

Document downloaded from:

<http://hdl.handle.net/10251/128735>

This paper must be cited as:

Romero-Cuellar, J.; Abbruzzo, A.; Adelfio, G.; Francés, F. (2019). Hydrological post-processing based on approximate Bayesian computation (ABC). *Stochastic Environmental Research and Risk Assessment*. 33(7):1361-1373. <https://doi.org/10.1007/s00477-019-01694-y>



The final publication is available at

<http://doi.org/10.1007/s00477-019-01694-y>

Copyright Springer-Verlag

Additional Information

# Hydrological post-processing based on approximate Bayesian computation (ABC)

Jonathan Romero-Cuellar\* · Antonino  
Abbruzzo · Giada Adelfio · Félix  
Francés

Received: date / Accepted: date

**Abstract** This study introduces a method to quantify the conditional predictive uncertainty in hydrological post-processing contexts when it is cumbersome to calculate the likelihood (intractable likelihood). Sometimes, it can be difficult to calculate the likelihood itself in hydrological modelling, specially working with complex models or with ungauged catchments. Therefore, we propose the ABC post-processor that exchanges the requirement of calculating the likelihood function by the use of some sufficient summary statistics and synthetic datasets. The aim is to show that the conditional predictive distribution is qualitatively similar produced by the exact predictive (MCMC post-processor) or the approximate predictive (ABC post-processor). We also use MCMC post-processor as a benchmark to make results more comparable with the proposed method. We test the ABC post-processor in two scenarios: i) the Aipe catchment with tropical climate and a spatially-lumped hydrological model (Colombia) and ii) the Oria catchment with oceanic climate and a spatially-distributed hydrological model (Spain). The main finding of the study is that the approximate (ABC post-processor) conditional predictive uncertainty is almost equivalent to the exact predictive (MCMC post-processor) in both scenarios.

**Keywords** Free-likelihood approach · probabilistic modelling · uncertainty analysis · hydrological forecasting · summary statistics

---

Jonathan Romero-Cuellar\* (corresponding author) and Félix Francés  
Research Institute on Water and Environmental Engineering, Universitat Politècnica de València, Spain, E-mail: jorocue1@upv.es

Antonino Abbruzzo and Giada Adelfio  
Department of Economics, Business and Statistics, Università degli Studi di Palermo, Italy

## 1 Introduction

Making unbiased, accurate and reliable streamflow predictions has regularly been one of the main goals for hydrologists. These hydrological predictions are valuable for risk assessment, water resources management, and ecological issues. Characterizing, quantifying, reducing, and communicating the uncertainty of predictions is essential for decision-making and water management under anthropogenic conditions (Butts et al., 2004; Liu and Gupta, 2007). Uncertainty is everywhere and is impossible to avoid it (Lindley and Smith, 1972). Indeed, uncertainty is a fact of hydrology (Wilby and Harris, 2006). Generally speaking, uncertainty analysis is a crucial part of hydrological modelling process (Schoups and Vrugt, 2010). It is particularly useful for modelling comparison and selection (Schoups et al., 2008), improving model predictions, supporting decision-making (Reichert et al., 2015) and advancing towards reliable measurement systems. Most significantly, uncertainty analysis plays a considerable role in applications and the dialogue with decision-makers (Montanari and Koutsoyiannis, 2012). Some sources of uncertainty include input errors (e.g. rainfall sampling, low gauge density, interpolation method), epistemic errors (e.g. model parameters, model structure), and output errors (e.g. associated with rating curve errors). The propagation of confidence bounds from different uncertainty sources to model output is crucial for hydrologic modelling (Liang et al., 2012). Applying statistical post-processing methods is a useful approach to quantify the joint effect of these uncertainties. Hydrologic post-processors are statistical models that relate observed variables of interest (streamflow, water level) to predictors derived from deterministic hydrologic model outputs (Ye et al., 2014). The hydrologic post-processing aim is to reduce biases and quantify the uncertainty of deterministic predictions (Wentao et al., 2017).

In the context of conditional predictive uncertainty, which means fixed hydrological predictions, several techniques have been developed to quantify total uncertainty. Early works included methods as Model Output Statistics (MOS) (Glahn and Lowry, 1972) and Hydrological Uncertainty Processor (HUP) (Krzysztofowicz and Kelly, 2000). More recent literature used the Bayes's theorem-based methods, such as Bayesian Model Average (BMA) (Raftery et al., 2005; Vrugt and Robinson, 2007), Model Conditional Processor (MCP) (Todini, 2008; Coccia and Todini, 2011) and Bayesian Joint Probability (BJP) (Wang et al., 2009). Moreover, there exists a variety of regression-based models, including a meta-Gaussian approach (Montanari and Brath, 2004; Montanari and Grossi, 2008), quantile regression (Weerts et al., 2011) and General Linear Model Post-Processor (GLMPP) (Zhao et al., 2011). Also, many other methods have been proposed, including non-parametric post-processor (Brown and Seo, 2010), machine learning (Solomatine and Shrestha, 2009), data-driven resampling techniques (Sikorska et al., 2015), Bayesian neural networks (Zhang and Zhao, 2012) and post-processing with error model (Evin et al., 2014). Several copula models have been proposed like a BMA-copulas (Madadgar

and Moradkhani, 2014), pair-copulas in a multi-model ensemble (Klein et al., 2016) and ensemble copula coupling (Scheffzik et al., 2013).

Although there are many approaches to improve hydrologic predictions by reducing uncertainties, they have not been standardized (Wagener and Gupta, 2005; Montanari and Koutsoyiannis, 2012; Wentao et al., 2017; van Oijen, 2017), and less attention has been paid in presence of intractable likelihood. The Approximate Bayesian Computation (ABC) (Nott et al., 2011; Vrugt and Sadegh, 2013; Fenicia et al., 2018; Kavetski et al., 2018) method deals with inferential problems with intractable likelihood. By intractable likelihood, we mean that the likelihood function is unavailable in closed form or by numerical derivation (Robert, 2016). In this context and for the sake of simplicity, we use a monthly error model that is useful for water resources management applications. To our best knowledge, up to now, this is the first study that has proposed a hydrological post-processor based on approximate Bayesian computation (ABC). This study introduces a method to quantify the conditional predictive uncertainty in hydrological post-processing contexts when it is cumbersome to calculate the likelihood (intractable likelihood). Sometimes, it can be difficult to calculate the likelihood itself in hydrological modelling, specially working with complex models or with ungauged catchments. Therefore, we propose the ABC post-processor that exchanges the requirement of calculating the likelihood function by the use of some sufficient summary statistics and synthetic datasets. The aim is to show that the conditional predictive distribution is qualitatively similar produced by the exact predictive (MCMC post-processor) or the approximate predictive (ABC post-processor). We test the ABC post-processor in two scenarios: i) The Aipe catchment (poor predictions) and ii) the Oria catchment (good predictions). Deterministic and probabilistic verification frameworks are used to compare the performance of the ABC post-processor with the Markov Chain Monte Carlo (MCMC) approach (Gelman et al., 2013), that works when the likelihood is tractable. The rest of the paper is structured as follows. The theory and methods are described in Section 2, applications in Section 3, followed by discussion and conclusions in Section 4.

## 2 Theory and methods

Biased, inaccurate, and unreliable predictions in hydrology are mainly consequence of several sources of uncertainties. A hydrologic post-processor is an approach to deal with uncertainties from deterministic hydrologic model outputs propagated from all upstream sources. Applying statistical post-processing methods is useful to quantify these uncertainties. In this section, we describe the theory of hydrologic post-processing focussing both on algorithms dealing with intractable likelihood (ABC post-processor) and tractable likelihood (MCMC post-processor). These two post-processors are compared through verification metrics to assess their performance.

## 2.1 Monthly streamflow post-processor

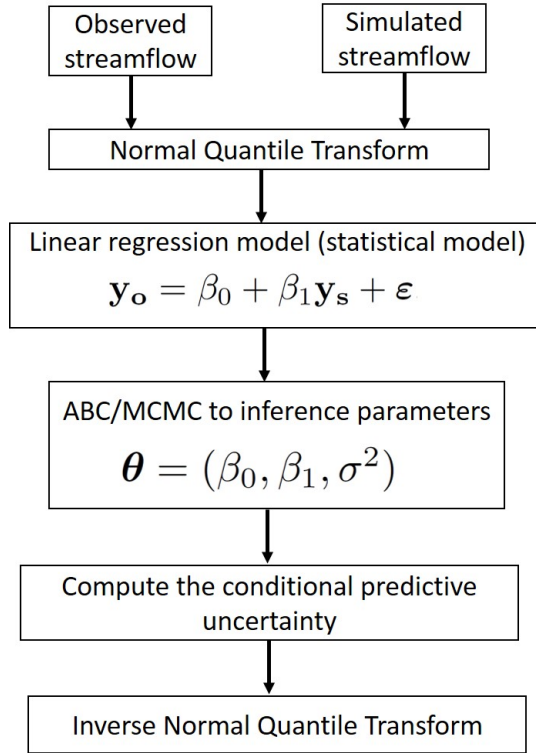
Let  $\mathbf{y}^s = (y_1^s, \dots, y_T^s)^\top$  be the output of a deterministic hydrological model and  $\mathbf{y}^o = (y_1^o, \dots, y_T^o)^\top$  the observations. The hydrologic post-processor works by relating model outputs (e.g., streamflow) to corresponding observations through a statistical model (Ye et al., 2014). It serves the purpose of removing model biases from all upstream uncertainty sources. In this paper, we assume a linear model between  $\mathbf{y}^o$  and  $\mathbf{y}^s$

$$\mathbf{y}^o = \beta_0 + \beta_1 \mathbf{y}^s + \boldsymbol{\varepsilon}, \quad (1)$$

where the vectors  $\mathbf{y}^o$  and  $\mathbf{y}^s$  are expressed in  $m^3 s^{-1}$ ,  $\beta_0$  and  $\beta_1$  are statistical parameters and  $\boldsymbol{\varepsilon}$  is a random variable that represents the error term in the statistical model. The three components on the right-hand side symbolise three distinct sources of estimation errors. The first term,  $\beta_0$ , represents constant deviation and can be called the displacement error. The second term, with an error parameter  $\beta_1$ , denotes a scale error or dynamic-range error. The third term is a random error, which is assumed as independent and identically distributed, with zero mean value and a standard deviation of  $\sigma$ . The first two terms describe a deterministic relationship between  $\mathbf{y}^o$  and  $\mathbf{y}^s$ , and they characterise the systematic or epistemic error with  $\beta_0$  and  $\beta_1$ . The error term expresses random fluctuations due to the effect of factors out of our control or measurement, and it is assumed that  $\boldsymbol{\varepsilon}_i \sim N(0, \sigma^2)$ , identical independent distributed (i.i.d.). This assumption is the most common, but it can be relaxed in favour of more general cases (e.g. Schoups and Vrugt (2010)). The linear model and the assumption of normality on the error term can be too restrictive for hydrological post-processors. Nevertheless, the linear error model is our first approximation although it is not always appropriate. We know that some circumstances require non-linear error models, but this linear error model worked for our water resources management application. In fact, Tian et al. (2016) proved that a linear model and three parameters are sufficient to fully capture the characteristic error of monthly predictions. The ABC is not strongly influenced by the model and the assumptions, since it is based on a distance measure between summary statistics of the observed and simulated data. This point will be clarified in Section 2.3. To complete the error assumptions ( $\boldsymbol{\varepsilon}$ ), the observed and simulated streamflows are transformed to the Normal space previously applying the Normal Quantile Transformation procedure (NQT). Waerden (1953) described the theory behind the NQT, and Krzysztofowicz and Kelly (2000) demonstrated its application in hydrology. Figure 1 shows the steps we used to derive the conditional predictive uncertainty distribution.

## 2.2 MCMC for monthly streamflow post-processor in Bayesian framework

In Bayesian Statistics, parameters are treated as random variables and inference is based on the posterior parameter distribution. The posterior parameter



**Fig. 1** Flow chart of the process we used to derive the conditional predictive uncertainty distribution.

distribution of model (1), can be written, from the Bayes theorem, as:

$$p(\theta | \mathbf{y}^o, \mathbf{y}^s) = \frac{p(\mathbf{y}^o | \theta, \mathbf{y}^s) p(\theta)}{\int p(\mathbf{y}^o | \theta, \mathbf{y}^s) p(\theta) d\theta}, \quad (2)$$

where  $p(\theta)$  indicates the prior parameter distribution,  $p(\mathbf{y}^o | \theta, \mathbf{y}^s)$  denotes the likelihood of  $\mathbf{y}^o$  conditional on the parameters  $\theta = (\beta_0, \beta_1, \sigma^2)$  and the deterministic output  $\mathbf{y}^s$ . We assume flat uniform priors for  $\theta = (\beta_0, \beta_1, \sigma^2)$  and from the assumptions on model (1) it follows that  $\mathbf{Y}^o | \theta, \mathbf{y}^s \sim N(\mu = \beta_0 + \beta_1 \mathbf{y}^s, \sigma^2)$ . Given the model it is not possible, as most of the time, to compute in closed form the integral in the denominator of the equation (2). So, we approximate the posterior distribution (2) by using the MCMC algorithms (Gelman et al., 2013). Specifically, we use the adaptive Metropolis-Hastings (Haario et al., 2001) to perform the Bayesian inference. This algorithm has been shown to perform adequately in hydrologic problems (Marshall et al., 2004). For detail on the implementation of adaptive Metropolis-Hastings algorithm to hydrologic modelling studies, relate to the research stated by Marshall et al. (2004).

In this paper we are most interested in conditional prediction uncertainty. Let  $\tilde{y}^o$  be a future observation for model (1), then the posterior predictive

density (which incorporate our uncertainty) of a future observation  $p(\tilde{y}_o|\mathbf{y}_s)$ , is given by

$$p(\tilde{y}_o|\mathbf{y}_s) = \int_{\Theta} p(\tilde{y}_o|\boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{y}_o, \mathbf{y}_s)d(\boldsymbol{\theta}). \quad (3)$$

In words, the posterior predictive density is an average of conditional predictions over the posterior distributions of parameters (Gelman et al., 2013), reflecting both the uncertainty of the model and the uncertainty due to variability in future observations (Yoon et al., 2010).

### 2.3 ABC post-processor

The idea behind the ABC approach was introduced in population and evolutionary genetics (Pritchard et al., 1999; Tavaré et al., 1997). Furthermore, Nott et al. (2011) were the first to introduce the ABC method in hydrology community. ABC is adequate for inference problems where sampling from the assumed probability model is much easier than evaluating its probability density function (e.g., intractable likelihood) (Fenicia et al., 2018). Using the ABC does not evade the requirement of Bayesian inference to stipulate a probability model of the data, but rather exchanges the requirement of calculating the likelihood function by the requirement of sampling model output realizations (Kavetski et al., 2018). ABC is a class of sampling methods that bypass exact likelihood calculations with a simulation of the model that produces synthetic datasets. The method then relies on some metric (distance) to compare simulated data to the data that were observed (Turner and Van Zandt, 2012). Then, the aim is to obtain an estimate of the posterior distribution of the parameter of the model. Recall that the posterior parameter  $\boldsymbol{\theta}$  is the distribution of that parameter conditioned on the observed data and the deterministic output. Without a likelihood, it is not possible to write down an expression for this posterior, or to estimate it using Monte Carlo methods. However, we can simulate data  $\mathbf{y}_{sim}$  using some  $\boldsymbol{\theta} = \boldsymbol{\theta}^*$  and retain  $\boldsymbol{\theta}^*$  as a sample from the posterior if some pre-defined distance  $d\{\mathbf{y}_{sim}, \mathbf{y}^o\}$  between the observed and the simulated data is less than some small value  $\epsilon_0$ . There are three main ABC algorithms: 1) accepted-rejected (Beaumont et al., 2002), 2) Markov Chain Monte Carlo ABC (Marjoram et al., 2003), and 3) Sequential Monte Carlo ABC (Sisson et al., 2007). Although some ABC algorithms are more efficient, whether the ABC algorithm is optimized (or not), is not relevant to the results; hence, we use the accepted-rejected algorithm of Beaumont et al. (2002) for our application. We developed the ABC post-processor in R (R Core Team, 2013) using the package abc (Csillery et al., 2012). The pseudo code for the ABC is summarized in Algorithm 1.

Algorithm 1 proceeds in the following way: first, we sample a candidate parameter value  $\boldsymbol{\theta}^*$  from the flat prior distribution. We then use this candidate to simulate a dataset  $\mathbf{y}_{sim}$  from the normal model of interest that has the same number of observations as the observed data  $\mathbf{y}^o$ . Thereafter, we compare the simulated data  $\mathbf{y}_{sim}$  to the observed data  $\mathbf{y}_o$  by computing a distance between

**Algorithm 1** ABC accepted/rejected algorithm

- 
- 1: Sample  $\theta_i^*$ ,  $i = 1, \dots, N$  from the prior:  $\theta^* \sim p(\theta)$
  - 2: Generate data  $\mathbf{y}_{sim}^i = (y_1^i, y_2^i, \dots, y_T^i)^\top$ ,  $i = 1, \dots, N$ , from the model,  $p(\cdot|\theta_i^*)$
  - 3: For  $0 \leq i \leq N$ , store  $\theta_i^*$  if:

$$d\{\eta(\mathbf{y}_{sim}^i), \eta(\mathbf{y}^o)\} \leq \epsilon_0$$

where  $\eta(\cdot)$  is a vector statistic,  $d\{\cdot, \cdot\}$  is a distance criterion, and, given  $N$ , the tolerance level  $\epsilon_0$  is chosen to be small.

---

them given by a distance function  $d\{\eta(\mathbf{y}_{sim}), \eta(\mathbf{y}^o)\} \leq \epsilon_0$ . For computational ease, it is often convenient to define  $d\{\cdot, \cdot\}$  as a distance between summary statistics  $S(\mathbf{y}_{sim})$  and  $S(\mathbf{y}^o)$ . Ideally, the summary statistics  $S(\cdot)$  should be sufficient for the parameter  $\theta$ . In this study, we consider five summary statistics including the sample mean, variance, skewness, kurtosis, and first sample autocorrelation (Fearnhead and Prangle, 2012) and run the algorithm 1 using the Euclidean distance between summary statistics and tolerance level  $\epsilon_0 = 0.01$ . Algorithm 1 produces the empirical posterior parameter distribution which we indicate with  $p_{\epsilon_0}^*(\theta|\eta(\mathbf{y}_{sim}), \eta(\mathbf{y}^o))$ . Estimates of the parameter  $\theta$  can be obtained by calculating the mean, mode or median of this empirical distribution. However, our interest in the paper is not on these estimates but on the predictive posterior uncertainty.

The predictive posterior uncertainty is formally defined as

$$g(\tilde{y}^o|\mathbf{y}^s) = \int_{\Theta} p(\tilde{y}^o|\theta, \mathbf{y}^o, \mathbf{y}^s)p(\theta|\mathbf{y}^o)d\theta, \quad (4)$$

and it is approximated by

$$g^*(\tilde{y}^o|\mathbf{y}^s) = \frac{1}{M} \sum_{i=1}^M p(\tilde{y}^o|\theta_i^*, \mathbf{y}^o),$$

where  $M$  is the number of retained  $\theta^*$ .

It has been shown that even though the posterior parameter predictive distribution  $p_{\epsilon_0}^*(\theta|\eta(\mathbf{y}_{sim}), \eta(\mathbf{y}^o))$  is not close to the true posterior distribution in (2) the posterior predictive distribution  $g^*(\tilde{y}^o|\mathbf{y}^s)$  under some regularity conditions can still be valid to approximate the (4) (Marin et al., 2012). In particular, there are three main regularity conditions: first, the data generating process (DGP) is correctly specified (probability model); second, both  $p_{\epsilon_0}^*(\theta|\eta(\mathbf{y}_{sim}), \eta(\mathbf{y}^o))$  and  $p(\theta|\mathbf{y}^o, \mathbf{y}^s)$  are Bayesian consistent for the true value of  $\theta$  and large samples. Blackwell and Dubins (1962) and Diaconis and Freedman (1986) proved that, if all three regularity conditions are satisfied, then the  $g(\tilde{y}^o|\mathbf{y}^s)$  and  $p(\tilde{y}^o|\mathbf{y}^s)$  yield the same forecasting asymptotically. In other words,  $g(\tilde{y}^o|\mathbf{y}^s)$  and  $p(\tilde{y}^o|\mathbf{y}^s)$  merge asymptotically. Frazier et al. (2019) recently demonstrated theoretically and numerically the previous conclusion in economic models. Finally, the motivation for the use of ABC in hydrological models is evident: in cases where the likelihood is not accessible, the parametric posterior distribution itself is inaccessible and the integral in (3) cannot be computed via the MCMC methods.



## 2.4 Verification metrics

Deterministic and probabilistic verification frameworks are used to assess outputs from the proposed ABC post-processor and the MCMC approach. We examine the accuracies, reliability, and robustness of the proposed method. These verification metrics are analysed during both the calibration and validation periods. In general, uncertainty analysis methods could be portrayed by the 95% uncertainty band that has to be as narrow as possible but still containing the largest amount of observations. Since verification metrics have been reported by a large and adequate body of literature, only a brief description of each is presented here.

As deterministic metrics, we include the Nash-Sutcliffe Efficiency (NSE), and the Kling-Gupta Efficiency (KGE) indices. NSE has been extensively applied to assess hydrological models. Likewise, KGE was presented as the modified version of NSE by Gupta et al. (2009). This metric involves the correlation, bias, and variability. Both NSE and KGE can range from  $-\infty$  to 1 with NSE or KGE = 1 as a perfect fit between observation and simulation. As probabilistic metrics, we include the reliability and precision. Reliability refers to the statistical consistence of predictions with observed data, and precision refers to the concentration of the predictive distribution (small uncertainty). Zero is the worst reliability value while one is the best. To evaluate the reliability and precision of predictive distributions, Laio and Tamea (2007) suggested the use of PQQ plots (Thyer et al., 2009). In the PQQ plot context, if the predictive distribution and observed data are consistent, the corresponding p-value distribution should be uniformly distributed over the interval  $[0,1]$ . We apply the Kolmogorov-Smirnov test (K-S) to check this uniformity.

Finally, to gain more insight into the probabilistic metrics and following Li et al. (2017), we also compute the containing ratio (CR), which is the percentage of the measurement bracketed by this band, the average bandwidth of 95% uncertainty band (B) and the average deviation amplitude (D). We use the 95% prediction interval based on the 5 and 95 percentiles. As a result, an adequate predictive uncertainty is achieved when the CR is close to 95%. The smallest values of B and D are preferred. These three indices quantify the degree of predictions deviating from observations. Our strategy to compute comparative performance metrics is compatible with others similar studies such as Shafii et al. (2014); Ye et al. (2014); Khajehei and Moradkhani (2017).

## 3 Applications

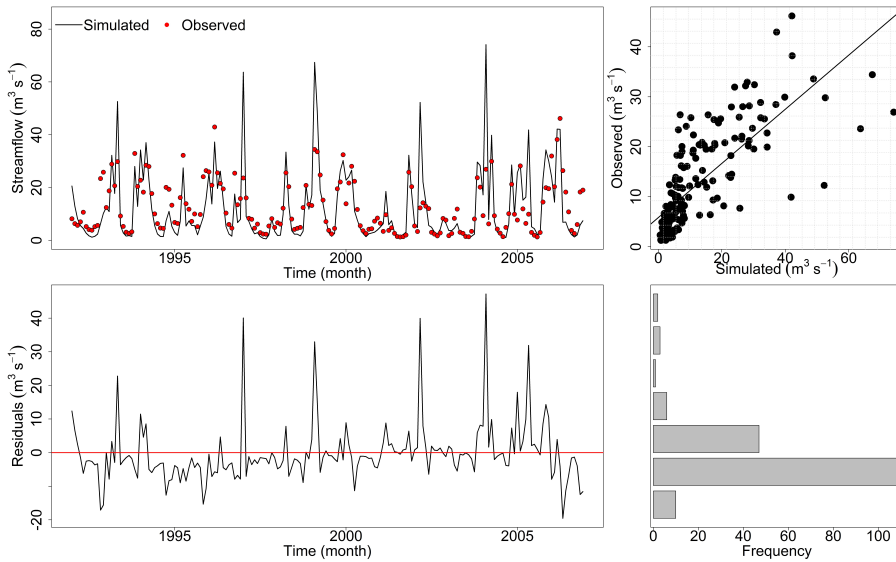
Both methods are applied to monthly streamflows in two scenarios: the Aipe catchment with poor predictions (Colombia) and the Oria catchment with good predictions (Spain). By poor and good predictions, we mean that the  $NSE < 0.5$  and  $NSE > 0.8$  respectively. These scenarios allow a contrast in hydrology. The essential hydrologic features of both scenario catchments are summarised in Table 1.

**Table 1** Hydrologic features of the two case studies catchments.

Catchment	Area ( $km^2$ )	P* (mm/year)	PET (mm/year)	Q (mm/year)	Run-off ratio (Q/P)	Aridity ratio (PET/P)
Aipe	688.9	1922.71	1981.54	706.89	0.377	1.031
Oria	73	1498	733.4	765	0.511	0.489

\* P: Mean areal precipitation, PET: potential evapotranspiration, and Q: streamflow.

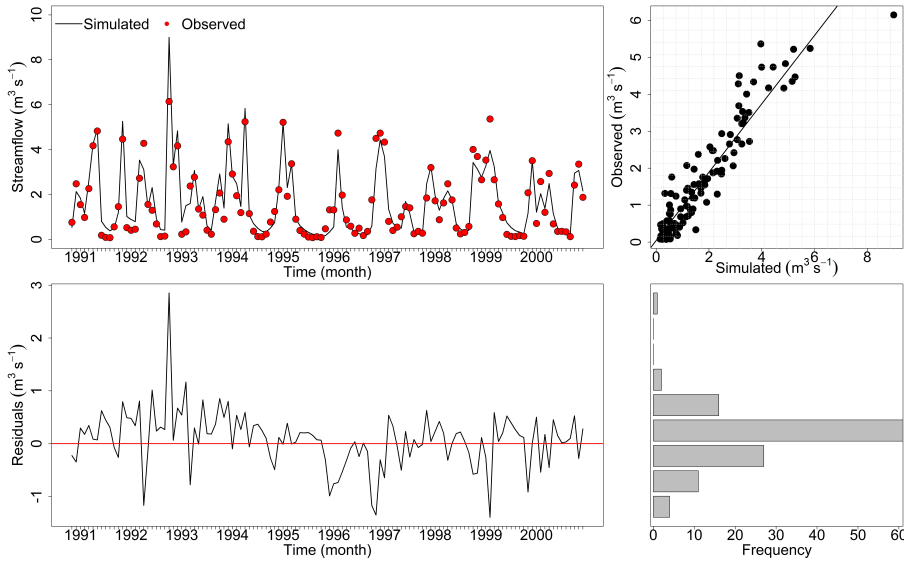
The first scenario (poor predictions) is the Aipe river catchment in Huila State, in southern Colombia. Precipitation, potential evaporation and streamflow monthly time series are available from 1992 to 2012. The first fourteen years of data are used for model calibration, while the last six years served as a validation dataset to assess predictive capability. We used the abcd water balance model to simulate streamflows (Thomas, 1981). This model is a well known conceptual spatially-lumped rainfall-runoff model which transforms precipitation and potential evapotranspiration data to streamflow at the catchment outlet. The hydrological model is selected for its conceptual simplicity and general usage. Figure 2 represents the output from the hydrological model. The abcd water balance model implemented in Aipe catchment is described in detail by Romero-Cuellar et al. (2018).



**Fig. 2** Monthly time series of deterministic streamflow predictions (solid line) and observations (red dots) from the Aipe river catchment. Scatter plot of simulated versus observed streamflows. Time series and histogram of the residuals.

The simulated streamflow time series in Fig. 2 shows that deterministic hydrological predictions overestimate the observed streamflows. Moreover, the

residuals time series shows high values, and the histogram of errors indicates a negative bias. The second scenario (good predictions) is the Oria river catchment located in the Basque Country Region, in northern Spain. The hydrological data for this catchment were collected from 1987 to 2000. We used ten years (i.e., 1990-2000) for calibration process and three years for validation (i.e., 1987-1990). Daily runoff simulations from the TETIS model aggregated to monthly values are used to estimate the predictive uncertainty at the C2Z1 Agauntza gauge. The TETIS model is a conceptual spatially-distributed hydrological model where each grid cell represents a tank model with six tanks connected among them. TETIS is a grid-based model, which takes advantage of all the spatially distributed information available. More details about the TETIS model and Oria application are in Francés et al. (2007); Vélez et al. (2009). Figure 3 represents the performance of the hydrological model in the Oria river.



**Fig. 3** Monthly time series of deterministic streamflow predictions (solid line) and observations (red dots) from the Oria river catchment. Scatter plot of simulated versus observed streamflows. Time series and histogram of the residuals.

In contrast to the Aipe river, in the Oria river the hydrological model has better performance. The simulated streamflow time series in Fig. 3 indicates that deterministic hydrological predictions correspond to the observed streamflows. Moreover, the error histogram is centred in zero. For each of the case studies, we calibrate a set of hydrological parameters. Calibrated parameters were achieved through optimisation which minimises the aggregated differences between simulated and observed streamflow values. Then, we use outputs from the hydrological model as inputs for the statistical model (hydro-

logical post-processing). Next, the adaptative Metropolis-Hastings algorithm is implemented and run until the convergence of the parameter posterior distribution (statistical model) is achieved. Convergence is determined by both the visual trace plot evaluation of the posterior chains and the Gelman and Rubin (1992) R statistic which considers convergence in terms of the variance with a single chain and the variance between multiple parallel chains. Finally, we compute the conditional predictive distribution.

The results for monthly streamflow forecast are now presented. They are presented independently for both scenarios catchment and calibration and validation period. In section 3.1 the performance of the Aipe catchment is evaluated, while in section 3.2 the performance of the Oria catchment is assessed. For each of the scenarios, the MCMC algorithm is first applied to estimate the predictive uncertainty in time series domain, and the ABC algorithm is then used to assess the predictive uncertainty in summary statistic domain and free-likelihood function. As mentioned before, this study shows that the predictive superiority of the exact predictive (MCMC post-processor), over the approximate (ABC post-processor) using some sufficient summary statistics and synthetic datasets, is minimal. To achieve this aim, we use MCMC post-processor as a benchmark to make results more comparable with the proposed method.

### 3.1 First Scenario: The Aipe Catchment

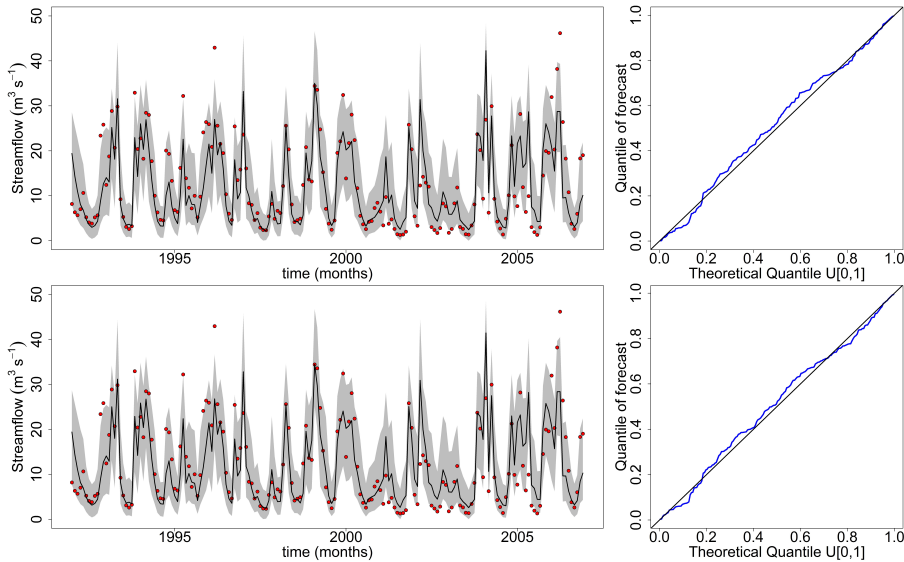
Figure 2 represents a general view of hydrological model performance (deterministic predictions). Note that the hydrological model does not mimic observed data quite well, it does not reproduce maximum streamflow events. In other words, the model over-predicts peak streamflows. Moreover, Fig. 2 shows that the error variance is heteroscedastic, and the error histogram is not normal. In summary, the hydrological model (deterministic predictions) has a poor performance. Regarding deterministic metrics, the most considerable improvement is the result of post-processing approaches (Table 2). For example, when post-processing is used, the NSE increases in as many as 74.63% for the calibration and 25.84% for the validation period. The acceptable range for NSE is considered to be above 0.5 (Moriassi et al., 2007); therefore, results in Table 2 point that the deterministic prediction does not have a satisfying skill as compared with the post-processing approaches, which frequently guides to valid performance metrics because they work directly to improve the errors in model outputs (Ye et al., 2014). Visible improvements are also inspected concerning KGE; for instance, KGE increases in as many as 30.3% for the calibration and 16% for the validation period (improvements are not as pronounced as for NSE).

Regarding probabilistic metrics, the predictive PQQ plot of the MCMC (upper-right) and ABC (lower-right) post-processing is presented in Fig. 4 during the calibration period. Fig. 4 shows realistic narrow predictive bounds because only some peak flows are not bracketed, and all low streamflows are

**Table 2** Deterministic and probabilistic performance metrics of the raw prediction, MCMC and ABC post-processor for the Aipe catchment.

Performance metric	Calibration			Validation		
	Deterministic prediction	Post-processing MCMC	Post-processing ABC	Deterministic prediction	Post-processing MCMC	Post-processing ABC
NSE	0.165	0.669	0.671	0.571	0.777	0.773
KGE	0.527	0.769	0.764	0.637	0.757	0.744
Reliability	-	0.996	0.996	-	0.993	0.993
Precision	-	2.403	2.306	-	2.581	2.500
K-S test (p-value)	-	0.465	0.750	-	0.132	0.223
B ( $m^3/s$ )	-	14.95	15.64	-	25.78	26.86
CR (%)	-	88.33	88.89	-	94.44	95.83
D ( $m^3/s$ )	-	6.82	6.92	-	12.23	12.42

bracketed. It is important to notice that the upper and lower predictive uncertainty appearing in the Fig. 4 are almost identical. Besides, it is clear from these figures that the curves closely follow the bisector. This means that the predictive distributions of post-processed streamflows are reliable. We also see that reliability indexes are close to one, further confirming that predictions are reliable (Table 2).

**Fig. 4** Conditional predictive uncertainty from MCMC (upper) and ABC (lower) post-processor on the Aipe catchment. PQQ-plot of the conditional predictive distribution (right). Dots indicate observations, line indicates median prediction and grey region indicates 90% uncertainty.

Generally speaking, we do not find crucial differences between the MCMC and ABC post-processing approaches. Reliability and precision metrics are

practically equal for both post-processing methods in calibration and validation periods (Table 2). Additionally, both techniques pass the K-S test. In the calibration period, the B and D metrics of ABC post-processor are slightly higher than the MCMC post-processor. Table 2 further reports the coverage rate (CR) for testing the sharpness of the conditional predictive uncertainty. Perfect predictive distribution would expect that the CR close to the assumed 90% prediction level. In the calibration period, the CR for the two post-processors are quite similar. In general, all verification metrics of both methods deteriorate in the validation period except for the CR that improves slightly. These results suggest that the ABC post-processor which uses just some sufficient summary statistics and a free-likelihood function may have similar performance to the MCMC post-processor that uses a likelihood function. Therefore, the ABC post-processor has a satisfactory performance. The second scenario (section 3.2) investigates whether these findings would hold for good hydrological predictions.

### 3.2 Second Scenario: The Oria Catchment

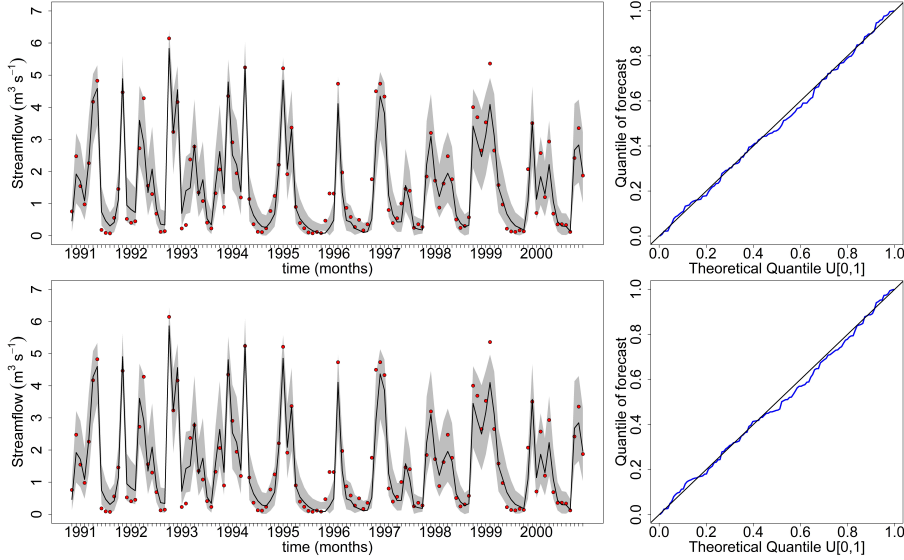
In the second scenario, we analyse an oceanic climate and spatially-distributed hydrological model. In contrast to the first catchment, the hydrological model (deterministic predictions) has a good performance as seen in verification metrics in Table 3.

**Table 3** Deterministic and probabilistic performance metrics of the raw prediction, MCMC and ABC post-processor for the Oria catchment.

Performance metric	Calibration			Validation		
	Deterministic prediction	Post-processing		Deterministic prediction	Post-processing	
		MCMC	ABC		MCMC	ABC
NSE	0.875	0.910	0.911	0.939	0.955	0.956
KGE	0.918	0.903	0.910	0.891	0.909	0.917
Reliability	-	0.995	0.995	-	0.982	0.982
Precision	-	2.950	2.870	-	2.280	2.190
K-S test (p-value)	-	0.972	0.923	-	0.868	0.872
B ( $m^3/s$ )	-	1.47	1.51	-	1.34	1.37
CR (%)	-	86.07	86.07	-	77.78	80.56
D ( $m^3/s$ )	-	0.85	0.86	-	0.66	0.67

In general, post-processing approaches improve performance forecasts. Nevertheless, improvements are not as pronounced as for the first scenario. For instance, when post-processing is used, the NSE increases barely 3.9% for the calibration and 1.8% for the validation period. Furthermore, we do not find improvements regarding the KGE. Concerning probabilistic metrics, Fig. 5 shows time series and predictive PQQ plots for streamflows predictions using the MCMC (upper) and ABC (lower) post-processing during the calibration period. As is evident in the Fig. 5, the exact (MCMC) and the approximate (ABC) conditional predictive uncertainty are seen to be an extremely close

match. In particular, prediction uncertainty bands supply an adequate description of observed values. The predictive PQQ plots in Fig. 5 confirm that predictions under post-processing are providing adequate representation of observed streamflows, as PQQ plots follow the bisector. This means that the predictive distributions of post-processed streamflows are reliable. Besides, we can confirm this by the reliability index (Table 3).



**Fig. 5** Conditional predictive uncertainty from MCMC (upper) and ABC (lower) post-processor on the Oria catchment. PQQ-plot of the conditional predictive distribution (right). Dots indicate observations, line indicates median prediction and grey region indicates 90% uncertainty.

As is consistent with the first scenario, the approximate (ABC) conditional predictive uncertainty is almost equivalent to the exact predictive (MCMC) (Fig. 5). Reliability and K-S test metrics are practically equal for both post-processing approaches in calibration and validation periods. Only a tiny difference is identified in the precision metric (Table 3). This finding is not surprising as the ABC uses summary statistics. Regularly summary statistics are linked with loss of information. This loss of information may be undesired for predicting objectives. Nevertheless, this problem could be avoided by using a set of sufficient summary statistics. In particular, during the calibration and validation periods, the B and D metrics of ABC post-processor are slightly higher than the MCMC post-processor, indicating wider predictive uncertainty bounds than the MCMC post-processor. For this reason, ABC method has 2% more of the observed samples in the 90% predictive uncertainty than the MCMC post-processor during the validation period (Table 3). Finally, a performance comparison between the scenarios based predictive

uncertainty shows that, the scenario 2 gave narrower prediction uncertainty bands than scenario 1. This is due to the quality of hydrological predictions influences the conditional predictive uncertainty. Contrary to scenario 1, the CR for both post-processors deteriorate in the validation period. To sum up, in both scenarios there is little visual distinction between the approximate (ABC post-processor) and the exact (MCMC post-processor) conditional predictive uncertainty.

#### 4 Discussion and Conclusions

The main aim of this study is to show that the conditional predictive distribution is qualitatively similar produced by the exact predictive (MCMC post-processor) or the approximate predictive (ABC post-processor) using some sufficient summary statistics and synthetic datasets. To achieve this aim, we use MCMC post-processor as a benchmark to make results more comparable with the proposed method. We apply both methods (ABC and MCMC) to two scenarios: the Aipe catchment (Colombia) and the Oria catchment (Spain). The advantage of the proposed method (ABC post-processor) is highlighted in the inferential problems with intractable likelihood. Sometimes, it can be difficult to calculate the likelihood itself in hydrological modelling, specially working with complex models or with ungauged catchments.

As well as Ye et al. (2014), Bogner et al. (2016) and Woldemeskel et al. (2018) we confirm that post-processing techniques can improve forecasts significantly when hydrological model predictions are especially poor (Table 2). Furthermore, we find through our numerical evidence that the MCMC and ABC post-processors provide similar predictive performance. This result was confirmed by Fenicia et al. (2018) and Kavetski et al. (2018), that the shape of the predictive distribution is qualitatively similar produced by the exact predictive (MCMC) or approximate predictive (ABC). These findings can be attributed to the correct specification of the probability model (data generating process). When the assumed probability model is correctly specified, just a little information is lost regarding predictive performance. In contrast, significant differences can appear when the probability model to represent observed streamflows is inadequate. In addition, Frazier et al. (2019) demonstrated theoretically that the ABC made forecasts which were asymptotically similar to those obtained from the exact Bayesian methods when the sample was large, the data generating process was correctly specified and, if the conditions for Bayesian consistency and asymptotic normality of both the exact (MCMC) and the approximate (ABC) posteriors were satisfied. Moreover, we recommend to check Drovandi and Pettitt (2011) for the use of ABC with synthetically case studies in the presence of intractable likelihood.

In both scenarios, we have proved that the predictive superiority of the exact predictive (MCMC post-processor) over the approximate (ABC post-processor) is minimal for hydrological models. There are two significant differences between this study and the previous works (e.g. Fenicia et al. (2018) and



Kavetski et al. (2018)). First, they used the ABC method to calibrate jointly hydrological models and to compute the predictive uncertainty. Instead of Bayesian calibration of hydrological models, we used the ABC method in hydrological post-processing context, so we calculated the conditional predictive uncertainty, while Kavetski et al. (2018) calculated the predictive uncertainty. Second, Fenicia et al. (2018) and Kavetski et al. (2018) did not use summary statistics, which is a tremendous difference with this study. In other words, we obtained a similar main conclusion but in a different context and method.

Although we used a comparative analysis between MCMC and ABC hydrologic post-processors, it should be noted though that this study does not aim to show that ABC post-processor has a better performance than MCMC post-processor. Thus, we use MCMC post-processor as a benchmark to make results more comparable with the proposed post-processor. Moreover, we know that the MCMC method has proven its ability in hydrological predictions. In addition, it is true that any comparison can be affected by different factors, but it must be emphasised the focus of this comparison is between two post-processors without particular emphasis on the hydrological model performance. Actually, they are just predictions, which are the input for both post-processors in two scenarios (poor and good predictions).

We also applied the NQT transformation to achieve assumptions of the error model, but any transformation produces information loss. Besides, we know that our analysis is conditioned to the linear regression model (hydrologic post-processor). In standard linear regression, the average link between observed and simulated streamflows is summarised with a single slope parameter expressing this relationship. However, Diks and Vrugt (2010) pointed out that a simple regression method could result in improvements equivalent to more complex methods. Furthermore, our idea is to show that the ABC approach can be used to compute the conditional predictive uncertainty rather than to perform a complex post-processor, and therefore, we can tolerate some of the less realistic assumptions. Although the examples that we have used in this paper are moderately simple, generally speaking, the ABC post-processor is highly flexible and can be used for more complex models. Future research should develop methods that relax any transformation, evaluate the impact of pre-processing and post-processing and explore the multi-regression model. We only scrutinize the ABC to compute the approximate conditional predictive uncertainty in hydrological post-processing context, but there are other approximations, for instance, Bayesian synthetic likelihood (Price et al., 2018), Bayesian empirical likelihood (Mengersen et al., 2013), variational Bayes (Tran et al., 2017) and bootstrap methods (Zhu et al., 2016).

The ABC post-processor has potential in areas such as operational hydrology, flood protection, drinking water production, risk assessment, irrigation management, water resources management and ecological issues. Besides, the ABC post-processor offers the opportunity to improve decision support with intractable likelihood function, e.g to predict in ungauged basins or evaluate climate change predictive uncertainty. In summary, we conclude that deterministic predictions (no post-processing) perform poorly concerning deterministic

verification metrics, and the MCMC and ABC post-processors provide similar predictive performance. Therefore, the approximate Bayesian computation may then be used as an alternative method to estimate the conditional predictive uncertainty of hydrological predictions with intractable likelihood.

## Acknowledgements

This study was partially supported by the Departamento del Huila Scholarship Program No. 677 (Colombia) and Colciencias, by the Spanish Research Project TETIS-MED (ref. CGL2014-58127-C3-3-R) and TETIS-CHANGE (ref. RTI2018-093717-B-I00). Also, G. Adelfio's research has been supported by the national grant of the Italian Ministry of Education University and Research (MIUR) for the PRIN-2015 program, 'Complex space-time modelling and functional analysis for probabilistic forecast of seismic events'. The authors also wish to thank the associate editor and two anonymous reviewers for their thoughtful comments for the revision of the manuscript.

## References

- Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate bayesian computation in population genetics. *Genetics*, 162(4):2025–2035.
- Blackwell, D. and Dubins, L. (1962). Merging of opinions with increasing information. *Ann. Math. Statist.*, 33(3):882–886.
- Bogner, K., Liechti, K., and Zappa, M. (2016). Post-processing of stream flows in switzerland with an emphasis on low flows and floods. *Water*, 8(4).
- Brown, J. D. and Seo, D.-J. (2010). A nonparametric postprocessor for bias correction of hydrometeorological and hydrologic ensemble forecasts. *Journal of Hydrometeorology*, 11(3):642–665.
- Butts, M. B., Payne, J. T., Kristensen, M., and Madsen, H. (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology*, 298(1):242–266.
- Coccia, G. and Todini, E. (2011). Recent developments in predictive uncertainty assessment based on the model conditional processor approach. *Hydrol. Earth Syst. Sci*, 15:3253–3274.
- Csillery, K., Francois, O., and Blum, M. G. B. (2012). abc: an r package for approximate bayesian computation (abc). *Methods in Ecology and Evolution*.
- Diaconis, P. and Freedman, D. (1986). On the consistency of bayes estimates. *Ann. Statist.*, 14(1):1–26.
- Diks, C. G. H. and Vrugt, J. A. (2010). Comparison of point forecast accuracy of model averaging methods in hydrologic applications. *Stochastic Environmental Research and Risk Assessment*, 24(6):809–820.
- Drovandi, C. C. and Pettitt, A. N. (2011). Likelihood-free bayesian estimation of multivariate quantile distributions. *Computational Statistics & Data Analysis*, 55(9):2541 – 2556.

- Evin, G., Thyer, M., Kavetski, D., McInerney, D., and Kuczera, G. (2014). Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity. *Water Resources Research*, 50(3):2350–2375.
- Fearnhead, P. and Prangle, D. (2012). Constructing summary statistics for approximate bayesian computation: semi-automatic approximate bayesian computation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(3):419–474.
- Fenicia, F., Kavetski, D., Reichert, P., and Albert, C. (2018). Signature-domain calibration of hydrological models using approximate bayesian computation: Empirical analysis of fundamental properties. *Water Resources Research*, 54(0).
- Francés, F., Vélez, J. I., and Vélez, J. J. (2007). Split-parameter structure for the automatic calibration of distributed hydrological models. *Journal of Hydrology*, 332(1):226 – 240.
- Frazier, D. T., Maneesoonthorn, W., Martin, G. M., and McCabe, B. P. (2019). Approximate bayesian forecasting. *International Journal of Forecasting*, 35(2):521 – 539.
- Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.
- Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.
- Glahn, H. R. and Lowry, D. A. (1972). The use of model output statistics (mos) in objective weather forecasting. *Journal of Applied Meteorology*, 11(8):1203–1211.
- Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. (2009). Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1):80 – 91.
- Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, 7(2):223–242.
- Kavetski, D., Fenicia, F., Reichert, P., and Albert, C. (2018). Signature-domain calibration of hydrological models using approximate bayesian computation: Theory and comparison to existing applications. *Water Resources Research*, 54(0).
- Khajehei, S. and Moradkhani, H. (2017). Towards an improved ensemble precipitation forecast: A probabilistic post-processing approach. *Journal of Hydrology*, 546:476–489.
- Klein, B., Meissner, D., Koblalka, H.-U., and Reggiani, P. (2016). Predictive Uncertainty Estimation of Hydrological Multi-Model Ensembles Using Pair-Copula Construction. *Water*, 8(4):125.
- Krzysztofowicz, R. and Kelly, K. S. (2000). Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resources Research*, 36(11):3265–3277.
- Laio, F. and Tamea, S. (2007). Verification tools for probabilistic forecasts of continuous hydrological variables. *Hydrology and Earth System Sciences*, 11(4):1267–1277.

- Li, B., Liang, Z., He, Y., Hu, L., Zhao, W., and Acharya, K. (2017). Comparison of parameter uncertainty analysis techniques for a topmodel application. *Stochastic Environmental Research and Risk Assessment*, 31(5):1045–1059.
- Liang, Z., Chang, W., and Li, B. (2012). Bayesian flood frequency analysis in the light of model and parameter uncertainties. *Stochastic Environmental Research and Risk Assessment*, 26(5):721–730.
- Lindley, D. V. and Smith, A. F. M. (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(1):1–41.
- Liu, Y. and Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, 43(7):W07401.
- Madadgar, S. and Moradkhani, H. (2014). Improved Bayesian multimodeling: Integration of copulas and Bayesian model averaging. *Water Resources Research*, 50(12):9586–9603.
- Marin, J.-M., Pudlo, P., Robert, C. P., and Ryder, R. J. (2012). Approximate bayesian computational methods. *Statistics and Computing*, 22(6):1167–1180.
- Marjoram, P., Molitor, J., Plagnol, V., and Tavaré, S. (2003). Markov chain monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26):15324–15328.
- Marshall, L., Nott, D., and Sharma, A. (2004). A comparative study of markov chain monte carlo methods for conceptual rainfall-runoff modeling. *Water Resources Research*, 40(2):W02501.
- Mengersen, K. L., Pudlo, P., and Robert, C. P. (2013). Bayesian computation via empirical likelihood. *Proceedings of the National Academy of Sciences*, 110(4):1321–1326.
- Montanari, A. and Brath, A. (2004). A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research*, 40(1).
- Montanari, A. and Grossi, G. (2008). Estimating the uncertainty of hydrological forecasts: A statistical approach. *Water Resources Research*, 44(12).
- Montanari, A. and Koutsoyiannis, D. (2012). A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research*, 48(9):W09555.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3):885–900.
- Nott, D. J., Marshall, L., and Brown, J. (2011). Generalized likelihood uncertainty estimation (glue) and approximate bayesian computation: What’s the connection? *Water Resources Research*, 48(12):W12602.
- Price, L. F., Drovandi, C. C., Lee, A., and Nott, D. J. (2018). Bayesian synthetic likelihood. *Journal of Computational and Graphical Statistics*, 27(1):1–11.
- Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., and Feldman, M. W. (1999). Population growth of human y chromosomes: a study of y chromo-

- some microsatellites. *Molecular Biology and Evolution*, 16(12):1791–1798.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). Using bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133(5):1155–1174.
- Reichert, P., Langhans, S. D., Lienert, J., and Schuwirth, N. (2015). The conceptual foundation of environmental decision support. *Journal of Environmental Management*, 154:316 – 332.
- Robert, C. P. (2016). Approximate bayesian computation: A survey on recent results. In Cools, R. and Nuyens, D., editors, *Monte Carlo and Quasi-Monte Carlo Methods*, pages 185–205, Cham. Springer International Publishing.
- Romero-Cuellar, J., Buitrago-Vargas, A., Quintero-Ruiz, T., and Francés, F. (2018). Modelling the potential impacts of climate change on the hydrology of the aipe river basin in huila, colombia. *Ribagua*, 5(1):63–78.
- Schefzik, R., Thorarinsdottir, T. L., and Gneiting, T. (2013). Uncertainty Quantification in Complex Simulation Models Using Ensemble Copula Coupling. *Statistical Science*. *Statistical Science*, 28(4):616–640.
- Schoups, G., van de Giesen, N. C., and Savenije, H. H. G. (2008). Model complexity control for hydrologic prediction. *Water Resources Research*, 44(12):W00B03.
- Schoups, G. and Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, 46(10):W10531.
- Shafii, M., Tolson, B., and Matott, L. S. (2014). Uncertainty-based multi-criteria calibration of rainfall-runoff models: a comparative study. *Stochastic Environmental Research and Risk Assessment*, 28(6):1493–1510.
- Sikorska, A. E., Montanari, A., and Koutsoyiannis, D. (2015). Estimating the uncertainty of hydrological predictions through data-driven resampling techniques. *Journal of Hydrologic Engineering*, 20(1):A4014009.
- Sisson, S. A., Fan, Y., and Tanaka, M. M. (2007). Sequential monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6):1760–1765.
- Solomatine, D. P. and Shrestha, D. L. (2009). A novel method to estimate model uncertainty using machine learning techniques. *Water Resources Research*, 45(12).
- Tavaré, S., Balding, D. J., Griffiths, R. C., and Donnelly, P. (1997). Inferring coalescence times from dna sequence data. *Genetics*, 145(2):505–518.
- Thomas, H. (1981). Improved methods for national water assessment, water resources contract: WR15249270. Technical report, Harvard University, Cambridge.
- Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S. W., and Srikanthan, S. (2009). Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resources Research*, 45(12).

- Tian, Y., Nearing, G. S., Peters-Lidard, C. D., Harrison, K. W., and Tang, L. (2016). Performance metrics, error modeling, and uncertainty quantification. *Monthly Weather Review*, 144(2):607–613.
- Todini, E. (2008). A model conditional processor to assess predictive uncertainty in flood forecasting. *International Journal of River Basin Management*, 6(2):123–137.
- Tran, M.-N., Nott, D. J., and Kohn, R. (2017). Variational bayes with intractable likelihood. *Journal of Computational and Graphical Statistics*, 26(4):873–882.
- Turner, B. M. and Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2):69–85.
- van Oijen, M. (2017). Bayesian Methods for Quantifying and Reducing Uncertainty and Error in Forest Models. *Current Forestry Reports*, 3(4):269–280.
- Vélez, J. J., Puricelli, M., López Unzu, F., and Francés, F. (2009). Parameter extrapolation to ungauged basins with a hydrological distributed model in a regional framework. *Hydrology and Earth System Sciences*, 13(2):229–246.
- Vrugt, J. A. and Robinson, B. A. (2007). Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging. *Water Resources Research*, 43(1):W01411.
- Vrugt, J. A. and Sadegh, M. (2013). Toward diagnostic model calibration and evaluation: Approximate Bayesian computation. *Water Resources Research*, 49:4335–4345.
- Waerden, B. V. D. (1953). Order tests for the two-sample problem and their power. *Indagationes Mathematicae (Proceedings)*, 56:80.
- Wagener, T. and Gupta, H. V. (2005). Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment*, 19(6):378–387.
- Wang, Q., Robertson, D., and Chiew, F. S. (2009). A bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. *Water Resources Research*, 45(5):W05407.
- Weerts, A. H., Winsemius, H. C., and Verkade, J. S. (2011). Estimation of predictive hydrological uncertainty using quantile regression: examples from the national flood forecasting system (england and wales). *Hydrology and Earth System Sciences*, 15(1):255–265.
- Wentao, L., Qingyun, D., Chiyuan, M., Aizhong, Y., Wei, G., and Zhenhua, D. (2017). A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *Wiley Interdisciplinary Reviews: Water*, 4(6):e1246.
- Wilby, R. L. and Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the river thames, uk. *Water Resources Research*, 42(2):W02419.
- Woldemeskel, F., McInerney, D., Lerat, J., Thyer, M., Kavetski, D., Shin, D., Tuteja, N., and Kuczera, G. (2018). Evaluating residual error approaches for post-processing monthly and seasonal streamflow forecasts. *Hydrology and Earth System Sciences*, 22(12):6257–6278.

- 
- Ye, A., Duan, Q., Yuan, X., Wood, E. F., and Schaake, J. (2014). Hydrologic post-processing of MOPEX streamflow simulations. *Journal of Hydrology*, 508:147–156.
- Yoon, S., Cho, W., Heo, J.-H., and Kim, C. E. (2010). A full bayesian approach to generalized maximum likelihood estimation of generalized extreme value distribution. *Stochastic Environmental Research and Risk Assessment*, 24(5):761–770.
- Zhang, X. and Zhao, K. (2012). Bayesian Neural Networks for Uncertainty Analysis of Hydrologic Modeling: A Comparison of Two Schemes. *Water Resources Management*, 26(8):2365–2382.
- Zhao, L., Duan, Q., Schaake, J., Ye, A., and Xia, J. (2011). A hydrologic post-processor for ensemble streamflow predictions. *Advances in Geosciences*, 29:51–59.
- Zhu, W., Marin, J. M., and Leisen, F. (2016). A bootstrap likelihood approach to bayesian computation. *Australian & New Zealand Journal of Statistics*, 58(2):227–244.