

# <sup>1</sup> Bayesian spectral likelihood for hydrological <sup>2</sup> parameter inference

Bettina Schaefli<sup>1</sup> and Dmitri Kavetski<sup>2</sup>

---

D. Kavetski, School of Civil, Environmental and Mining Engineering, University of Adelaide, Adelaide, SA, Australia

B. Schaefli, Institute of Earth Surface Dynamics, Faculty of Geosciences and Environment, University of Lausanne, 1015 Lausanne, Switzerland (bettina.schaefli@unil.ch)

<sup>1</sup>Institute of Earth Surface Dynamics,  
Faculty of Geosciences and Environment,  
University of Lausanne, Switzerland

<sup>2</sup>School of Civil, Environmental and  
Mining Engineering, University of Adelaide,  
Adelaide, SA, Australia

**Abstract.** This paper proposes a spectral-domain likelihood function for the Bayesian estimation of hydrological model parameters from a time series of model residuals. The spectral-domain error model is based on the Power-Density-Spectrum (PDS) of the stochastic process assumed to describe residual errors. The Bayesian Spectral Likelihood (BSL) is mathematically equivalent to the corresponding Bayesian Time-domain Likelihood (BTL) and yields the same inference when all residual error assumptions are satisfied (and all residual error parameters are inferred). However, the BSL likelihood function does not depend on the residual error distribution in the original time-domain, which offers a theoretical advantage in terms of robustness for hydrological parameter inference. The theoretical properties of BSL are demonstrated and compared to BTL and a previously proposed spectral likelihood by Montanari and Toth (2007), using a set of synthetic case studies and a real case study based on the Leaf River catchment in the US. The empirical analyses confirm the theoretical properties of BSL when applied to heteroscedastic and autocorrelated error models (where heteroscedasticity is represented using the log-transformation and autocorrelation is represented using an AR(1) process). Unlike MTL, the use of BSL did not introduce additional parametric uncertainty compared to BTL. Future work will explore the application of BSL to challenging modeling scenarios in arid catchments and "indirect" calibration with non-concomitant input/output time series.

## 1. Introduction

Bayesian and other likelihood-based inference methods have a strong tradition in hydrological modeling, with the overall goal of providing reliable hydrological predictions and uncertainty estimates [e.g., Kuczera, 1983; Beven and Binley, 1992; Kuczera and Parent, 1998; Bates and Campbell, 2001, and many others]. The key ingredient of likelihood-based inference is the likelihood function, which should provide a probabilistic description of the uncertainty in the model predictions [e.g., Box and Tiao, 1992]. In the simplest case, the likelihood function aims to describe the statistical properties of the model residual errors, i.e., the time series of differences between observed responses (e.g., streamflow) and corresponding model predictions [e.g., Box and Tiao, 1992; Kuczera and Parent, 1998].

A major concern is that, in hydrology, probabilistic inference methods have often been used with ostensibly wrong assumptions [e.g., as noted by Beven and Binley, 1992; Kavetski et al., 2006; Honti et al., 2013, and others]. For example, it is still common for hydrological calibration applications to assume independent and identically distributed Gaussian model residuals, and relatively few studies rigorously assess how well these assumptions are actually satisfied [e.g., Engeland et al., 2005; Schaeffli et al., 2007]. Recent work is addressing these shortcomings, contributing more statistically reliable error models and likelihood functions [e.g., Kuczera, 1983; Kavetski et al., 2006; Schaeffli et al., 2007; Thyer et al., 2009; Schoups and Vrugt, 2010; Smith et al., 2010; Pianosi and Raso, 2012; Evin et al., 2014; McInerney et al., 2017, and many others].

In this paper, we explore new perspectives for hydrological parameter inference by introducing a Bayesian Spectral Likelihood (BSL), based on a statistical description of

45 the model residuals in the spectral-domain, i.e., in the Fourier-transformed-domain rather  
46 than in the time-domain. The term "spectral-domain" rather than "frequency-domain" is  
47 used to emphasize that the proposed likelihood is based on the power-density spectrum.

48 This work is not the first attempt to use spectral methods in hydrological modeling.  
49 For example, Montanari and Toth [2007] applied the Whittle likelihood [Whittle, 1953]  
50 to calibrate hydrological model parameters. Other studies used spectral likelihoods or  
51 spectral signatures in more informal settings. Quets et al. [2010] used the sum of squared  
52 differences between the Fourier amplitudes, or between the Fourier amplitudes and phases  
53 of observed and simulated streamflow time series, to calibrate the SWAT model. A sim-  
54 ilar approach was followed by Pauwels and De Lannoy [2011] and De Vleeschouwer and  
55 Pauwels [2013]. Moussu et al. [2011] used the root-mean-squared difference between the  
56 estimated autocorrelation functions of the observed and simulated streamflow series to  
57 calibrate two conceptual rainfall-runoff models of a karst system. Winsemius et al. [2009]  
58 and Hartmann et al. [2013] used the streamflow autocorrelation function as a signature to  
59 assess the model performance in a multi-criteria model identification setting. Schaefli and  
60 Zehe [2009] proposed to assess hydrologic model performance in terms of the Kolmogorov-  
61 Smirnov distance between the estimated wavelet power spectra of observed and simulated  
62 streamflow series. Several studies proposed to use spectral calibration for non-concomitant  
63 (or indirect) calibration, where input and output observations are not available over the  
64 same time period [Montanari and Toth, 2007; De Vleeschouwer and Pauwels, 2013].

65 The main motivation for this paper is to present the key theoretical aspects of spectral  
66 parameter inference, especially in light of recent interest in spectral model calibration and  
67 performance assessment. Our paper addresses the current research gap that the major-

68 ity of hydrological calibration approaches based on spectral techniques do not explicitly  
69 articulate the probabilistic assumptions underlying their choice of objective (likelihood)  
70 function; this limitation complicates the derivation of probability limits on the estimated  
71 model parameters and predictions. The properties of BSL are investigated using a series  
72 of synthetic and real data case studies, and are compared to the properties of the corre-  
73 sponding (standard) time-domain likelihood and of the spectral-domain likelihood used  
74 previously by Montanari and Toth [2007]. The method of Montanari and Toth [2007]  
75 is of particular relevance to this work, because to our knowledge it is the only spectral  
76 calibration method in the hydrological literature that uses a spectral likelihood function  
77 with an explicit probabilistic interpretation.

78 The BSL approach introduced in this work is obtained by expressing the probability  
79 density function (pdf) of residual errors of a hydrological model in terms of their Fourier  
80 power-density spectrum [e.g., Jenkins and Watts, 1968]. This spectral-domain probabilis-  
81 tic characterization is presented in considerable detail because, despite spectral analysis  
82 being widely used in time series analysis, it remains relatively rare in hydrology, and ex-  
83 isting literature generally does not describe the pdf of the entire power-density spectrum.

84 The remainder of the paper is structured as follows. Section 2 presents all required  
85 definitions and the derivation of BSL for common stochastic error models. This section  
86 also briefly outlines the relationship of BSL to the likelihood presented by Montanari and  
87 Toth [2007]. Section 3 details the case studies and the analysis methodology. Section 4  
88 presents and discusses the case study results. Section 6 summarizes the key conclusions of  
89 the paper, outlines some important open questions and suggests future research directions.

90

## 2. Theoretical development

### 2.1. Bayesian time-domain likelihood (BTL)

Consider a hydrological model  $\mathbf{H}$

$$\widehat{\mathbf{Y}} = \mathbf{H}(\boldsymbol{\theta}, \mathbf{X}) \quad (1)$$

where  $\mathbf{X} = (\mathbf{X}_t)_{t=1,\dots,N}$  are the system inputs (e.g., rainfall and potential evapotranspiration) at time steps  $t = 1, \dots, N$ ,  $\widehat{\mathbf{Y}} = (\widehat{\mathbf{Y}}_t)_{t=1,\dots,N}$  is the system output predicted by the model (e.g. streamflow), and  $\boldsymbol{\theta}$  is a vector of model parameters. On overview of all used mathematical notations is given in Table 1.

In practice, the true system input  $\mathbf{X}$  is unknown, and we only have observed inputs  $\widetilde{\mathbf{X}}$ , which are affected by sampling and measurements errors. In this paper, we represent total predictive uncertainty using residual errors, which are assumed to aggregate the effects of all sources of error including data uncertainty and model structural errors. We do not attempt error decomposition, i.e., to model individual sources of error using separate error models [e.g., Kavetski et al., 2006; Renard et al., 2011].

The simulated system output  $\widehat{\mathbf{Y}} = (\widehat{\mathbf{Y}}_t)_{t=1,\dots,N}$  differs from the observed system output,  $\widetilde{\mathbf{Y}} = (\widetilde{\mathbf{Y}}_t)_{t=1,\dots,N}$ , for several reasons: i) errors in the observed system inputs, e.g., raingauge sampling errors [Renard et al., 2011; McMillan et al., 2011]; ii) errors in the observed system output, e.g., rating curve errors [Thyer et al., 2009; McMillan and Westerberg, 2015]; iii) structural errors in the model equations, e.g., due to the inability of lumped models to represent spatially distributed processes, incomplete representation of dominant hydrological processes, etc. [Beven and Binley, 1992; Kuczera et al., 2006]; (iv) parameter errors, including those due to the uncertainty arising from finite-length and uncertain calibration data, due to limitations of parameter optimization algorithms, etc.

Consider the vector of model residuals  $\boldsymbol{\varepsilon}$ ,

$$\boldsymbol{\varepsilon} = q(\tilde{\mathbf{Y}}) - q(\hat{\mathbf{Y}}) \quad (2)$$

110 where we allow for a response transformation  $q()$  [e.g., logarithmic, see McInerney et al.,  
111 2017].

By construction, Equation 2 lumps all sources of error in the residual error term. Therefore, in this case, likelihood-based inference requires the specification of a statistical model of the residual errors, i.e., a "residual error model",

$$\boldsymbol{\varepsilon} \sim \mathcal{E}(\hat{\mathbf{Y}}, \boldsymbol{\vartheta}) = \mathcal{E}(\mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\vartheta}), \quad (3)$$

112 where  $\boldsymbol{\vartheta}$  denotes the error model parameters (which can be inferred or fixed a priori).

Hydrological model residuals are often well-described by Gaussian AR(1) processes ("red noise") [Schaefli et al., 2007; Evin et al., 2013; Li et al., 2013].

$$\boldsymbol{\varepsilon}_t = \mu_\varepsilon + \rho(\boldsymbol{\varepsilon}_{t-1} - \mu_\varepsilon) + \delta_t, \delta_t \sim \text{NID}(\mu_\delta, \sigma_\delta^2) \quad (4)$$

113 where  $\delta_t$  is the innovation at time step  $t$ ,  $\rho$  is the (lag-1) autoregressive parameter,  
114  $\mu_\delta$  is the innovation mean,  $\sigma_\delta^2$  the innovation variance,  $\mu_\varepsilon$  the residual mean and NID  
115 denotes the independent Gaussian distribution. To assist in the residual error analysis, it  
116 is convenient to define the innovation mean,  $\mu_\delta = \mu_\varepsilon(1 - \rho)$ .

The BTL corresponding to AR(1) residuals with Gaussian innovations is:

$$p(\tilde{\mathbf{Y}} \mid \mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \frac{\sqrt{1 - \rho^2}}{(2\pi\sigma_\delta^2)^{N/2}} \exp \left[ -\frac{1}{2\sigma_\delta^2} \left\{ (1 - \rho^2)(\boldsymbol{\varepsilon}_1 - \mu_\varepsilon)^2 + \sum_{t=2}^N (\boldsymbol{\varepsilon}_t - \mu_\varepsilon - \rho(\boldsymbol{\varepsilon}_{t-1} - \mu_\varepsilon))^2 \right\} \right]. \quad (5)$$

117 where  $\boldsymbol{\varepsilon}_t$  are the raw residuals computed from Equation 2 [see e.g. Priestley, 1981, for  
118 the derivation of the probability density of Gaussian AR(1) processes].

119 In many cases, the assumption of Gaussian errors is not supported by residual analysis  
 120 [e.g., Schoups and Vrugt, 2010]. An alternative assumption that we consider in this  
 121 study is that the innovations follow a Laplace distribution, with pdf  $f_{\text{Laplace}}(x|\mu, b) =$   
 122  $1/(2b)\exp(-|x - \mu|/b)$ , where the standard deviation is  $\sqrt{2}b$ .

The BTL corresponding to AR(1) residual errors with Laplacian innovations is

$$p(\tilde{\mathbf{Y}} | \mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \sqrt{1 - \rho^2} \left( \frac{\sqrt{2}}{2\sigma_\delta} \right)^N \exp \left[ -\frac{\sqrt{2}}{\sigma_\delta} \left\{ \sqrt{1 - \rho^2} |\boldsymbol{\varepsilon}_1 - \mu| + \sum_{t=2}^N |\boldsymbol{\varepsilon}_t - \mu - \rho(\boldsymbol{\varepsilon}_{t-1} - \mu)| \right\} \right], \quad (6)$$

123 For details of more general AR( $n$ ) processes, see Box and Jenkins [1976, p. 274ff].

## 2.2. Spectral-domain: basic concepts

124 To derive the likelihood of the model residuals in the spectral-domain rather than in the  
 125 time-domain, we need the same key ingredients as for BTL: (i) a residual error model in  
 126 the spectral-domain (e.g., based on the power-density spectrum of the stochastic process  
 127 assumed to describe the residuals), (ii) a parametric description of the probability distri-  
 128 bution function associated with this error model, and (iii) spectral-domain realizations of  
 129 the model residuals (either obtained directly in the spectral-domain or from a transform  
 130 of time-domain realizations). The derivation of these ingredients is presented next.

131 In the derivations to follow, it is important to distinguish between the application of  
 132 Fourier transforms to deterministic vs stochastic processes. In general, we use lower  
 133 case symbols (e.g.,  $\mathbf{z}$ ) to denote deterministic processes and realizations (samples) from  
 134 stochastic processes and upper case symbols (e.g.,  $\mathbf{Z}$ ) to denote stochastic processes (for  
 135 example, the process  $\mathbf{Z}$  that generated the realization  $\mathbf{z}$ ).

Consider a deterministic process  $\mathbf{z} = (z_t)_{t=-\infty, \dots, \infty}$ , defined over discrete time steps  $t$  of length  $\Delta t$ . If the process is absolutely summable [Oppenheim and Schaffer, 1989, p. 47],



its discrete time Fourier transform can be written as

$$\mathbb{F}_\omega[\mathbf{z}] = \Delta t \sum_{t=-\infty}^{\infty} z_t \exp(-it\omega\Delta t) \quad (7)$$

136 where  $i = \sqrt{-1}$ .

137 The discrete time Fourier transform  $\mathbb{F}_\omega$  is a vector of complex numbers. Its components  
138 are indexed by the angular frequency  $\omega$  [rad/T] [e.g., Oppenheim and Schafer, 1989, p.  
139 698].

140 Next, consider a stochastic process  $\mathbf{Z} = (Z_t)_{t=-\infty, \dots, \infty}$ . In this case, the Fourier transform  
141  $\mathbb{F}_\omega[\mathbf{Z}]$  is itself a stochastic process, obtained as a derived distribution. In particular, the  
142 Fourier transform maps a time series of random variables  $\mathbf{Z}$  into a set of frequency-ordered  
143 random variables  $\mathbb{F}_\omega[\mathbf{Z}]$ . This can be seen by considering the application of the Fourier  
144 transform to a set of individual time series  $\mathbf{z}$  sampled from a stochastic process  $\mathbf{Z}$ , and  
145 then considering the distribution of the set of transformed time series  $\mathbb{F}_\omega[\mathbf{z}]$ .

Stochastic processes can be analyzed using the power-density spectrum (PDS), defined  
as the Fourier transform of the autocovariance function of  $\mathbf{Z}$  [Oppenheim and Schafer,  
1989, p. 843]:

$$\mathbb{P}_\omega[\mathbf{Z}] = \mathbb{F}_\omega \left[ v_\ell[\mathbf{Z}] \right] \quad (8)$$

146 where  $v_\ell[\mathbf{Z}] = \mathbb{E}[Z_t \text{ conj}(Z_{t+\ell})]$  is the autocovariance function of process  $\mathbf{Z}$ ,  $\mathbb{E}[\cdot]$  is the  
147 expectation operator,  $\ell$  is the lag [T] and  $\text{conj}(\cdot)$  denotes complex conjugation. Note that  
148 the autocovariance function is often referred to as the "autocorrelation sequence" in the  
149 signal processing literature [e.g., Oppenheim and Schafer, 1989, p. 743].

Assuming process ergodicity, the PDS of a stochastic process can be related to the  
expectation of the Fourier transform of the stochastic process. For a finite-domain process

$\mathbf{Z}^{(N)} = (Z_t)_{t=1,\dots,N}$ , it holds that [Oppenheim and Verghese, 2015, chapter 11]

$$\mathbb{P}_\omega[\mathbf{Z}^{(N)}] = \frac{1}{N} \mathbb{E} \left[ \left| \mathbb{F}_\omega[\mathbf{Z}^{(N)}] \right|^2 \right], \quad (9)$$

150 where  $|\cdot|$  denotes the absolute value (or, more generally, the complex modulus). The  
 151 PDS of an infinite-domain process is obtained by taking the limit  $N \rightarrow \infty$  in Equation 9.

152 In the signal processing literature, both formulations of the PDS are attributed to  
 153 Wiener-Khinchin; in recent literature, the formulation in Equation 9 is referred to as the  
 154 Einstein-Wiener-Khinchin theorem [Oppenheim and Verghese, 2015, chapter 11].

155 We stress that, unlike the Fourier transform of a stochastic process, the PDS of a  
 156 stochastic process is a deterministic quantity: it is defined either in terms of the autocorrelation  
 157 variance function (Equation 8) or in terms of expectations (Equation 9).

### 2.3. The PDS variate, PDSV

158 Given a finite-length sample  $\mathbf{z}^{(N)} = (z_t)_{t=1,\dots,N}$ , the Fourier transform  $\mathbb{F}_\omega[\mathbf{z}^{(N)}]$  provides  
 159 a spectral-domain sample of the discrete-time Fourier transform of the entire infinite-  
 160 domain process  $\mathbf{Z}$  [Oppenheim and Schaffer, 1989, p. 695].

Due to the finite length of the sample, the mapping between the time- and spectral-  
 domain is possible only at a finite number of frequencies  $\omega_j$ . These frequencies are given  
 by integer multiples of the fundamental frequency  $\omega_f$ :

$$\omega_j = j\omega_f; j = 0, \dots, N - 1 \quad (10)$$

$$\omega_f = \frac{2\pi}{N\Delta t} \quad (11)$$

161 Accordingly, we will use the subscript  $j$  as the index of the discrete-time finite-sample  
 162 Fourier transform. Since the exponential basis functions used by the Fourier transform are

163 orthogonal and complex-valued,  $N/2$  frequencies are sufficient to describe the  $N$  elements  
 164 of  $\mathbf{z}^{(N)}$  and hence we have  $j = 0, \dots, N/2 - 1$ .

We now define a transformation of a stochastic process such that the expected value of this transformation is the PDS of the original stochastic process. To this end, we define the "power-density spectrum variate" (PDSV) corresponding to a finite-length process  $\mathbf{Z}^{(N)} = (Z_t)_{t=1, \dots, N}$ , as:

$$\mathbb{Q}_j[\mathbf{Z}^{(N)}] = \frac{1}{N\Delta t} \left| \mathbb{F}_j \left[ \mathbf{Z}^{(N)} \right] \right|^2. \quad (12)$$

165 The following observations can be made:

166 1. The probability distribution  $p(\mathbb{Q}[\mathbf{Z}^{(N)}])$  is a derived distribution that depends on  
 167  $p(\mathbf{Z}^{(N)})$ ;

168 2. The power-density spectrum  $\mathbb{P}_j[\mathbf{Z}^{(N)}]$  is the expectation of  $\mathbb{Q}_j[\mathbf{Z}^{(N)}]$ , i.e.  $\mathbb{P}_j[\mathbf{Z}^{(N)}] =$   
 169  $\mathbb{E} \left[ \mathbb{Q}_j[\mathbf{Z}^{(N)}] \right]$ ;

3. A sample from the distribution of  $\mathbb{Q}_j[\mathbf{Z}^{(N)}]$  can be obtained by applying the transformation in Equation 12 to a time-domain sample  $\mathbf{z}^{(N)}$  from the stochastic process  $\mathbf{Z}^{(N)}$ .

$$P_j[\mathbf{z}^{(N)}] = \frac{1}{N\Delta t} \left| \mathbb{F}_j \left[ \mathbf{z}^{(N)} \right] \right|^2. \quad (13)$$

170 The quantity  $P_j[\mathbf{z}^{(N)}]$  is often called the periodogram of  $\mathbf{z}^{(N)}$  (usually defined with  
 171  $\Delta t = 1$ ) [Oppenheim and Schaffer, 1989]. The periodogram can be seen to represent a  
 172 "single sample" estimator of the mean of the power-density spectrum.

## 2.4. Statistical properties of quantities in the spectral-domain

173 We now turn our attention to the probability distributions of quantities in the spectral-  
 174 domain.

175 For  $j \geq 1$ , the quantity  $\left| \mathbb{F}_j \left[ \mathbf{Z}^{(N)} \right] \right|^2$  is known to have the  $\chi^2$ -distribution with two  
 176 degrees of freedom, i.e. an exponential distribution with pdf  $f_{\text{exp}}(x|\beta) = 1/\beta \exp(-x/\beta)$ ,  
 177 where  $E[x] = \beta$  and  $\text{var}[x] = \beta^2$  [e.g. Bartlett, 1950]. The periodogram at  $j = 0$  follows the  
 178  $\chi^2$  distribution with 1 degree of freedom (see below). These distributional properties of  
 179 the PDS variate have important implications for parameter inference, and are elaborated  
 180 in further detail in Section 2.10.

Since  $P_j[\mathbf{z}^{(N)}]$  represents a sample from  $\mathbb{Q}_j[\mathbf{Z}^{(N)}]$ , and  $\mathbb{P}_j[\mathbf{Z}^{(N)}]$  is by definition the  
 expected value of  $\mathbb{Q}_j[\mathbf{Z}]$ , we can express the probability density of  $P_j[\mathbf{z}^{(N)}]$  for  $j > 0$  using  
 the exponential pdf with mean  $\mathbb{P}_j[\mathbf{Z}] = E[\mathbb{Q}_j[\mathbf{Z}^{(N)}]]$ :

$$p(P_j[z]|\mathbb{P}_j[\mathbf{Z}]) = f_{\text{exp}}(P_j[z]|\mathbb{P}_j[\mathbf{Z}]); j = 1, \dots, N/2 - 1, \quad (14)$$

181 where, for simplification, we omitted the time-domain superscript ( $N$ ) on  $Z$ .

The probability distribution of  $\mathbb{Q}_j[\mathbf{Z}^{(N)}]$  at  $j = 0$  is the scaled  $\chi^2$  distribution with 1  
 degree of freedom,

$$p(P_j[z]|\mathbb{P}_j[\mathbf{Z}]) = f_{\chi_1^2}(P_j[z]|\mathbb{P}_j[\mathbf{Z}]); j = 0, \quad (15)$$

182 where  $f_{\chi_1^2}(x|\beta) = \frac{1}{\beta} \frac{1}{\sqrt{2\pi(x/\beta)}} \exp[-(x/\beta)/2]$ .

183 Unlike the exponential distribution, for the  $\chi^2$  distribution with 1 degree of freedom,  
 184 we have  $E[x] = \beta$  and  $\text{var}[x] = 2\beta^2$ . Inspection of the  $\chi^2$  distribution with 1 degree  
 185 of freedom indicates that it corresponds to the distribution of the square of a Gaussian  
 186 variate (hence it can be referred to as a "squared-Gaussian" distribution).

187 Having derived the distribution of the PDSV for all frequencies ( $p(P_j[z]|\mathbb{P}_j[\mathbf{Z}])$ ) in equa-  
 188 tions 14 and 15, and having a method to compute samples from the PDSV (the peri-

189 odogram in Equation 13), the remaining step is to obtain an expression for the PDS  
 190  $\mathbb{P}_j[\mathbf{Z}]$  of specific stochastic processes used to describe residual errors.

## 2.5. PDS of uncorrelated processes

For simple uncorrelated processes, the following relations hold:

$$\mathbb{P}_0[\mathbf{Z}] = \Delta t \sigma_Z^2 + N \Delta t \mu_Z^2 \quad (16)$$

$$\mathbb{P}_j[\mathbf{Z}] = \Delta t \sigma_Z^2; j = 1, \dots, N/2 - 1. \quad (17)$$

191 Equation 17, obtained from Parseval's theorem [Jenkins and Watts, 1968], is a well-  
 192 known result and generally referred to as the "mean value of the periodogram".

193 Equation 16 is less known. In fact, the PDS spike at zero frequency of any process  
 194 with non-zero constant mean is often discarded, e.g., in the Whittle estimator [Whittle,  
 195 1953] used by Montanari and Toth [2007], and in the mixed time-domain spectral-domain  
 196 calibration presented by Morlando et al. [2016]. An example where it is explicitly included  
 197 is the analysis of De Vleeschouwer and Pauwels [2013].

## 2.6. Bayesian spectral likelihood for uncorrelated processes

The Bayesian spectral likelihood (BSL) of time-domain observations  $\tilde{\mathbf{Y}}$  is obtained as the joint probability of the corresponding spectral-domain residual realizations  $P[\boldsymbol{\varepsilon}]$  at all frequencies. When the residual errors are assumed to be uncorrelated, we obtain,

$$\begin{aligned} p(\tilde{\mathbf{Y}} \mid \mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) &= p(P_0[\boldsymbol{\varepsilon}] \mid \mathbb{P}_0[\boldsymbol{\varepsilon}]) \prod_{j=1}^{N/2-1} p(P_j[\boldsymbol{\varepsilon}] \mid \mathbb{P}_j[\boldsymbol{\varepsilon}]) \\ &= f_{\chi_1^2}(P_0[\boldsymbol{\varepsilon}] \mid \mathbb{P}_0[\boldsymbol{\varepsilon}]) \prod_{j=1}^{N/2-1} f_{\text{exp}}(P_j[\boldsymbol{\varepsilon}] \mid \mathbb{P}_j[\boldsymbol{\varepsilon}]), \end{aligned} \quad (18)$$

198 where  $\mathbb{P}_0[\boldsymbol{\varepsilon}]$  is given in Equation 16 and  $\mathbb{P}_j[\boldsymbol{\varepsilon}]$  is given in Equation 17.

199 Note that the PDSV is obtained from the actual realization of residuals  $\boldsymbol{\varepsilon}$  (Equation 2),  
 200 whereas the PDS is a property of the underlying stochastic process  $\boldsymbol{\mathcal{E}}$  (e.g., Gaussian).

201 Substituting the expressions for  $\mathbb{P}_0[\boldsymbol{Z}]$  and  $\mathbb{P}_j[\boldsymbol{Z}]$  from Equation 16 and Equation 17,  
 202 and assuming  $\Delta t = 1$  and  $\mu_{\boldsymbol{\mathcal{E}}}^2 = 0$ , yields the BSL for zero-mean white noise with variance  
 203  $\sigma_{\delta}$ :

$$p(\tilde{\boldsymbol{Y}} \mid \boldsymbol{X}, \boldsymbol{\theta}, \sigma_{\delta}) = f_{\chi_1^2}(P_0[\boldsymbol{\varepsilon}]|\sigma_{\delta}^2) \prod_{j=1}^{N/2-1} f_{\text{exp}}(P_j[\boldsymbol{\varepsilon}]|\sigma_{\delta}^2), \quad (19)$$

204 where the definitions of  $f_{\text{exp}}()$  and  $f_{\chi_1^2}()$  are as given in Section 2.3.

## 2.7. PDS of autocorrelated processes

205 As the residuals of hydrological errors are typically highly autocorrelated, it is of interest  
 206 to consider the PDS of autocorrelated processes.

207 Regardless of the autocorrelation structure, Equation 16 holds for frequency  $j = 0$ .  
 208 However, for  $j > 0$ , the PDS of an autocorrelated process depends on  $j$ .

For Gaussian red noise, AR(1), it can be shown that [Brockwell and Davis, 1987]

$$\mathbb{P}_j[\boldsymbol{Z}] = \frac{\Delta t \sigma_Z^2}{\rho^2 \sin^2(\omega_j) + [1 - \rho \cos(\omega_j)]^2}. \quad (20)$$

More generally, the PDS of any stationary process  $\boldsymbol{Z}$  with finite variance and linear au-  
 tocorrelation structure (e.g., with moving-average components, non-Gaussian innovations,  
 etc.) can be expressed using a "profile" function,  $\zeta_j^Z(\boldsymbol{\vartheta}_{\zeta})$  [Fox and Taqqu, 1986]:

$$\mathbb{P}_j[\boldsymbol{Z}^{(N)}] = \Delta t \sigma_Z^2 \zeta_j^Z(\boldsymbol{\vartheta}_{\zeta}). \quad (21)$$

209 where  $\boldsymbol{\vartheta}_{\zeta}$  are the parameters of the autocorrelation structure.

210 The profile function  $\zeta_j^Z$  is a function of the frequency index  $j$  and depends solely on the  
 211 autocorrelation structure of the process; it depend neither on the probability distribution  
 212 e innovations, nor on the variance of the innovations.

213 Equation 20 can be derived from the general Equation 21, with parameters  $\boldsymbol{\vartheta}_\zeta = \rho$   
 214 [Box et al., 1994]. The corresponding expressions for other autoregressive processes can  
 215 be found in references such as Box et al. [1994] and others.

## 2.8. General Bayesian spectral likelihood

216 The general BSL formulation is obtained by expressing the joint probability distribution  
 217 of the PDS variate (PDSV) at all frequencies  $j = 0, \dots, N/2 - 1$ :

$$p(\tilde{\mathbf{Y}} \mid \mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \prod_{j=0}^{N/2-1} f_j(P_j[\boldsymbol{\varepsilon}] \mid \mathbb{P}_j[\boldsymbol{\mathcal{E}}]). \quad (22)$$

218 Recalling the different form of the probability distribution for frequency  $j = 0$  (Equa-  
 219 tion 15) than for frequencies  $j > 0$  (Equation 14), the above equation becomes:

$$p(\tilde{\mathbf{Y}} \mid \mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = f_{\chi_1^2}(P_0[\boldsymbol{\varepsilon}] \mid \mathbb{P}_0[\boldsymbol{\mathcal{E}}]) \prod_{j=1}^{N/2-1} f_{\text{exp}}(P_j[\boldsymbol{\varepsilon}] \mid \mathbb{P}_j[\boldsymbol{\mathcal{E}}]), \quad (23)$$

$\mathbb{P}_0$  is given by Equation 16 and  $\mathbb{P}_j$  by Equation 21:

$$\mathbb{P}_0[\boldsymbol{\mathcal{E}}] = \Delta t \sigma_\delta^2 + N \Delta t \mu_\varepsilon^2 \quad (24)$$

$$\mathbb{P}_j[\boldsymbol{\mathcal{E}}] = \Delta t \sigma_\delta^2 \zeta_j^\boldsymbol{\varepsilon}[\boldsymbol{\vartheta}_\zeta]. \quad (25)$$

220 The above equations hold for any homoscedastic (constant-variance) residual model,  
 221 independent of the residual distribution. For hydrological model residuals, the main  
 222 focus will be on AR(1) residual models, for which  $\zeta_j^\boldsymbol{\varepsilon}$  is given by Equation 20. Note

223 that the homoscedasticity assumption can be addressed by using transformations such as  
 224 logarithmic or Box-Cox when calculating the residuals in Equation 2 (see below).

## 2.9. Incorporation of BSL into a full Bayesian framework

The preceding Section 2.6 and Section 2.8 derived the likelihood function  $p(\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}, \boldsymbol{\theta}, \boldsymbol{\vartheta})$  for Bayesian spectral-domain inference. The Bayesian posterior distribution  $p(\boldsymbol{\theta}, \boldsymbol{\vartheta}|\tilde{\mathbf{Y}}, \tilde{\mathbf{X}})$  is then obtained by specifying a prior distribution for all inferred quantities,  $p(\boldsymbol{\theta}, \boldsymbol{\vartheta})$ ,

$$p(\boldsymbol{\theta}, \boldsymbol{\vartheta}|\tilde{\mathbf{Y}}, \tilde{\mathbf{X}}) \propto p(\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) \times p(\boldsymbol{\theta}, \boldsymbol{\vartheta}) \quad (26)$$

225 The specification of the prior distribution allows incorporating existing (approximate)  
 226 knowledge of hydrological model parameters, eg, based on previous investigations [e.g.,  
 227 Viglione et al., 2013], theoretical constraints, as well as estimates of error model param-  
 228 eters from auxiliary studies such as rainfall and rating curve error analysis [e.g., Renard  
 229 et al., 2011]. In the simplest instance where such additional information is not avail-  
 230 able, such as in the case studies of this paper, a uniform prior distribution can be used,  
 231  $p(\boldsymbol{\theta}, \boldsymbol{\vartheta}) \propto \text{const.}$

In general, the likelihood function must account for any data transformations, such as the logarithmic or Box-Cox transformations often used to stabilize the error variance,

$$p(\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \det J_q(\tilde{\mathbf{Y}}) \times p(\boldsymbol{\mathcal{E}}|\tilde{\mathbf{X}}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) \quad (27)$$

232 where  $\det J_q$  denotes the Jacobian determinant of transformation  $q$ , e.g., in the case  
 233 of the logarithmic transformation  $q(y) = \log y$  used in Section 3.4, we have  $\det J_q(\tilde{y}) =$   
 234  $\prod_{t=1}^N 1/\tilde{y}$  [e.g., see McInerney et al., 2017]. Unless the data transformation includes fitted  
 235 parameters (e.g., the Box-Cox transformation applied with a fitted rather than fixed value



236 of the power parameter  $\lambda$ ), the Jacobian term is constant with respect to the inferred  
237 quantities  $(\boldsymbol{\theta}, \boldsymbol{\vartheta})$ , and can be treated as part of the proportionality constant in Equation 26.

### 2.10. Theoretical advantages of BSL

238 An interesting and useful feature of BSL is that its likelihood function, given in Equa-  
239 tion 23, is "almost" (asymptotically for large  $N$ ) independent from the distribution of  
240 residuals in the original (time) domain. This behavior arises from the  $\chi^2$  form of the  
241 probability distribution of the PDS variate (PDSV), used to derive BSL (see equations 14  
242 and 15).

243 Intuitively, the asymptotic properties of the PDSV can be related to the mathemati-  
244 cal form of the Fourier transform, which is defined as a sum of a series of variables (see  
245 Equation 7). When these variables are random, as is the case when the Fourier transform  
246 is applied to a stochastic process, the Central Limit Theorem results in an asymptotic  
247 convergence to a Gaussian distribution almost irrespective of the distribution of the in-  
248 dividual terms in the sum [see Brillinger, 1981; Cohen, 1998, for details]. Next, when we  
249 consider the definition of the power-density spectrum variate (PDSV,  $\mathbb{Q}$ ) in Equation 12,  
250 we see that, for  $j > 0$ , the complex modulus operation results in the sum of squares  
251 of two Gaussian terms, which by definition yields the  $\chi^2$  distribution with 2 degrees of  
252 freedom, i.e., the exponential distribution. A slightly different result holds for  $j = 0$  (see  
253 Equation 14).

254 A more formal derivation of the (asymptotic) distributional properties of spectral quan-  
255 tities is provided by Brillinger [1981, theorem 4.4.2] and [Cohen, 1998]. This behavior and  
256 associated advantages are illustrated and discussed in Section 4.

257 It is emphasized that, strictly speaking, the Power-density spectrum variate (PDSV) fol-  
 258 lows the exact  $\chi^2$  distribution only for white noise (uncorrelated homoscedastic Gaussian  
 259 processes). For strongly non-Gaussian, heteroscedastic and/or auto-correlated processes,  
 260 the distribution of the PDSV converges to the  $\chi^2$  distribution asymptotically as  $N \rightarrow \infty$   
 261 [e.g., Duchon and Robert Hale, 2012, chapter 1]. The greater the departure from white  
 262 noise, the longer data period (larger value of  $N$ ) is needed before the  $\chi^2$  distribution be-  
 263 comes a reasonable approximation (similar to Central Limit Theorem converging slower  
 264 when summing highly non-Gaussian, heteroscedastic and/or autocorrelated random vari-  
 265 ables). Given the long times series used in this paper (e.g.,  $N = 2^{15}$  in case study 2),  
 266 convergence of the PDSV to the  $\chi^2$  distribution is not a limiting factor.

267 Another useful theoretical feature of BSL is that it can quite readily accommodate vir-  
 268 tually any residual autocorrelation structure, as long as the autocorrelation decays to zero.  
 269 This can be achieved by substituting the appropriate parameterization for  $\zeta_j^{\mathcal{E}}$  in Equa-  
 270 tion 25. Examples of estimating the parameters of the well-known "1/ $f$ " noise [e.g. West  
 271 and Shlesinger, 1990; Ward and Greenwood, 2007a] and of a process with an exponentially  
 272 decaying PDS are provided in Section 3. In contrast, it might be difficult to derive the  
 273 corresponding autocorrelation functions in the time-domain, and indeed impossible in the  
 274 case of "1/ $f$ " noise [Ward and Greenwood, 2007b]. The practical advantages afforded by  
 275 this flexibility of BSL are discussed further in Section 5.

### 2.11. Relationship to the estimator of Montanari and Toth (2007)

Montanari and Toth [2007] have previously investigated the calibration of hydrologi-  
 cal models using the maximum likelihood estimator introduced by Whittle [1953]. The

likelihood proposed by Montanari and Toth [2007] is

$$p(\tilde{\mathbf{Y}} \mid \mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \prod_{j=1}^{N/2} f_{\text{exp}}(P_j[\tilde{\mathbf{Y}}] | P_j[\mathbf{H}(\boldsymbol{\theta}, \tilde{\mathbf{X}})] + \mathbb{P}_j[\boldsymbol{\mathcal{E}} | \boldsymbol{\vartheta}]) \quad (28)$$

where  $P_j[\mathbf{H}(\boldsymbol{\theta}, \tilde{\mathbf{X}})]$  is the periodogram of the model simulation and is used as an estimate of the PDS of the hydrological model  $\mathbb{P}_j[\mathbf{H}(\boldsymbol{\theta})]$ . The term  $\mathbb{P}_j[\boldsymbol{\mathcal{E}} | \boldsymbol{\vartheta}]$  is the PDS of the residual model as in Equation 23, and  $P_j[\tilde{\mathbf{Y}}]$  is the periodogram of the observed output. Note that, as discussed in Section 2.11, Equation 28 does not include the PDS value for  $j = 0$ . Furthermore, both periodograms (of the model simulation and of the observed output) are obtained via the Fast Fourier transform without windowing [Montanari and Toth, 2007, and the R-code available from the authors].

The approach suggested by Montanari and Toth [2007], which will be referred to as the "Montanari-Toth likelihood" (MTL), requires two important assumptions:

1. The PDS of the observations  $\tilde{\mathbf{Y}} = \mathbf{H}(\boldsymbol{\theta}, \tilde{\mathbf{X}}) + \boldsymbol{\mathcal{E}}$  can be approximated as  $\mathbb{P}[\tilde{\mathbf{Y}}] = \mathbb{P}[\mathbf{H}(\boldsymbol{\theta}, \tilde{\mathbf{X}})] + \mathbb{P}[\boldsymbol{\mathcal{E}}]$  or, more generally,  $\mathbb{P}[q(\tilde{\mathbf{Y}})] = \mathbb{P}[q(\mathbf{H}(\boldsymbol{\theta}, \tilde{\mathbf{X}}))] + \mathbb{P}[\boldsymbol{\mathcal{E}}^*]$  when the residuals  $\boldsymbol{\mathcal{E}}^*$  are defined in transformed space. This approximation holds only if the hydrological model and the residual model are independent, i.e., if their cross-spectrum is zero [e.g., see Brockwell and Davis, 1987].

2. The generally unknown PDS of the hydrological model  $\mathbb{P}[\mathbf{H}(\boldsymbol{\theta}, \tilde{\mathbf{X}})]$  can be approximated by the periodogram of a model simulation,  $P[\mathbf{H}(\boldsymbol{\theta}, \tilde{\mathbf{X}})]$ . As the periodogram is a "single-sample" estimator of the PDS, this approximation introduces additional noise into the estimation procedure. More stable (less noisy) PDS estimators exist [e.g., Welch, 1967] but have not been investigated in the context of the Whittle estimator for hydrological model calibration.

296 As shown by Montanari and Toth [2007], assumptions 1-2 above can often provide  
297 useful practical results. However, the additional approximation errors introduced by these  
298 assumptions can be expected to inflate posterior parameter uncertainty compared to BSL,  
299 which does not make these assumptions. These theoretical considerations are investigated  
300 empirically in Section 4 and discussed in Section 5.

301 Note also that MTL is formulated in terms of the observed streamflows given a model  
302 simulation and a residual error model, whereas BSL is formulated in terms of the observed  
303 residuals given a residual error model. In this respect, the convenience of the BSL formu-  
304 lation presented in this work is that it does not require dealing with spectral analysis of  
305 the hydrological model equations.

### 3. Case studies

306 Four case studies are presented: (i) illustration of the properties of the periodogram  
307 (used in the BSL inference), using pure random processes; (ii) inference of parameters  
308 of autocorrelation functions of pure random processes; (iii) synthetic hydrological cali-  
309 bration, where we investigate the inference of hydrological and error model parameters  
310 under controlled conditions; and (iv) real hydrological calibration, where we investigate  
311 parameter inference when model assumptions are not fully met.

312 A summary of the case studies is given in Table 3 and a summary of notations in Table 4.

#### 3.1. Case study 1 (synthetic): Properties of the periodogram / BSL

313 One of the interesting properties of the periodogram of a random process is that its ele-  
314 ments follow (approximately) the  $\chi^2$  probability distribution regardless of the probability  
315 distribution of the original random process (see Section 2.8). This property is illustrated

316 for uncorrelated non-Gaussian processes, with innovations from the following four proba-  
 317 bility distributions: i) uniform distribution in  $[-1, 1]$ , ii) Laplace distribution with  $\mu = 0$ ,  
 318  $\sigma = 1$ , iii) bimodal Gaussian distribution with  $\mu_1 = 1$ ,  $\mu_2 = -1$ ,  $\sigma_1 = \sigma_2 = 0.5$  and weight  
 319 0.2 of the first component; (iv) an AR(1) process with the parameters of the Gaussian  
 320 error model of Table 4, selected based on hydrological experience.

321 The methodology employed to empirically confirm the probability distribution proper-  
 322 ties of the periodogram is given in Appendix A1.

### 3.2. Case study 2 (synthetic): Inference of pure random processes

323 The ability of BSL to retrieve the parameters of the process that generated the "ob-  
 324 served" data are first illustrated using a synthetic case study based on pure random  
 325 processes (i.e., without a deterministic component).

The following stochastic processes are investigated:

$$f_{\mathbb{P}1}(\omega; A_1, B_1) = A_1 \exp(-B_1\omega) \quad (29)$$

$$f_{\mathbb{P}2}(\omega; A_2, B_2) = A_2/\omega^{B_2} \quad (30)$$

326 with reference parameter values  $(A_1, B_1) = (100, 0.001)$  and  $(A_2, B_2) = (10.5, 3)$ , re-  
 327 spectively.

328 The methodology for generating the synthetic data for this case study is detailed in  
 329 Appendix A2. This analysis allows establishing the theoretical properties of BSL under  
 330 idealized conditions. As we do not carry out a Markov Chain Monte Carlo (MCMC)  
 331 analysis of the posterior distribution, we are limited to examining the properties of the  
 332 optimal BSL estimate, rather than of the entire BSL distribution. Note that optimization

333 of the likelihood corresponds to optimization of the Bayesian posterior under uniform  
 334 prior assumptions.

### 3.3. Case study 3 (synthetic): simple hydrological model

335 Following the basic verification of the BSL using pure random processes, we investigate  
 336 its properties when applied to hydrological models with synthetic rainfall-runoff data.  
 337 The synthetic data is generated using a simple rainfall generator and hydrological model,  
 338 in order for the synthetic streamflow data to generally resemble real observations; see  
 339 Appendix A3 for a detailed description.

340 The rainfall generator used is a Poisson rectangular pulse model with an exponential  
 341 distribution for both the rain cell intensity and the duration [e.g. Bierkens and Puente,  
 342 1990]. The model has three parameters: the arrival rate  $\lambda$ , the mean intensity  $\bar{i}_r$  and the  
 343 mean duration  $\bar{t}_r$ .

344 The hydrological model used is a simple model with two linear reservoirs in series,  
 345 described by three parameters. The reservoir outflow is  $q = ks$ , where  $q$  is the outflow,  
 346  $s$  is the storage and  $k^{-1}$  is the residence time. The residence time of the first reservoir,  
 347  $k_1^{-1}$ , is shorter than the residence time of the second reservoir,  $k_2^{-1}$ . The leaching from  
 348 reservoir 1 to reservoir 2 is assumed to be constant and equal to  $l_g$ . We refer to this model  
 349 as the *simple linear HM* ("linearity" here refers to the flux formulation).

350 The following residual error models are used: (i) Gaussian AR(1) process (Section 2.1);  
 351 (ii) Laplacian AR(1) process (Section 2.1); and (iii) a process with the autocorrelation  
 352 structure  $v_\ell = f(\ell; \varrho_1, \varrho_2) = \exp(-\varrho_1 \ell - \varrho_2 \ell^{0.5})$ , which corresponds to the autocorrelation  
 353 structure of the model residuals of Schaepli et al. [2007]. We consider cases where the error  
 354 model is specified correctly and cases where it is misspecified. The analyses of misspecified

355 error models include erroneous distributional assumptions and erroneous autocorrelation  
 356 structure assumptions.

357 The model parameters and the selected reference values for the synthetic case studies  
 358 are summarized in Table 4. The maximum likelihood parameter set is estimated using  
 359 the Nelder-Mead simplex algorithm [e.g. Press et al., 2007] in Matlab Version 2010b.

360 The statistical reliability of the predictions in this case study is assessed against multiple  
 361 realizations of synthetic data using a predictive quantile-quantile plot, constructed as a  
 362 generalization of the predictive qq-plot plot proposed by Thyer et al. [2009] for the case  
 363 of a single reference realization (the observed data) (see Appendix A3).

### 3.4. Case study 4 (real data): Leaf River modeling

364 The behavior of BSL under real data conditions is investigated by calibrating the hy-  
 365 drological model HYMOD [Boyle, 2000] to the well-known Leaf River basin near Collins,  
 366 Mississippi [e.g. Sorooshian et al., 1993; Vrugt et al., 2005; Smith et al., 2008]. The catch-  
 367 ment has an area of about 1950 km<sup>2</sup>. Daily area-average precipitation, evapotranspiration  
 368 and streamflow estimates are available from the Hydrologic Research Laboratory of the  
 369 National Weather Service. The calibration period ranges from October 1948 to September  
 370 1951. The validation period ranges from January 1951 to December 1969.

To stabilize the variance of the model residuals, we apply a log-transformation to the  
 observed and simulated streamflow

$$\varepsilon_t = \log(\tilde{y}_t + A) - \log(\hat{y}_t + A), \quad (31)$$

371 where  $A$  is a small fixed offset to avoid numerical problems when applying the trans-  
 372 formation to zero and near-zero flows. Here, we use  $A = 10^{-4}$  (mm/d).

373 We assume that the residuals of log-transformed responses can be described by a random  
 374 vector  $\boldsymbol{\mathcal{E}}$  that (approximately) follows an AR(1) process with Gaussian innovations. In this  
 375 case, BTL and BSL can be applied without further modification to the log-transformed  
 376 residuals (see Section 2.9). The posterior parameter distribution with BSL or BTL are  
 377 sampled using the Metropolis algorithm described in Schaepli et al. [2007], which was used  
 378 to produce 1000 samples from a stable chain (no update of the sampling distribution).

379 All error model parameters are sampled jointly with the hydrological model parameters.  
 380 We use uniform priors for all parameters except the error model innovation variance  $\sigma_\delta^2$ ,  
 381 for which Jeffreys prior is used ( $p(\sigma_\delta^2) = 1/\sigma_\delta^2$ ) [see Schaepli et al., 2007]. Note that  
 382 the mean of the innovations  $\mu_\delta$ , which effectively acts as a mass balance parameter, is  
 383 inferred jointly with all other parameters. In principle this estimation approach can lead  
 384 to non-robust predictions, as shown empirically by Evin et al. [2014]. Although such non-  
 385 robustness was not seen in the current case study, we note that joint inference of mass  
 386 balance parameters, error variance and error autocorrelation should be undertaken with  
 387 care to avoid poor inference and predictions.

## 4. Results

### 4.1. Case study 1: Theoretical properties

388 An important property of BSL, arising from its use of the periodogram, is that its  
 389 formulation does not depend on the process distribution in the original (time-)domain  
 390 (Section 2.10). This is illustrated in Figure ??, which shows Gaussian qq-plots of the  
 391 realizations from three different non-Gaussian random processes, and the  $\chi^2$  qq-plots of  
 392 the corresponding periodograms. Figure ??a-b show three theoretical process examples,  
 393 while Figure ??c-d applies this analysis to the (highly non-Gaussian) residuals time series



394 obtained from the inferred maximum likelihood parameter set of the Leaf River case study  
395 (Section 4.4). In all cases, irrespective of the process distribution in the original domain,  
396 the periodograms follow a  $\chi^2$  distribution with two degrees of freedom for all frequencies  
397  $j > 0$ . This findings provides empirical confirmation of the theoretical considerations  
398 given in Section 2.10.

## 4.2. Case study 2: Inference of PDS parameters

399 Another important property of BSL is that it can be readily used to infer the parameters  
400 of processes with virtually any PDS. This is illustrated in Figure ??, which shows, for  
401 each parameter of the two pure random processes given in Table 3, the distributions of  
402 optimal estimates obtained by maximizing the likelihood function over multiple process  
403 realizations with the same underlying true parameters (see Section 3.2).

404 Figure ?? show that the distributions are (correctly) centered on the true parameter  
405 values used to generate the original process realizations. The variability of the optimal  
406 estimates is indicative of the parametric uncertainty associated with fitting the model  
407 to finite-length realizations (in this particular example, parametric uncertainty is quite  
408 small, less than 5% in both cases, due to the relatively long realization used (see Appendix  
409 A2).

## 4.3. Case study 3: Synthetic hydrological calibration

### 410 4.3.1. Correct versus misspecified error models

411 The mathematical equivalence of BSL and BTL is illustrated in Figure ??, which shows  
412 the parameter distributions obtained for the simple linear HM with a Gaussian AR(1)  
413 residual model where all model assumptions are respected. The inferred parameter dis-  
414 tributions are almost indistinguishable for the BSL and BTL likelihoods.

415 Figure ?? shows the same experiment, except that the residuals are generated from  
416 a Laplace AR(1) distribution. We consider three likelihoods: Gaussian BTL, Laplace  
417 BTL and BSL (which remains unchanged because it does not depend on the assumed  
418 distribution of residual errors).

419 The parameter distributions inferred with BSL, Gaussian BTL and Laplace BTL are al-  
420 most identical, and are centered on the true parameter values. This finding demonstrates  
421 the general robustness of the Gaussian BTL with respect to the underlying distribution of  
422 model residuals. This robustness is confirmed by repeating the same experiment (results  
423 not shown) with different residual error parameters (namely  $\rho_\varepsilon \in 0, 0.5, 0.8$ ,  $\mu_\delta \in 0, 0.23, 1$ )  
424 and  $\sigma_\delta \in 0.06, 0.13, 0.5$ ), and for log-normal residual distributions. This robustness ex-  
425 presses the fact that in any of these experiments, BTL-Gauss is maximized for almost  
426 the same parameter set as BTL-Laplace or BTL-lognormal (differences in optimal pa-  
427 rameter values of a few percent). The robustness of least squares parameter estimates  
428 to moderate departures from Gaussian distribution assumptions is fairly well established  
429 in the statistical literature [e.g., White, 1981]. Note that here we are concerned with  
430 departures from the overall shape of the error distribution rather than to the presence of  
431 strong outliers, as in the latter case least squares estimates can indeed deteriorate very  
432 rapidly [Press et al., 2007].

433 The robustness of BTL does start to break down if wrong assumptions are made about  
434 the residual autocorrelation structure (rather than about the residual distribution). This  
435 is illustrated in Figure ??, which shows the parameter distributions obtained for BTL,  
436 BSL and MTL for the synthetic case when the residuals have an exponential rather than  
437 AR(1) autocorrelation structure (see Section 3.2). The parameter distributions show

438 that both BTL-Gauss-AR1 and BSL-AR1 yield unbiased parameter estimates. However,  
439 compared to the distributions obtained under the correct likelihood (called BSL-nonAR1),  
440 BTL-Gauss-AR1 yields too wide hydrological parameter distributions, in particular for  
441 parameter  $k_2$  (compare Figure ?? top row and bottom row). The distributions of the  
442 parameters  $k_1$ ,  $k_2$  and  $l_g$  under BSL-AR1 are similar to the distributions obtained under  
443 the correct likelihood BSL-nonAR1 (compare Figure ?? 2nd row and bottom row). This  
444 finding suggests that BSL is slightly more robust than BTL to violations of assumptions  
445 describing the residual error autocorrelation.

446 The differences between the distributions become more visible when comparing their  
447 reliability using the predictive qq-plots shown in Figure ?. These predictive qq-plots  
448 show the probability distribution of the underlying true reference simulations within the  
449 model simulations. The predictive qq-plot for BSL-nonAR1 is clearly closer to the 1:1 line  
450 (corresponding to a perfectly reliable probabilistic model) than for the other likelihoods.  
451 Finally, Figure ? shows the spectral-domain differences between the PDS inferred under  
452 the assumptions of BTL-Gauss-AR1, BSL-AR1 and BSL-nonAR1 versus the true PDS. It  
453 can be seen that there is a relatively pronounced difference between the PDS inferred with  
454 the correct likelihood BSL-nonAR1 and the PDS obtained with the likelihoods assuming  
455 (wrongly) a AR1 process.

#### 456 4.3.2. Comparison of MTL versus BSL and BTL

457 Figure ? compares the parameter distributions for the simple linear HM obtained using  
458 the MTL likelihood to the parameter distributions obtained using BTL and BSL.

459 The MTL inference of all model parameters (hydrological and error model) is unbiased  
460 despite the fact that the error innovation mean cannot be inferred (Section 2.11). However,

461 the posterior distributions (of all parameters) have a larger variance under MTL than  
462 under BTL and BSL, in particular for the hydrological model parameter  $k_2$  and the error  
463 model autocorrelation  $\rho$ .

464 In terms of sensitivity to residual error assumptions, MTL does not, in theory, depend on  
465 the residual distribution in the original domain (this property is similar to BSL). For the  
466 previously discussed case of Laplace distributed residuals, MTL indeed results in unbiased  
467 parameter distributions (Figure ??), despite the fact that the mean of the innovations,  $\mu_\delta$ ,  
468 cannot be inferred with MTL (the zero frequency does not enter the computation of the  
469 Whittle likelihood). This example shows that non-zero-mean residuals do not necessarily  
470 lead to biased MTL estimates.

471 In contrast, MTL is highly sensitive to wrong autocorrelation assumptions as demon-  
472 strated with the experiment with non-AR(1) residual realizations: the resulting parameter  
473 distributions are biased (Figure ??), the prediction range does not correspond to the range  
474 of reference simulations (Figure ??) and, compared to the periodogram of the residuals,  
475 the PDS does not show enough power for high frequencies (Figure ??).

#### 4.4. Case study 4: Leaf river case study

476 The posterior distributions of HYMOD and residual error model parameters inferred in  
477 the Leaf River case study using BTL-Gauss, BSL and MTL are shown in Figure ?. The  
478 corresponding maximum likelihood parameter sets are listed in Table 5.

479 The distributions produced using BTL-Gauss and BSL are very similar. Notable dif-  
480 ferences arise between the inferred residual innovation means  $\mu_\delta$ , the HYMOD parameter  
481  $b_H$  (which parameterizes the degree of spatial variability of the soil moisture) and the  
482 HYMOD parameter  $\alpha$  (which parameterizes the distribution of flow between the slow

483 and quick reservoirs). MTL gives a very different range of posterior parameter values for  
484 the two reservoir parameters  $\nu_s$  and  $\nu_q$  resulting in lower base flow, slower recessions and  
485 a slower response to rainfall events. Overall, this leads to more strongly autocorrelated  
486 residuals, as reflected in the distribution of  $\rho$  for MTL.

487 Figure ?? shows the streamflow simulation during the validation period corresponding  
488 to the parameter set with the highest BSL value. Included are plots of the log-transformed  
489 streamflow to show the model performance during low flows, a plot of the corresponding  
490 residual time series, a plot of the residuals against the rank of the simulated streamflow,  
491 a predictive qq-plot and a plot of the partial autocorrelation of the residuals.

492 The diagnostic plots in Figure ?? show that, in the case of BSL, the logarithmic transfor-  
493 mation stabilizes the variance of residual errors and the assumption of constant-variance  
494 Gaussian residuals holds at least approximately. The autocorrelation of the residual errors  
495 is reasonably approximated by the AR(1) process (Figure ??f).

496 In contrast, the corresponding diagnostic plots for MTL clearly show that the results  
497 obtained with this likelihood do not comply with the underlying assumptions. In par-  
498 ticular, the residuals are strongly non-symmetric (Figure ??d), do not have an AR(1)  
499 autocorrelation structure (Figure ??c,f) and are non-Gaussian (Figure ??e).

500 The results for the maximum likelihood simulation with BTL are very similar to BSL  
501 Figure ?. However, the residual time series computed in the BSL and BTL inferences  
502 are not identical. In particular, the distributions of residual model parameters are slightly  
503 different (especially for the innovation mean  $\mu_\delta$ , see Figure ??), which translates into a  
504 different mean and total variance in the AR(1) residual model.

505 Accordingly, the total prediction limits for BSL and BTL are also different (Figure ??).  
506 These limits are obtained from 500 random model realizations (hydrologic model simu-  
507 lation plus residual error realization) corresponding to random draws from the posterior  
508 parameter distributions. As can be seen in Figure ??, the parametric uncertainty is rela-  
509 tively small for BSL, BTL and MTL (because of the length of the calibration data), and  
510 the majority of the predictive uncertainty is due to residual errors.

511 Overall, the total 90% prediction limits obtained with BSL and BTL from the 1000  
512 samples span, respectively, around 93% and 90% of observed values in the calibration  
513 period and around 94% and 87% of observed values in the entire simulation period (cal-  
514 ibration and validation). The predictive qq-plots for the validation period (Figure ??a  
515 and d) show that both likelihoods lead to very similar statistical reliability, with minor  
516 deviations from the uniform distribution. Considering high flow and low flow separately,  
517 (Figure ??b-c and e-f) suggests that BSL gives more reliable results than BTL for low  
518 flow simulations (Figure ??c and f).

## 5. Discussion

519 Given the theoretical aspects presented in Section 2 and the results of the empirical  
520 case studies reported in Section 3, we are now in a position to discuss the advantages and  
521 limitations of the BSL approach, relate it to the existing techniques for parameter inference  
522 in the time- and spectral-domains, and outline directions for further investigations.

523 We begin by comparing BSL to other spectral calibration methods proposed in the  
524 hydrological literature, with a particular focus on the MTL approach [Montanari and  
525 Toth, 2007], and then make a broader comparison to traditional time-domain calibration  
526 (BTL).

527 The majority of spectral-domain calibration methods in the hydrological literature are  
528 heuristic, in the sense that they do not explicitly articulate a probabilistic model of the  
529 system of interest (here, the catchment and observation systems) when constructing the  
530 objective function. For example, consider the case of parameter calibration that searches  
531 for the hydrological parameter set to match the autocorrelation function of the simulated  
532 discharge and the autocorrelation function of the observed discharge, using the root-mean-  
533 squared-error as a distance metric [Moussu et al., 2011]. This approach is useful from the  
534 point of view of maximizing particular model fit features, and establishing the sensitivity  
535 of the fit to parameter values, but cannot provide *probabilistic* estimates of uncertainty in  
536 the estimated parameters and predictions.

537 In addition, heuristic approaches, such as matching the autocorrelation function, hide a  
538 number of assumptions, such as the distributional properties of the errors. For example,  
539 using the sum-of-squared differences between the autocorrelation functions of observed  
540 versus simulated streamflow implies an assumption that these differences follow an inde-  
541 pendent Gaussian distribution. Unless these assumptions are stated and tested explicitly,  
542 the ability of the inference to provide meaningful probabilistic estimates is questionable.

543 For these reasons, our interest in this study is on likelihood functions explicitly derived  
544 from probabilistic models of the hydrological system and observations systems.

545 To the best of our knowledge, MTL is the only approach in the hydrological literature  
546 where a likelihood function is formally articulated from a probabilistic description of the  
547 data. The Whittle likelihood employed in the MTL approach is used outside of hydrology,  
548 in particular to infer the parameters of time series models [Ives et al.] or to estimate the  
549 power-density spectrum of time series [Choudhuri et al., 2004]. However, it is usually used

550 in contexts where the PDS of the model can be computed directly [e.g., Montanari et al.,  
551 1997], whereas in the MTL approach, the PDS of the (error) model is approximated (with  
552 potentially large errors) by subtracting the periodogram of model simulations from the  
553 periodogram of observed times series (Section 2.11).

554 In this work, we use a different strategy when deriving the BSL approach - we articulate  
555 the spectral-domain error model by computing the distribution of residual errors in the  
556 spectral-domain from the time series of residuals in the time-domain.

557 What are the advantages and limitations of the BSL approach versus the MTL ap-  
558 proach? By explicitly computing the residual error time series and then transforming to  
559 the spectral-domain, BSL avoids the approximations and ensuing noise incurred by MTL.  
560 This behavior can be seen in the empirical case studies 3.1 and 3.2 (Table 3, Figure ??,  
561 Figure ??) where parametric uncertainty in the BSL approach was similar to the BTL  
562 inference. Especially under synthetic conditions with the correct error model (case study  
563 3.1, Figure ??), we can take BTL as the reference solution because it works directly with  
564 the raw residuals without any spectral-domain transformations.

565 In contrast, MTL inference yields parameter distributions that are clearly wider than  
566 those of BTL and BSL, as can be seen in case studies 3.1, 3.2, 3.3 (Figure ??, Figure ??,  
567 Figure ??). We can attribute the additional noise in MTL to at least two potential  
568 reasons: (i) MTL uses the periodogram to estimate the PDS of the hydrological model.  
569 As the periodogram is a single-sample estimator of the PDS, it increases the variance of  
570 the resulting parameter estimates; (ii) MTL excludes the 0th frequency of the PDS from  
571 the likelihood function and hence loses information about the mean of the stochastic  
572 process assumed to describe residual errors.



573 This paper has not investigated the individual impact of these approximations on the  
574 MTL inference. It is possible that the use of more robust PDS estimators, for example,  
575 Welch's method [Welch, 1967] could reduce the impact of some of the limitations. That  
576 said, testing the theoretical properties of approximations to spectral properties of hydro-  
577 logical models is difficult because, except for special cases, the true PDS of hydrological  
578 models is unknown. For example, Bierkens and Puente [1990] proposed analytical expres-  
579 sions for the autocorrelation of the outputs from a simple hydrological model forced with  
580 stochastic inputs. However, their derivations hold only for the specific model used and  
581 only for small lags; these results are hence of limited value with respect to more general  
582 analyses of the MTL inference scheme.

583 It is worth adding that, in BSL, the periodogram of residual errors is used *not* as an  
584 estimator of their power-density spectrum (PDS), but to compute samples of the power-  
585 density spectrum variate (PDSV). Hence, replacing the periodogram  $P(\boldsymbol{\varepsilon})$  used in BSL by  
586 a more stable estimator of the PDS [e.g., the method of Welch, 1967] would be detrimental  
587 to the probabilistic properties of BSL, because it would (by construction) under-estimate  
588 the variability of residual errors in the spectral domain.

589 The advantage of the BSL approach in avoiding the approximations of MTL does not  
590 come free. In particular, the MTL approach appears simpler to apply to indirect cali-  
591 bration problems such as non-concomitant calibration, because the residual errors do not  
592 have to be computed explicitly. BSL will require further development to be applicable to  
593 this problem, including an approximation of the PDS of the residual errors that does not  
594 ignore the cross-spectrum (Section 2.11).

595 These extensions to the MTL and BSL approaches lie beyond the scope of current work  
596 and will be explored in follow-up studies.

597 We now shift our attention to a broad comparison of BSL and time-domain estimation  
598 (BTL). In view of the close correspondence of BSL and BTL, does BSL represents a new  
599 error model, or a new solution approach for the same error model as BTL? In principle,  
600 BSL is obtained via Fourier transform of the time-domain realization and, therefore,  
601 could be seen to rely on the same initial assumptions as BTL (here, that residuals follow  
602 a Gaussian AR(1) process). However, due to the properties of the Fourier transform and  
603 the PDS definition listed in Section 2, the influence of assumptions such as that the errors  
604 are Gaussian is greatly diminished. In this respect, the BSL approach could be viewed to  
605 represent a more robust error model, especially with respect to distributional assumptions.

606 The advantages / attractive features of BSL versus BTL can be summarizes as follows.

607 a) Theoretical robustness against violation of distributional assumptions. As noted in  
608 the theory section (Section 2.3), the power-density spectrum variate central to the BSL  
609 approach has a  $\chi^2$  distribution almost independently from the (residual error) time series  
610 it is computed from. This is an attractive theoretical property because it can be expected  
611 to reduce the impact of violating distributional assumptions such as Gaussian errors, etc.

612 We note that least squares methods are often robust against departures from Gaussian  
613 assumptions [e.g., White, 1981, , and case study 3.2]. For this reason, a comparison  
614 of BSL and BTL in catchments with strongly non-Gaussian errors is of interest and  
615 recommended for future work. The impact of strong outliers is of particular interest given  
616 the susceptibility of least squares estimation to this particular departure from Gaussianity  
617 [Press et al., 2007].

618 b) Flexibility in representing the autocorrelation structure of the model residuals. This  
619 flexibility arises due to the structure of BSL where the autocorrelation profile function  
620 ((21)) is formulated directly in the spectral-domain. In case study 2.2, we considered  
621 processes with autocorrelation structures that cannot be readily formulated in the time-  
622 domain. For example, the correlation structure of "1/f" (pink) noise has no simple rep-  
623 resentation in the time-domain. Using BTL for such problems would require approxima-  
624 tions, e.g., by AR( $n$ ) processes, and does not appear robust (e.g., case study 3.3).

625 The theoretical flexibility of BSL over BTL in representing error autocorrelation be-  
626 comes particularly attractive when modeling environmental processes with strong cyclic  
627 behavior. For example, water temperature time series typically exhibit a pronounced di-  
628 urnal cycle [e.g., Comola et al., 2015]. A water temperature model that does not capture  
629 this diurnal cycle will generally yield residuals with a cyclic autocorrelation structure  
630 peaking every 24 hours. This type of autocorrelation structure is difficult to represent in  
631 the time-domain, but is relatively easier to represent in the spectral-domain (e.g., with  
632 the profile function, equation (21)). This example illustrates a case where BSL offers use-  
633 ful practical advantages for modeling environmental systems where the autocorrelation of  
634 the model and/or observation errors is (much) easier in the spectral-domain than in the  
635 time-domain.

636 c) Opportunities to extend the method to indirect calibration problems, by taking  
637 advantage of working in the spectral-domain (see discussion above).

638 The practicality of a calibration scheme such as BSL depends not only on its inferential  
639 properties, but also on its computational cost. A major computational feature of spectral-  
640 domain methods is their use of Fourier transformed quantities. In practical work, the

641 Fourier transform is invariably implemented using the Fast Fourier Transform (FFT)  
642 algorithm, which requires of the order of  $N \log N$  operations [Rao et al., 2010]. BTL  
643 does not require any FFT operations (it operates exclusively in the time-domain), MTL  
644 requires a single FFT operation per likelihood evaluation (FFT of the hydrological model  
645 simulations, assuming the FFT of observed data is pre-computed once) and BSL also  
646 requires a single FFT operation per likelihood evaluation (FFT of the residual error time  
647 series). While the cost of FFT can be appreciable for very long time series, in most cases  
648 we expect it to be dominated by the cost of running the hydrological model, which in  
649 general requires the solution of differential and algebraic equations at each time step.  
650 Consequently, it is unlikely that the computational cost of FFT within the BSL (and  
651 MTL) approaches could be a major limiting factor in practical work.

652 Finally, in terms of future work, we note that many aspects of hydrological calibration  
653 in the spectral-domain remain poorly understood. Based on the findings reported in this  
654 paper, the following specific directions deserve focused investigation:

655 (a) Investigate the robustness of BSL versus MTL and BTL under conditions of strongly  
656 non-Gaussian errors. In principle, BSL and MTL should provide practical robustness,  
657 which should be established using both "realistically constructed" synthetic data and real  
658 data case studies, in particular in arid/semi-arid catchments;

659 (b) Apply BSL to environmental modeling problems where the error time series ex-  
660 hibit nontrivial persistence patterns (for example, the diurnal water temperature models  
661 mentioned earlier);

662 (c) Indirect calibration, including calibration using non-concomitant input-output data  
663 time series, and calibration in ungauged catchments;

664 (d) Using BSL in studies exploring error decomposition. In this paper, we focused  
665 exclusively on aggregated treatment of errors using a single residual error model. The  
666 alternative paradigm of error decomposition is of tremendous interest, as it allows esti-  
667 mating dominant sources of uncertainty and devising strategies for reducing these errors.  
668 In principle, BSL can be incorporated directly into the likelihood terms of hierarchical  
669 Bayesian approaches such as BATEA [Kavetski et al., 2006; Renard et al., 2011], but the  
670 advantages, limitations and practicalities of doing so remain to be established empirically.

## 6. Conclusions

671 This paper presents the derivation of the spectral-domain counterpart of the widely used  
672 time-domain likelihood for Bayesian inference of environmental models. The theoretical  
673 and empirical properties of the proposed Bayesian spectral likelihood (BSL) are compared  
674 to the properties of the Bayesian time-domain likelihood (BTL), and to the Whittle-type  
675 spectral-domain likelihood (MTL) previously proposed by Montanari and Toth [2007].

676 The key conclusions of this paper are as follows:

677 1. The Bayesian spectral-domain likelihood derived in this work is mathematically  
678 equivalent to its time-domain counterpart in the case when the residual errors are as-  
679 sumed to be Gaussian (and autocorrelated). However, the spectral formulation offers two  
680 theoretical benefits: (i) its likelihood function is (asymptotically) independent from the  
681 probability distribution of residual errors, and (ii) it can accommodate residual errors  
682 with more complicated autocorrelation structure (for which time-domain representations  
683 are difficult or impossible).

684 2. At least under synthetic conditions, the time-domain likelihood is relatively robust  
685 to departures from the assumption of Gaussian residuals, but this robustness breaks down

686 for departures from the assumed autocorrelation structure. This is an aspect in which the  
687 spectral-domain inference might offer practical benefits over time-domain inference.

688 3. For the Whittle-type spectral-domain likelihood proposed by Montanari and Toth  
689 [2007], the synthetic and real data studies suggest that the simplifying assumptions made  
690 in this likelihood tend to produce parameter distributions that are too wide compared to  
691 inference in the time-domain, and potentially biased when autocorrelation assumptions  
692 are violated. The Bayesian spectral-domain likelihood introduced in this work does not  
693 appear to suffer from these limitations and does not incur a loss of information compared  
694 to the corresponding time-domain inference.

695 4. The real data case study based on the Leaf River and the hydrological model HY-  
696 MOD reinforces the (relative) robustness of the time-domain and spectral-domain infer-  
697 ence for a typical hydrological setting. Both inferences produced similar results, despite  
698 some moderate departures from the residual error model assumptions.

699 The theoretical derivations and analyses presented in this paper represent the first step  
700 towards formal Bayesian inference in the spectral-domain. Further work is required to  
701 better understand the properties of spectral-domain inference and its potential advan-  
702 tages in environmental model calibration. Future studies will include: (i) investigation  
703 of the robustness of BTL and BSL in cases where the model residuals are strongly non-  
704 Gaussian, e.g., as common in models of arid and semi-arid catchments; (ii) a wider range  
705 of hydrological case studies to gain more general insights into the practical performance of  
706 spectral-domain inference, including for models with non-trivial / cyclic autocorrelation  
707 structures; (iii) extensions of BSL to parameter inference with non-concomitant input-

708 output time series; and (iv) extensions of BSL to more comprehensive inference setups  
709 with individual treatment or sources of uncertainty (error decomposition).

710 A Matlab implementation of BSL for AR(1) error models is available in the Supporting  
711 Information.

## 7. Acknowledgements

712 The research of the first author was supported by research grants of the Swiss National  
713 Science Foundation (SNF, PZ00P2\_147366, PP00P2\_157611). A Matlab implementation  
714 of the HYMOD model, including the Leaf river data set, was provided by the model  
715 identification toolbox of Hoshin Gupta (University of Arizona). No other data sets were  
716 used for this theoretical study. An R implementation of the MTL likelihood was provided  
717 by Alberto Montanari (University of Bologna). We thank two anonymous reviewers, Anna  
718 Sikorska and Chief Editor Alberto Montanari for their constructive comments during the  
719 review process.

## Appendix A: Details of case study methodology

### A1. Analysis of periodogram properties in case study 1

720 This section details the methodology employed in case study 1 to empirically confirm  
721 the probability distribution properties of the periodogram:

722 1. Select a time-domain process (e.g. from the list of case studies summarized in Table  
723 3);

724 2. Generate a sample of multiple independent realizations from the stochastic process.

725 Here we generate a sample  $\mathbf{z}$  of length  $N = 2^{13}$ .

726 3. Compute the periodogram  $\mathbb{P}_j(\mathbf{z})$  using the FFT operation [Welch, 1967];

727 4. Produce a Gaussian qq-plot of the sample  $z$  from in Step 2. This plot is used to  
728 illustrate that the distribution of the process  $\mathbf{Z}$  in the time-domain is clearly non-Gaussian.

729 5. Produce a  $\chi^2$  qq-plot of the values of the periodogram  $\mathbb{P}_j(\mathbf{z})$  from Step 3. This  
730 plot is used to demonstrate that distribution of the process in the spectral-domain follows  
731 the  $\chi^2$  distribution. Note that for an uncorrelated process, the distribution of the PDS  
732 element  $\mathbb{P}_j[\mathbf{Z}]$  does not depend on the element index  $j$ .

### A2. Generation of synthetic data and inference verification in case study 2

733 This section describes the methodology employed in case study 2 to generate the syn-  
734 thetic data and use it to verify the results of the inference und all tested likelihoods.

735 1. Select a stochastic process, defined by its PDS function  $f_{\mathbb{P}}(\omega; \boldsymbol{\gamma})$ , where  $\boldsymbol{\gamma}$  is a set of  
736 parameters. We also select a reference set of parameters,  $\boldsymbol{\gamma}_r$ .

2. Generate a random process realization  $\mathbf{z}_{\mathbb{P}(\boldsymbol{\gamma}_r)}^{(i)}$  of length  $N$  with PDS  $f_{\mathbb{P}}(\boldsymbol{\gamma}_r)$ , as follows

$$\mathbf{z}_{\omega}^{(i)} \leftarrow N(0, 1^2) \tag{A1}$$



$$\mathbf{z}_{\mathbb{P}(\gamma_r)}^{(i)} = f_{\text{FFT}}^{-1} \left[ f_{\mathbb{P}}(\omega; \gamma_r)^{0.5} f_{\text{FFT}}(\mathbf{z}_{\omega}^{(i)}) \right], \quad (\text{A2})$$

737 where  $f_{\text{FFT}}$  denotes the fast Fourier transform (FFT) operation,  $f_{\text{FFT}}^{-1}$  is the inverse FFT  
 738 operation, and  $\mathbf{z}_{\omega}^{(i)}$  is a standard Gaussian white noise realization (but any other probabil-  
 739 ity distribution could be used here).  $f_{\mathbb{P}}(\omega; \gamma_r)^{0.5} f_{\text{FFT}}(\mathbf{z}_{\omega}^{(i)})$  corresponds to an element by  
 740 element (i.e. frequency by frequency) multiplication of the square-root of the PDS  $f_{\mathbb{P}}(\gamma_r)$   
 741 with  $f_{\text{FFT}}(\mathbf{z}_{\omega}^{(i)})$ .

742 3. Infer the maximum BSL estimate  $\hat{\gamma}^{(i)}$  by maximizing the log-BSL,  $\log p(\mathbf{z}_{\mathbb{P}(\gamma_r)}^{(i)} \mid \gamma)$   
 743 with respect simplex direct search algorithm [Lagarias et al., 1998] for this maximization.  
 744 direct search algorithm [Lagarias et al., 1998] for this maximization.

745 4. Repeat steps 2-3 for  $i = 1, \dots, m$  to obtain an empirical distribution of optimized  
 746 parameter estimates  $\hat{\gamma}^{(i)}$ . All presented results use  $m = 400$  replicates each of length  
 747  $N = 2^{15}$ .

748 5. Compare the empirical distribution of  $\hat{\gamma}^{(i)}$  to the reference value  $\gamma_r$ .

749 This analysis allows establishing the theoretical properties of BSL under idealized con-  
 750 ditions. As we do not carry out a Markov Chain Monte Carlo analysis of the posterior  
 751 distribution, we are limited to examining the properties of the optimal BSL estimate,  
 752 rather than of the entire BSL distribution. Note that optimization of the likelihood cor-  
 753 responds to optimization of the Bayesian posterior under uniform prior assumptions.

### A3. Generation of synthetic data in the hydrological case study 3

754 This section describes the methodology employed in case study 3 to investigate the BSL  
 755 inference under synthetic hydrological conditions.

756 1. Select a rainfall generator  $\mathbf{G}$ , and a reference parameter set  $\varphi_r$ .

- 757 2. Select a rainfall-runoff model  $\mathbf{H}$ , and a reference parameter set  $\boldsymbol{\theta}_r$ .
- 758 3. Select a stochastic process to represent residual errors  $\boldsymbol{\mathcal{E}}$ , and a reference parameter  
759 set  $\boldsymbol{\vartheta}_r$ .
- 760 4. Generate a realization of the rainfall,  $\mathbf{x}^{(i)} \leftarrow \mathbf{G}(\boldsymbol{\varphi}_r)$  of length  $N$ .
- 761 5. Compute the synthetic "true" streamflow realization,  $\mathbf{y}^{(i)} = \mathbf{H}(\boldsymbol{\theta}_r, \mathbf{x}^{(i)})$  of length  
762  $N$ .
- 763 6. Generate a realization of residuals,  $\boldsymbol{\varepsilon}^{(i)} \leftarrow \boldsymbol{\mathcal{E}}(\boldsymbol{\vartheta}_r)$  of length  $N$ .
- 764 7. Compute the synthetic "observed" streamflow realization,  $\tilde{\mathbf{y}}^{(i)} = \mathbf{y}^{(i)} + \boldsymbol{\varepsilon}^{(i)}$ , of length  
765  $N$ .
- 766 8. Select a likelihood formulation (BTL, BSL or MTL) and infer the maximum like-  
767 lihood estimates  $(\hat{\boldsymbol{\theta}}^{(i)}, \hat{\boldsymbol{\vartheta}}^{(i)})$  by maximizing the log-likelihood,  $\log p(\tilde{\mathbf{y}}^{(i)} \mid \mathbf{x}^{(i)}, \boldsymbol{\theta}, \boldsymbol{\vartheta})$  with  
768 respect to  $(\boldsymbol{\theta}, \boldsymbol{\vartheta})$ . Note that, similar to case study 1, this procedure corresponds to maxi-  
769 mizing the Bayesian posterior under uniform prior assumptions.
- 770 9. Repeat steps 4-8 for  $i = 1, \dots, m$  to obtain an empirical distribution of the optimized  
771 rainfall-runoff model parameters  $\hat{\boldsymbol{\theta}}^{(i)}$  and residual model parameters  $\hat{\boldsymbol{\vartheta}}^{(i)}$ . All presented  
772 results use  $m = 400$  replicates.

773 The statistical reliability of the predictions obtained in this study is assessed using a  
774 predictive quantile-quantile (qq) plot constructed as follows: (i) compute the median of the  
775 reference simulations at time step  $t$ ; (ii) estimate the quantiles of this median simulation  
776 within the predictions at time step  $t$ ; (iii) estimate the frequency of these quantiles for  
777 selected bins; (iv) repeat steps (i) - (iii) but swapping the reference simulations and the  
778 predictions; (v) plot the frequencies against each other; this should plot on a 1:1 line.  
779 Note that the predictive qq-plot constructed using the approach above is a generalization

780 of the predictive qq-plot proposed by [Thyer et al., 2009], where only a single reference  
781 realization (the observed data) was used.

## References

782 M. S. Bartlett. Periodogram analysis and continuous spectra. *Biometrika*, 37:1–16, 1950.

783 B. C. Bates and E. P. Campbell. A markov chain monte carlo scheme for parameter  
784 estimation and inference in conceptual rainfall-runoff modeling. *Water Resour. Res.*,  
785 37(4):937–947, 2001.

786 K. J. Beven and A. Binley. The future of distributed models: model calibration and un-  
787 certainty prediction. *Hydrol. Proc.*, 6(3):279–298, 1992. doi: 10.1002/hyp.3360060305.

788 M. F. P. Bierkens and C. E. Puente. Analytically derived runoff models based on rainfall  
789 point processes. *Water Resour. Res.*, 26(11):2653–2659, 1990.

790 G. Box, G. M. Jenkins, and G. C. Reinsel. *Time series analysis : forecasting and control*.  
791 Prentice Hall, Englewood Clifs, NJ, 1994.

792 G. E. Box and G. Jenkins. *Time series analysis : forecasting and control*. Holden-Day,  
793 San Francisco, 1976.

794 G. E. P. Box and G. Tiao. *Bayesian inference in statistical analysis*. Wiley, 1992.

795 D. P. Boyle. *Multicriteria calibration of hydrological models*. PhD thesis, University of  
796 Arizona, 2000.

797 P. Brockwell and R. Davis. *Time Series: Theory and Methods*. Springer series in statistics.  
798 Springer, New York, 1987.

799 N. Choudhuri, S. Ghosal, and A. Roy. Bayesian estimation of the spectral density of a  
800 time series. *Journal of the American Statistical Association*, 99(468):1050–1059, 2004.

- 801 L. Cohen. Generalization of the wiener-khinchin theorem. *IEEE Signal Processing Letters*,  
802 5(11):292–294, 1998.
- 803 F. Comola, B. Schaeffli, A. Rinaldo, and M. Lehning. Thermodynamics in the hydro-  
804 logic response: Travel time formulation and application to alpine catchments. *Water*  
805 *Resources Research*, 51(3):1671–1687, 2015. doi: 10.1002/2014wr016228.
- 806 N. De Vleeschouwer and V. R. N. Pauwels. Assessment of the indirect calibration of a  
807 rainfall-runoff model for ungauged catchments in flanders. *Hydrol. Earth Syst. Sci.*, 17  
808 (5):2001–2016, 2013. doi: 10.5194/hess-17-2001-2013.
- 809 C. Duchon and R. Robert Hale. *Time Series Analysis in Meteorology and Climatology:*  
810 *An Introduction*. Wiley-Blackwell, Oxford, 2012.
- 811 K. Engeland, C. Y. Xu, and L. Gottschalk. Assessing uncertainties in a conceptual water  
812 balance model using bayesian methodology. *Hydrol. Sci. J.*, 50(1):45–63, 2005.
- 813 G. Evin, D. Kavetski, M. Thyer, and G. Kuczera. Pitfalls and improvements in the joint  
814 inference of heteroscedasticity and autocorrelation in hydrological model calibration.  
815 *Water Resour. Res.*, 49:4518–4524, 2013. doi: 10.1002/wrcr.20284.
- 816 G. Evin, M. Thyer, D. Kavetski, D. McInerney, and G. Kuczera. Comparison of joint  
817 versus postprocessor approaches for hydrological uncertainty estimation accounting for  
818 error autocorrelation and heteroscedasticity. *Water Resour. Res.*, 50:2350–2375, 2014.  
819 doi: 10.1002/2013wr014185.
- 820 R. Fox and M. S. Taqqu. Large-sample properties of parameter estimates for strongly  
821 dependent stationary gaussian time-series. *Ann. Stat.*, 14(2):517–532, 1986.
- 822 A. Hartmann, M. Weiler, T. Wagener, J. Lange, M. Kralik, F. Humer, N. Mizyed, A. Rim-  
823 mer, J. A. Barber, B. Andreo, C. Butscher, and P. Huggenberger. Process-based karst

- 824 modelling to relate hydrodynamic and hydrochemical characteristics to system proper-  
825 ties. *Hydrol. Earth Syst. Sci.*, 17(8):3305–3321, 2013. doi: 10.5194/hess-17-3305-2013.
- 826 M. Honti, C. Stamm, and P. Reichert. Integrated uncertainty assessment of discharge  
827 predictions with a statistical error model. *Water Resour. Res.*, 49:4866–4884, 2013.  
828 doi: 10.1002/wrcr.20374.
- 829 A. R. Ives, K. C. Abbott, and N. L. Ziebarth. Analysis of ecological time series with  
830 arma(p,q) models. *Ecology*, 91(3):858–871. doi: 10.1890/09-0442.1.
- 831 G. Jenkins and D. G. Watts. *Spectral analysis and its applications*. Holden-Day Series in  
832 Time Series Analysis. Holden-Day, San Francisco, 1968.
- 833 D. Kavetski, G. Kuczera, and S. W. Franks. Bayesian analysis of input uncertainty  
834 in hydrological modeling: 1. theory. *Water Resour. Res.*, 42(3):W03407, 2006. doi:  
835 10.1029/2005WR004368.
- 836 G. Kuczera. Improved parameter inference in catchment models. 1. evaluating parameter  
837 uncertainty. *Water Resour. Res.*, 19(5):1151–1162, 1983.
- 838 G. Kuczera and E. Parent. Monte carlo assessment of parameter uncertainty in conceptual  
839 catchment models: The metropolis algorithm. *J. Hydrol.*, 211(1-4):69–85, 1998. doi:  
840 10.1016/S0022-1694(98)00198-X.
- 841 G. Kuczera, D. Kavetski, S. Franks, and M. Thyer. Towards a bayesian total er-  
842 ror analysis of conceptual rainfall-runoff models: Characterising model error us-  
843 ing storm-dependent parameters. *Journal of Hydrology*, 331(5):61–177, 2006. doi:  
844 10.1029/WR019i005p01163.
- 845 J. Lagarias, J. Reeds, M. Wright, and P. Wright. Convergence properties of the nelder-  
846 mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1):112–147,

- 847 1998.
- 848 L. Li, C. Y. Xu, and K. Engeland. Development and comparison in uncertainty assessment  
849 based bayesian modularization method in hydrological modeling. *Journal of Hydrology*,  
850 486:384–394, 2013. doi: 10.1016/j.jhydrol.2013.02.002.
- 851 D. McInerney, M. Thyer, D. Kavetski, J. Lerat, and G. Kuczera. Improving probabilistic  
852 prediction of daily streamflow by identifying pareto optimal approaches for modeling  
853 heteroscedastic residual errors. *Water Resources Research*, pages n/a–n/a, 2017. doi:  
854 10.1002/2016WR019168.
- 855 H. McMillan and I. K. Westerberg. Rating curve estimation under epistemic uncertainty.  
856 *Hydrol. Process.*, 29:1873–1882, 2015. doi: 10.1002/hyp.10419.
- 857 H. McMillan, B. Jackson, M. Clark, D. Kavetski, and R. Woods. Rainfall uncertainty in  
858 hydrologic modelling: An evaluation of multiplicative error models. *Journal of Hydrol-*  
859 *ogy*, 400:83–94, 2011.
- 860 A. Montanari, R. Rosso, M. S. Taqqu. Fractionally differenced ARIMA models applied to  
861 hydrologic time series: Identification, estimation, and simulation. *Water Resour. Res.*,  
862 33(5):1035–1044, 1997. doi: 10.1029/97WR00043.
- 863 A. Montanari and E. Toth. Calibration of hydrological models in the spectral domain:  
864 an opportunity for ungauged basins? *Water Resour. Res.*, 43(5):W05434, 2007. doi:  
865 10.1029/2006WR005184.
- 866 F. Morlando, L. Cimorelli, L. Cozzolino, G. Mancini, D. Pianese, and F. Garofalo. Shot  
867 noise modeling of daily streamflows: A hybrid spectral- and time-domain calibration ap-  
868 proach. *Water Resources Research*, page Early View, 2016. doi: 10.1002/2015wr017613.

- 869 F. Moussu, L. Oudin, V. Plagnes, A. Mangin, and H. Bendjoudi. A multi-objective  
870 calibration framework for rainfall-discharge models applied to karst systems. *J. Hydrol.*,  
871 400(3-4):364–376, 2011. doi: DOI: 10.1016/j.jhydrol.2011.01.047.
- 872 A. Oppenheim and Schafer. *Discrete-Time Signal Processing, 1st ed.* Prentice-Hall, Upper  
873 Saddle River, NJ, 1989.
- 874 A. Oppenheim and G. Verghese. *Signals, Systems and Inference.* Prentice Hall,, 2015.
- 875 V. R. N. Pauwels and G. J. M. De Lannoy. Multivariate calibration of a water and energy  
876 balance model in the spectral domain. *Water Resour. Res.*, 47(7):W07523, 2011. doi:  
877 10.1029/2010wr010292.
- 878 F. Pianosi and L. Raso. Dynamic modeling of predictive uncertainty by regression on  
879 absolute errors. *Water Resour. Res.*, 48(3):W03516, 2012. doi: 10.1029/2011wr010603.
- 880 W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical recipes: The art of*  
881 *scientific computing (3rd ed.)*. Cambridge University Press, New York, 2007.
- 882 M. Priestley. *Spectral Analysis and Time Series.* Academic Press, London, 1981.
- 883 J. J. Quets, G. J. M. De Lannoy, and V. R. N. Pauwels. Comparison of spectral and time  
884 domain calibration methods for precipitation-discharge processes. *Hydrol. Proc.*, 24(8):  
885 1048–1062, 2010. doi: 10.1002/hyp.7546.
- 886 K. R. Rao, D. N. Kim, J. J. Hwang. *Fast Fourier transform: algorithms and applications.*  
887 Springer, Dordrecht, 423 p., 2010.
- 888 B. Renard, D. Kavetski, M. Leblois, Etienneand Thyer, G. Kuczera, and S. W. Franks. To-  
889 wards a reliable decomposition of predictive uncertainty in hydrologic modeling: Char-  
890 acterizing rainfall errors using conditional simulation. *Water Resour. Res.*, 47:W11516,  
891 2011. doi: 10.1029/2011WR010643.

- 892 B. Schaeffi and E. Zehe. Hydrological model performance and parameter estimation in  
893 the wavelet-domain. *Hydrology and Earth System Sciences*, 13:1921–1936, 2009. doi:  
894 10.5194/hess-13-1921-2009.
- 895 B. Schaeffi, D. Balin Talamba, and A. Musy. Quantifying hydrological modeling errors  
896 through a mixture of normal distributions. *J. Hydrol.*, 332(3-4):303–315, 2007. doi:  
897 10.1016/j.jhydrol.2006.07.005.
- 898 G. Schoups and J. A. Vrugt. A formal likelihood function for parameter and predictive  
899 inference of hydrologic models with correlated, heteroscedastic, and non-gaussian errors.  
900 *Water Resour. Res.*, 46(10):W10531, 2010. doi: 10.1029/2009WR008933.
- 901 P. J. Smith, K. J. Beven, and J. A. Tawn. Detection of structural inadequacy in process-  
902 based hydrological models: A particle-filtering approach. *Water Resour. Res.*, 44(1):  
903 W01410, 2008.
- 904 T. Smith, A. Sharma, L. Marshall, R. Mehrotra, and S. Sisson. Development of a formal  
905 likelihood function for improved bayesian inference of ephemeral catchments. *Water*  
906 *Resour. Res.*, 46:W12551, 2010. doi: doi:10.1029/2010WR009514.
- 907 S. Sorooshian, Q. Duan, and V. K. Gupta. Calibration of rainfall-runoff models: Appli-  
908 cation of global optimization to the sacramento soil moisture accounting model. *Water*  
909 *Resour. Res.*, 29(4):1185–1194, 1993.
- 910 D. Brillinger. *Time series - Data analysis and theory, expanded edition*. Holden-Day Inc.,  
911 San Francisco, 1981.
- 912 M. Thyer, B. Renard, D. Kavetski, G. Kuczera, S. Franks, and S. S. Srikanthan. Critical  
913 evaluation of parameter consistency and predictive uncertainty in hydrological modeling:  
914 A case study using bayesian total error analysis. *Water Resour. Res.*, 45:W00B14, 2009.



915 doi: 10.1029/2008WR006825.

916 A. Viglione, R. Merz, J. L. Salinas, and G. Blschl. Flood frequency hydrology:  
917 3. a bayesian analysis. *Water Resources Research*, 49(2):675–692, 2013. doi:  
918 10.1029/2011wr010782.

919 J. A. Vrugt, C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten. Im-  
920 proved treatment of uncertainty in hydrologic modeling: Combining the strengths of  
921 global optimization and data assimilation. *Water Resour. Res.*, 41:W01017, 2005. doi:  
922 10.1029/2004WR003059.

923 L. M. Ward and P. E. Greenwood. 1/f noise. *Scholarpedia*, 2(12):1537, 2007a.

924 L. M. Ward and P. E. Greenwood. 1/f noise, scholarpe-  
925 dia, 2(12):1537. doi:10.4249/scholarpedia.1537, revision 90924,  
926 [www.scholarpedia.org/article/1/f\\_noise](http://www.scholarpedia.org/article/1/f_noise) (accessed on 29 march 2017). 2007b.

927 P. D. Welch. Use of fast fourier transform for estimation of power spectra - a method  
928 based on time averaging over short modified periodograms. *Ieee Transactions on Audio*  
929 *and Electroacoustics*, AU15(2):70–73, 1967. doi: 10.1109/tau.1967.1161901.

930 B. J. West and M. Shlesinger. The noise in natural phenonema. *American Scientist*, 78  
931 (1):40–45, 1990.

932 H. White. Consequences and detection of misspecified nonlinear regression models. *J.*  
933 *Am. Stat. Assoc.*, 76:419433, 1981.

934 P. Whittle. Estimation and information in stationary time series. *Ark. Mat.*, 2:423434,  
935 1953.

936 H. Winsemius, B. Schaeffli, A. Montanari, and H. Savenije. On the calibration of hydrolog-  
937 ical models in ungauged basins: a framework for integrating hard and soft hydrological

938 information. *Water Resour. Res.*, 45:W12422, 2009. doi: 10.1029/2009WR007706.

939 I. Zhurbenko. Periodogram, <http://www.encyclopediaofmath.org/index.php/periodogram>.

940 In M. Hazewinkel, editor, *This text originally appeared in: Encyclopedia of mathematics*.

941 Kluwer, Dordrecht, 1991.

**Table 1.** Mathematical notations used for the methods (for case studies see Table 1) We use capital letters for stochastic processes, lower case letters for process realizations (samples) or deterministic processes. Bounded domain continuous-time stochastic processes are written as  $(Z_t)_{t=1,\dots,N}$ , bounded domain stochastic discrete-time processes are written as  $(Z_k)_{k=1,\dots,N}$ .

| Notation                 | Type           | Meaning   |
|--------------------------|----------------|---|
| $p()$                    | Function       | Probability density   |
| $E[]$                    | Function       | Expectation   |
| $t$                      | Variable       | Time step of a discrete-time process  |
| $\Delta t$               | Parameter      | Length of the discrete time step  |
| $\ell$                   | Variable       | Time lag (autocorrelation function)   |
| $N$                      | Parameter      | Number of time steps  |
| $\omega$                 | Variable       | Angular frequency index of continuous Fourier transform                       |
| $\omega_f$               | Parameter      | Fundamental frequency, $= \frac{2\pi}{N\Delta t}$                             |
| $\omega_j$               | Variable       | Frequency discrete-time Fourier transform ( $j - th$ multiple of $\omega_f$ ) |
| $j$                      | Variable       | Frequency index of discrete Fourier transform                                 |
| $\mathbf{H}$             | Process        | Hydrological process model  |
| $\mathbf{X}$             | Process        | True hydrol. system input   |
| $\widetilde{\mathbf{X}}$ | Process        | Observed system input   |
| $\mathbf{Y}$             | Process        | True system output  |
| $\widetilde{\mathbf{Y}}$ | Process        | Observed system output  |
| $\mathcal{E}$            | Stoch. Process | Model residual process  |
| $\mathbf{Z}$             | Stoch. Process | General stochastic process  |
| $z$                      | Det. Process   | General deterministic process   |
| $\mathbf{Z}^{(N)}$       | Stoch. Process | Discrete process $Z$ in the bounded domain $N$                                |
| $\varepsilon$            | Det. Process   | Error process realisation   |
| $\delta$                 | Det. Process   | Realisation of error process innovations                                      |
| $\rho$                   | Parameter      | AR(1) autoregressive parameter  |
| $\mu$                    | Parameter      | Error process mean  |
| $\sigma$                 | Parameter      | Error process variance  |
| $\mu_\delta$             | Parameter      | Innovation process mean   |
| $\sigma_\delta$          | Parameter      | Innovation process variance   |
| $\mathbf{y}$             | Det. Process   | Realisation of hydrol. system output  |
| $\widetilde{\mathbf{Y}}$ | Det. Process   | Realisation of observed system output   |
| $\mathbf{x}$             | Det. Process   | Realisation of rainfall input   |
| $\widehat{\mathbf{y}}$   | Det. Process   | Realisation of hydrol. process model (simulation)                             |
| $\mathbb{F}$             | Variable       | Fourier transform (discrete- and continuous-time)                             |
| $\mathbb{P}$             | Variable       | Power density spectrum (PDS)  |
| $\mathbb{Q}$             | Variable       | Stochastic process having PDS as expected value (PDSV)                        |
| $P$                      | Variable       | Sample of the PDSV, Periodogram   |
| $\mathbb{P}_j$           | Variable       | Mean of the periodogram   |

**Table 2.** Table 1 (continued)

| Notation              | Type      | Meaning                                       |
|-----------------------|-----------|---|
| $\zeta_j^Z$           | Variable  | Profile function (Equation 21)                |
| $\zeta_j^\varepsilon$ | Variable  | Spectral profile function of residuals        |
| $f_{\text{FFT}}$      | Operation | Fast Fourier transform (FFT) operation        |
| $\theta$              | Parameter | Parameter vector of hydrol. process model     |
| $\vartheta$           | Parameter | Error model parameters                        |
| $\vartheta_\zeta$     | Parameter | Autocorrelation parameters of error model     |
| $\mu_\varepsilon$     | Parameter | Mean of error process                         |
| $f_j$                 | Function  | Probability density function at frequency $j$ |
| $f_{\chi_1^2}$        | Function  | $\chi^2$ probability density function         |
| $f_{\text{exp}}$      | Function  | Exponential probability density function      |
| $q$                   | Function  | Transformation function, e.g. log             |
| $J_q$                 | Function  | Jacobian of the transformation $q$            |
| $v_\ell$              | Function  | Autocovariance function at lag $\ell$         |

**Table 3.** Summary of all the used case studies.

| Name           | Type      | Purpose   |
|----------------|-----------|---|
| Case study 1.1 | synthetic | uncorrelated process with innovations from a uniform distribution in $[-1, 1]$  |
| Case study 1.2 | synthetic | as 1.1 but Laplace distribution with $\mu = 0, \sigma = 1$  |
| Case study 1.3 | synthetic | as 1.1 but bimodal Gaussian distrib.: $\mu_1 = 1, \mu_2 = -1, \sigma_1 = \sigma_2 = 0.5$ , weight of 1st component: 0.2                   |
| Case study 1.4 | synthetic | AR(1) process with parameters of Gaussian error model of Table 4  |
| Case study 2.1 | synthetic | pure random process with $f_{\mathbb{P}_1}(\omega; (A_1, B_1)) = A_1 \exp(-B_1\omega)$ , $A_1 = 100, B_1 = 0.001$                         |
| Case study 2.2 | synthetic | pure random process with $f_{\mathbb{P}_2}(\omega; (A_2, B_2)) = A_2/\omega^{B_2}$ , $A_2 = 10.5, B_2 = 3$                                |
| Case study 3.1 | synthetic | simple hydrologic model + Gaussian AR(1) error process (Table 4)  |
| Case study 3.2 | synthetic | as 3.1 but Laplacian AR(1) error process  |
| Case study 3.3 | synthetic | as 3.1 but error process with autocorrel. structure $v_\ell = f(\ell; \varrho_1, \varrho_2) = \exp(-\varrho_1\ell - \varrho_2\ell^{0.5})$ |
| Case study 4   | real data | HYMOD model for Leave river + AR(1) Gaussian error process in log-transformed space   |

**Table 4.** Mathematical notations used for the case studies (for methods see Table 1), including the reference values for the model parameters.

| Notation               | Type      | Meaning  | Reference value                    |
|------------------------|-----------|--|------------------------------------|
| $m$                    | Parameter | Number of experiment repetitions                                 |                                    |
| $\mathbf{G}$           | Process   | Rainfall generator model   |                                    |
| $\boldsymbol{\varphi}$ | Parameter | Parameter vector of rainfall generator                           |                                    |
| $\lambda$              | Parameter | Arrival rate (rainfall generator)                                | 0.5 d <sup>-1</sup>                |
| $\bar{i}_r$            | Parameter | Rainfall event mean intensity (rainfall generator) [L/T]         | 3.3 mm/d                           |
| $\bar{t}_r$            | Parameter | Rainfall event mean duration (rainfall generator) [T]            | 0.8 d                              |
| $\rho$                 | Parameter | Autocorrelation parameters (synthetic rainfall-runoff)           |                                    |
| $\ell$                 | Variable  | Lag (synthetic rainfall-runoff)                                  |                                    |
| $k_1$                  | Parameter | Linear reservoir coefficient 1 [1/T] (synthetic rainfall-runoff) | 0.1 d <sup>-1</sup>                |
| $k_2$                  | Parameter | Linear reservoir coefficient 2 [1/T] (synthetic rainfall-runoff) | 0.05 d <sup>-1</sup>               |
| $l_g$                  | Parameter | Leaching parameter [L/T] (synthetic rainfall-runoff)             | $\lambda \bar{i}_r \bar{t}_r$ mm/d |
| $\rho_\varepsilon$     | Parameter | AR1 parameter (error model)                                      | 0.8                                |
| $\mu_\delta$           | Parameter | Mean of innovations (error model)                                | 0.25                               |
| $\sigma_\delta$        | Parameter | Standard deviation of innov. (error model)                       | 0.015                              |
| $s_{\max}$             | Parameter | Max. storage (HYMOD)[L]  |                                    |
| $b_H$                  | Parameter | Spatial variability (HYMOD)[-]                                   |                                    |
| $\alpha$               | Parameter | Flow splitting (HYMOD)[-]  |                                    |
| $\nu_s$                | Parameter | Residence time slow reservoir (HYMOD) [T]                        |                                    |
| $\nu_q$                | Parameter | Residence time fast reservoirs (HYMOD) [T]                       |                                    |
| $m_H$                  | Parameter | Number of fast reservoirs (HYMOD) [-]                            |                                    |
| $A$                    | Parameter | Numerical offset for log computation                             |                                    |

**Table 5.** Leaf River case study: limits of the uniform priors, parameter values inferred using BTL and BSL (columns denoted with *inf*), and empirical quantities estimated from the computed residuals (columns denoted with *emp*). The number of fast reservoirs,  $m_H$ , is fixed to 2 after initial optimization. The prior for  $\sigma_\delta$  is obtained according to the method used in [Schaeffli et al., 2007]. NSE stands for the Nash-Sutcliffe efficiency.

|           | $s_{\max}$<br>mm | $b_H$ | $\alpha$ | $\nu_s$ | $m_H$ | $\rho$ inf | $\mu_\delta$ inf<br>mm/d | $\sigma_\delta$ inf<br>mm/d | $\rho$ emp | $\mu_\delta$ emp<br>mm/d | $\sigma_\delta$ emp<br>mm/d | NSE  |
|-----------|------------------|-------|----------|---------|-------|------------|--------------------------|-----------------------------|------------|--------------------------|-----------------------------|------|
| Prior min | 50               | 0.05  | 0.01     | 0.001   | 0.001 | 0          | -0.25                    | 0                           | -          | -                        | -                           | -    |
| Prior max | 800              | 1.95  | 1        | 0.20    | 0.95  | 0.99       | 0.25                     | -                           | -          | -                        | -                           | -    |
| BTL-Gauss | 149              | 0.51  | 0.11     | 0.16    | 0.01  | 0.90       | -0.02                    | 0.19                        | 0.71       | -0.05                    | 2.08                        | 0.70 |
| BSL       | 174              | 0.38  | 0.10     | 0.15    | 0.01  | 0.89       | 0.01                     | 0.20                        | 0.73       | -0.02                    | 2.10                        | 0.68 |
| MTL       | 173              | 0.34  | 0.08     | 0.08    | 0.84  | 0.97       | -0.01                    | 0.21                        | 0.80       | -0.02                    | 2.13                        | 0.59 |