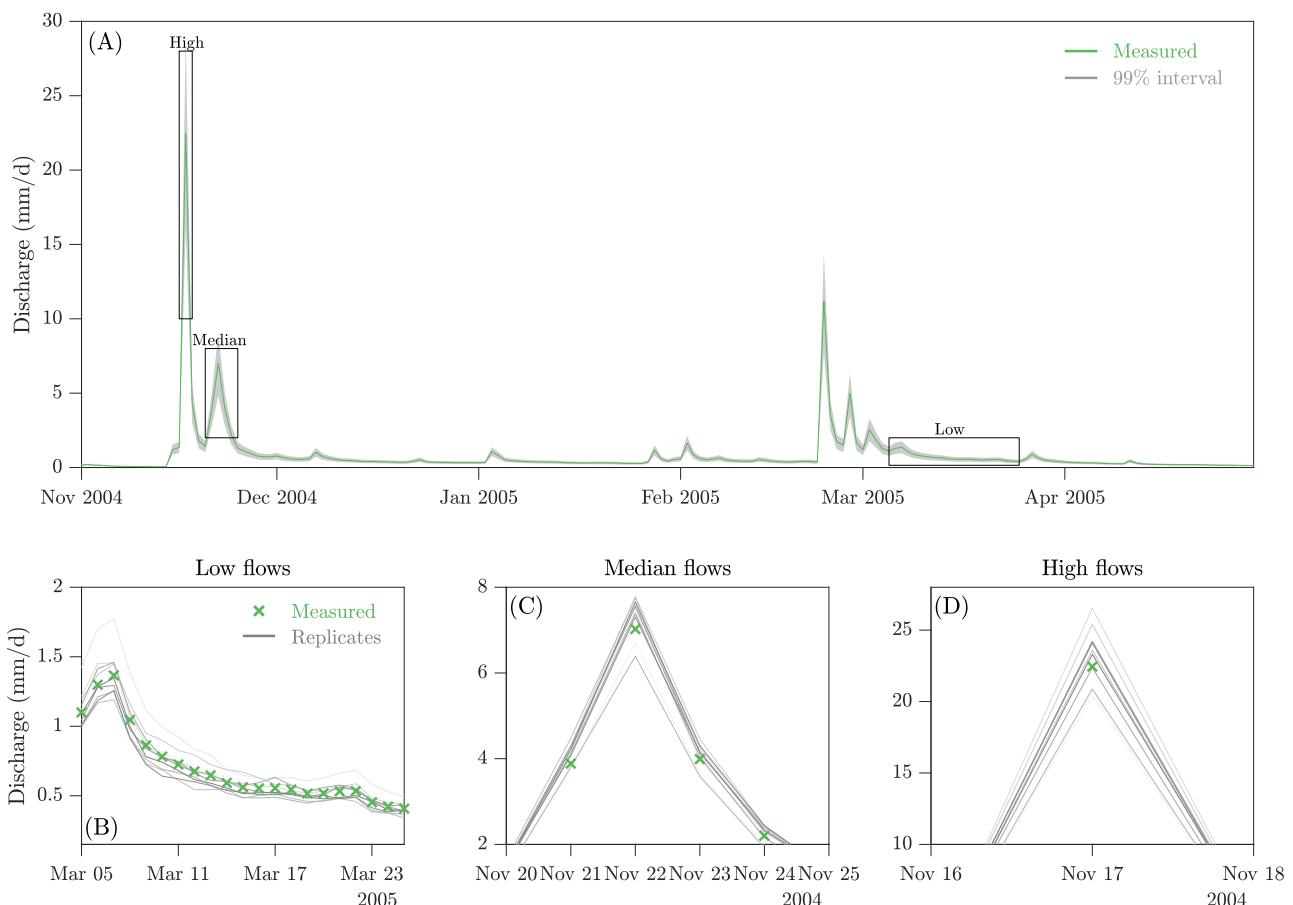


Figure S25: Illustration of the streamflow replicates generated using daily streamflow data from the Cowhouse Creek at Pidcock, TX (USGS 08101000), an example of catchment with the *intermittent regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid green line. The top panel only visualizes percentiles of the discharge uncertainty without recourse to the underlying replicates. Therefore, the bottom panel displays a selection of the replicates for small excerpts of the discharge record with (b) low, (c) median and (d) high flows, respectively.

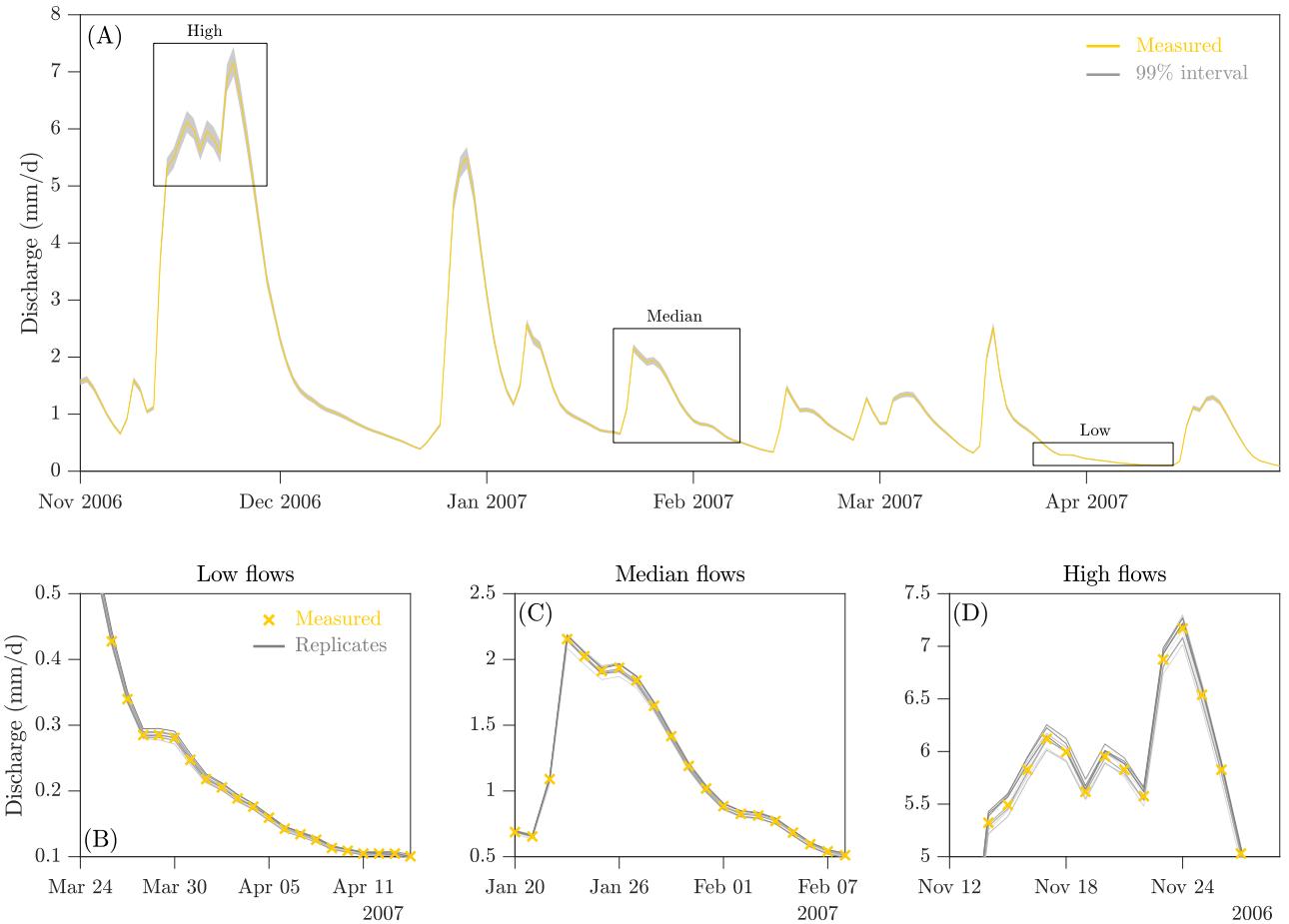


Figure S26: Illustration of the streamflow replicates generated using daily streamflow data from the Potecasi Creek near Union, NC (USGS 02053200), an example of catchment with the *weak winter regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid yellow line. The top panel only visualizes percentiles of the discharge uncertainty without recourse to the underlying replicates. Therefore, the bottom panel displays a selection of the replicates for small excerpts of the discharge record with (b) low, (c) median and (d) high flows, respectively.

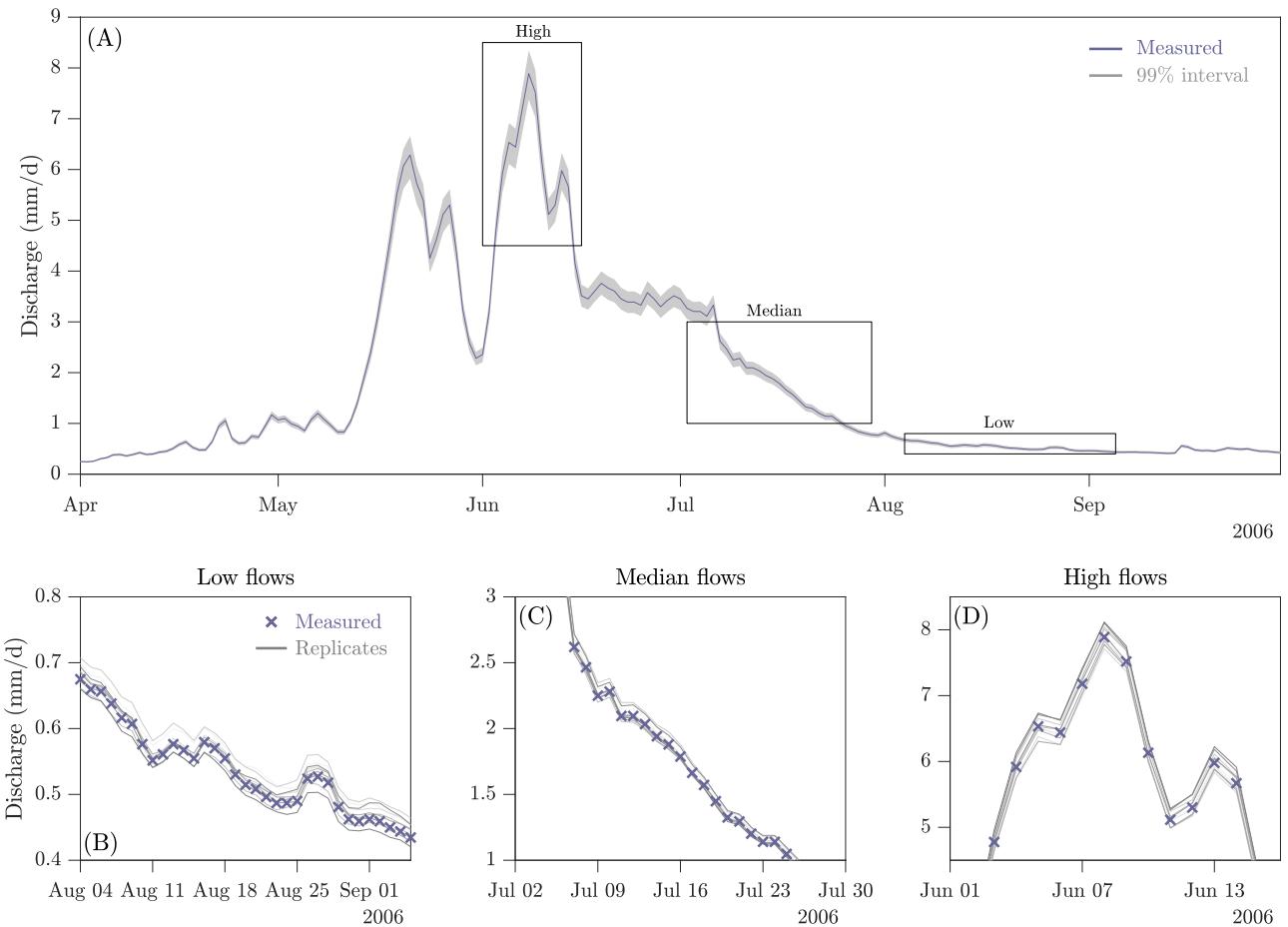


Figure S27: Illustration of the streamflow replicates generated using daily streamflow data from the South Fork Shoshone River near Valley, WY (USGS 06280300), an example of catchment with the *melt regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid purple line. The top panel only visualizes percentiles of the discharge uncertainty without recourse to the underlying replicates. Therefore, the bottom panel displays a selection of the replicates for small excerpts of the discharge record with (b) low, (c) median and (d) high flows, respectively.

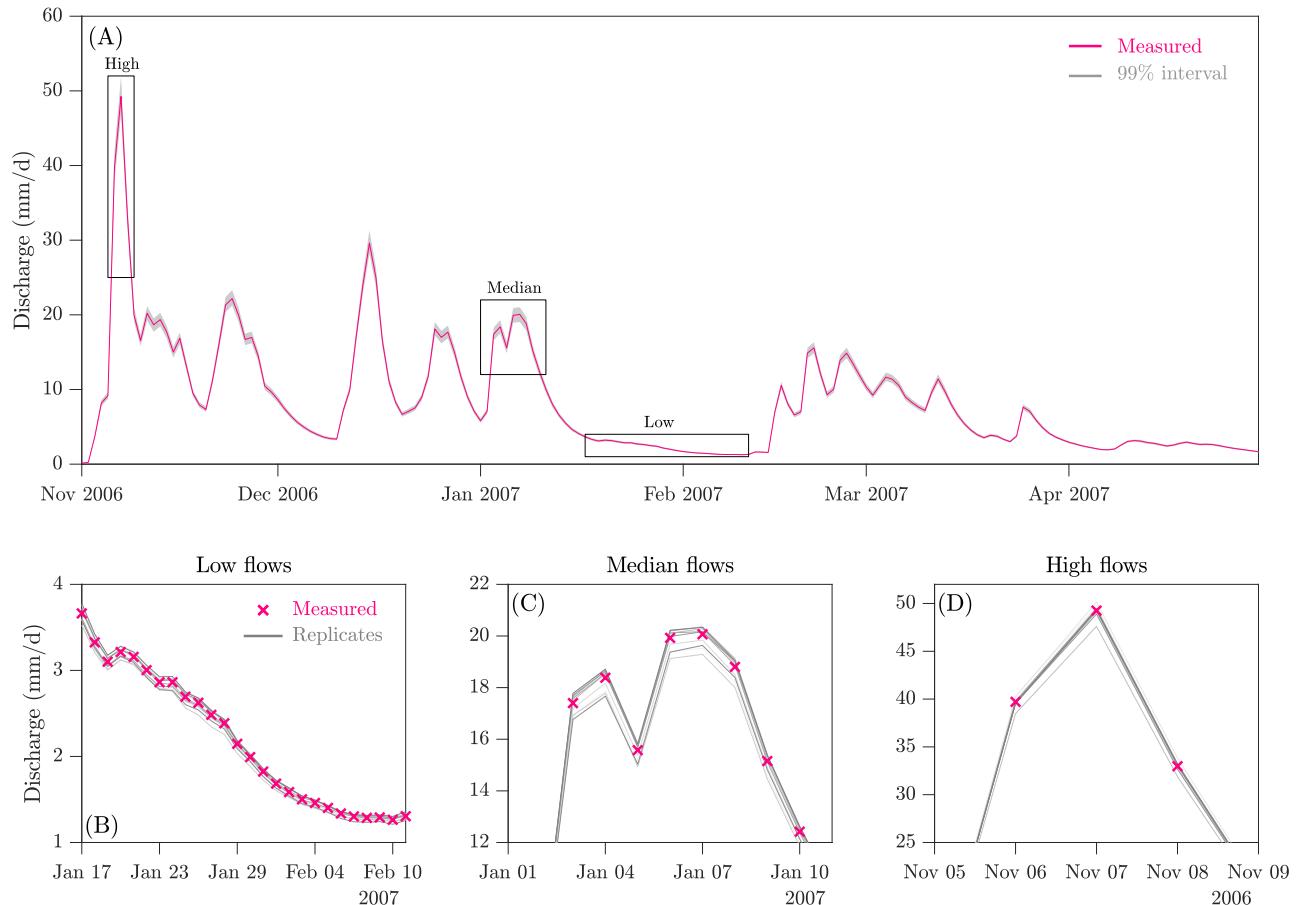


Figure S28: Illustration of the streamflow replicates generated using daily streamflow data from the Nehalem River near Foss, OR (USGS 14301000), an example of catchment with the *New Year's regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid pink line. The top panel only visualizes percentiles of the discharge uncertainty without recourse to the underlying replicates. Therefore, the bottom panel displays a selection of the replicates for small excerpts of the discharge record with (b) low, (c) median and (d) high flows, respectively.

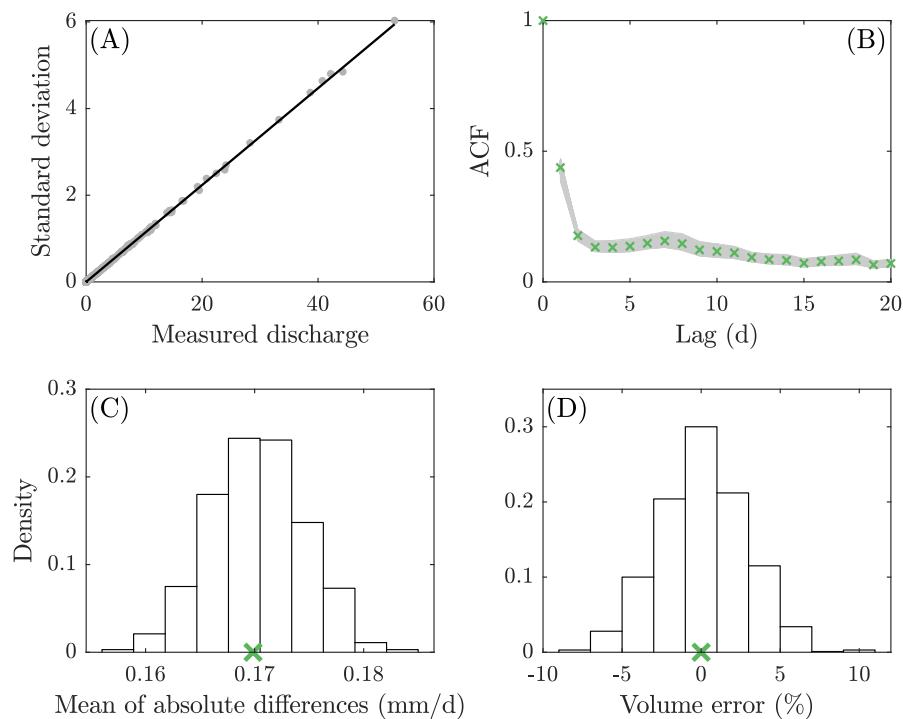


Figure S29: Characteristics of the streamflow replicates for the Cowhouse Creek at Pidcocke, TX (USGS 08101000), an example of catchment with the *intermittent regime*. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the discharge record (green crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The green crosses highlight the values computed from the original record.

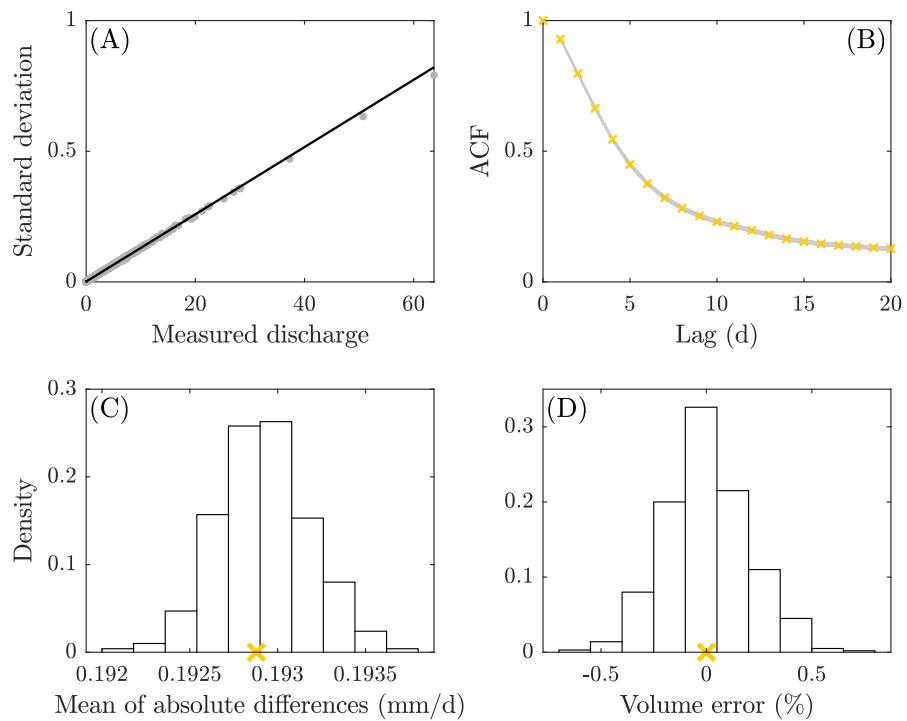


Figure S30: Characteristics of the streamflow replicates for the Potecasi Creek near Union, NC (USGS 02053200), an example of catchment with the *weak winter regime*. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the original record (yellow crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The yellow crosses highlight the values computed from the original record.

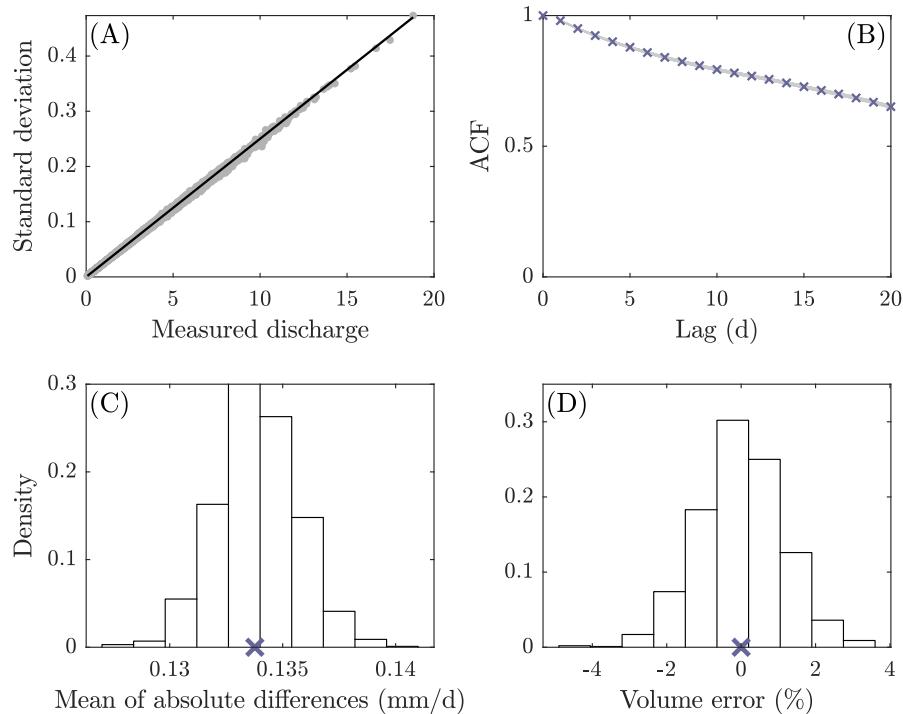


Figure S31: Characteristics of the streamflow replicates for the South Fork Shoshone River near Valley, WY (USGS 06280300), an example of catchment with the *melt regime*. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the original record (purple crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The purple crosses highlight the values computed from the original record.

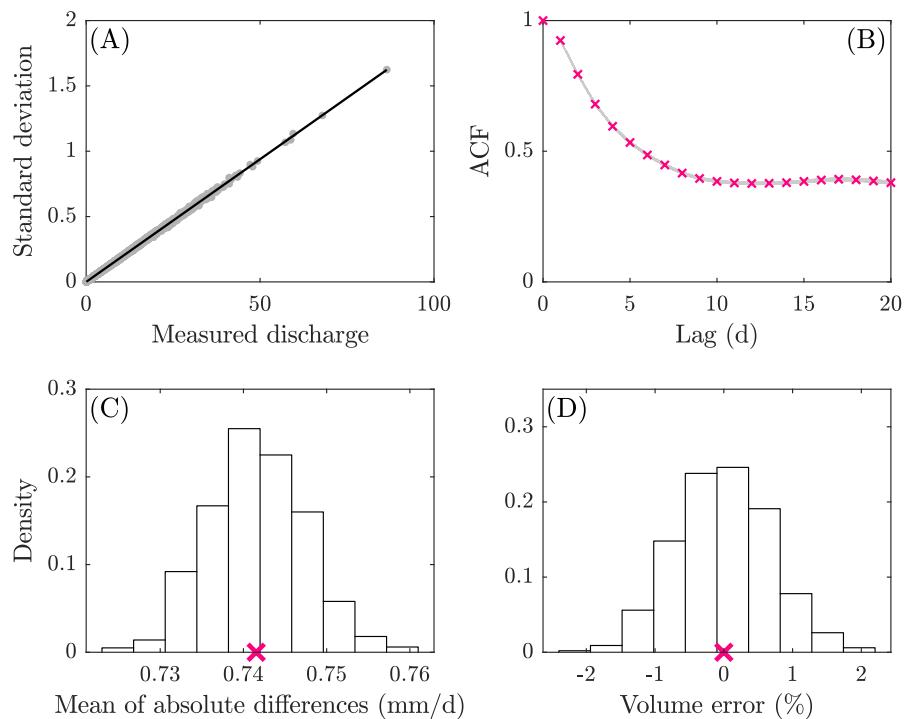


Figure S32: Characteristics of the streamflow replicates for the Nehalem River near Foss, OR (USGS 14301000), an example of catchment with the *New Year's regime*. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the original record (pink crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The pink crosses highlight the values computed from the original record.

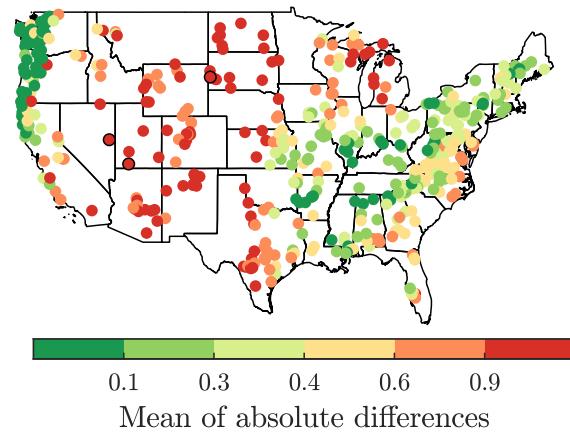


Figure S33: Mean of absolute differences of the discharge time series of the CAMELS catchments. Black circles correspond to catchments for which the value computed from the original record is not inside the sampled distribution, i.e., the mean of absolute differences of the replicates is consistently lower/higher than the mean of absolute differences of the original record.

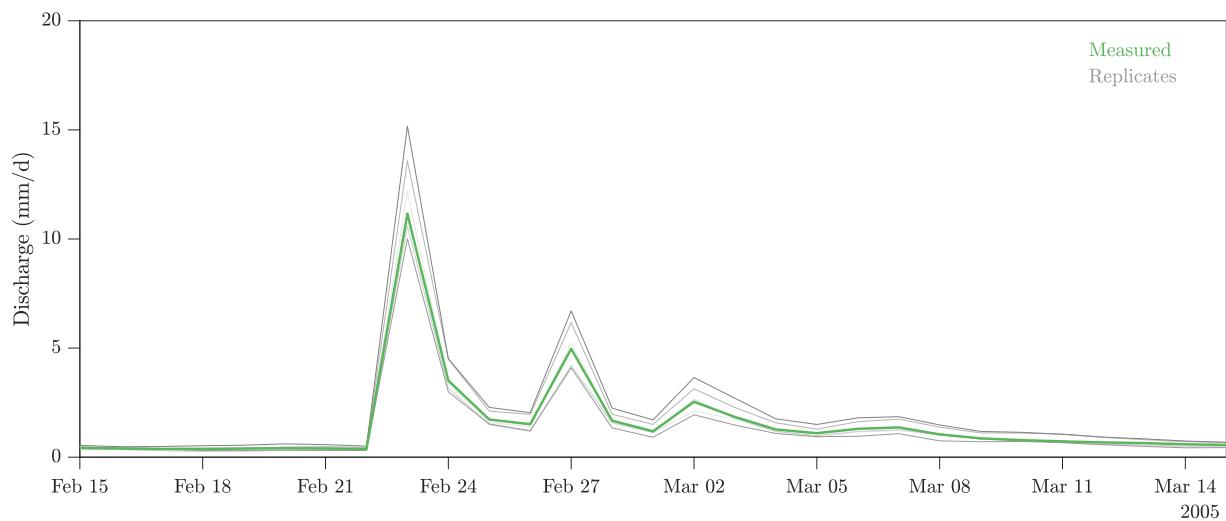


Figure S34: Illustration of the streamflow replicates using $\sigma_{td} = 0.20\tilde{y}_{td}$ and daily streamflow data from the Cowhouse Creek at Pidcock, TX (USGS 08101000), an example of catchment with the *intermittent regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid green line.

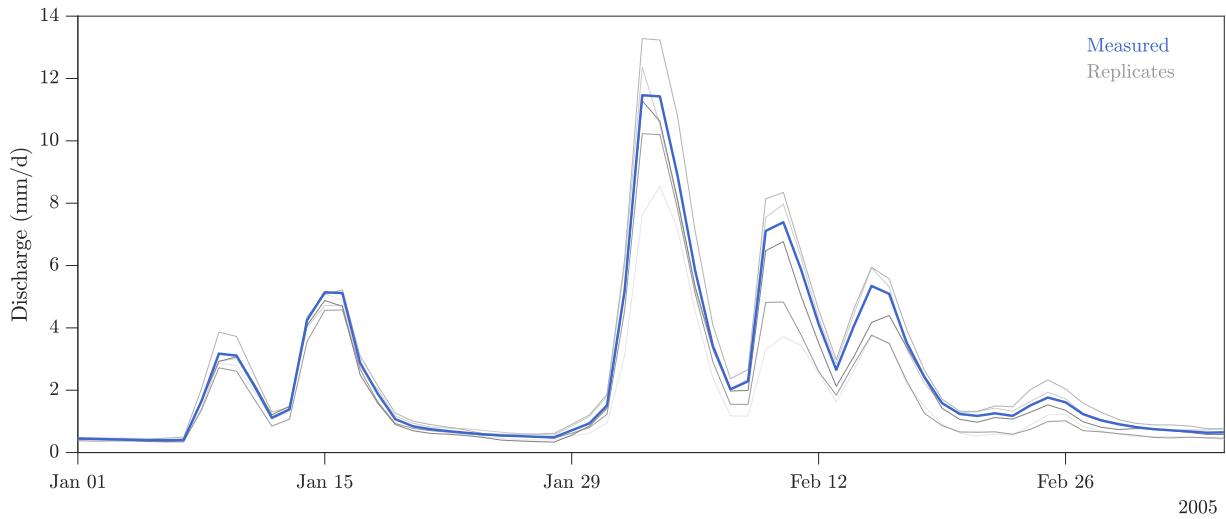


Figure S35: Illustration of the streamflow replicates generated using $\sigma_{td} = 0.20\bar{y}_{td}$ and daily streamflow data from the Leaf River near Collins, MS (USGS 02472000), an example of catchment with a *strong winter regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid blue line.

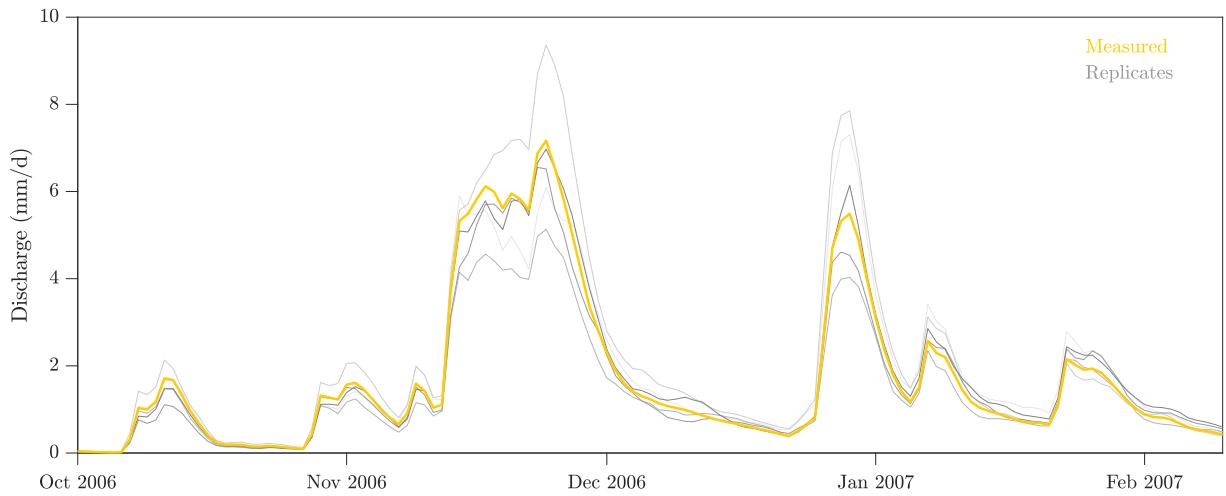


Figure S36: Illustration of the streamflow replicates generated using $\sigma_{td} = 0.20\bar{y}_{td}$ and daily streamflow data from the Potecasi Creek near Union, NC (USGS 02053200), an example of catchment with the *weak winter regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid yellow line.

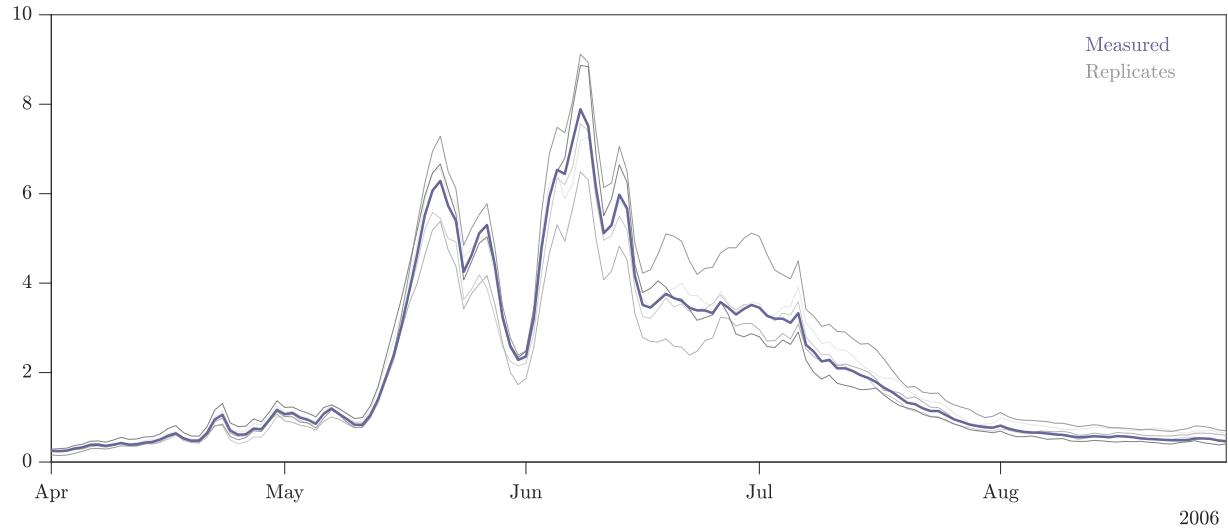


Figure S37: Illustration of the streamflow replicates generated using $\sigma_{td} = 0.20\tilde{y}_{td}$ and daily streamflow data from the South Fork Shoshone River near Valley, WY (USGS 06280300), an example of catchment with the *melt regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid purple line.

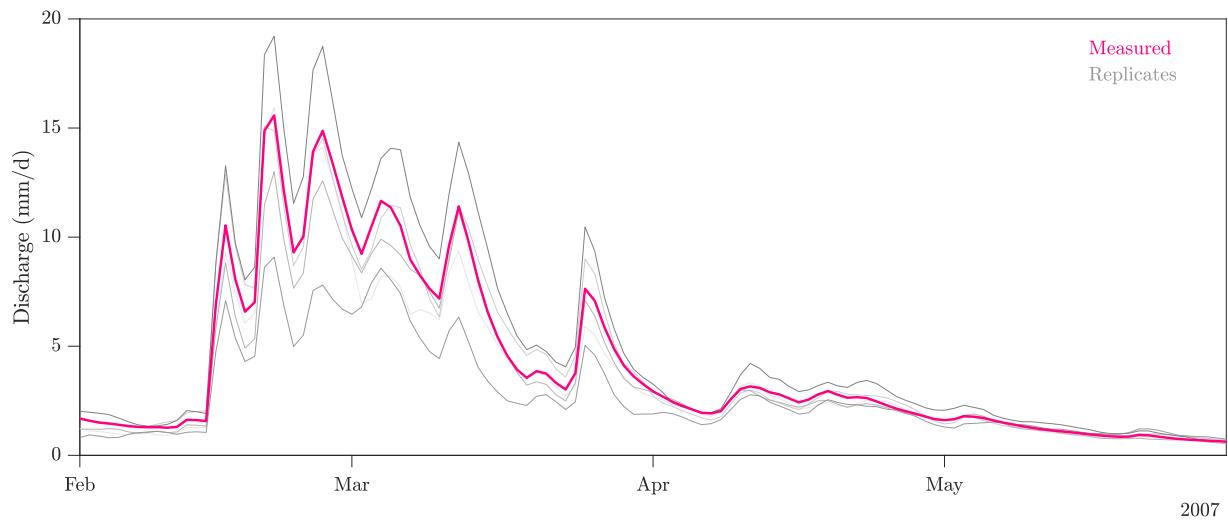


Figure S38: Illustration of the streamflow replicates generated using $\sigma_{td} = 0.20\tilde{y}_{td}$ and daily streamflow data from the Nehalem River near Foss, OR (USGS 14301000), an example of catchment with the *New Year's regime*. (a) 99% confidence intervals (gray region) of the $N = 1,000$ replicates of the discharge record for a representative portion of the 34-year data set. The discharge data are separately indicated with a solid pink line.

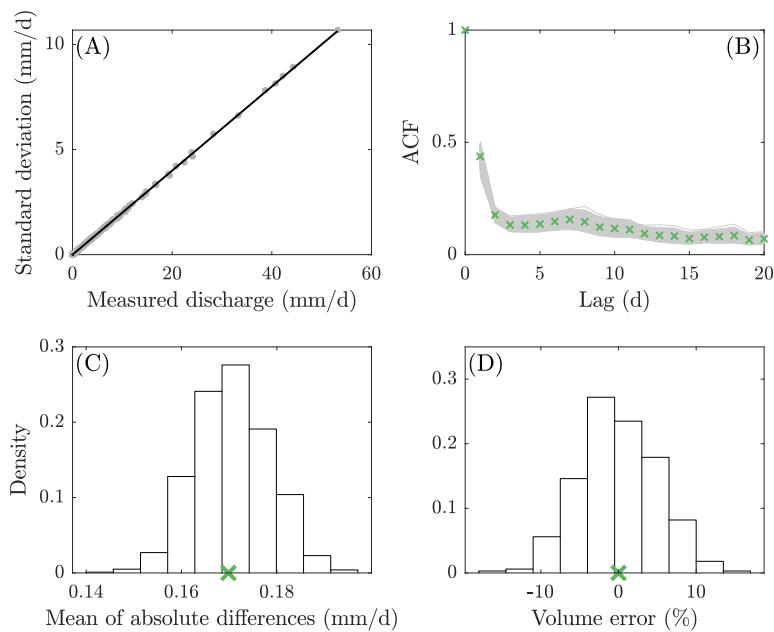


Figure S39: Characteristics of the streamflow replicates for the Cowhouse Creek at Pidcocke, TX (USGS 08101000), an example of catchment with the *intermittent regime*. Replicates of the discharge record are generated using $\sigma_{td} = 0.20\bar{y}_{td}$. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the original record (blue crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The blue crosses highlight the values computed from the original record.

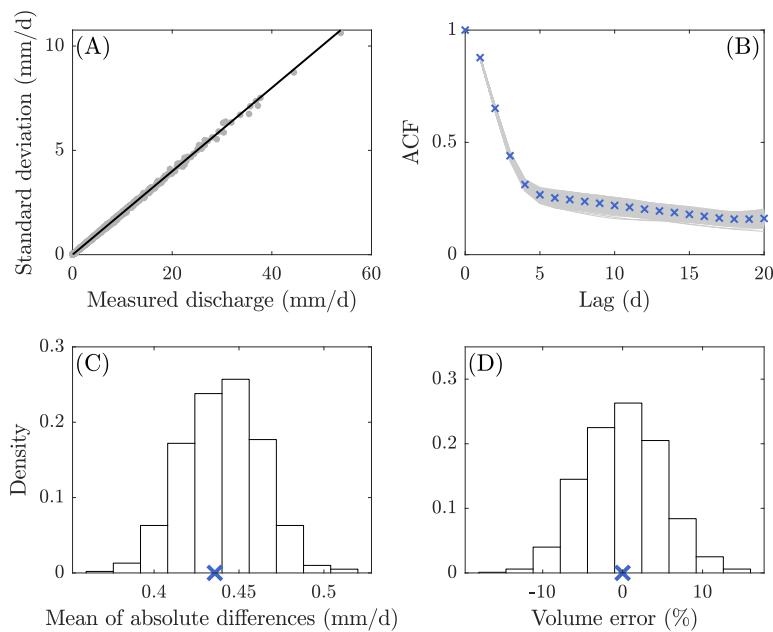


Figure S40: Characteristics of the streamflow replicates for the Leaf River near Collins, MS (USGS 02472000), an example of catchment with a *strong winter regime*. Replicates of the discharge record are generated using $\sigma_{td} = 0.20\bar{y}_{td}$. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the original record (blue crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The blue crosses highlight the values computed from the original record.

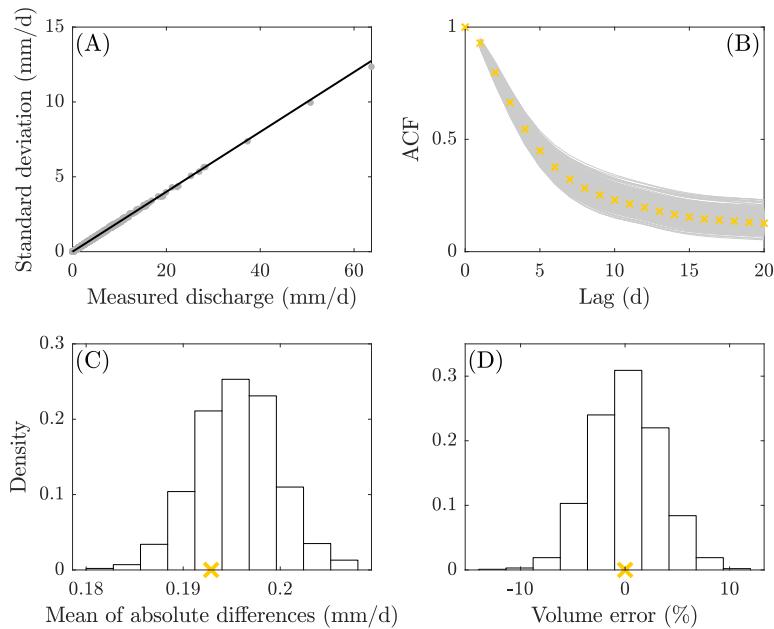


Figure S41: Characteristics of the streamflow replicates for the Potecasi Creek near Union, NC (USGS 02053200), an example of catchment with the *weak winter regime*. Replicates of the discharge record are generated using $\sigma_{td} = 0.20\bar{y}_{td}$. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the original record (blue crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The blue crosses highlight the values computed from the original record.

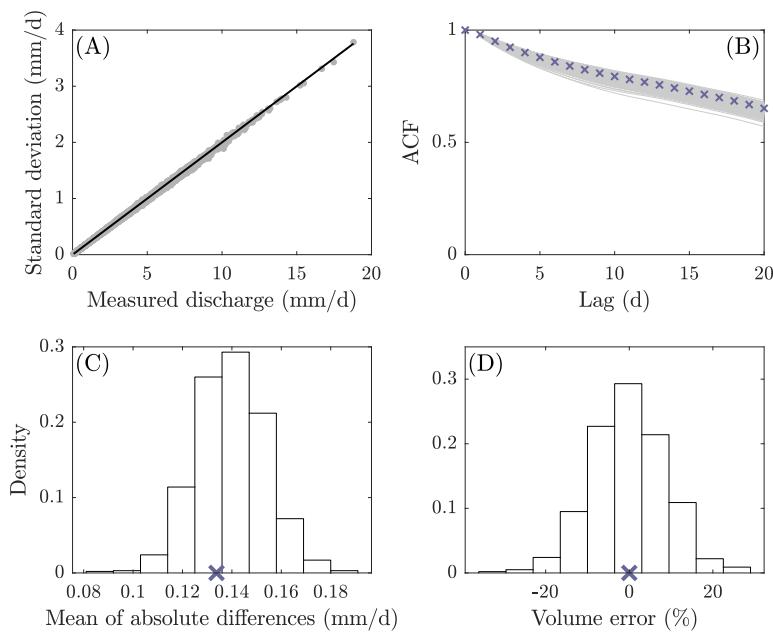


Figure S42: Characteristics of the streamflow replicates for the South Fork Shoshone River near Valley, WY (USGS 06280300), an example of catchment with the *melt regime*. Replicates of the discharge record are generated using $\sigma_{td} = 0.20\tilde{y}_{td}$. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the original record (blue crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The blue crosses highlight the values computed from the original record.

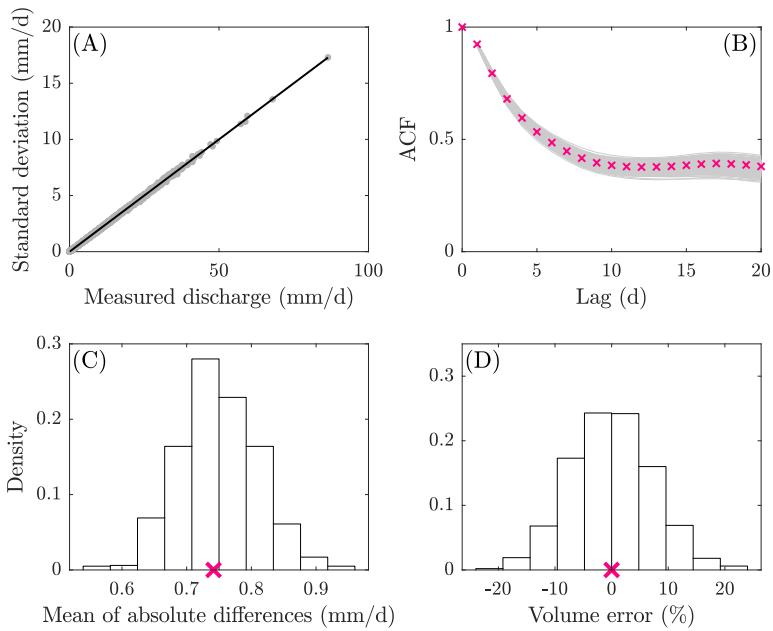


Figure S43: Characteristics of the streamflow replicates for the Nehalem River near Foss, OR (USGS 14301000), an example of catchment with the *New Year's regime*. Replicates of the discharge record are generated using $\sigma_{td} = 0.20\bar{y}_{td}$. (a) Standard deviation of the $N = 1,000$ replicates as a function of discharge. Each gray dot signifies a different data pair. The solid black line signifies the heteroscedastic error model that was used to create the discharge replicates. (b) The ACF of the replicates (gray lines) and the original record (blue crosses). (c) and (d) Frequency distributions of the mean absolute discharge differences and volume error of the thousand replicates. The blue crosses highlight the values computed from the original record.

References

- Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017a). The camels data set: catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 21(10), 5293–5313. Retrieved from <https://www.hydrol-earth-syst-sci.net/21/5293/2017/> doi: 10.5194/hess-21-5293-2017
- Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017b). Catchment attributes for large-sample studies. *Boulder, CO: UCAR/NCAR*. Retrieved from <https://doi.org/10.5065/D6G73C3Q/> doi: 10.5065/D6G73C3Q
- Box, G., Jenkins, G., Reinsel, G., & Ljung, G. (2015). *Time series analysis: Forecasting and control*. Wiley. Retrieved from <https://books.google.com/books?id=rNt5CgAAQBAJ>
- Brunner, M. I., Melsen, L. A., Newman, A. J., Wood, A. W., & Clark, M. P. (2020). Future streamflow regime changes in the united states: assessment using functional classification. *Hydrology and Earth System Sciences*, 24(8), 3951–3966. Retrieved from <https://hess.copernicus.org/articles/24/3951/2020/> doi: 10.5194/hess-24-3951-2020
- Clark, M. P., & Kavetski, D. (2010). Ancient numerical daemons of conceptual hydrological modeling: 1. fidelity and efficiency of time stepping schemes. *Water Resources Research*, 46(10). Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009WR008894> doi: 10.1029/2009WR008894
- Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., ... Hay, L. E. (2008). Framework for understanding structural errors (fuse): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, 44(12). Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007WR006735> doi: 10.1029/2007WR006735
- Falcone, J. A. (2011). *GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow* (Tech. Rep.). USGS. Retrieved 2021-02-28, from <http://pubs.er.usgs.gov/publication/>

70046617 doi: 10.3133/70046617

Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., & Hochreiter, S. (2020). Data for “rainfall-runoff prediction at multiple timescales with a single long short-term memory network”. *Zenodo*. Retrieved from <https://doi.org/10.5281/zenodo.4072701> doi: 10.5281/zenodo.4072701

Hall, P., Kay, J. W., & Titterington, D. M. (1990). Asymptotically optimal difference-based estimation of variance in nonparametric regression. *Biometrika*, 77(3), 521–528. Retrieved from <http://www.jstor.org/stable/2336990> doi: 10.2307/2336990

Kavetski, D., & Clark, M. P. (2010). Ancient numerical daemons of conceptual hydrological modeling: 2. impact of time stepping schemes on model analysis and prediction. *Water Resources Research*, 46(10). Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009WR008896> doi: 10.1029/2009WR008896

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., ... Duan, Q. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous usa: data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrology and Earth System Sciences*, 19(1), 209–223. Retrieved from <https://www.hydrol-earth-syst-sci.net/19/209/2015/> doi: 10.5194/hess-19-209-2015

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of Hydrology*, 303(1), 290-306. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0022169404004056> doi: <https://doi.org/10.1016/j.jhydrol.2004.08.026>

Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). Using bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133(5), 1155 - 1174.

: Retrieved from <https://journals.ametsoc.org/view/journals/mwre/133/5/mwr2906.1.xml> doi: 10.1175/MWR2906.1

Schoups, G., Vrugt, J. A., Fenicia, F., & van de Giesen, N. C. (2010). Corruption of accuracy and efficiency of markov chain monte carlo simulation by inaccurate numerical implementation of conceptual hydrologic models. *Water Resources Research*, 46(10). Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009WR008648> doi: 10.1029/2009WR008648

Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., & Verstraten, J. M. (2005). Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resources Research*, 41(1). Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004WR003059> doi: 10.1029/2004WR003059

Zhou, Y., Cheng, Y., Wang, L., & Tong, T. (2015). Optimal difference-based variance estimation in heteroscedastic nonparametric regression. *Statistica Sinica*, 25, 1377–1397. Retrieved from <http://dx.doi.org/10.5705/ss.2013.010> doi: 10.5705/ss.2013.010