



Optimization of rain gauge sampling density for river discharge prediction using Bayesian calibration

Alexandre M.J.-C. Wadoux^{1,4}, Gerard B.M. Heuvelink¹, Remko Uijlenhoet² and Sytze de Bruin³

¹ Soil Geography and Landscape group, Wageningen University and Research, Wageningen, the Netherlands

² Hydrology and Quantitative Water Management group, Wageningen University and Research, Wageningen, the Netherlands

³ Laboratory of Geo-Information Science and Remote Sensing, Wageningen University and Research, Wageningen, the Netherlands

⁴ Current affiliation: Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Australia

ABSTRACT

River discharges are often predicted based on a calibrated rainfall-runoff model. The major sources of uncertainty, namely input, parameter and model structural uncertainty must all be taken into account to obtain realistic estimates of the accuracy of discharge predictions. Over the past years, Bayesian calibration has emerged as a suitable method for quantifying uncertainty in model parameters and model structure, where the latter is usually modelled by an additive or multiplicative stochastic term. Recently, much work has also been done to include input uncertainty in the Bayesian framework. However, the use of geostatistical methods for characterizing the prior distribution of the catchment rainfall is underexplored, particularly in combination with assessments of the influence of increasing or decreasing rain gauge network density on discharge prediction accuracy. In this article we integrate geostatistics and Bayesian calibration to analyze the effect of rain gauge density on river discharge prediction accuracy. We calibrated the HBV hydrological model while accounting for input, initial state, model parameter and model structural uncertainty, and also taking uncertainties in the discharge measurements into account. Results for the Thur basin in Switzerland showed that model parameter uncertainty was the main contributor to the joint posterior uncertainty. We also showed that a low rain gauge density is enough for the Bayesian calibration, and that increasing the number of rain gauges improved model prediction until reaching a density of one gauge per 340 km². While the optimal rain gauge density is case-study specific, we make recommendations on how to handle input uncertainty in Bayesian calibration for river discharge prediction and present the methodology that may be used to carry out such experiments.

Submitted 4 March 2020

Accepted 25 June 2020

Published 30 July 2020

Corresponding author

Alexandre M.J.-C. Wadoux,
alexandre.wadoux@sydney.edu.au

Academic editor

Tal Svoray

Additional Information and
Declarations can be found on
page 26

DOI 10.7717/peerj.9558

© Copyright
2020 Wadoux et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Ecohydrology, Spatial and Geographic Information Science

Keywords Sampling design, Bayesian uncertainty analysis, HBV model, Kriging, Input uncertainty

INTRODUCTION

Uncertainty analysis has garnered considerable attention in hydrological modelling during the past decades (e.g., *Pappenberger & Beven, 2006*; *Han & Coulibaly, 2017*). There is agreement on the necessity to provide (realistic) uncertainty bounds to end-users and practitioners (*Beven, 2006*; *Andréassian et al., 2007*; *Bal et al., 2014*). *Brown & Heuvelink (2005)* defines uncertainty as an expression of confidence about how well we know the “truth”. Similarly, *Maskey (2004)* defines uncertainty as a measure of the information about an unknown quantity to be measured or a situation to be forecast, and discusses the nature of different potential sources of uncertainty and their effect on flood forecasting.

It is generally recognized that three principal sources of uncertainty cause uncertainty in model output: model input uncertainty (including initial state and boundary conditions), model parameter uncertainty and model structural uncertainty (*Refsgaard et al., 2007*; *Van der Keur et al., 2008*; *Houska et al., 2015*). For example, *Højberg & Refsgaard (2005)* found that parameter uncertainty cannot fully cover model structural uncertainty, while *Tian, Booij & Xu (2014)* showed that parameter uncertainty for three rainfall-runoff models tested on two catchments has a larger contribution to model output uncertainty than model structural uncertainty. *Kavetski, Kuczera & Franks (2006)* found that input (rainfall) uncertainty has a considerable effect on the predicted outflow and output prediction intervals. In addition to these three main sources, there is usually also uncertainty in the measurements of the model output (*Di Baldassarre & Montanari, 2009*). This source of uncertainty must be taken into account if these measurements are used to calibrate the model.

Explicit integration of all sources of uncertainty is not an easy task. This problem has been tackled using approaches such as the pseudo-Bayesian Generalized Likelihood Uncertainty Estimation (GLUE) methodology (*Beven et al., 2000*), the Integrated Bayesian Uncertainty Estimator (IBUNE) (*Ajami, Duan & Sorooshian, 2007*) and by using Bayesian Total Error Analysis (BATEA) (*Kavetski, Kuczera & Franks, 2006*). In general, Bayesian analysis has received wide attention because it provides a comprehensive and general framework to specify uncertainty explicitly using probability distributions. It also fosters easy updating of distributions when additional information comes available. The main steps of a Bayesian uncertainty framework are summarized as follows (*Kennedy & O'Hagan, 2001*): (1) an explicit probability model is specified for each uncertainty source (input, model parameters, model structure), based on prior information; (2) measurements of the model output are used to update the prior distributions to posterior distributions, typically using Markov chain Monte Carlo techniques; (3) the posterior distributions are used to propagate uncertainty in model input, model parameters and model structure to model output for (future) cases where model output is not observed; (4) results are tested against independent validation data to evaluate whether the assumptions made as part of the procedure are realistic.

Numerous studies on Bayesian uncertainty analysis for distributed and physically based hydrological models have been conducted and published. There is general agreement that rainfall-runoff data are often insufficient for supporting reliable inference for complex

models involving many spatially distributed physical catchment processes (Beven, 2006; Renard et al., 2010; Laiolo et al., 2016). Wagener et al. (2001) refers to “non-identifiability” leading to “ill-posed” inference of the parameters, which can be avoided by using simpler (lumped) hydrological models with fewer parameters. Lumped hydrological models consider the quantity of interest (e.g. discharge) to be derived from catchment-averaged inputs (e.g. rainfall, potential evapotranspiration). Model inputs often contain substantial error which affect model output. In general, input uncertainty in lumped models is mainly caused by measurement and interpolation errors. For instance, rainfall measurements obtained using rain gauges are not error-free (Habib, Krajewski & Kruger, 2001), while interpolation error is added when rain gauge measurements at point locations are aggregated to spatial averages as needed in lumped rainfall-runoff models. In the case of rainfall, radar images provide time series of spatial rainfall fields (Ciabatta et al., 2016), thus avoiding interpolation error, but these often suffer from complex spatio-temporal errors which make them inaccurate in some circumstances (Cecinati et al., 2017). Thus, rainfall point observations remain a major source for estimating catchment-average rainfall.

There is a recent trend towards a decrease of hydrometric network density (Mishra & Coulibaly, 2009; Keum & Kaluarachchi, 2015). Yet, the uncertainty in average rainfall strongly depends on rain gauge sampling density (Xu et al., 2013; Terink et al., 2018). Hence, a reduction of the rain gauge density will increase the uncertainty about the discharge predicted by the rainfall-runoff model. Renard et al. (2011) used a geostatistical model to infer the catchment-average rainfall and the associated uncertainty from the rain gauges using block kriging. Next, they used the block kriging conditional distribution as a prior in the Bayesian calibration of a lumped rainfall-runoff model. Taking prior knowledge on input uncertainty into account overcomes ill-posedness and significantly improved the accuracy of the runoff predictions (Renard et al., 2010). Clearly, the higher the rain gauge density the narrower the block kriging prior. Thus, a different sampling density leads to a different prior and posterior and ultimately to a different output uncertainty distribution. To the best of our knowledge, little has been done to investigate the effect of rain gauge density on the model output uncertainty within a Bayesian framework.

The objective of this work is to evaluate the effect of rain gauge sampling density on uncertainty in the output of a lumped rainfall-runoff model. The methodology used relies on geostatistics to quantify prior input uncertainty and on Bayesian calibration for model parameter and model structural uncertainty quantification. We calibrate the lumped HBV model (Lindström et al., 1997) using a Bayesian uncertainty framework that accounts for input, parameter, output observation, initial state and model structural uncertainty. Model residuals comprising model structural uncertainty and discharge error are characterized using a time series model, while Markov chain Monte Carlo methods are used to obtain posteriors of the input, model and initial state parameters. The propagation of uncertainties associated with the model input, model parameters and model structure is then analyzed using regular Monte Carlo methods. Several rain gauge density scenarios were tested, each time recalibrating the model and providing discharge predictions with uncertainty

intervals. Rainfall posterior intervals as well as model predictive abilities were assessed and discussed. The approach was tested in a case study using ten-day average rainfall and discharge data of the 1,696 km² Thur basin in Switzerland for the years 2004 to 2011.

METHODS

Rainfall-runoff model

Consider a hydrological model H that predicts river discharge from catchment average rainfall. Let $\mathbf{y} = [y_1 \ y_2 \dots y_T]^T$ and $\bar{\mathbf{z}} = [\bar{z}_1 \ \bar{z}_2 \dots \bar{z}_T]^T$ be time series of measured discharge and (known) catchment average rainfall, respectively. We assume that the relation between \mathbf{y} and $\bar{\mathbf{z}}$, which is governed by the model H , is affected by multiplicative measurement and model structural errors, which after log-transformation gives:

$$\log(\mathbf{y}) = \log(H(\bar{\mathbf{z}}, \boldsymbol{\varphi})) + \boldsymbol{\varepsilon} + \boldsymbol{\eta} \quad (1)$$

where $\boldsymbol{\varphi}$ is a vector comprising model parameters and the initial state, $\boldsymbol{\varepsilon} = [\varepsilon_1 \ \varepsilon_2 \dots \varepsilon_T]^T$ is log-transformed model structural uncertainty and $\boldsymbol{\eta} = [\eta_1 \ \eta_2 \dots \eta_T]^T$ is log-transformed discharge measurement error. Uncertainty in model input (i.e. $\bar{\mathbf{z}}$) and model parameters (i.e. $\boldsymbol{\varphi}$) will be introduced in the next section. It is common ([Weerts & El Serafy, 2006](#); [Rakovec et al., 2015](#)) to assume that the η_t ($t = 1 \dots T$) are independent and identically distributed normal variates, with constant mean μ_η and variance σ_η^2 .

It is unrealistic to have temporal independence for $\boldsymbol{\varepsilon}$ and hence we represent it by a first-order autoregressive model (AR(1)):

$$\varepsilon_t = \beta_0 + \beta_1 \cdot \varepsilon_{t-1} + \delta_t, \quad t = 1 \dots T \quad \varepsilon_0 \sim \mathcal{N}(\mu_0, \sigma_0^2) \quad (2)$$

where the δ_t ($t = 1 \dots T$) are independent and identically distributed normal variates, with zero mean and constant variance σ_δ^2 . The parameters that characterize $\boldsymbol{\varepsilon}$ are merged into a parameter vector $\boldsymbol{\theta} = \{\beta_0, \beta_1, \sigma_\delta^2, \mu_0, \sigma_0^2\}$. In this study, μ_0 and σ_0^2 are assumed known since their effect is typically negligible after a few time steps. Since model structural uncertainty and discharge measurement error are not physically related, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ are statistically independent. We further require $E[e^\varepsilon] = E[e^\eta] = 1$ to avoid bias in the multiplicative model structural uncertainty and discharge measurement error term, so that their expectation is forced to 1. This is achieved by setting $\mu_\eta = -\frac{1}{2}\sigma_\eta^2$ and by imposing that $\frac{2\beta_0}{1-\beta_1} + \frac{\sigma_\delta^2}{1-\beta_1^2} = 0$, which implies that $\beta_0 = -\frac{1}{2} \frac{1-\beta_1}{1-\beta_1^2} \sigma_\delta^2$. Thus, parameters μ_η and β_0 are determined by the other parameters and the set of free parameters of $\boldsymbol{\varepsilon}$ is reduced to $\boldsymbol{\theta} = \{\beta_1, \sigma_\delta^2, \mu_0, \sigma_0^2\}$.

To simplify notation we define $\mathbf{u} = \log(\mathbf{y}) - \log(H(\bar{\mathbf{z}}, \boldsymbol{\varphi}))$ and obtain:

$$\mathbf{u} = \boldsymbol{\varepsilon} + \boldsymbol{\eta}. \quad (3)$$

Bayesian uncertainty framework

Conventional estimation of parameters $\boldsymbol{\varphi}$, $\boldsymbol{\theta}$ and σ_η^2 using Bayesian calibration ([Beven & Freer, 2001](#); [Kavetski, Kuczera & Franks, 2006](#)) starts by using Bayes' law to derive that the

posterior distribution of the parameters is proportional to the product of the likelihood of the observed residuals \mathbf{u} and the prior distribution of the parameters:

$$p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2 | \mathbf{u}) = \frac{p(\mathbf{u} | \boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2) \cdot p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2)}{p(\mathbf{u})} \propto p(\mathbf{u} | \boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2) \cdot p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2) \quad (4)$$

where $p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2)$ is a joint prior for the rainfall-runoff model parameters and the model structural and measurement uncertainty parameters. Since there were no physical reasons to include mutual dependencies, the priors for $\boldsymbol{\varphi}$, $\boldsymbol{\theta}$, and σ_{η}^2 were treated as statistically independent, i.e. $p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2) = p(\boldsymbol{\varphi}) \cdot p(\boldsymbol{\theta}) \cdot p(\sigma_{\eta}^2)$.

We extend the conventional approach by considering the case where the model input (i.e. the catchment-average rainfall) is also uncertain. Since model output and input are not independent (as is obvious from Eq. (1)) we have $p(\bar{\mathbf{z}} | \mathbf{y}) \neq p(\bar{\mathbf{z}})$, which implies that model output observations (or observed residuals) should be used to update the probability distribution of the model input, just as these are used to update the distributions of the parameters of the rainfall-runoff model and the model structural and measurement uncertainty. This can be achieved by adding $\bar{\mathbf{z}}$ to the parameters, so that Eq. (4) is replaced by:

$$p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2, \bar{\mathbf{z}} | \mathbf{u}) \propto p(\mathbf{u} | \boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2, \bar{\mathbf{z}}) p(\boldsymbol{\varphi}) p(\boldsymbol{\theta}) p(\sigma_{\eta}^2) p(\bar{\mathbf{z}}). \quad (5)$$

We test three different approaches toward including input rainfall uncertainty by estimating the whole set of parameters twice. In the first approach we calibrate $\bar{\mathbf{z}}$ while in the second and third we ignore that discharge measurements inform catchment-averaged rainfall and exclude $\bar{\mathbf{z}}$ from the Bayesian calibration. More information on the approaches tested is given in the next section. Note that, unconditional to \mathbf{u} , there are no physical reasons to include mutual dependencies between $\boldsymbol{\varphi}$, $\boldsymbol{\theta}$, σ_{η}^2 and $\bar{\mathbf{z}}$. The model parameter priors $p(\boldsymbol{\varphi})$ can be derived from expert judgment and may be centered around deterministically calibrated parameter values, while uninformative (i.e. wide) priors are typically chosen for $\boldsymbol{\theta}$ and σ_{η}^2 . The remaining terms of the right-hand side of Eq. (5) are the likelihood $p(\mathbf{u} | \boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2, \bar{\mathbf{z}})$ and the prior for $\bar{\mathbf{z}}$. We work these out in the next two subsections, starting with the latter.

Rainfall prior

The rainfall priors for all time instances were derived using a geostatistical approach. Let $z_t(s)$ denote the rainfall at location s in the catchment \mathcal{A} at time $t \in \{1 \dots T\}$. We treat $z_t = \{z_t(s) | s \in \mathcal{A}\}$ as a realization of a random field Z_t . We further assume that $\log(Z_t)$ is a stationary normally distributed random field, characterized by a (constant) mean and isotropic variogram $\gamma(h)$ (where h is geographical distance). Since it is unrealistic to assume that $\log(Z_t)$ has the same statistical properties for all times t , in the case study we classified all times into a finite number of classes that are judged sufficiently homogeneous with respect to rainfall intensity, and assumed constant statistical properties within each class. The rainfall intensity classes are ordered by rainfall intensity level (the first class having the lowest intensity and the final class the highest). The class boundaries are based on equidistant quantiles from the cumulative distribution function of the rainfall intensity.

Note that while we include spatial autocorrelation, we ignore temporal autocorrelation. In other words, we assume that the correlation between $\log(Z_t(s))$ and $\log(Z_{t+v}(s+h))$ is zero if $v \neq 0$. Ignoring temporal autocorrelation is acceptable if the temporal support of rainfall data is sufficiently large (Cecinati et al., 2017) (in the case study we consider rainfall accumulated over ten days). Estimation of the variogram $\gamma(h)$ for each rainfall intensity class may be done using the conventional Methods of Moments estimator and by pooling sample variograms derived for all time instants within the same rainfall intensity class (Muthusamy et al., 2017).

To sample from the distribution of Z_t we use conditional Sequential Gaussian Simulation (cSGS) (Cressie, 2015). For each time instant t , we first simulate fields of log-transformed rainfall, conditional to the observations $\log(z_t(s_i)), i = 1 \dots n$, where n is the number of rain gauge locations. Next we back-transform these fields to the original scale to obtain conditional rainfall simulation fields $z_t^l = \{z_t^l(s) | s \in \mathcal{A}\}, l = 1 \dots L$, where L is the number of simulated fields. Finally, each simulated field is spatially aggregated to obtain catchment average rainfall simulations:

$$\bar{z}_t^l = \frac{1}{|\mathcal{A}|} \int_{s \in \mathcal{A}} z_t^l(s) ds. \quad (6)$$

In practice, the integral in Eq. (6) is approximated by a summation over a (sufficiently dense) spatial grid. The simulations z_t^l are also generated on this same grid. We used a simulation approach to derive the probability distribution of the catchment average rainfall, because analytical solutions are generally not available.

The set of L simulations of catchment average rainfall provides an empirical representation of the prior distribution of \bar{z}_t for all t , which is an accurate approximation of the true prior if L is sufficiently large. The empirical prior cumulative distribution of \bar{z}_t is then given by:

$$F_{\bar{z}_t}(a) = \frac{1}{L} \sum_{l=1}^L I(\bar{z}_t^l \leq a) \quad (7)$$

where I is an indicator function equal to 1 if its argument is true and 0 otherwise. The probability density function derived from this empirical rainfall distribution is used to obtain density values for each new jump in the (rainfall) parameter space during MCMC.

Likelihood function

Deriving the likelihood $p(\mathbf{u} | \boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_\eta^2, \bar{\mathbf{z}})$ is a major step in the Bayesian model calibration. For notational convenience we will drop the conditioning information in this subsection and write \mathbf{u} instead of $\{\mathbf{u} | \boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_\eta^2, \bar{\mathbf{z}}\}$, but note that because of the conditioning information \mathbf{u} satisfies Eq. (3), with all parameters of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ known.

We start by writing the joint distribution of \mathbf{u} as a product of conditional distributions:

$$p(\mathbf{u}) = p(u_0) \cdot p(u_1 | u_0) \cdot p(u_2 | u_0, u_1) \dots p(u_T | u_0 \dots u_{T-1}). \quad (8)$$

Because \mathbf{u} is the sum of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ it cannot be written as an AR(1) model and does not satisfy the Markov property (Durbin & Koopman, 2012). The conditional distributions can

therefore not be reduced to a simple form. Instead, we use a Kalman filter approach (*Kalman et al., 1960*) to evaluate the conditional distribution in Eq. (8). Since all stochastic variables involved are normal the conditional distributions are also normal, leaving only their means and variances to be determined. These are derived in a recursive way ($t = 1 \dots T$):

$$\hat{\varepsilon}_0^+ = E[\varepsilon_0] = \mu_0, \quad \sigma_0^{2+} = \text{var}(\varepsilon_0) = \sigma_0^2 \quad (9)$$

$$\hat{\varepsilon}_t^- = \beta_0 + \beta_1 \hat{\varepsilon}_{t-1}^+ \quad (10)$$

$$\sigma_t^{2-} = \beta_1^2 \sigma_{t-1}^{2+} + \sigma_\delta^2 \quad (11)$$

$$\hat{\varepsilon}_t^+ = \hat{\varepsilon}_t^- + k_t(u_t - \hat{\varepsilon}_t^-) - \mu_\eta \quad (12)$$

$$\sigma_t^{2+} = (1 - k_t) \sigma_t^{2-} \quad (13)$$

$$k_t = \frac{\sigma_t^{2-}}{\sigma_t^{2-} + \sigma_\eta^2}. \quad (14)$$

Here, $\hat{\varepsilon}_t^- = E[\varepsilon_t | u_0 \dots u_{t-1}]$ is the time update, $\hat{\varepsilon}_t^+ = E[\varepsilon_t | u_0 \dots u_t]$ the measurement update and k_t is the Kalman gain. The prediction error variances associated with $\hat{\varepsilon}_t^-$ and $\hat{\varepsilon}_t^+$ are given by σ_t^{2-} and σ_t^{2+} , respectively. Obtaining the mean and variance of $\{u_t | u_0 \dots u_{t-1}\}$ is now easy and given by:

$$E[u_t | u_0 \dots u_{t-1}] = \hat{u}_t = \hat{\varepsilon}_t^- + \mu_\eta \quad (15)$$

$$\text{var}(u_t | u_0 \dots u_{t-1}) = \sigma_t^{2-} + \sigma_\eta^2. \quad (16)$$

The log-transformed conditional distribution at time $t > 0$ is thus given by:

$$\log(p(u_t | u_0, \dots, u_{t-1})) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma_t^{2-} + \sigma_\eta^2) - \frac{1}{2} \left\{ \frac{(u_t - \hat{\varepsilon}_t^- + \mu_\eta)^2}{\sigma_t^{2-} + \sigma_\eta^2} \right\} \quad (17)$$

while for $t = 0$ we have $u_0 \sim \mathcal{N}(\mu_0 + \mu_\eta, \sigma_0^2 + \sigma_\eta^2)$. Taking the logarithm of Eq. (8) implies that we must sum over all time steps so that the log-likelihood is given by (defining $\hat{\mu}_0 = \mu_0$ and $\sigma_0^{2-} = \sigma_0^2$):

$$\log(p(\mathbf{u})) = -\frac{T+1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=0}^T (\log(\sigma_t^{2-} + \sigma_\eta^2)) - \frac{1}{2} \sum_{t=0}^T \left(\frac{(u_t - \hat{\varepsilon}_t^- + \mu_\eta)^2}{\sigma_t^{2-} + \sigma_\eta^2} \right). \quad (18)$$

Markov chain Monte Carlo

The proportionality sign in Eq. (5) means that the posterior on its left-hand side differs from the right-hand side by a multiplicative constant. This constant is unknown (that is to say, it can only be obtained using formidable computer power) and hence the posterior cannot be determined explicitly. To overcome this problem a common approach is to sample from the posterior distribution $p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2, \bar{\mathbf{z}}|\mathbf{u})$ using Markov chain Monte Carlo (MCMC). In this paper we use the Metropolis algorithm, which may not be the most efficient approach but perfectly valid and relatively easy to implement. It is described in detail in [Brooks et al. \(2011\)](#). Thus, a large sample N of the joint posterior distribution of $(\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_{\eta}^2, \bar{\mathbf{z}})$ is generated, where N is typically taken in the range 10^4 to 10^5 . Convergence can be assessed by running several independent Markov chains and checking that the sample distributions are sufficiently similar ([Gelman et al., 2014](#)). Another important performance indicator is the acceptance rate, i.e. the number of accepted proposals divided by the total number of proposals. We manually tuned the size of the jump in the parameter space to obtain an acceptance rate between 0.25 and 0.5 ([Rosenthal et al., 2011](#)) to obtain sufficient exploration of the parameter space without grossly deteriorating efficiency. The acceptance rate was calculated after removing the first set of proposals, also called the burn-in phase.

Prediction

Once the joint conditional posterior distribution of all parameters and the catchment-average rainfall (Eq. (5)) has been obtained, it can be used for discharge prediction. To derive the prediction, it is important to distinguish between the true discharge \mathbf{d} and the measured discharge \mathbf{y} . From the model defined in Eq. (1) it follows that the log-transformed true discharge $\log(\mathbf{d}) = \log(\mathbf{y}) - \boldsymbol{\eta}$ is the sum of the log-transformed model output ($H(\bar{\mathbf{z}}, \boldsymbol{\varphi})$) and the log-transformed model structural uncertainty $\boldsymbol{\varepsilon}$. Using the law of total probability, the probability distribution of the discharge can be written as:

$$p(\mathbf{d}) = \iiint p(\mathbf{d}|\boldsymbol{\varphi}, \boldsymbol{\theta}, \bar{\mathbf{z}}) \cdot p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \bar{\mathbf{z}}) d\boldsymbol{\varphi} d\boldsymbol{\theta} d\bar{\mathbf{z}}. \quad (19)$$

This multi-dimensional integral is usually solved numerically using Monte Carlo sampling. Since $\log(\mathbf{d}) = \log(H(\bar{\mathbf{z}}, \boldsymbol{\varphi})) + \boldsymbol{\varepsilon}$, sampling from $p(\mathbf{d}|\boldsymbol{\varphi}, \boldsymbol{\theta}, \bar{\mathbf{z}})$ involves a deterministic run of the rainfall-runoff model and simulating a realization from the AR(1) model of $\boldsymbol{\varepsilon}$. One might think that realizations of the posterior $p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \bar{\mathbf{z}})$ are already available from the MCMC sampling, but this is not the case. The problem is that realizations of this posterior are only available for the calibration period, i.e. a time period with discharge measurements. Since prediction is only needed for time periods without discharge measurements, we consider a case in which there are no discharge measurements at or near the prediction time. Hence, there are no realizations of the posterior $p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \bar{\mathbf{z}})$ for a time period without discharge measurements either. To make the distinction between the calibration and prediction periods explicit, let the prediction period be from $t = T + V + 1$ to $t = T + V + W$, where V is typically larger than the catchment response time to ensure that the validation data are independent from the calibration data. We denote the catchment-average rainfall for the prediction period by $\bar{\mathbf{z}}_+$. Thus, we derive

$p(\mathbf{d}_+)$ from the (posterior) distribution of $\boldsymbol{\varphi}$, $\boldsymbol{\theta}$ and $\bar{\mathbf{z}}_+$ using Eq. (19), with \mathbf{d} replaced by \mathbf{d}_+ and $\bar{\mathbf{z}}$ replaced by $\bar{\mathbf{z}}_+$.

We consider three approaches to derive a “posterior” distribution $p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \bar{\mathbf{z}}_+)$:

1. For $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$, use the MCMC sample of their joint posterior as explained in the Section on MCMC. For $\bar{\mathbf{z}}_+$, apply a linear correction to its prior mean as follows:

$$\mu_{\bar{\mathbf{z}}_+^{po}} = a\mu_{\bar{\mathbf{z}}_+^{pr}} + b \quad (20)$$

where the coefficients a and b are derived by fitting a linear regression between the means of the rainfall posterior distribution $\mu_{\bar{\mathbf{z}}_+^{po}}$ and the means of the rainfall prior distribution $\mu_{\bar{\mathbf{z}}_+^{pr}}$ for the calibration period. A similar approach is used in [Huard & Mailhot \(2008\)](#). Thus, realizations \bar{z}_t^l of the rainfall prior are converted to realizations $\bar{z}_t^{l,po}$ of the “posterior” by shifting the means, while keeping the same shape and standard deviation:

$$\bar{z}_t^{l,po} = \bar{z}_t^l - \mu_{\bar{\mathbf{z}}_+^{pr}} + \mu_{\bar{\mathbf{z}}_+^{po}} = \bar{z}_t^l + (a - 1)\mu_{\bar{\mathbf{z}}_+^{pr}} + b. \quad (21)$$

This approach has the disadvantage that only the mean is corrected, using a simple linear transform. It is not obvious how the correction can be improved. Another disadvantage is that the posteriors of $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$ are decoupled from that of $\bar{\mathbf{z}}_+$. In other words, they are made statistically independent.

2. For $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$, use the MCMC sample from their joint posterior as explained in the section on MCMC. For $\bar{\mathbf{z}}_+$, let the posterior distribution be identical to its prior. This approach has the disadvantage that the posteriors of $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$ are again decoupled from that of $\bar{\mathbf{z}}_+$.
3. Ignore that discharge measurements inform catchment-averaged rainfall and exclude $\bar{\mathbf{z}}$ from the Bayesian calibration. Thus, the prior distribution of $\bar{\mathbf{z}}_+$ as derived using block kriging is used as in Approach 2, but unlike in Approach 2 the parameters $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$ are calibrated without including $\bar{\mathbf{z}}$ in the calibration procedure. The disadvantage of this approach is that it ignores that rainfall and discharge are dependent, but it has two important advantages. First, there is no interference between calibration of model parameters and rainfall input, which causes problems in the prediction period. Second, the overall number of parameters to be calibrated is much smaller compared to the case in which rainfall input is to be calibrated as well. While this approach essentially boils down to Eq. (4), rainfall uncertainty during the calibration period must be taken into account. This is achieved by integrating the likelihood in Eq. (4) over all realizations of the rainfall $\bar{\mathbf{z}}$:

$$p(\mathbf{u}|\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_\eta^2) = \int p(\mathbf{u}, \bar{\mathbf{z}}|\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_\eta^2) d\bar{\mathbf{z}} = \int p(\mathbf{u}|\boldsymbol{\varphi}, \boldsymbol{\theta}, \sigma_\eta^2, \bar{\mathbf{z}}) \cdot p(\bar{\mathbf{z}}) d\bar{\mathbf{z}}. \quad (22)$$

Note that here we used the fact that $\bar{\mathbf{z}}$ is independent of the model parameters.

The three approaches are different in the way they account for input rainfall uncertainty and in the resultant computational complexity. Approach 1 is the most rigorous way to include input uncertainty in a Bayesian uncertainty analysis. It has the largest flexibility but involves calibration of a rainfall parameter (the catchment average rainfall) for each time step of the calibration period. This increases dramatically the overall number of calibration parameters and the Bayesian uncertainty analysis becomes quickly intractable when the temporal resolution is high. A second problem arises during prediction because in this approach it is not possible to update the rainfall prior derived from geostatistical analysis of the rainfall. This problem may be addressed by using a linear correction (described above), or more simply to use Approach 2 and not correct the prior distribution of the rainfall during prediction. Going one step further, Approach 3 tests if the computational burden of calibrating the rainfall is necessary and whether propagating the input uncertainty can be done by simply sampling from the empirical distribution of the rainfall uncertainty. Approach 3 is the most straightforward way to include rainfall uncertainty into the Bayesian uncertainty analysis.

Validation measures

To evaluate the methodology it is necessary to statistically validate the discharge predictions and associated prediction uncertainty using independent discharge measurements. We do so for the prediction period, since discharge measurements during this period were not used for calibration and prediction. However, since discharge measurements are not error-free, the validation must take discharge measurement error into account. Discharge measurement error was defined by η in the Methods section. The distribution of η is characterized by a single parameter σ_η^2 , the calibration of which was explained in the section on MCMC.

With little modification, prediction equation Eq. (19) can be rewritten to include discharge measurement error:

$$p(\mathbf{y}_+) = \iiint p(\mathbf{y}_+ | \boldsymbol{\varphi}, \boldsymbol{\theta}, \bar{\mathbf{z}}_+, \sigma_\eta^2) \cdot p(\boldsymbol{\varphi}, \boldsymbol{\theta}, \bar{\mathbf{z}}_+, \sigma_\eta^2) d\boldsymbol{\varphi} d\boldsymbol{\theta} d\bar{\mathbf{z}}_+ d\sigma_\eta^2. \quad (23)$$

Here, $\mathbf{y}_+ = [y_{T+V+1}, y_{T+V+2}, \dots, y_{T+V+W}]^T$ denotes the modelled discharge measurements for the prediction period. The predictions and prediction intervals of \mathbf{y}_+ can now be compared to the actual discharge observations to assess the quality of the model. To distinguish between modelled and observed discharge measurements, in this section we denote the first by \mathbf{y}_+^m and the second by \mathbf{y}_+^o .

Several measures are employed to assess the quality of the model and the effect of the rain gauge density on the discharge prediction intervals. These are the Mean Error (ME), the Root Mean Squared Error (RMSE), the Nash-Sutcliffe model efficiency coefficient (NSE) and the Prediction Intervals Coverage Probability (PICP). The PICP is the percentage of observations covered by a defined prediction interval (Shrestha & Solomatine, 2008).

$$\text{ME} = \frac{1}{W} \sum_{t=T+V+1}^{T+V+W} (\bar{y}_t^m - y_t^o), \quad (24)$$

where \bar{y}_t^m is the arithmetic mean of the simulated discharges at time t .

$$\text{RMSE} = \sqrt{\frac{1}{W} \sum_{t=T+V+1}^{T+V+W} (\bar{y}_t^m - y_t^o)^2}, \quad (25)$$

$$\text{NSE} = 1 - \frac{\sum_{t=T+V+1}^{T+V+W} (\bar{y}_t^m - y_t^o)^2}{\sum_{t=T+V+1}^{T+V+W} (y_t^o - \bar{y}^o)^2}, \quad (26)$$

where $\bar{y}^o = \frac{1}{W} \sum_{t=T+V+1}^{T+V+W} y_t^o$.

$$\text{PICP} = \frac{100}{W} \sum_{t=T+V+1}^{T+V+W} I_t \quad (27)$$

with

$$I_t = \begin{cases} 1 & \text{if } y_t^m(\text{low}) \leq y_t^o \leq y_t^m(\text{up}) \\ 0 & \text{otherwise} \end{cases} \quad (28)$$

where $y_t^m(\text{low})$ and $y_t^m(\text{up})$ are the lower and upper limits of the prediction interval for y_t as computed by the model. In the case study, we will compute the PICP both for the 50% and 90% prediction intervals.

Sampling density scenarios

To investigate the effect of rain gauge density on the uncertainty of the discharge predictions, several scenarios were developed. Each scenario comprises a number of rain gauges that are optimally selected from the existing rain gauge locations in the study area. The optimal locations were derived using Spatial Simulated Annealing (SSA), by minimizing the time-averaged block kriging prediction error variance:

$$\frac{1}{T+W} \left\{ \sum_{t=1}^T \sigma_{OK}^2(\mathcal{A}, t) + \sum_{t=T+V+1}^{T+V+W} \sigma_{OK}^2(\mathcal{A}, t) \right\} \quad (29)$$

where $\sigma_{OK}^2(\mathcal{A}, t)$ is the rainfall ordinary block-kriging variance at time t and using the entire study area \mathcal{A} as a “block”. Because there is no obvious analytical means to compute the block kriging variance of a lognormally distributed variable, the block kriging variance was approximated by computing the variance of 200 rainfall fields simulated using cSGS, in a similar way as described before. For more details about SSA we refer to [Van Groenigen & Stein \(1998\)](#) and [Wadoux et al. \(2017\)](#).

DATA AND MODEL

Study area

The study area is the Thur basin (1,696 km²), located in the North-East of Switzerland ([Fig. 1](#)). The Thur river is a tributary of the Rhine river and is the largest non-regulated river in Switzerland. The elevation within the basin ranges from 357 to 2437 m above sea level (m.a.s.l.) with an average height of 765 m.a.s.l. The climate is submontane, relatively

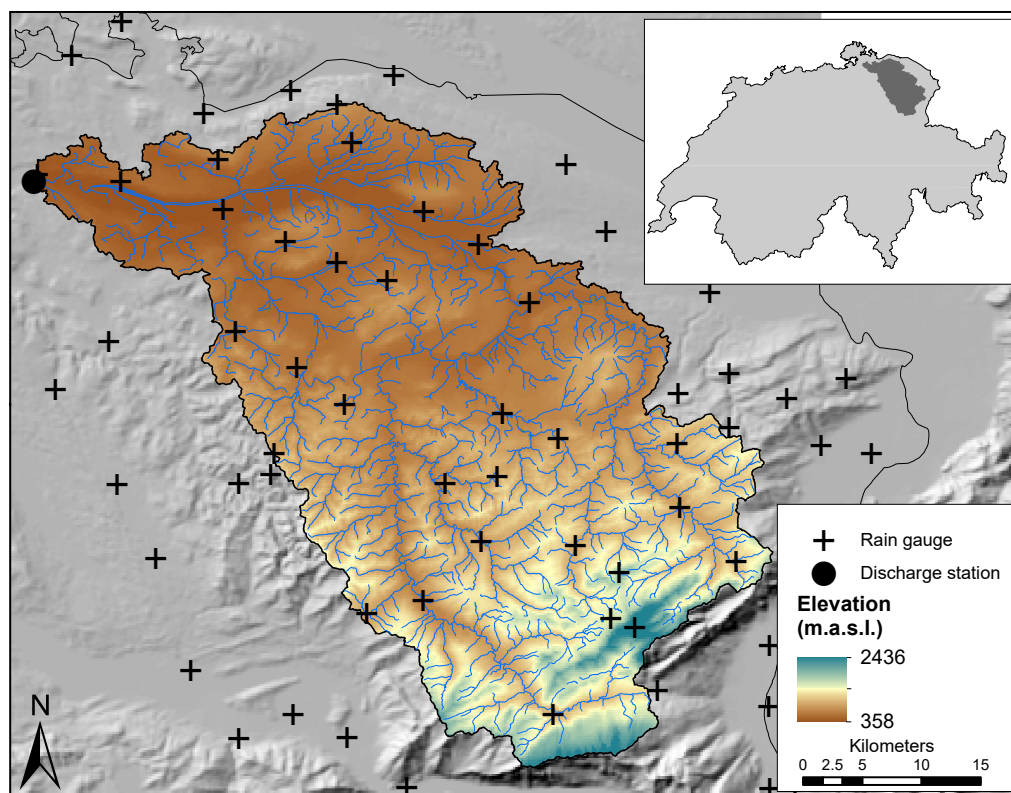


Figure 1 Map of the Thur River Basin with locations of rain gauges and discharge station.

Full-size [DOI: 10.7717/peerj.9558/fig-1](https://doi.org/10.7717/peerj.9558/fig-1)

cool and dominated by high precipitation, most of which falls during the summer months (June-August) in the form of rain, whereas the winter months (December-February) are characterized by snow. Precipitation is mainly orographic-induced but convective rainfall events also occur in the summer months. The Thur basin has been subject of several previous studies (e.g. [Melsen et al., 2016](#)) and data availability is large. Three datasets are used in this study:

- Daily average temperature data for the period 2004-2011 from the Swiss Federal Office for Meteorology and Climatology (MeteoSwiss). Daily temperature is provided as a spatial grid of about $2,300 \text{ m} \times 2,300 \text{ m}$ resolution based on an interpolation between meteorological stations ([Frei, 2014](#)).
- Daily tipping bucket rain gauge data from MeteoSwiss. Combining manual and automatic gauges, a total of 29 rain gauges also measuring snowfall are available for the period 2004-2011. For the purpose of this study, we included another set of 40 gauges that are within a maximum distance of 20 km from the basin boundary. The rain gauges are distributed evenly in space, with relatively few gauges at high elevation ranges.
- Daily cumulative discharge data for the period 2004-2011 from the Swiss Federal Office for the Environment (FOEN). The discharge measuring station is located at the outlet of the basin at Andelfingen, at an altitude of 356 m.a.s.l.

Meteorological data can be accessed free of charge for teaching and research purposes by request to the IDAweb data internet portal of MeteoSwiss for Science & Research (gate.meteoswiss.ch/idaweb/login.do). Catchment data can be requested at geodata@swisstopo.ch or in the Swisstopo website (swisstopo.ch). Discharge data can be downloaded through the website (bafu.admin.ch) of the Swiss Federal Office of the Environment (FOEN).

The HBV model

The HBV model ([Lindström et al., 1997](#)) is a conceptual lumped rainfall-runoff model developed by the Swedish Meteorological and Hydrological Institute. We chose the HBV model because of its low input data requirement and because it includes a snow melt routine. The required input data consist of time series of catchment-averaged rainfall and air temperature. The model is structured in different routines such as snow melt, evaporation, soil moisture and groundwater. Channel routing is described by a triangular hydrograph. For more detailed information about HBV, we refer to the original publication of [Lindström et al. \(1997\)](#) and to [Heistermann & Kneis \(2011\)](#) for the specific version used in the case study.

Application to the case study

Rainfall-runoff model—We decided to implement a simplified version of the HBV model from the R package RHydro ([Reusser, Buytaert & Vitolo, 2017](#)). Time series of rainfall and temperature were split into calibration (2004–2007) and validation (2008–2011) periods. The first year of the two periods (2004 and 2008) was considered as a warm-up period and discarded from the results. The daily time series were aggregated to 10-day averages to keep the number of rainfall parameters that must be calibrated in Approaches 1 and 2 manageable and to avoid convergence problems in the MCMC analysis. The HBV model was run on time steps of 10 days. The impact of the time-steps on the results is discussed more extensively in the Discussion. Prior to the Bayesian calibration and uncertainty analysis, a deterministic calibration was performed. We used a differential evolution algorithm to minimize the Mean Squared Error (MSE) between measured and predicted discharge for the calibration period. The estimated parameters are shown in [Table 1](#) and were used to help define plausible ranges for the priors of the model parameters.

Rainfall input—We defined ten rainfall intensity classes based on the 10-day catchment averaged rainfall amounts and fitted exponential variogram models for each class. Rainfall intensity increases with class number. Variogram fitting was based on all rainfall observations that are in the same class, both using rain gauges inside and outside the basin and using an approach described in [Muthusamy et al. \(2017\)](#). A plot of the fitted variograms is presented in [Fig. 2](#). Periods with an average rainfall of less than 0.1 mm were not interpolated and considered as dry. Log rainfall was simulated 500 times on a 1 km × 1 km resolution grid using the class-specific variogram and were conditioned on the rain gauges inside the catchment only. Processing was done using the R package gstat

Table 1 Model parameters and error model parameters with initial values and prior distributions. The implementation of the HBV model is based on *Heistermann & Kneis (2011)*.

Parameter name	Definition	Initial value through deterministic calibration ^a	Prior distribution of parameter ^b
Model parameters			
CFMAX	Degree day factor for snow melt [mm/°C/d]	0.324	<i>Beta</i> [3, 5, 0, 10]
TT	Temperature threshold below which precipitation falls as snow [°C]	0.387	<i>Beta</i> [3, 2, -3, 3]
FC	Field capacity [mm]	106.4	<i>Beta</i> [1, 4, 50, 200]
MINSM	Minimum soil moisture for storage [mm]	44.66	<i>Beta</i> [1, 4, 0, 200]
BETA	Parameter to control the fraction of rain and snow melt partitioned for groundwater recharge [-]	1.888	<i>Beta</i> [1, 1, 1, 5]
LP	Fraction of soil moisture-field capacity-ratio above which actual evapotranspiration equals potential evapotranspiration [-]	0.845	<i>Beta</i> [1, 1, 0, 1]
CET	Correction factor for potential evapotranspiration [-]	0.570	<i>Beta</i> [1, 1, 0, 20]
KPERC	Percolation coefficient [1/d]	2.670	<i>Beta</i> [1, 1, 0, 5]
K0	Fast storage coefficient of soil upper zone [1/d]	0.498	<i>Beta</i> [1, 1, 0, 0.5]
UZL	Threshold above which soil upper zone storage empties at rate computed by storage coefficient K0 [mm]	1.226	<i>Beta</i> [3, 4, 0, 60]
K1	Slow storage coefficient of soil upper zone [1/d]	0.356	<i>Beta</i> [1, 4, 0, 0.5]
K2	Storage coefficient of soil lower zone [1/d]	9.7×10^{-4}	<i>Beta</i> [1, 4, 0, 0.1]
MAXBAS	Length of (triangular) unit hydrograph [d]	3.887	<i>Beta</i> [1, 4, 0, 6]
etpmean	Mean evaporation [mm/d]	4.446	<i>Beta</i> [1, 3, 0, 50]
tmean	Mean temperature [°C]	4.549	<i>Beta</i> [4, 4, -20, 30]
n	(Real) number of storages in linear storage cascade	1.901	<i>Beta</i> [1, 2, 0, 7.5]
k	Decay constant for linear storage cascade	1.481	<i>Beta</i> [1, 2, 0, 5]
Initial state parameters			
snow	Snow storage [mm]	178	<i>Beta</i> [1, 3, 0, 200]
sm	Soil moisture storage [mm]	369	<i>Beta</i> [1, 3, 0, 200]
suz	Soil upper zone storage [mm]	117	<i>Beta</i> [1, 3, 0, 200]
slz	Soil lower zone storage [mm]	44	<i>Beta</i> [1, 3, 0, 200]
Error model parameters			
β_1	Coefficient for the AR(1) in Eq. (2)	-	<i>U</i> [0,1]
σ_δ^2	Variance for the AR(1), Eq. (2)	-	<i>U</i> [0,1]
σ_η^2	Measurement error variance, Eq. (1)	-	<i>U</i> [0,1]
\bar{z}	Vector of length $T + W$ for the rainfall parameters	-	$F_{\bar{z}}$, see Eq. (7)

Notes.

^aA deterministic calibration is performed prior to the Bayesian calibration. The parameters are optimized with differential evolution. The objective function is the MSE between the predicted and measured discharge.

^b*Beta* [a,b,c,d] represents a beta distribution in the interval [c,d] with shape parameters a and b. *U*[a,b] is a uniform distribution over the interval [a, b].

(*Gräler, Pebesma & Heuvelink, 2016*). Log rainfall simulations were back-transformed at point locations and spatially averaged (*Heuvelink & Pebesma, 1999*).

Five rain gauge scenarios were considered, comprising 1, 2, 5, 15 and 29 rain gauges, respectively. For each scenario, the rain gauges were selected using SSA by thinning the

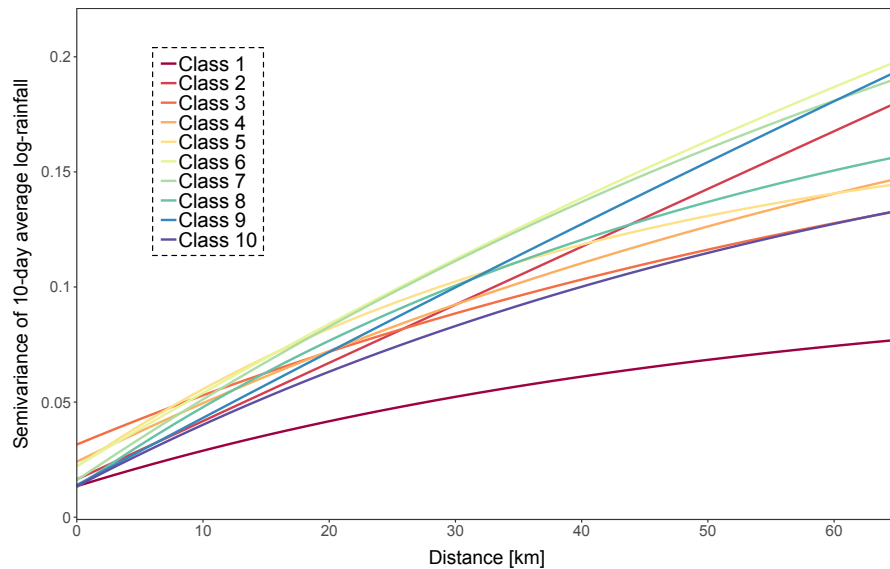


Figure 2 Fitted exponential variograms for each of the ten rainfall classes.

Full-size  DOI: [10.7717/peerj.9558/fig-2](https://doi.org/10.7717/peerj.9558/fig-2)

existing network. The minimization criterion was the average block kriging variance computed by discretization of the area into 500 sub-areas. Implementation was done with the R package *spsann* (Samuel-Rosa, 2017). The initial SSA temperature was set to 0.1 and the cooling factor was set to 0.8. The total number of SSA iterations was fixed at 10,000. Five out of the eight tested scenarios are reported in this study.

Bayesian inference—Model parameters and their priors are shown in Table 1. Prior parameter distributions were chosen based on expert knowledge, previous work in the same basin and optimized parameter values of the deterministic calibration. For each rain gauge scenario the Bayesian inference was performed. The number of MCMC iterations was fixed at 10^6 for Approaches 1 and 2 and to 10^4 for Approach 3. The process was repeated several times to ensure convergence of the parameter estimates.

RESULTS

Rain gauge optimization

Figure 3 shows maps of the study area with the locations of the selected rain gauges, for different rain gauge density scenarios. Recall that during the optimization with SSA, the locations of the rain gauges for a given density are selected from the existing locations. Figure 3 shows that the spread of rain gauges is fairly homogeneous in the geographic space. For the scenario with a single rain gauge, the optimal location is in the center of the study area, while with two rain gauges the optimal positions are near the boundaries. This is a known phenomenon when optimizing sampling locations for mapping with ordinary kriging (Wadoux, Marchant & Lark, 2019).

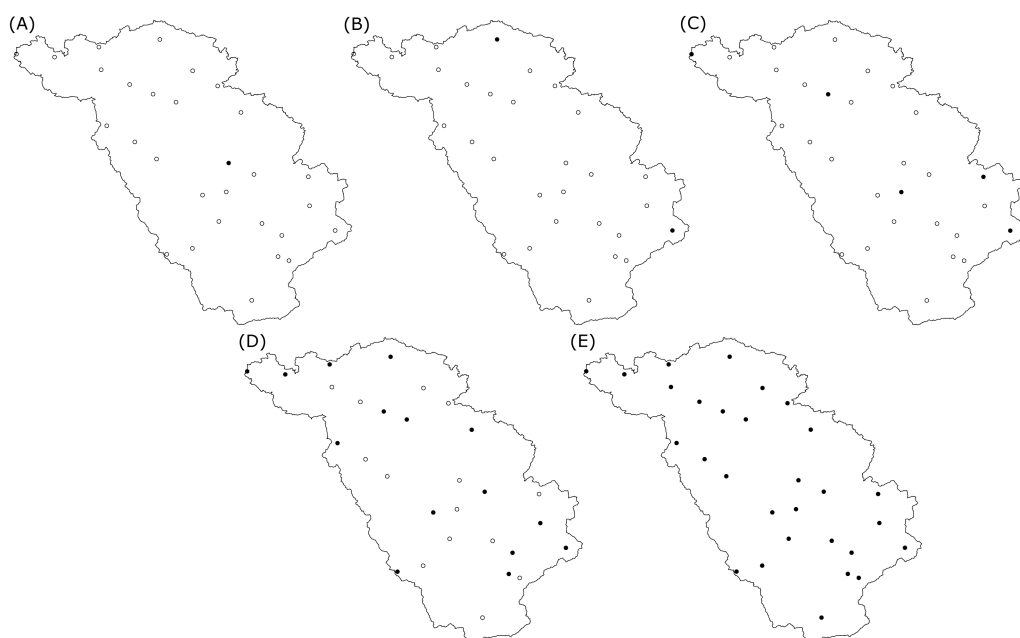


Figure 3 Spatial location of the optimized rain gauges for scenarios of (A) one, (B) two, (C) five, (D) fifteen and (E) all (29) rain gauges. The solid black dots are the selected optimal locations for the rain gauges, empty black dots are the possible candidate locations during the SSA optimization.

Full-size DOI: [10.7717/peerj.9558/fig-3](https://doi.org/10.7717/peerj.9558/fig-3)

Calibration

Figure 4 shows the posterior distribution of the calibrated model parameters for Approaches 1 and 2, for different rain gauge density scenarios, along with their prior distribution. For most parameters the posterior distribution is narrower than the prior distribution. This particularly holds for parameters associated with the routing routine (MAXBAS, n , k), the initial state (snow, sm , suz , slz) and the error model (β_1 , σ_δ^2 , σ_η^2). Note that while the posterior distribution of some parameters (e.g. LP, KPERC, K0, MAXBAS and n) is comparable for different rain gauge density scenarios, all other parameter distributions show (large) differences between different rain gauge density scenarios. The distributions do not seem to be narrower in case of higher rain gauge density.

Figure 5 shows the prior and posterior distributions of the calibrated parameters for Approach 3, for different rain gauge scenarios. The posterior distributions are much narrower than the prior distributions, particularly for the parameters of the snow routine (CFMAX, TT, FC), the routing routine (MAXBAS, n , k), initial state and error model. For parameters LP, K0 and K1, the posterior distributions are very similar to the priors. In contrast to **Fig. 4**, all parameters have very similar posterior distributions for different rain gauge densities.

Prediction

Figure 6 shows the rainfall prior and corrected prior (“posterior”) for the prediction period. Recall that the prior distribution was directly sampled for Approaches 2 and 3 while the rainfall “posterior” distribution was sampled for Approach 1. As expected, the

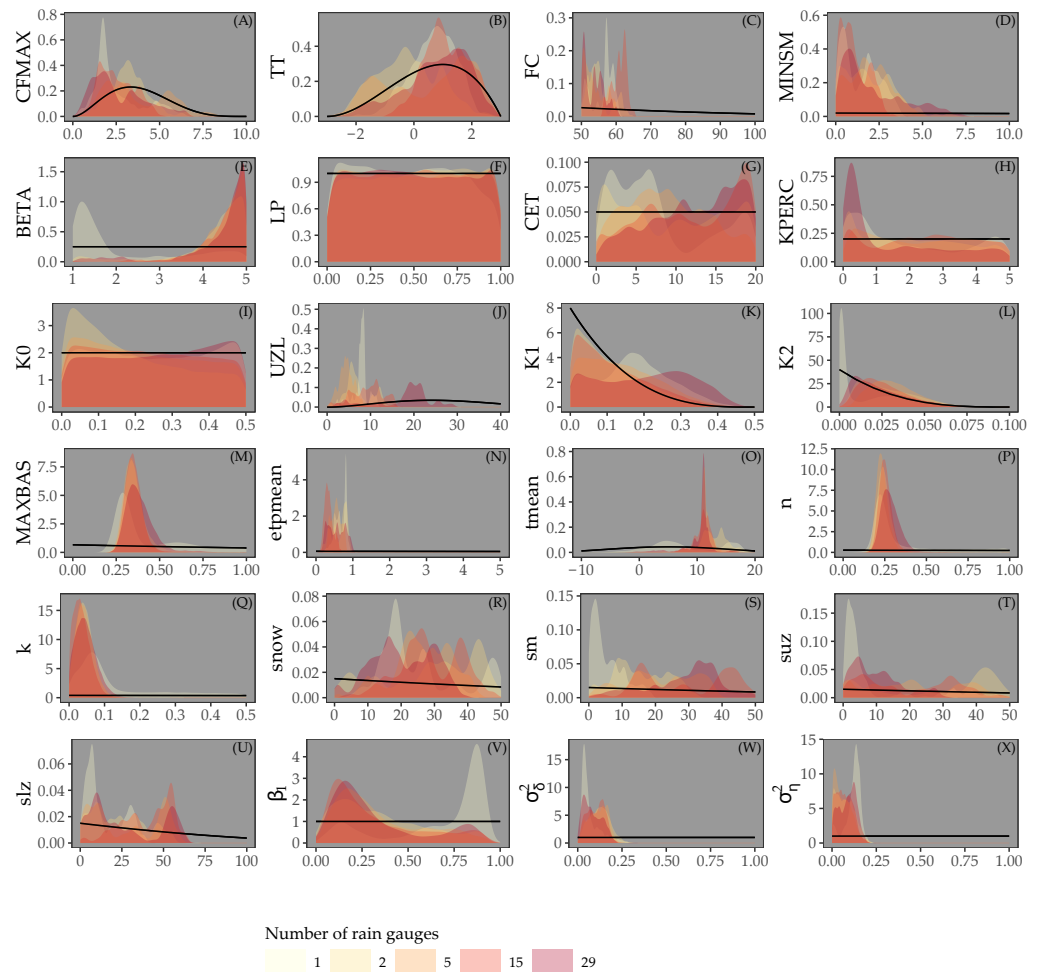


Figure 4 Parameters estimated by Bayesian calibration for Approaches 1 and 2. Black lines in A-X represent prior distributions and colored shapes posterior densities for different rain gauge density scenarios. Rainfall parameter results not shown.

Full-size [DOI: 10.7717/peerj.9558/fig-4](https://doi.org/10.7717/peerj.9558/fig-4)

prediction interval width increases when using a smaller number of rain gauges. Both rainfall prior and posterior distributions showed very similar prediction interval widths and mean values. Small differences can be observed when using a small number of rain gauges (i.e. 1 or 2 rain gauges).

Figure 7 shows the 90% prediction intervals of the discharge for the three approaches and including/excluding various uncertainty sources. For the case where all uncertainty sources are accounted for (plots (A, D, G)), there is a clear pattern towards a smaller width of the prediction intervals with an increase of the number of rain gauges. While there is no clear difference in terms of prediction intervals between the three approaches, Approach 3 provides a larger interval width at a certain time period (e.g., events at 2009-07 and 2010-08) for the scenario involving 15 rain gauges. Note that the differences between rain gauge scenarios are most pronounced for high flows and become negligible for low flows.

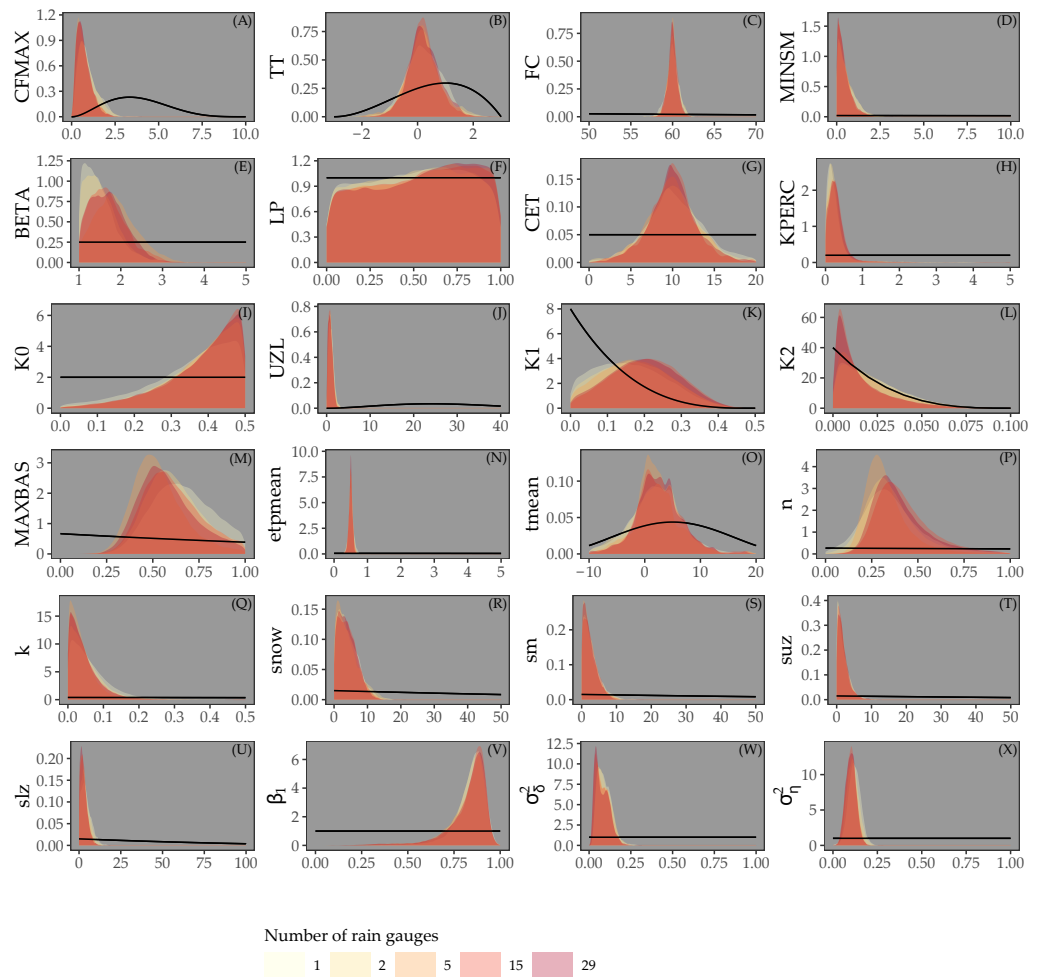


Figure 5 Parameters estimated by Bayesian calibration for Approach 3. Black lines in (A-X) represent prior distributions and colored shapes posterior densities for different rain gauge density scenarios.

Full-size [DOI: 10.7717/peerj.9558/fig-5](https://doi.org/10.7717/peerj.9558/fig-5)

For example, the scenario involving a single rain gauge shows large uncertainty for all approaches at the discharge peak for the event at 2010–08.

Plots A, D and G in [Figures 7A, 7D and 7G](#) do not provide information about the separate effect of input, model parameters and model structural uncertainty on the joint predictive uncertainty. Therefore we also performed an uncertainty propagation analysis that includes/excludes the various uncertainty sources for the cases where model structural uncertainty was ignored and where model structural and model parameter (comprising initial state and error model parameters) uncertainty were both ignored. When model structural uncertainty is ignored ([Fig. 7B, 7E and 7H](#)) the uncertainty decreases. With the additional effect of model parameter uncertainty removed ([Fig. 7C, 7F and 7I](#)), the prediction intervals become much narrower, with large differences for different rain gauge densities, i.e. the larger the number of rain gauges, the smaller the prediction interval width. The latter reduction is particularly visible when increasing the density from 1 to 5 rain

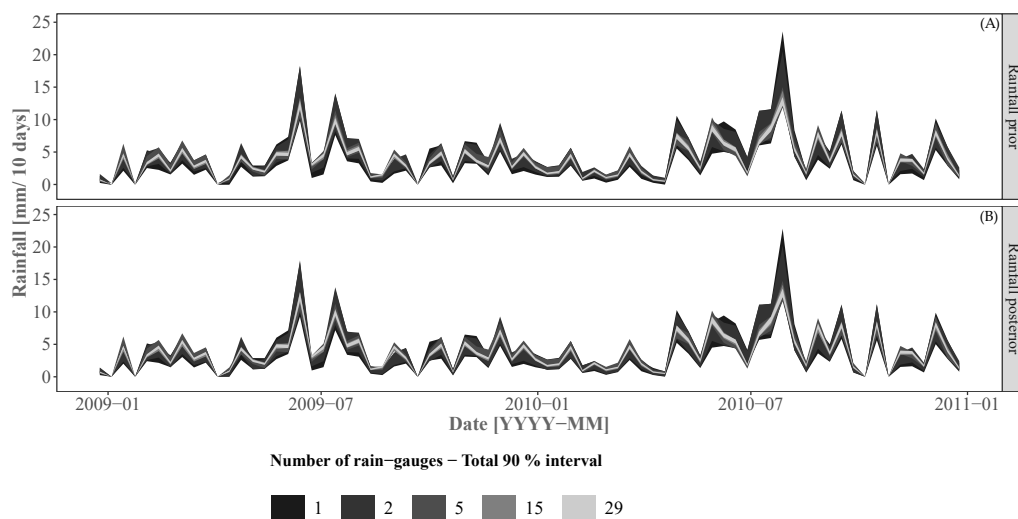


Figure 6 Rainfall priors (A) and posteriors (B) for the prediction/validation period. The rainfall posterior is the corrected prior for Approach 1 for the prediction/validation period. Approaches 2 and 3 sample from the rainfall prior for prediction and validation.

Full-size [DOI: 10.7717/peerj.9558/fig-6](https://doi.org/10.7717/peerj.9558/fig-6)

gauges, and becomes marginal when using between 5 to 29 rain gauges. As noted before, the largest difference is obtained for high flow periods.

Validation

Table 2 shows the validation statistics for the three approaches and the five tested rain gauge scenarios. For the three approaches, increasing the number of rain gauges led to an increase of the predictive power of the model (increase of the NSE). This increase was generally modest, except for Approach 3. It was accompanied by a modest decrease of the residual as characterized by the RMSE. The model predictions were practically unbiased (largest ME deviation from zero equals -0.42), which shows that the prediction inaccuracy was mainly due to random error rather than systematic error. There is no clear pattern regarding the PICP with increasing rain gauge density. For both intervals (50% and 90%), the percentage of observations covered by the interval was within a reasonable range of variation. There is a slight overestimation of the uncertainty for Approaches 1 and 2, while there is a small underestimation of uncertainty for Approach 3 at the 90% interval. Approaches 1 and 2 were very similar in terms of validation statistics, particularly when a large number of rain gauges was used. This is according to expectations, as the effect of the linear correction on the prior diminishes with an increasing number of rain gauges. Approach 2 is more accurate for a small number of rain gauges, while Approach 1 has the largest NSE of all scenarios when using all rain gauges.

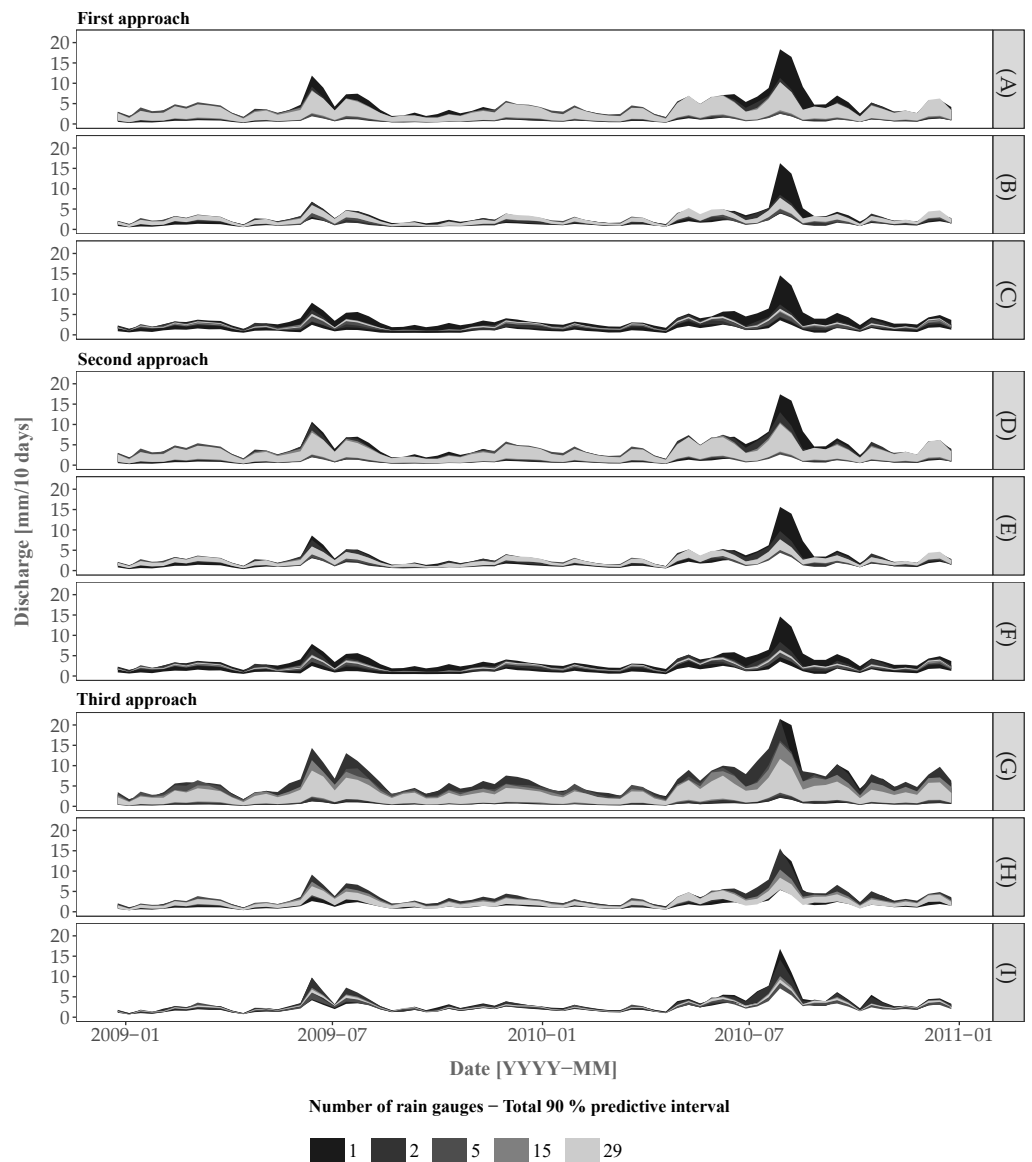


Figure 7 Prediction of the discharge using the three approaches for the cases (A, D, G) where all sources of error are accounted for, (B, E, H) model structural uncertainty is ignored and (C, F, I) model structural uncertainty and model parameter uncertainty (excluding the rainfall input parameters for Approaches 1 and 2) are ignored.

Full-size DOI: [10.7717/peerj.9558/fig-7](https://doi.org/10.7717/peerj.9558/fig-7)

DISCUSSION

Consistency of parameter estimates

Our experiments suggest that several model parameters estimated in Approaches 1 and 2 might be weakly identified because of their wide posterior distribution. *Gelman et al. (2014)* stressed that the concept of identification is not so important in the Bayesian perspective and that one must rather look at how much information is supplied by the data, i.e. that

Table 2 Validation measures for the three approaches and five rain gauge densities, computed over the period 2009–2011.

	Number of rain gauges				
	1	2	5	15	29
Approach 1					
NSE	0.28	0.28	0.48	0.54	0.56
RMSE (mm/10 days)	1.25	1.25	1.06	1.01	0.99
ME (mm/10 days)	−0.42	−0.23	−0.31	−0.15	−0.16
PICP 50%	52.72	58.18	53.63	61.64	59.69
PICP 90%	90.00	90.63	94.54	92.72	94.54
Approach 2					
NSE	0.39	0.48	0.45	0.49	0.50
RMSE (mm/10 days)	1.15	1.06	1.09	1.04	1.01
ME (mm/10 days)	−0.19	0.08	−0.02	−0.03	0.09
PICP 50%	54.54	60.00	59.99	62.63	60.00
PICP 90%	94.54	92.72	92.72	93.63	94.54
Approach 3					
NSE	0.17	0.24	0.36	0.38	0.42
RMSE (mm/10 days)	1.37	1.27	1.18	1.15	1.12
ME (mm/10 days)	0.28	0.17	−0.11	0.10	0.06
PICP 50%	55.45	56.36	51.81	52.72	53.63
PICP 90%	88.18	86.36	84.54	86.36	85.45

the joint parameter posterior must occupy less space than the joint prior distribution. In our case, Figs. 4 and 5 show that posteriors were narrower than the priors. This indicates that information was supplied by the data and explains why the parameter posteriors were actually accurate predictors, as shown by the NSE value being higher for Approaches 1 and 2 than for Approach 3, despite the wider posterior distributions of parameters of the third approach.

The posterior of parameter β_1 of the AR(1) model suggests that temporal correlation of the model structural error was weak to moderate, which agrees with findings of [Huard & Mailhot \(2008\)](#). Parameters of the model structural error were well identified. Realistic assumptions regarding the model structural error model formulation play a major role to distinguish between model structural and input uncertainty. Since model structural uncertainty is incorporated explicitly it is unlikely that input uncertainty compensates for deficits in the model structure ([Thyer et al., 2009](#)). We followed the approach of [Beven & Freer \(2001\)](#) that model structural uncertainty can be described by a first order autoregressive model. We acknowledge that more complex structures of model residuals can be formulated, such as by using an ARMA or ARIMA model, but this would be at the expense of increasing parameter space dimensionality.

Prediction

In contrast to many studies reported in the literature (e.g. [Kavetski, Kuczera & Franks, 2006](#)), analysis of the predictive uncertainty shows that in the case study the contribution

of rainfall uncertainty is relatively small and that discharge predictive uncertainty is mainly dominated by model structural and model parameter uncertainty. As a consequence, the effect of the rain gauge density diminishes if model parameter and model structural uncertainty are accounted for. This was different in the case of using a single rain gauge for calibration. In this case, rainfall uncertainty was an important contributor to the total predictive uncertainty, particularly for the peak flow. Large model parameter and model structural uncertainty offsets small input rainfall uncertainty. In particular for Approach 3, model structural uncertainty is clearly the largest contributor to the total predictive uncertainty. Several investigations obtained similar results. For example, [Engeland, Xu & Gottschalk \(2005\)](#) showed that model structural uncertainty is larger than model parameter uncertainty for simple conceptual models with few well-defined parameters. Our results also confirm the study of [Talamba, Parent & Musy \(2010\)](#) for a lumped hydrological model. [Talamba, Parent & Musy \(2010\)](#) showed, for a fully distributed hydrological model, that accounting for input rainfall uncertainty did not lead to a substantial change in terms of estimated parameters and model performance, because other sources of uncertainty dominated the total predictive uncertainty. This is similar to our case study, where in all three approaches the prediction intervals have a similar width and range. The validation measures show that Approaches 1 and 2, i.e. the case where rainfall parameters are calibrated, outperform Approach 3.

Differences between the three approaches

Comparison of the three approaches revealed that Approach 3 has larger prediction intervals and poorer model performance, despite the fact that in Approach 3 most parameters have a well-defined unimodal posterior distribution. Note also that the choice of approach leads to different posterior ranges of parameter estimates. The smaller number of parameters to calibrate in Approach 3 (i.e. the input rainfall parameters are not calibrated) suggests that inference benefits from the reduced dimensionality of the parameter space. The validation results show that this was not the case when a small number of rain gauges is used (for instance, $NSE = 0.17$ using 1 rain gauge). [Thyer et al. \(2009\)](#) and [Huard & Mailhot \(2008\)](#) reported similar results when calibrating time-dependent rainfall input parameters. They showed how calibrating input rainfall parameters for each time step compensates for the situation where a rainfall event is not recorded by a small number of rain gauges, and how this can lead to a near-perfect match between the observed and predicted discharge. In the latter case, [Huard & Mailhot \(2008\)](#) demonstrated that, since model and input rainfall parameters are estimated jointly, it is likely that the input rainfall parameters compensate for structural deficits of the model. In our case this was avoided by: (i) explicitly accounting for model structural uncertainty; and (ii) defining meaningful priors for the input rainfall parameters using geostatistical analysis. Note that these conclusions could be balanced with the use of radar remote sensing images in combination with the rain gauges to estimate the rainfall. Using one single rain gauge could be sufficient, provided that the radar image detected the rainfall event ([Di Piazza et al., 2011](#)).

From a numerical perspective, a major difference between Approaches 1 and 2 and Approach 3 is the number of parameters to calibrate. Hydrologists tend to shy-away from

high-dimensional rainfall input parameter space, because it often leads to question the statistical significance of the inferred parameters (e.g. [Vrugt et al., 2008](#)). A solution is to resort to MCMC search algorithms, which are efficient to explore the multi-dimensional and correlated parameter space. This has been recently tackled in the hydrological literature (e.g. by [Laloy & Vrugt, 2012](#)) and in the more general statistical literature (e.g. [Ter Braak, 2006](#)). Another solution is to reduce the dimension of the parameter space. Alternative methods to Approach 1 (i.e. the case where each time step is an input rainfall parameter to calibrate) exist. For example, [Kavetski, Kuczera & Franks \(2006\)](#) use storm-event multipliers under the assumption of perfect dependence of input errors within single storm events. By letting these multipliers vary according to the plausible range of hydrological variation, they correct for systematic error in the rainfall input. The major limitation is the need to define hydrological ranges in which the calibrated multiplier is kept constant.

In practical terms, one would not make a serious flaw by taking Approach 3 and not calibrating the input rainfall uncertainty as additional parameters in the Bayesian uncertainty analysis. This conclusion is particularly valid when the number of rain gauges is larger than five in our case study, but this might not hold for all cases. In all approaches, the uncertainty was accurately quantified. This means that despite poorer predictive performance, Approach 3 still yields valuable information because the prediction uncertainty is fairly low. In most cases, practitioners are interested in reliable prediction of the discharge, but given the computational complexity of Approaches 1 and 2, modellers might rightfully opt for Approach 3 with negligible consequences in terms of uncertainty quantification and a relatively small decrease of predictive performance.

Implications for rain gauge density

The impact of the rain gauge density on parameter posterior distributions was modest. Parameter posteriors for Approaches 1 and 2 show that there was typically little difference between the rain gauge scenarios, while there was almost no difference in the case of Approach 3. A difference is found in Approaches 1 and 2, where using a single rain gauge led to a different posterior distribution of some parameters. This suggests that parameter estimation was robust to the density of the rain gauges, provided that the network is composed of more than a single rain gauge. This is an important finding, as the necessary condition for the regionalization of a rainfall-runoff model is that the parameters are insensitive to the choice of the number and locations of rain gauges ([Thyer et al., 2009](#)). We acknowledge that these results may be different in different circumstances, such as in a case where rainfall uncertainty has a larger impact on the parameter posterior distributions, or in the case of using remote sensing images and environmental predictors as covariate in the kriging step ([Bostan, Heuvelink & Akyurek, 2012](#)). This study, however, contradicts the findings of [Zeng et al. \(2018\)](#), who found that parameter posterior distributions vary considerably under different rain gauge densities. However, [Zeng et al. \(2018\)](#) did not propagate rainfall input uncertainty and simply sampled a large number of possible rain gauge combination for a given density and analyzed the differences between rain gauge densities in terms of the model parameter posterior distribution. Thus, they ignored a substantial proportion of the uncertainty, which potentially caused model parameter

uncertainty to compensate for the unaccounted rainfall uncertainty (*Kavetski, Kuczera & Franks, 2006*).

In our experiment, low rain gauge densities already produced accurate model predictions. This particularly applied to Approaches 1 and 2 where only five gauges led to a NSE greater than 0.45 (two gauges are sufficient in Approach 2 for a NSE of 0.48). This threshold is low compared to other studies. *Dong, Dohmen-Janssen & Booij (2005)* reported that five rain gauges were enough to calibrate the HBV model in a 17,000 km² basin in China, using the expected variance of the areal rainfall as a measure of input uncertainty. In a 3,234 km² catchment in France, *Anctil et al. (2006)* concluded that ten rain gauges (out of 23) is an absolute minimum to predict discharge using a neural network model. However, the rainfall spatial variation was not modeled explicitly in this study. *Bárdossy & Das (2008)* found that the overall model performance worsened radically with an excessive reduction of rain gauges in the upper Neckar catchment of about 4,000 km². They optimized the rain gauge locations for different rain gauge densities using simulated annealing and kriging with external drift. They showed a significant reduction of the rainfall input variance with increasing density, which paired with a decrease of the discharge prediction error. However, the cited studies did not use Bayesian calibration. In a Bayesian framework, *Zeng et al. (2018)* found that 10–15 rain gauges were necessary to obtain stable parameter estimates for medium-size sub-basins, but the propagation of the input rainfall uncertainty was not analyzed in this study. In our study, the input rainfall uncertainty was estimated using geostatistics and propagated in a Bayesian framework. A fairly accurate estimate of the catchment average 10-day rainfall was obtained using just two or five rain gauges, which explains why a surprisingly small number of rain gauges was enough to calibrate the hydrological model. The rainfall posterior parameters adjust for the missing information using the discharge data. In Approach 3, i.e. when the rainfall is not updated by the discharge data, the model performance was worse than the other approaches in all cases, and particularly for cases with a small number of rain gauges (fewer than five rain gauges).

Although using a small number of rain gauges led to accurate model prediction, using more rain gauges improves the model predictions. The results of our case study showed that a density larger than five rain gauges led to a marginal improvement of the prediction accuracy. This is equivalent to one rain gauge per 340 km². It should be noted that this result cannot easily be generalized because it is likely case-dependent. In particular, when optimizing the number of rain gauges we assumed that these are placed optimally so that the block kriging variance is minimal, while in practice the gauges might be clustered at some parts (e.g. at lower altitude) of the catchment. In such case, we may need more gauges to reach the same block kriging variance. The surprisingly low rain gauge density is also likely related to the 10-day time step that we used in the case study. Aggregating rainfall over 10-day periods automatically leads to a decrease of the rainfall spatial variation. [Figure 2](#) shows that the spatial variation of the 10-day average rainfall in the Thur basin is relatively small: the variogram sill was small and the variogram range was large. Note also that we assumed a log-normal distribution of the rainfall amount at point support, but that more complex models can be used to account for the typically skew and zero-inflated rainfall distribution (*Engeland et al., 2016*). We emphasize that predicting 10-day average discharge

also leads to smoothing and hence will miss peak discharges. For certain applications a smaller time step will be required, although the time step that we used is suitable for many applications, e.g. for total discharge prediction over long time periods or agricultural water management. Despite data availability, we did not use subcatchments and/or elevation zones but focused on the total river discharge prediction. This study could theoretically be extended to investigate the effect of rain gauge density and rainfall variability on several of the Thur gauged sub-catchments, with a distributed or semi-distributed model. In this way, it would be possible to assess the effect of rainfall spatial distribution on the discharge at the catchment outlet. This would undoubtedly have an impact on the optimal rain gauge sampling design and density, but this approach would be at the expense of numerical efficiency.

Limitations, extensions and improvements

Finally, we emphasize that in this study we made several assumptions and modelling choices that may not hold for other case studies and/or other model applications.

The log-transformed discharge measurement error was assumed normally distributed and constant over time. This is a common assumption in hydrological modelling (e.g. [Huard & Mailhot, 2008](#)). In fact, several authors (e.g. [Weerts & El Serafy, 2006](#); [Mazzoleni et al., 2018](#)) model discharge error by a normal distribution with standard deviation fixed at 10% of the measured discharge. In this study we did not fix the measurement error variance but estimated it from the data, together with other model parameters. Further, we avoided the common assumption of temporal independence of log-transformed model structural uncertainty, made for example by [Rakovec et al. \(2015\)](#) and [Mazzoleni et al. \(2018\)](#). We judged that it was unrealistic to assume that deficiencies in model structure are independent in time and therefore accounted for temporal dependence using a first-order auto-regressive model. In future research more elaborate ARIMA models could be used to represent model structural uncertainty.

For several hydrological applications, such as flood forecasting, daily or sub-daily rainfall information is required to capture the high variability of peak flows. The conclusions drawn in this study regarding rain gauge density will likely not hold when rainfall variability increases or rainfall and discharge data are aggregated over smaller time steps. Likely there is also an important effect of catchment size and/or location. The minimum average density of rain gauges required for modelling (in our case study one rain gauge per 340 km²) may not hold for smaller catchments or sub-catchments, while it might be a too high density for very large, continental-scale catchments. As mentioned before, the effect of elevation on rainfall and discharge prediction error was not included in this study either, although it may have an impact on the optimal locations of the rain gauges. The use of radar/satellite derived rainfall fields (e.g., [Huza et al., 2014](#); [Cecinati et al., 2017](#)) may also have a significant impact on the uncertainty of the rainfall fields and therefore on the joint posterior uncertainty of the discharge. This might be particularly useful if combined with spatially-distributed hydrological models, such as the WaSIM-ETH ([Schulla & Jasper, 2007](#)) and wflow_sbm ([Imhoff et al., 2020](#)) models. Our methodology also applies to spatially distributed modelling approaches, although computing time will likely increase.

Since our conclusions are case-study specific, this means that for each case study, one should design an experiment to derive the minimum number and locations of rain gauges to be used in hydrological model calibration. In this work we have presented a methodology that may be used to carry out such experiments.

CONCLUSION

We calibrated the HBV rainfall-runoff model accounting for input, parameter, initial state and model structural uncertainty using a Bayesian framework for a 1,700 km² basin in Switzerland. Prior input rainfall distributions were derived using a geostatistical approach. We tested several scenarios for incorporating the input uncertainty and assessed the effect of rain gauge density on calibration. The main conclusions are:

- Assumptions regarding the formulation of the model structural uncertainty play a major role in distinguishing between model structural and input uncertainty. Since model structural uncertainty is incorporated explicitly it is unlikely that input uncertainty compensates for deficits in the model structure.
- In our case study, input uncertainty was small compared to model structural and parameter uncertainty.
- Calibrating the rainfall parameters (Approach 1) led to more accurate model performance compared to the case where rainfall uncertainty was not updated using discharge data (Approaches 2 and 3). The increased dimensionality of the parameter space when calibrating rainfall did not lead to computational intractability.
- Parameter estimates were robust to rain gauge density. This is important, as this enables regionalization of the rainfall-runoff model.
- In our case study, using only two rain gauges did not seriously deteriorate discharge prediction.
- Adding up to five rain gauges improved the model prediction. Adding even more only produced a marginal improvement of the prediction accuracy. For our study area, five rain gauges is equivalent to one rain gauge per 340 km².

ACKNOWLEDGEMENTS

We thank MeteoSwiss, Swisstopo and FOEN for providing the data used for this research.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This project received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 607000. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

European Union's Seventh Framework Programme for research, technological development and demonstration: 607000.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Alexandre M.J.-C. Wadoux conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Gerard B.M. Heuvelink, Remko Uijlenhoet and Sytze de Bruin conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

This work uses publicly available data.

Meteorological data can be accessed free of charge for teaching and research purposes upon request to the IDAweb data internet portal of MeteoSwiss for Science & Research (gate.meteoswiss.ch/idaweb/login.do) after registration, and by searching the database name: RdissaggH.

Catchment data can be requested at geodata@swisstopo.ch or in the Swisstopo website (swisstopo.ch), accession number: GWN07 of VECTOR25.

Discharge data can be downloaded through the website (bafu.admin.ch) of the Swiss Federal Office of the Environment (FOEN), database number 136, station code S1 and FOEN code 2044.

REFERENCES

- Ajami NK, Duan Q, Sorooshian S. 2007.** An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research* **43**:W01403.
- Anctil F, Lauzon N, Andréassian V, Oudin L, Perrin C. 2006.** Improvement of rainfall-runoff forecasts through mean areal rainfall optimization. *Journal of Hydrology* **328**(3–4):717–725 DOI [10.1016/j.jhydrol.2006.01.016](https://doi.org/10.1016/j.jhydrol.2006.01.016).
- Andréassian V, Lerat J, Loumagne C, Mathevet T, Michel C, Oudin L, Perrin C. 2007.** What is really undermining hydrologic science today? *Hydrological Processes* **21**(20):2819–2822 DOI [10.1002/hyp.6854](https://doi.org/10.1002/hyp.6854).
- Bal G, Rivot E, Baglinière J-L, White J, Prévost E. 2014.** A hierarchical Bayesian model to quantify uncertainty of stream water temperature forecasts. *PLOS ONE* **9**(12):e115659.
- Bárdossy A, Das T. 2008.** Influence of rainfall observation network on model calibration and application. *Hydrology and Earth System Sciences* **12**(1):77–89 DOI [10.5194/hess-12-77-2008](https://doi.org/10.5194/hess-12-77-2008).

- Beven K. 2006.** On undermining the science? *Hydrological Processes* **20(14)**:3141–3146
[DOI 10.1002/hyp.6396](https://doi.org/10.1002/hyp.6396).
- Beven K, Freer J. 2001.** Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* **249(1)**:11–29
[DOI 10.1016/S0022-1694\(01\)00421-8](https://doi.org/10.1016/S0022-1694(01)00421-8).
- Beven KJ, Freer J, Hankin B, Schulz K. 2000.** The use of generalised likelihood measures for uncertainty estimation in high order models of environmental systems. In: *Nonlinear and nonstationary signal processing*. Cambridge: CUP, 115–151.
- Bostan PA, Heuvelink GBM, Akyurek SZ. 2012.** Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. *International Journal of Applied Earth Observation and Geoinformation* **19**:115–126
[DOI 10.1016/j.jag.2012.04.010](https://doi.org/10.1016/j.jag.2012.04.010).
- Brooks S, Gelman A, Jones G, Meng X-L. 2011.** *Handbook of Markov chain Monte Carlo*. Boca Raton: CRC press.
- Brown JD, Heuvelink GBM. 2005.** Assessing uncertainty propagation through physically based models of soil water flow and solute transport. In: *Encyclopedia of Hydrological Sciences*. Hoboken: Wiley Online Library, 1181–1195.
- Cecinati F, Rico-Ramirez MA, Heuvelink GBM, Han D. 2017.** Representing radar rainfall uncertainty with ensembles based on a time-variant geostatistical error modelling approach. *Journal of Hydrology* **548**:391–405 [DOI 10.1016/j.jhydrol.2017.02.053](https://doi.org/10.1016/j.jhydrol.2017.02.053).
- Ciabatta L, Brocca L, Massari C, Moramarco T, Gabellani S, Puca S, Wagner W. 2016.** Rainfall-runoff modelling by using SM2RAIN-derived and state-of-the-art satellite rainfall products over Italy. *International Journal of Applied Earth Observation and Geoinformation* **48**:163–173 [DOI 10.1016/j.jag.2015.10.004](https://doi.org/10.1016/j.jag.2015.10.004).
- Cressie N. 2015.** *Statistics for spatial data*. Chichester: John Wiley & Sons.
- Di Baldassarre G, Montanari A. 2009.** Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences* **13(6)**:913
[DOI 10.5194/hess-13-913-2009](https://doi.org/10.5194/hess-13-913-2009).
- Di Piazza A, Conti FL, Noto LV, Viola F, La Loggia G. 2011.** Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. *International Journal of Applied Earth Observation and Geoinformation* **13(3)**:396–408
[DOI 10.1016/j.jag.2011.01.005](https://doi.org/10.1016/j.jag.2011.01.005).
- Dong X, Dohmen-Janssen CM, Booij MJ. 2005.** Appropriate spatial sampling of rainfall or flow simulation/échantillonnage spatial de la pluie approprié pour la simulation d'écoulements. *Hydrological Sciences Journal* **50(2)**:279–298
[DOI 10.1623/hysj.50.2.279.61801](https://doi.org/10.1623/hysj.50.2.279.61801).
- Durbin J, Koopman SJ. 2012.** *Time series analysis by state space methods*. Vol. 38. Oxford: OUP.
- Engeland K, Steinsland I, Johansen SS, Petersen-Øverleir A, Kolberg S. 2016.** Effects of uncertainties in hydrological modelling. A case study of a mountainous catchment in Southern Norway. *Journal of Hydrology* **536**:147–160
[DOI 10.1016/j.jhydrol.2016.02.036](https://doi.org/10.1016/j.jhydrol.2016.02.036).

- Engeland K, Xu C-Y, Gottschalk L. 2005. Assessing uncertainties in a conceptual water balance model using Bayesian methodology/estimation bayésienne des incertitudes au sein d'une modélisation conceptuelle de bilan hydrologique. *Hydrological Sciences Journal* **50**(1):45–63.
- Frei C. 2014. Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. *International Journal of Climatology* **34**(5):1585–1605 DOI [10.1002/joc.3786](https://doi.org/10.1002/joc.3786).
- Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. 2014. *Bayesian data analysis*. Vol. 2. Boca Raton: CRC press.
- Gräler B, Pebesma E, Heuvelink GBM. 2016. Spatio-temporal interpolation using gstat. *The R Journal* **8**:204–218 DOI [10.32614/RJ-2016-014](https://doi.org/10.32614/RJ-2016-014).
- Habib E, Krajewski WF, Kruger A. 2001. Sampling errors of tipping-bucket rain gauge measurements. *Journal of Hydrologic Engineering* **6**(2):159–166 DOI [10.1061/\(ASCE\)1084-0699\(2001\)6:2\(159\)](https://doi.org/10.1061/(ASCE)1084-0699(2001)6:2(159)).
- Han S, Coulibaly P. 2017. Bayesian flood forecasting methods: a review. *Journal of Hydrology* **551**:340–351 DOI [10.1016/j.jhydrol.2017.06.004](https://doi.org/10.1016/j.jhydrol.2017.06.004).
- Heistermann M, Kneis D. 2011. Benchmarking quantitative precipitation estimation by conceptual rainfall-runoff modelling. *Water Resources Research* **47**(6):W06514.
- Heuvelink GBM, Pebesma EJ. 1999. Spatial aggregation and soil process modelling. *Geoderma* **89**(1):47–65 DOI [10.1016/S0016-7061\(98\)00077-9](https://doi.org/10.1016/S0016-7061(98)00077-9).
- Højberg AL, Refsgaard JC. 2005. Model uncertainty–parameter uncertainty versus conceptual models. *Water Science and Technology* **52**(6):177–186.
- Houska T, Kraft P, Chamorro-Chavez A, Breuer L. 2015. Spotting model parameters using a ready-made python package. *PLOS ONE* **10**(12):e0145180.
- Huard D, Mailhot A. 2008. Calibration of hydrological model GR2M using Bayesian uncertainty analysis. *Water Resources Research* **44**(2):W02424.
- Huza J, Teuling AJ, Braud I, Grazioli J, Melsen LA, Nord G, Raupach TH, Uijlenhoet R. 2014. Precipitation, soil moisture and runoff variability in a small river catchment (Ardèche, France) during HyMeX Special Observation Period 1. *Journal of Hydrology* **516**:330–342 DOI [10.1016/j.jhydrol.2014.01.041](https://doi.org/10.1016/j.jhydrol.2014.01.041).
- Imhoff RO, Van Verseveld WJ, Van Osnabrugge B, Weerts AH. 2020. Scaling point-scale (pedo) transfer functions to seamless large-domain parameter estimates for high-resolution distributed hydrologic modelling: an example for the Rhine river. *Water Resources Research* **56**(4):e2019WR026807.
- Kalman RE. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* **82**(1):35–45 DOI [10.1115/1.3662552](https://doi.org/10.1115/1.3662552).
- Kavetski D, Kuczera G, Franks SW. 2006. Bayesian analysis of input uncertainty in hydrological modelling: 1. Theory. *Water Resources Research* **42**:W03408.
- Kennedy MC, O'Hagan A. 2001. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B* **63**(3):425–464 DOI [10.1111/1467-9868.00294](https://doi.org/10.1111/1467-9868.00294).
- Keum J, Kaluarachchi JJ. 2015. Development of a decision-making methodology to design a water quality monitoring network. *Environmental Monitoring and Assessment* **187**(7):466 DOI [10.1007/s10661-015-4687-z](https://doi.org/10.1007/s10661-015-4687-z).

- Laiolo P, Gabellani S, Campo L, Silvestro F, Delogu F, Rudari R, Pulvirenti L, Boni G, Fascetti F, Pierdicca N. 2016.** Impact of different satellite soil moisture products on the predictions of a continuous distributed hydrological model. *International Journal of Applied Earth Observation and Geoinformation* **48**:131–145 DOI [10.1016/j.jag.2015.06.002](https://doi.org/10.1016/j.jag.2015.06.002).
- Laloy E, Vrugt JA. 2012.** High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing. *Water Resources Research* **48**(1):W01526.
- Lindström G, Johansson B, Persson M, Gardelin M, Bergström S. 1997.** Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology* **201**(1-4):272–288 DOI [10.1016/S0022-1694\(97\)00041-3](https://doi.org/10.1016/S0022-1694(97)00041-3).
- Maskey S. 2004.** *Modelling uncertainty in flood forecasting systems*. Boca Raton: CRC Press.
- Mazzoleni M, Noh SJ, Lee H, Liu Y, Seo D-J, Amaranto A, Alfonso L, Solomatine DP. 2018.** Real-time assimilation of streamflow observations into a hydrological routing model: effects of model structures and updating methods. *Hydrological Sciences Journal* **63**(3):386–407 DOI [10.1080/02626667.2018.1430898](https://doi.org/10.1080/02626667.2018.1430898).
- Melsen L, Teuling A, Torfs P, Zappa M, Mizukami N, Clark M, Uijlenhoet R. 2016.** Representation of spatial and temporal variability in large-domain hydrological models: case study for a mesoscale pre-Alpine basin. *Hydrology and Earth System Sciences* **20**(6):2207–2226 DOI [10.5194/hess-20-2207-2016](https://doi.org/10.5194/hess-20-2207-2016).
- Mishra AK, Coulibaly P. 2009.** Developments in hydrometric network design: a review. *Reviews of Geophysics* **47**(2):RG2001.
- Muthusamy M, Schellart A, Tait S, Heuvelink GBM. 2017.** Geostatistical up-scaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models. *Hydrology and Earth System Sciences* **21**:1077–1091 DOI [10.5194/hess-21-1077-2017](https://doi.org/10.5194/hess-21-1077-2017).
- Pappenberger F, Beven KJ. 2006.** Ignorance is bliss: or seven reasons not to use uncertainty analysis. *Water Resources Research* **42**(5):W05302.
- Rakovec O, Weerts AH, Sumihar J, Uijlenhoet R. 2015.** Operational aspects of asynchronous filtering for flood forecasting. *Hydrology and Earth System Sciences* **19**(6):2911 DOI [10.5194/hess-19-2911-2015](https://doi.org/10.5194/hess-19-2911-2015).
- Refsgaard JC, Van der Sluijs JP, Højberg AL, Vanrolleghem PA. 2007.** Uncertainty in the environmental modelling process—a framework and guidance. *Environmental Modelling & Software* **22**(11):1543–1556 DOI [10.1016/j.envsoft.2007.02.004](https://doi.org/10.1016/j.envsoft.2007.02.004).
- Renard B, Kavetski D, Kuczera G, Thyer M, Franks SW. 2010.** Understanding predictive uncertainty in hydrologic modelling: the challenge of identifying input and structural errors. *Water Resources Research* **46**(5):W05521.
- Renard B, Kavetski D, Leblois E, Thyer M, Kuczera G, Franks SW. 2011.** Toward a reliable decomposition of predictive uncertainty in hydrological modelling: characterizing rainfall errors using conditional simulation. *Water Resources Research* **47**(11):W11516.

- Reusser D, Buytaert W, Vitolo C. 2017.** RHydro: Classes and methods for hydrological modelling and analysis. R package version 2014-04.1/r185. Available at <https://R-Forge.R-project.org/projects/r-hydro/>.
- Rosenthal JS. 2011.** Optimal proposal distributions and adaptive MCMC. In: *Handbook of Markov Chain Monte Carlo*. Vol. 4. Boca Raton: CRC Press.
- Samuel-Rosa A. 2017.** spsann: optimization of sample configurations using spatial simulated annealing. R package version 2.1-0. Available at <https://CRAN.R-project.org/package=spsann>.
- Schulla J, Jasper K. 2007.** *Model description WaSIM-ETH*. Zürich: Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology.
- Shrestha DL, Solomatine DP. 2008.** Data-driven approaches for estimating uncertainty in rainfall-runoff modelling. *International Journal of River Basin Management* 6(2):109–122 DOI 10.1080/15715124.2008.9635341.
- Talamba DB, Parent E, Musy A. 2010.** Bayesian multiresponse calibration of TOP-MODEL: application to the Haute-Mentue catchment, Switzerland. *Water Resources Research* 46(8):W08524.
- Ter Braak CJF. 2006.** A Markov Chain Monte Carlo version of the genetic algorithm differential evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing* 16(3):239–249 DOI 10.1007/s11222-006-8769-1.
- Terink W, Leijnse H, Van den Eertwegh G, Uijlenhoet R. 2018.** Spatial resolutions in areal rainfall estimation and their impact on hydrological simulations of a lowland catchment. *Journal of Hydrology* 563:319–335 DOI 10.1016/j.jhydrol.2018.05.045.
- Thyer M, Renard B, Kavetski D, Kuczera G, Franks SW, Srikanthan S. 2009.** Critical evaluation of parameter consistency and predictive uncertainty in hydrological modelling: a case study using Bayesian total error analysis. *Water Resources Research* 45(12):W00B14.
- Tian Y, Booi MJ, Xu Y-P. 2014.** Uncertainty in high and low flows due to model structure and parameter errors. *Stochastic Environmental Research and Risk Assessment* 28(2):319–332 DOI 10.1007/s00477-013-0751-9.
- Van der Keur P, Henriksen H-J, Refsgaard JC, Brugnach M, Pahl-Wostl C, Dewulf ARPJ, Buiteveld H. 2008.** Identification of major sources of uncertainty in current IWRM practice. Illustrated for the Rhine Basin. *Water Resources Management* 22(11):1677–1708 DOI 10.1007/s11269-008-9248-6.
- Van Groenigen JW, Stein A. 1998.** Constrained optimization of spatial sampling using continuous simulated annealing. *Journal of Environmental Quality* 27(5):1078–1086.
- Vrugt JA, Ter Braak CJF, Clark MP, Hyman JM, Robinson BA. 2008.** Treatment of input uncertainty in hydrologic modelling: doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research* 44(12):W00B09.
- Wadoux AMJC, Brus DJ, Rico-Ramirez MA, Heuvelink GBM. 2017.** Sampling design optimisation for rainfall prediction using a non-stationary geostatistical model. *Advances in Water Resources* 107:126–138 DOI 10.1016/j.advwatres.2017.06.005.
- Wadoux AMJ-C, Marchant BP, Lark RM. 2019.** Efficient sampling for geostatistical surveys. *European Journal of Soil Science* 70(5):975–989.

- Wagener T, Boyle DP, Lees MJ, Wheater HS, Gupta HV, Sorooshian S. 2001.** A framework for development and application of hydrological models. *Hydrology and Earth System Sciences* 5(1):13–26 DOI [10.5194/hess-5-13-2001](https://doi.org/10.5194/hess-5-13-2001).
- Weerts AH, El Serafy GYH. 2006.** Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. *Water Resources Research* 42(9):W09403.
- Xu H, Xu C-Y, Chen H, Zhang Z, Li L. 2013.** Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China. *Journal of Hydrology* 505:1–12 DOI [10.1016/j.jhydrol.2013.09.004](https://doi.org/10.1016/j.jhydrol.2013.09.004).
- Zeng Q, Chen H, Xu C-Y, Jie M-X, Chen J, Guo S-L, Liu J. 2018.** The effect of rain gauge density and distribution on runoff simulation using a lumped hydrological modelling approach. *Journal of Hydrology* 563:106–122 DOI [10.1016/j.jhydrol.2018.05.058](https://doi.org/10.1016/j.jhydrol.2018.05.058).